

HEC MONTRÉAL

École affiliée à l'Université de Montréal

**PARAMETER ESTIMATION AND FILTERING IN MULTIVARIATE
STOCHASTIC VOLATILITY MODELS**

par

Gabrielle Trudeau

Thèse présentée en vue de l'obtention du grade de Ph. D.
en administration (Spécialisation ingénierie financière)

Décembre 2024

© Gabrielle Trudeau, 2024

HEC MONTRÉAL

École affiliée à l'Université de Montréal

Cette thèse intitulée:

PARAMETER ESTIMATION AND FILTERING IN MULTIVARIATE STOCHASTIC VOLATILITY MODELS

Présentée par:

Gabrielle Trudeau

a été évaluée par un jury composé des personnes suivantes :

Pascal François
HEC Montréal
Président-rapporteur

Geneviève Gauthier
HEC Montréal
Directrice de recherche

Christian Dorion
HEC Montréal
Codirecteur de recherche

David Ardia
HEC Montréal
Membre du jury

Alexandre F Roch
Université du Québec à Montréal
Membre du jury

Maciej Augustyniak
Université de Montréal
Membre externe du comité conjoint

Giacomo Candian
HEC Montréal
Représentant du directeur de HEC Montréal

Résumé

Cette thèse se concentre sur les applications liées à l'estimation des paramètres des modèles stochastiques multivariés et à la filtration de la variance stochastique. De plus, les divergences de prix entre différents produits dérivés sur les marchés de l'électricité sont explorées.

Le premier essai effectue une comparaison empirique des prix de deux catégories de produits dérivés sur les marchés de l'électricité NYISO, PJM et ISO-NE : les Financial Transmission Rights (FTRs) et les contrats à terme. L'objectif est d'évaluer si ces deux catégories de produits dérivés sont évaluées de manière cohérente, étant donné que leurs paiements se chevauchent partiellement. Nous calculons des statistiques empiriques qui révèlent que les prix implicites pour la composante de perte du prix de l'électricité, obtenus en combinant les informations fournies par les prix des FTRs et des contrats à terme, sont plus volatils que les valeurs de perte réalisées correspondantes. Cela contredit l'idée qu'un prix devrait être moins volatil que le paiement et indique donc la présence d'un désalignement des prix entre les marchés FTR et les contrats à terme des trois marchés de l'électricité pour la période couverte par l'échantillon de données.

Dans le deuxième essai, nous développons une approximation analytique précise pour filtrer la variance instantanée dans un modèle à plusieurs facteurs de variance dans un cadre multi-actifs. Nous utilisons les statistiques intra-journalières, telles que la covariance et la variance réalisées, pour filtrer les quatre premiers moments conditionnels de l'état latent, assurant robustesse contre la non-linéarité et les erreurs de mesure, tout en restant efficace avec une grande dimensionnalité. Nous validons la précision du filtre par des simulations

de Monte Carlo et une analyse en panel. En outre, nous étendons son applicabilité en intégrant l'erreur de mesure des variances réalisées et proposons une correction pour les moments conditionnels à des fréquences plus basses de ces statistiques intra-journalières, offrant ainsi un estimateur caractérisant les quatre premiers cumulants de la variance, incluant les co-moments.

Le dernier essai porte sur les défis de l'estimation des paramètres dans les modèles de volatilité stochastique avec plusieurs composantes de variance et des effets de levier. Les données intra-journalières, comme la variance réalisée, sont utilisées pour cette étude avec une méthode des moments généralisés (MMG). Deux procédures sont comparées : les moments conditionnels (CM) et les moments inconditionnels (UM). L'étude constate que les méthodes CM offrent une autocorrélation plus faible dans les moments utilisés dans le MMG et une plus grande extensibilité grâce aux variables instrumentales. Cependant, elles sont plus vulnérables au bruit des récursions et aux erreurs de mesure liées à la variance réalisée. En revanche, les méthodes UM sont plus robustes, en particulier dans le modèle à deux facteurs de variance, mais sont moins flexibles car elles nécessitent des dérivations supplémentaires pour obtenir le même nombre de conditions que la méthode CM dans le MMG.

Mots clés: Marchés de l'électricité, Tarification de la congestion, Produits dérivés, NY-ISO, PJM, ISO-NE, Filtre, Modèle de volatilité stochastique multivarié, Variance réalisée, Covariance réalisée, Données intra-journalières, Méthode des moments généralisés.

Méthodes de recherche: Évaluation des produits dérivés, Filtre, Méthode des moments généralisés, Estimation des paramètres, Analyse multivariée, Analyse en panel.

Abstract

This thesis concentrates on applications related to multivariate stochastic models parameter estimation and variance filtration. In addition, it explores pricing discrepancies between products in electricity markets.

The first essay conducts an empirical comparison of prices for two categories of financial derivatives from the NYISO, PJM and ISO-NE power markets, namely Financial Transmission Right (FTR) and futures contracts. The objective is to assess whether these two categories of derivatives are priced consistently, as their payoffs partially overlap. Statistical metrics reveal that implied prices for the loss component of the power price, which are obtained by combining information provided by both FTR and futures prices, are more volatile than the corresponding realized loss values. This contradicts the idea that a price should be less volatile than the payoff and therefore indicates the presence of a pricing misalignment between the FTR and futures markets of the three power markets during the period covered by the data sample.

In the second essay, we develop a precise analytical approximation for filtering a multi-factor variance model within a multi-asset framework, specifically applied to realized variance data to estimate instantaneous variance. Our approach leverages intraday statistics, such as realized covariance, to filter the first four conditional moments of the latent state, ensuring robustness against nonlinearity and non-Gaussian measurement and state equations while remaining tractable even as the dimensionality of volatility components and assets increases. Through Monte Carlo simulations and panel analysis, we validate the fil-

ter's unbiasedness and precision. Additionally, we extend the filter's applicability beyond affine classes by incorporating measurement error of the realized variances and propose a correction for conditional moments at lower frequencies, resulting in a comprehensive estimator for the first four cumulants, including co-moments.

The final essay investigates the challenges of parameter estimation in stochastic volatility models with multiple variance components and leverage effects, using intraday data such as the realized variance. It compares two GMM procedures: conditional moments (CM) and unconditional moments (UM) for the one and two variance factor models. The study finds that CM methods offer lower autocorrelation in moment conditions and greater extensibility through instrumental variables but are more susceptible to recursion noise and measurement error. In contrast, UM methods are more robust, particularly in the two-variance factor model, but are less flexible as they require additional moment derivations.

Keywords: Electricity markets, Congestion pricing, Financial derivatives, NYISO, PJM, ISO-NE, Filter, Multivariate stochastic volatility model, Realized variance, Realized covariance, Intraday data, Generalized Methods of Moments.

Research methods: Derivative pricing, Filter, Generalized Methods of Moments, Parameter estimation, Multivariate analysis, Panel analysis.

Contents

Résumé	iii
Abstract	v
Contents	vii
List of figures	xiv
List of tables	xvi
Abbreviations	xix
Remerciements	xxii
1 Introduction	1
2 Pricing Inconsistency between the Futures and Financial Transmission Right Markets in North America	4
2.1 Introduction	5
2.2 Background on ISO-driven Electricity Markets	8
2.2.1 Day-ahead and real-time markets	8
2.2.2 Futures market	11
2.2.3 Financial Transmission Rights	13
2.2.3.1 Auction process	14
2.2.4 Data	16

2.3	Pricing the Electricity Derivatives	17
2.3.1	Contract pricing framework	18
2.3.2	Price consistency tests using zonal FTR prices	19
2.3.3	Price consistency tests using the most liquid FTR contracts	29
2.4	Conclusion	42
	References	44
3 Computationally-Efficient Variance Filtering in Multidimensional Affine Models		48
3.1	Introduction	49
3.2	The Model	54
3.2.1	Discretization	55
3.3	The Filter	59
3.3.1	An overview	59
3.3.2	Conditional moments of $\mathbf{V}_{(t+1)\Delta}$ and $\mathbf{O}_{(t+1)\Delta}$ with respect to $\mathcal{F}_{t\Delta}$	60
3.3.3	Conditional moments of $\mathbf{V}_{(t+1)\Delta}$ and $\mathbf{O}_{(t+1)\Delta}$ with respect to $\mathcal{G}_{t\Delta}$	62
3.3.4	Updating the latent variable $\mathbf{V}_{(t+1)\Delta}$ with respect to $\mathcal{G}_{(t+1)\Delta}$	63
3.4	Simulation Study	66
3.4.1	Simulation procedure	66
3.4.1.1	High-frequency	66
3.4.1.2	Intraday frequency	67
3.4.1.3	Daily frequency	67
3.4.2	Filtering error	68
3.4.3	The experiment	71
3.5	Conclusion	80
	References	81
4 Enhancing Parameter Estimation in Stochastic Volatility Models Using Intra-day Data: Generalized Method of Moments Comparative Analysis		86
4.1	Introduction	87
4.2	Theoretical Background	92

4.2.1	Continuous time model	92
4.2.2	Discrete time model	93
4.2.2.1	The returns	93
4.2.2.2	The realized variance	94
4.3	Simulation Procedure	96
4.3.1	High-frequency	96
4.3.2	Intraday frequency	97
4.3.3	Daily frequency	97
4.4	GMM Estimation with One Variance	98
4.4.1	GMM derivation	99
4.4.1.1	Without a drift and a leverage effect	99
4.4.1.2	With a drift and leverage effect	105
4.4.2	Simulation study	106
4.4.2.1	Without drift and leverage effect	108
4.4.2.2	With drift and leverage effect	110
4.5	GMM Estimation with Two Variances	123
4.5.1	GMM derivation	123
4.5.2	Simulation study	124
4.6	Conclusion	128
	References	129
5	Concluding Remarks	133
	Bibliography	135
A	Appendices of Pricing Inconsistency between the Futures and Financial Transmission Right Markets in North America	i
A.1	Supplementary Material	i
A.1.1	Price consistency tests using nodal FTR prices	i
A.1.2	Price consistency tests using the most liquid FTR contracts	ii
A.2	Stochastic Risk Premium	xiii

B Appendices of *Computationally-Efficient Variance Filtering in Multidimensional Affine Models* xix

B.1	Measurement Error	xix
B.1.1	Asymptotic theory	xix
B.1.2	Moments of $\xi_{t,t+1}$ with respect to the model filtration $\mathcal{F}_{t\Delta}$	xxi
B.1.3	Moments of $\xi_{t,t+1}$ with respect to the observed filtration $\mathcal{G}_{t\Delta}$	xxii
B.2	First two Conditional Moments	xxiii
B.3	Derivation of the Update Matrix	xxv
B.4	Third and Fourth Conditional Cumulant	xxv
B.4.1	Third cumulant	xxvi
B.4.1.1	Definitions	xxvi
B.4.1.2	Stochastic integrals	xxvii
B.4.1.3	Model components	xxviii
B.4.1.4	Latent variable update	xxix
B.4.2	Fourth cumulant	xxx
B.4.2.1	Definitions	xxxi
B.4.2.2	Stochastic integrals	xxxii
B.4.2.3	Model components	xxxiii
B.4.2.4	Latent variable update	xxxiv
B.5	Adjustment for Intraday Frequency	xxxv
B.6	Autocorrelation of Filtering Errors	xxxix
B.7	Panel Analysis	xli
B.7.1	Model	xli
B.7.2	Covariance estimates	xliii
B.7.3	Heteroskedastic errors with serial and spatial autocorrelation	xliv
	References	xlvii

C Appendices of *Enhancing Parameter Estimation in Stochastic Volatility Models*

<i>Using Intraday Data: Generalized Method of Moments Comparative Analysis</i>		xlix
C.1	Measurement Error Variance	xlix
C.2	Measurement Error Empirical Property	1

C.2.1	Model assumptions regarding the measurement error	li
C.2.2	Covariance with model components	li
C.2.3	Interaction with instrumental variables	liii
C.3	Conditional Moment Derivation	lvii
C.3.1	One-variance factor	lvii
C.3.2	Two-variance factor	lx
C.4	Unconditional Moment Derivation	lxii
C.5	Linear Regression Study	lxiii
C.5.1	One variance	lxiii
C.5.2	Two variances	lxvi
C.6	Optimization	lxx
C.6.1	GMM procedure	lxx
C.6.2	Scaling procedure	lxxii
C.6.3	Rescaling the parameters based on the frequency Δ	lxxii
C.6.4	Parameters optimization constraints	lxxiii
C.7	Supplementary Results	lxxiv
	References	lxxxii

D Online Appendices of *Computationally-Efficient Variance Filtering in Multidimensional Affine Models* lxxxiv

D.1	The Forward Price	lxxxiv
D.2	Realized Quarticity and the Measurement Error	lxxxvii
D.3	Measurement Error Additional Derivation	lxxxix
D.3.1	Derivation of $\mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}}$	lxxxix
D.3.2	Derivation of $E \left[\int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} ds \mathcal{F}_{t\Delta} \right]$	lxxxix
D.3.3	Derivation of $E \left[\int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} V_{\hat{k},s} ds \mathcal{F}_{t\Delta} \right]$	xcii
D.4	Kalman Filter Comparison	xciii
D.5	Derivation of Third and Fourth Conditional Cumulant	xcvi
D.5.1	Measurement error third cumulant with respect to the observed filtration	xcvii
D.5.2	Adding the measurement error third cumulant	xcviii

D.5.3	Measurement error fourth cumulant with respect to the observed filtration	xcix
D.5.4	Adding the measurement error fourth cumulant	ci
D.6	Higher Moments of Stochastic Integrals	cii
D.6.1	General result	cii
D.6.2	Fourth moment calculations	civ
D.7	Measurement Error under Assumption 3.4.2	cxvi
D.7.1	General results	cxvi
D.7.2	Approximations	cxviii
D.7.3	Application to second central moments	cxx
D.7.4	Application to moments of $\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1}$	cxxii
D.7.5	Application to moments of $\xi_{\hat{i},\hat{j},t,t+1}$	cxxiv
D.7.6	Application to moments of $\xi_{\tilde{i},\tilde{j},t,t+1}\xi_{\hat{i},\hat{j},t,t+1}$	cxxx
D.8	Second Update	cxxxiv
D.9	Unconditional Moments	cxxxvii
D.10	Additional Results	cxxxviii
D.10.1	Stochastic integral sample moments	cxxxviii
D.10.2	Sample correlation between the measurement error and stochastic integral	cxliii
D.10.3	Second update results under the asymptotic theory	cxliii
D.11	Sample Moments Comparison	cxlix
D.12	Filtering Error Derivation	clii
	References	clv

E Online Appendices of *Enhancing Parameter Estimation in Stochastic Volatility Models Using Intraday Data: Generalized Method of Moments Comparative Analysis* **clvi**

E.1	Risk-Neutralization Measure and the Drift	clvi
E.2	Moment Selection	clviii
E.3	Constant Measurement Error Variance	clxiii
E.4	Moments under the Model Filtration $\{\mathcal{F}_t\}_{t \geq 0}$	clxx

E.4.1	Summary	clxx
E.4.2	Proofs	clxxi
E.5	BZ Replication	clxxvii
E.6	Conditional Moment Derivation	clxxxii
E.6.1	One-variance factor	clxxxii
E.6.2	Two-variance factor	clxxxix
E.7	Unconditional Moment Derivation	cxc
E.8	Robustness Checks	cci
E.9	Background on the Generalized Methods of Moments	ccvii
E.10	Confidence Interval and Ellipse	ccix
E.10.1	Parameter average	ccix
E.10.2	Coverage ratio	ccx
E.10.3	Sargan–Hansen J-test	ccxi
	References	ccxii

List of Figures

2.1	NYISO zones map	9
2.2	ISO-NE zones map	9
2.3	PJM zones map	10
2.4	Time series of excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ in \$/MWh and their nearest-auction- implied price - NYISO	30
2.5	Time series of excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ in \$/MWh and their nearest-auction- implied price - PJM	31
2.5	Time series of excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ in \$/hour and their nearest-auction- implied price - PJM (cont.)	32
2.6	Time series of excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ in \$/MWh and their nearest-auction- implied price - ISO-NE	33
4.1	GMM results comparison - Scenario A	113
4.2	GMM results without error measurement - Scenario D	115
4.3	GMM results with error measurement - Scenario D	117
4.4	Results for specification (4.16) using CM moment selection and parameter of scenario E	120
4.5	Results for specification (4.16) using UM moment selection and parameter of scenario F	121
4.6	Results for specification (4.21) using UM moment selection and parameter of scenario G	127
A.1	Normalized Variance difference from Equation (A.3) as a function of η	xvi

B.1	Sample skewness and kurtosis of $\xi_{t,t+1}$	xxxvi
C.1	Empirical distribution of $\xi_{t,t+1}$ with parameter of scenario A	lii
C.2	Sample correlation between $\xi_{t,t+1}$, the instantaneous, integrated variance and the returns with parameter of scenario A	liv
C.3	Correlation between $\xi_{t,t+1}$ and instrumental variables	lvi
C.4	Regressions (C.13) to (C.15) R^2 comparison	lxiv
C.5	Importance of the measurement error conditional variance in the GMM moments compared to the moment residuals	lxvii
C.6	Regressions (C.16) to (C.18) R^2 comparison	lxix
D.1	Autocorrelation of $\varepsilon_{t,\ell}^{(k,n)}$ for case 2 assets and 3 variances	cxl
D.2	Sample skewness of stochastic integrals	cxli
D.3	Sample kurtosis of stochastic integrals	cxlii
D.4	Sample correlation between $\xi_{11,t,t+1}$ and stochastic integrals	cxliv
E.1	Results for specification (4.14) ($\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$) using parameters of sce- nario A where γ^2 estimated	clxviii
E.2	CM moments IV)-VI) with $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$	cciii
E.3	CM moments VII)-IX) with $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$	cciv
E.4	CM moments VII)-IX) with $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$	ccv
E.5	UM moments with $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$	ccvi

List of Tables

2.1	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - NYISO 24-hour contract	23
2.2	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - PJM on-peak contract	24
2.3	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - PJM off-peak contract	25
2.4	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - PJM 24-hour contract	26
2.5	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - ISO-NE on-peak contract	27
2.6	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ and of their nearest-auction-implied prices - ISO-NE off-peak contract	28
2.7	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - NYISO 24-hour contract	36
2.8	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM on-peak contract	37
2.9	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM off-peak contract	38
2.10	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM 24-hour contract	39
2.11	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - ISO-NE on-peak contract	40

List of tables

2.12	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - ISO-NE off-peak contract	41
3.1	Parameters for the simulation study	76
3.2	Filtering error for case 1 asset and 1 variance	77
3.3	Filtering error for case 2 assets and 2 variances	78
3.4	Filtering error for case 2 assets and 3 variances	79
4.1	GMM moment selection	102
4.2	Additional moments for specification (4.20)	106
4.3	Parameters for the simulation study	107
4.4	Comparison results for parameter set E for specifications (4.14)-(4.16) . .	118
4.5	Comparison results for parameter set F for specification (4.20)	122
4.6	GMM moments selection for specification (4.21)	124
A.1	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - NYISO 24-hour contract	iii
A.2	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM on-peak contract	iv
A.3	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM off-peak contract	v
A.4	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM 24-hour contract	vi
A.5	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - ISO-NE on-peak contract	vii
A.6	Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - ISO-NE off-peak contract	viii
A.7	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their earlier-auction-implied prices for the most liquid FTRs - NYISO 24-hour contract	ix
A.8	Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their earlier-auction-implied prices for the most liquid FTRs - PJM on-peak contract	x

A.9 Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their earlier-auction- implied prices for the most liquid FTRs - PJM off-peak contract	xi
A.10 Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their earlier-auction- implied prices for the most liquid FTRs - PJM 24-hour contract	xii
C.1 Comparison results for parameter set A for specifications (4.14)-(4.16) . .	lxxvi
C.2 Comparison results for parameter set B for specifications (4.14)-(4.16) . .	lxxvii
C.3 Comparison results for parameter set C for specifications (4.14)-(4.16) . .	lxxviii
C.4 Comparison results for parameter set A for specification (4.20)	lxxix
C.5 Comparison results for parameter set D for specification (4.20)	lxxx
C.6 Comparison results for parameter set E for specification (4.20)	lxxxix
D.1 Filtering error for case 2 assets and 3 variances ($N = 390$)	cxxxix
D.2 Filtering error for case 1 asset and 1 variance under Assumption 3.4.1 . .	cxlvi
D.3 Filtering error for case 2 assets and 2 variances under Assumption 3.4.1 .	cxlvii
D.4 Filtering error for case 2 assets and 3 variances under Assumption 3.4.1 .	cxlviii
D.5 Sample moments comparison for case 2 assets and 3 variances	cli
E.1 GMM moment selection	clx
E.2 GMM comparison results for specification (4.14): $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$	clxi
E.3 GMM comparison results for specification (4.16): $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$	clxii
E.4 GMM estimations results for specification (4.14) ($RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} +$ $\gamma \epsilon_{t,t+1}^{(N)}$) with BZ moment selection	clxv
E.5 GMM estimations results for specification (4.14) ($RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} +$ $\gamma \epsilon_{t,t+1}^{(N)}$) with CM1 moment selection	clxvi
E.6 Theoretical and empirical expected value for the measurement error γ^2 related to specification (4.14)	clxix
E.7 Table 1 from Bollerslev and Zhou, 2002	clxxviii
E.8 Replication of Table 1 from Bollerslev and Zhou, 2002	clxxix
E.9 Table 2 from Bollerslev and Zhou, 2002	clxxx
E.10 Replication of Table 2 from Bollerslev and Zhou, 2002	clxxxix

Abbreviations

AML	Approximate Maximum Likelihood
APF	Auxiliary Particle Filter
CfD	Contracts for Difference
CIR	Cox-Ingersoll-Ross
CM	Conditional moment
CRR	Congestion Revenue Rights
DA	Day-Ahead
DAM	Day-Ahead Market
DFM	Dynamic Factor Model
EKF	Extended Kalman Filter
FTR	Financial Transmission Right
GMM	Generalized methods of moments
HAC	Heteroskedasticity and autocorrelation consistent
HAR	Heterogeneous Autoregressive
HARQ	Heterogeneous Autoregressive Quarticity
HARK	Heterogeneous Autoregressive Kalman
ICE	Intercontinental Exchange
ISO	Independent System Operators
ISO-NE	Independent System Operators New England
KF	Kalman Filter
LBMP	Locational-Based Marginal Pricing
LMP	Locational Marginal Price

MCMC	Markov Chain Monte Carlo
MC	Monte Carlo
MCP	Market Clearing Price
MMSE	Minimizing the Mean Squared Error
MW	Megawatt
MWh	Megawatt-Hour
NYISO	New York Independent System Operator
OLS	Ordinary Least Square
OPF	Optimal Power Flow
PCA	Principal Component Analysis
PF	Particle Filter
PJM	Pennsylvania-New Jersey-Maryland
POI	Point of Injection
POW	Point of Withdrawal
QCV	Quadratic Covariance
QKF	Quadratic Kalman Filter
QV	Quadratic Variance
RCV	Realized Covariance
RMSE	Root Mean Square Error
RQ	Realized Quarticity
RSV	Realized stochastic volatility
RT	Real-Time
RTM	Real-Time Market
RTO	Regional Transmission Organization
RV	Realized Variance
SI	Sink
SO	Source
TCC	Transmission Congestion Contracts
UKF	Unscented Kalman Filter
UM	Unconditional moment

Abbreviations

VAR Vector Autoregression

Remerciements

J'aimerais exprimer ma profonde gratitude à mes directeurs de thèse, Geneviève Gauthier et Christian Dorion, pour leur accompagnement tout au long de ce parcours et pour m'avoir constamment poussé à approfondir mon raisonnement. Leurs expertises et conseils ont été essentiels à la réalisation de ce projet. Je tiens également à remercier Frédéric Godin d'avoir accepté de collaborer avec nous sur un projet en lien avec l'électricité, un domaine particulièrement fascinant, ainsi qu'à Nathalie Bilodeau pour son tout son aide et son écoute à travers ce parcours.

Je remercie chaleureusement toute ma famille et mes amis, mais plus particulièrement mon conjoint, Louis, dont le soutien incommensurable m'a permis de maintenir la discipline nécessaire pour franchir la ligne d'arrivée aujourd'hui. Je tiens aussi à exprimer ma reconnaissance à ma mère, qui a toujours été d'un soutien inconditionnel dans mon parcours académique, et à qui je dois une très grande partie de mes réussites. Enfin, je remercie mes amis Rémi, Samuel, Mamadou, Isabelle, Laurence, Lydia, Gabrielle, Jacquie, Eva, Ella et Delphine pour leur écoute, soutien et encouragements constants.

Chapter 1

Introduction

Accurately filtering and estimating stochastic volatility remains a key challenge due to the complex nature of financial markets and the limitations of available data. Volatility modeling is important for understanding asset price movements, managing risk, and pricing financial derivatives. This thesis explores methods for volatility filtering and estimation, focusing particularly on multiple latent factors.

The concept of commonality in financial markets has been recognized for some time and is even more pronounced during significant events like the 2008 financial crisis and the COVID-19 pandemic. This increased interconnectedness extends to various asset classes, including commodities. Understanding these shared factors is increasingly important during such periods of high commonality.

Isolating these common volatility factors is vital for practical financial applications, such as portfolio management. The last two essays of this thesis focus on developing methods to filter out these common factors and improve the parameter estimations of these models.

The first essay is on ISO-driven power markets like those in New York (NYISO), ISO New England (ISO-NE), and PJM, where electricity prices fluctuate significantly across zones, with congestion costs often spiking during periods of high demand. Financial Transmission Rights (FTRs) are critical instruments for hedging against these congestion risks, providing payouts linked to the price differences between grid nodes. While FTRs are ef-

fective hedging tools, their auction-based allocation and embedded constraints make them less accessible or transparent than traditional derivatives like futures contracts.

This study examines the pricing consistency between FTRs and futures contracts in these markets. Since FTRs cover only the congestion component, while futures contracts encompass all three components of the energy price (energy, loss, and congestion), we first compare the payoffs of FTRs and futures contracts. By also subtracting the average price of the derivatives, which accounts for the energy component that is consistent across the market, we succeed in isolating the implied price of the excess loss component. This approach allows us to compare the implied prices with the realized prices of the loss component across various zones, seasons and markets. Based on these comparisons, our empirical analysis reveals significant pricing discrepancies, highlighting inefficiencies and inconsistencies in these derivative markets.

The second essay focuses on developing a filter for multi-asset and multi-factor variance analysis. The presence of commonality in asset volatilities is well-documented, showing that not only do stocks within a market share systematic risk, but other markets, such as commodities. To address this, we propose a filtering method that incorporates multiple latent variance factors within a multi-asset framework. This new analytical filter overcomes the curse of dimensionality and effectively manages nonlinear frameworks, providing precise approximations for the first four conditional moments of latent variance components.

While stochastic volatility models and their extensions are widely used for single-asset analyses, their practical application for parameter estimation and filtering to multi-asset frameworks faces challenges. Our paper builds upon existing literature in analytical filtering to offer a fast and practical solution for multivariate filtering problems. Our approach uses intraday data and two types of observable variables, daily forward log-returns and realized variances and covariances, to identify latent factors driving common volatility movements.

This methodology also includes a correction for conditional moments of intraday statistics at lower frequencies when they have not yet reached their asymptotic distributions. The filter extends beyond the affine class, as the conditional moments of the measurement errors of realized variances are not linear in the instantaneous variance, the latent state.

Furthermore, our filter calculates the first four conditional moments of observables and latent variables, offering a comprehensive analysis of instantaneous variance components. Monte Carlo simulations validate the precision and robustness of our filter, demonstrating its effectiveness in managing increased dimensionality and non-Gaussian frameworks.

The third essay compares two Generalized Method of Moments (GMM) procedures for stochastic volatility models with one or two variance components: the conditional moments (CM) method and the unconditional moments (UM) method. Given the challenge of filtering latent factors, estimating parameters in multi-variance factor models is also a complex task. This chapter aims to enhance our understanding of GMM by using intraday data.

Our study uses realized variance, which converges to quadratic variation, as commonly done in the literature to estimate variance parameters. We compare the CM approach, as outlined by Bollerslev and Zhou, 2002, with the UM approach through Monte Carlo simulations.

We find that the CM method generally performs better when measurement error around realized variance is low, leading to more accurate estimates and lower residual autocorrelation in the moment conditions used in the GMM. However, it struggles with high measurement error and in multi-factor models. The UM method is more robust in these situations but is more complex to extend because it cannot use instrumental variables as in the CM method. Consequently, achieving the same number of conditions with UM requires deriving additional moments.

Chapter 2

Pricing Inconsistency between the Futures and Financial Transmission Right Markets in North America

Abstract¹

An empirical comparison of prices for two categories of financial derivatives of the NY-ISO, PJM and ISO-NE power markets, namely Financial Transmission Right (FTR) and futures contracts, is performed. The objective is to assess whether these two categories of derivatives are priced consistently, as their payoffs partially overlap. Statistical metrics reveal that implied prices for the loss component of the power price, which are obtained by combining information provided by both FTR and futures prices, are more volatile than corresponding realized loss values. This contradicts the idea of a price being a probability-weighted average of possible realized values, and therefore indicates the presence of the pricing misalignment between the FTR and futures markets of the three power markets during the period covered by the data sample.

¹Joint work with Geneviève Gauthier and Frédéric Godin. Gauthier is affiliated with HEC Montréal and Godin is affiliated with Concordia university. This chapter has been published in Energy Economics.

2.1 Introduction

In power markets driven by Independent System Operators (ISO), power prices vary substantially throughout their various zones and are subject to spikes, especially when demand is high. Three of such ISO-driven power markets are considered in this study: New York (NYISO), ISO New England (ISO-NE) and PJM. They all publish the power prices and their decomposition into three components: energy cost, thermal loss cost and congestion cost. When electricity demand is high in some of the zones, bottlenecks and capacity constraints on the power grid can make congestion costs explode. Such scenarios create hedging needs for market participants seeking to protect themselves from such congestion cost surges. A Financial Transmission Right (FTR)² provides a payoff equivalent to the congestion spread (congestion price difference) between two nodes of the power grid, making it a suitable instrument for hedging congestion risk, as highlighted in Deng and Oren, 2006. Lyons et al., 2000 argue that FTRs are a superior solution to Physical Transmission Rights, which provide exclusive rights to carry electricity between two nodes of a grid.

FTRs and congestion management have attracted some attention in the literature, as demonstrated by the survey papers of Kristiansen, 2005, Kumar et al., 2005, Sarkar and Khaparde, 2008 and Singla et al., 2014. Congestion-related contracts respond to the need of having to price inter-nodal risk, an idea first put forward in Hogan, 1992. Such derivatives are now well integrated into the operations of multiple power markets. These contracts are sold by the market operator via auction and are funded by congestion revenues, as explained in Alsac et al., 2004. Revenue adequacy is obtained if the contract allocation satisfies the network constraints, as detailed in Hogan, 2002, that is, the allocation is simultaneously feasible, therefore allowing the FTRs to be fully funded, as discussed in Hogan et al., 2013. While the purpose of such financial products is to provide hedging instruments for market participants, speculators may also profit from them while facilitating price discovery and enhancing market efficiency and liquidity. For instance, Adamson and Parker, 2013 show

²These contracts are also referred to by other names in the various electricity markets, such as Transmission Congestion Contracts (TCC) in the NYISO, Contracts for Difference (CfD) in Nord Pool, and Congestion Revenue Rights (CRR) in the California ISO.

that financial firms account for an increasing proportion of transactions in the NYISO's FTR market. Celebi et al., 2010, Ledgerwood and Pfeifenberger, 2013 and Prete et al., 2018 mention that virtual traders can also participate in transactions involving arbitrage with FTRs, which can unfortunately lead to cross-market manipulations.

The auction mechanism underlying the allocation of FTRs by the ISO, along with particular embedded allocation constraints, make FTRs harder to access than other power-related financial derivatives. Indeed, conventional derivatives such as futures and options are also traded in electricity markets, often through an exchange. The discrepancy between the trading procedure for FTRs and for conventional derivatives raises the question as to whether the pricing of FTRs is consistent with that of other electricity derivatives. This study investigates this topic by comparing prices observed for two categories of electricity derivatives, namely FTRs and futures contracts. To the best of our knowledge, no article in the existing literature studies consistency between the FTR and futures markets. It is precisely at this level that our contribution comes into play.

In other literature peripheral to our research question, several papers study the consistency between the FTRs' contract price and their realized payoff, as the price discovery mechanism and liquidity issues are characteristics that could potentially cause inefficiencies. Such literature investigates the distinct topic of standalone FTR market efficiency instead of considering consistency across derivative types as we do. Bartholomew et al., 2003, Adamson and Englander, 2005 and Siddiqui et al., 2005 question the efficiency of the TCC auction process in the NYISO through empirical studies or models. The TCC auction process could benefit from additional competition, which could increase its efficiency, as discussed by Hadsell and Shawky, 2009. The subject has also been studied by Deng et al., 2010 who argue that price discovery alone cannot explain the difference between FTR auction prices and associated realized revenues in the NYISO, and that auction mechanisms produce inefficiencies. The most recent research paper by Opgrand et al., 2022 on FTRs in PJM establishes that this difference is caused by a trading premium rather than by flaws in the auction mechanism. Zhang, 2009 reports significant underpricing of the TCCs relatively to their realized revenues in the early 2000s for the NYISO. He explains that this mispricing comes mostly from asymmetry of information and from risk aversion

among bidders. Adamson et al., 2010 argue that efficiency in the NYISO FTR market has improved over time. Through an ex-ante analysis using VAR models, Mount and Ju, 2014 found no proof of underpricing in six-month TCCs for the specific summer 2006 period. Leslie, 2021 asserts that a substantial portion of profits are carved out by financial traders in the TCC auction market placing bids on less liquid nodal TCCs, since zonal TCCs are more competitively priced. They explain that financial traders earn profit while providing price discovery and liquidity. However, the TCC prices converge to a fair price as the several rounds of auctions occur. The efficiency of the FTR auction process has also been challenged in other markets such as in Ontario, where systematic price biases are identified by Olmstead, 2018, or the California market where Baltaduonis et al., 2017 observe persistent abnormal return on FTRs.

Both futures contracts and FTRs provide payoffs related to future electricity prices. However, whereas FTRs cover only the price's congestion component, the payoff of futures is linked to the entire price, which also includes the energy price and the loss price components. Thus, the overlap in payoffs is only partial. Nevertheless, considering suitable pricing differentials between the various zones makes it possible to cancel out the implied pricing of the energy component, and thus deduce implied prices for the loss component when comparing FTR and futures prices. The basis of the analysis described in this paper therefore lies in studying whether implied prices for the loss component are commensurate with their realized values. In our empirical study, we compare the variance of the implied prices and the realized values of the loss component for all three power markets over various seasonal buckets. If the futures and FTRs are priced consistently, the variance of implied prices should be lower than that of realized values. Our empirical experiments highlight that this inequality is violated for 271 of the 272 zone/seasonal bucket combinations studied, which clearly indicates the presence of pricing inconsistencies between the futures and FTR markets.

This paper is subdivided as follows. Section 2 presents background information on electricity prices and associated financial derivatives in ISO-driven markets in North America. Section 3 explains and performs several tests to assess whether FTRs and futures contracts are priced consistently in the NYISO, the ISO-NE and PJM. Section 4 concludes.

2.2 Background on ISO-driven Electricity Markets

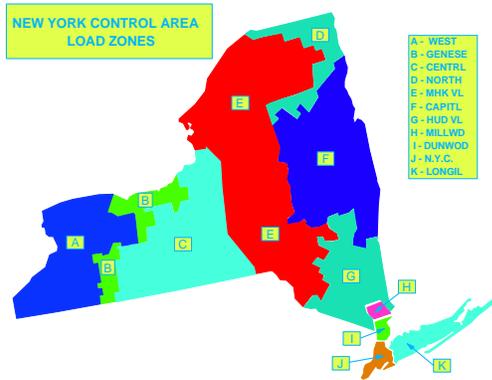
The NYISO and ISO-NE are the ISOs responsible for electricity flow-related operations in New York State and New England, respectively. PJM is a Regional Transmission Organization (RTO) which manages electricity-related operations in an area covering thirteen states in the eastern United States, as well as the District of Columbia. ISOs and RTOs serve similar purposes, with an RTO covering more than a single state. Responsibilities of an ISO (or an RTO) include determining electricity prices by periodically conducting auctions. This involves collecting supply and demand bids from the market's participants, and then setting prices by equating supply and demand in the most economical way by means of an optimization procedure, while ensuring the security and integrity of the electrical grid. Various types of financial derivatives have been developed for the ISO-driven power market. The payoff of such derivatives typically depends on electricity prices determined by the ISO. The two main classes of derivatives considered in this study are electricity futures contracts and FTRs. This section provides a brief overview of the price-setting procedure used by ISO-driven markets and the functioning of related financial derivatives.

2.2.1 Day-ahead and real-time markets

The ISO publicly reports electricity prices using a so-called Locational Marginal Price (LMP)³ methodology in two main markets referred to as the real-time market (RTM) and the day-ahead market (DAM). The LMP approach is meant to reflect the marginal cost of consumption of an additional megawatt-hour (MWh) of power on a given period and at a given node of the transmission grid. The DAM, which represents the bulk of electricity transactions, fixes the price of electricity one day ahead based on expected consumption and production. Conversely, the RTM acts as a balancing market addressing real-time deviations in consumption from planned volumes in the DAM. The LMP is published every five minutes in the RTM and for each hour in the DAM, for every node and zone of the power grid. This paper focuses on the day-ahead (DA) prices, because the payoff of FTRs is related to the DA prices and not real-time (RT) prices. Moreover, since futures are mainly

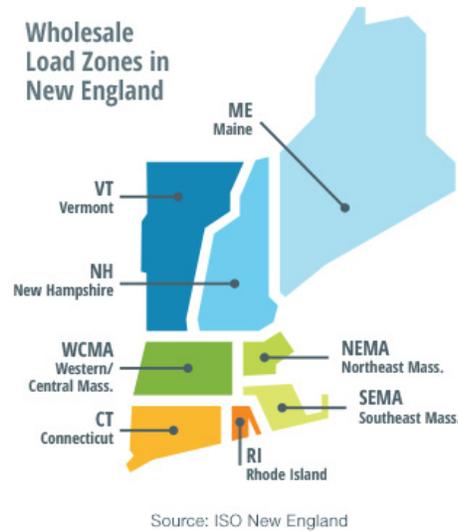
³In the NYISO, it is referred to as the Locational Based Marginal Price (LBMP).

available for zones and not for more granular nodes, only zonal LMPs are considered. Maps illustrating the various zones are provided in Figure 2.1 for the NYISO, in Figure 2.2 for the ISO-NE and in Figure 2.3 for PJM.



Source: <https://www.nyiso.com>.

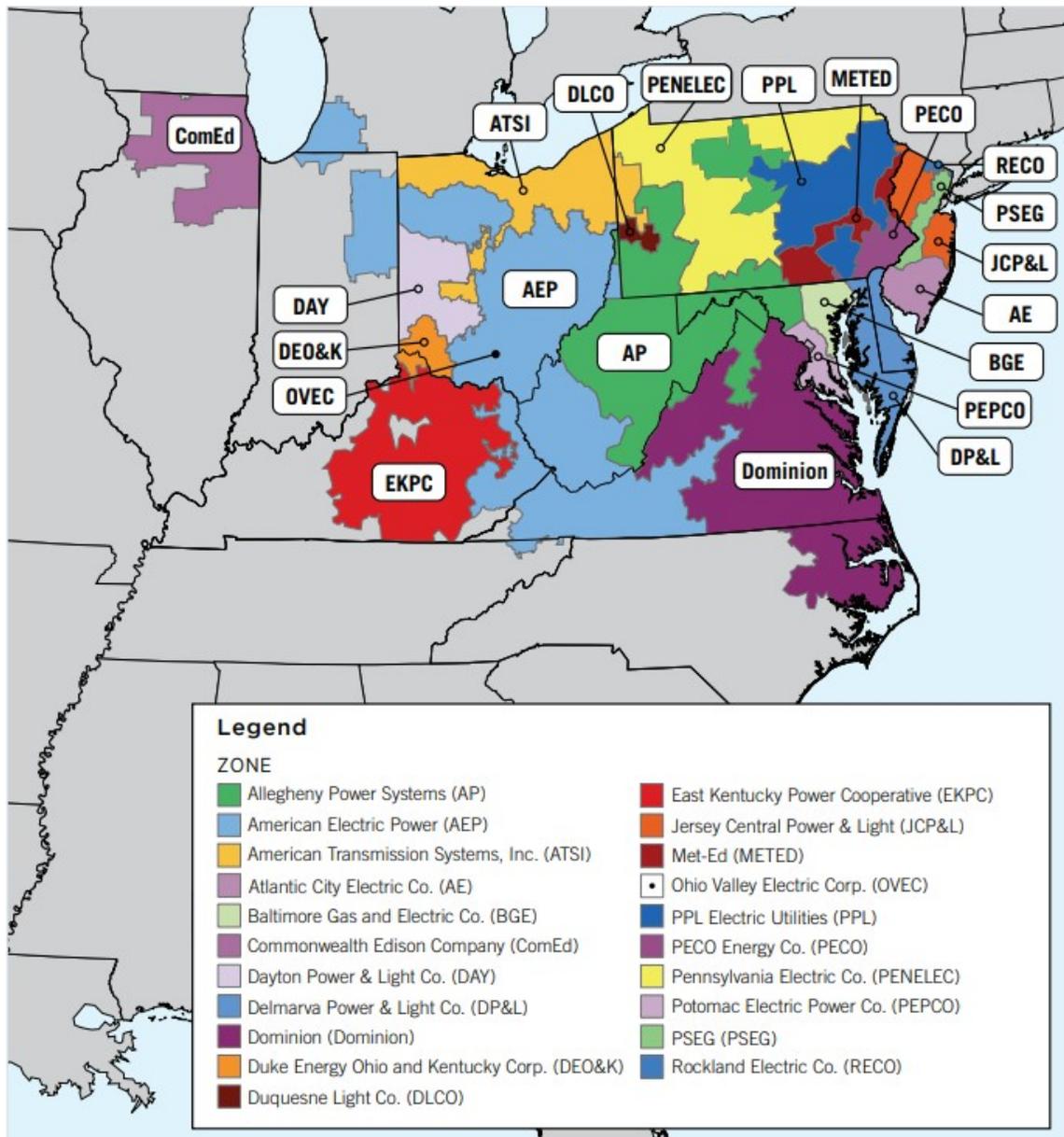
Figure 2.1: NYISO zones map



Source: <https://www.iso-ne.com/about/key-stats/maps-and-diagrams/>.

Figure 2.2: ISO-NE zones map

The LMP, in \$/MWh, is decomposed into three components when reported by the ISO. The first component is the energy price, which is the same for all nodes of the grid at any point in time. The second one is the energy loss price, which is a price markup resulting from thermal losses occurring during transmission of power on the grid. The third is the congestion component representing the extra cost stemming from the physical constraints of the power grid. For example, consumers located in high-demand locations may need to buy energy at a higher price from local generators if buying cheaper energy from generators in other areas of the grid proves impossible due to transmission bottlenecks, i.e. the lines are at full transmission capacity. The DA LMP for zone i and hour h and its three



Source: <https://www.pjm.com/library/maps>. In the data retrieved from the PJM website, some zones have an alternative associated abbreviation: Atlantic City Electric Company (AECO) and Duke Energy Ohio and Kentucky (DEOK).

Figure 2.3: PJM zones map

components are denoted by⁴

$$\text{LMP}_{h,i}^{\text{DA}} = \text{Energy}_h^{\text{DA}} + \text{Loss}_{h,i}^{\text{DA}} + \text{Congestion}_{h,i}^{\text{DA}}. \quad (2.1)$$

2.2.2 Futures market

A futures contract is a contract between two counterparties stipulating that the holder of the long position must buy the underlying asset from the holder of the short position at a certain price (the futures price) and a specified time in the future. To mitigate counterparty risk, futures contracts are typically marked-to-market daily by the exchange's clearinghouse where they are traded. The clearinghouse calculates required margins for both long and short positions, and daily compensation, i.e. the transfer of funds between the two accounts, is performed to reflect fluctuations of futures prices on the market.

A futures market is available on the ICE exchange for the New York, New England and PJM power markets. Although a variety of contracts are available, the analysis is focused on monthly futures contracts in the DA market. The underlying asset of such a contract is the monthly average electricity price for a given zone. In the NYISO, there are futures contracts that are traded for all zones, with the exception of zone H (Milwood). In the ISO-NE, only zones VT (Vermont) and CT (Connecticut) do not have traded futures contracts. Lastly, in PJM, all zones have traded futures contracts with the exception of PEPCO, RECO, EKPC and OVEC.

Two different types of futures contracts are traded to separate high energy demand hours from low ones. The first type is the on-peak hours futures contract, which covers the hours from 8 a.m. to 11 p.m., Monday to Friday, for the entire month of the contract. Conversely, the off-peak hours futures contract covers all remaining hours of the month. On-peak futures contracts are typically more liquid than off-peak contracts. For a contract size of 1 MW, the payoff of an on-peak hours ($\mathcal{T} = \text{on}$) or an off-peak hours ($\mathcal{T} = \text{off}$) futures contract is the arithmetic average of the DA price observed over the set of all covered hours $\mathcal{H}^{\mathcal{T}}$, $\mathcal{T} \in \{\text{on}, \text{off}\}$. The combined synthetic and not traded futures contract ($\mathcal{T} = \text{all}$)

⁴For the NYISO market, a slightly different convention is used, with the sign in front of the congestion component being reversed: $\text{LBMP}_{h,i}^{\text{DA}} = \text{Energy}_h^{\text{DA}} + \text{Loss}_{h,i}^{\text{DA}} + (-\text{Congestion}_{h,i}^{\text{DA}})$.

combines both on-peak and off-peak hours, thereby covering all hours in a given month: $\mathcal{H}^{\text{all}} = \mathcal{H}^{\text{on}} \cup \mathcal{H}^{\text{off}}$. For the NYISO and PJM, the construction of synthetic contracts covering period $\mathcal{T} \in \{\text{all}\}$ is performed since some of the FTRs apply to all hours rather than only to on-peak or off-peak hours separately. Therefore, for any zone i and contract type $\mathcal{T} \in \{\text{all, on, off}\}$, the futures payoff can be expressed as

$$\text{Futures payoff}_{i,\mathcal{H}\mathcal{T}} = \frac{1}{|\mathcal{H}\mathcal{T}|} \sum_{h \in \mathcal{H}\mathcal{T}} \text{LMP}_{h,i}^{\text{DA}}, \quad (2.2)$$

where $|\mathcal{H}|$ denotes the number of hours in set \mathcal{H} .⁵ Using the notation⁶

$$\begin{aligned} E_{\mathcal{H}} &\equiv \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} \text{Energy}_h^{\text{DA}}, & C_{i,\mathcal{H}} &\equiv \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} \text{Congestion}_{h,i}^{\text{DA}}, \\ L_{i,\mathcal{H}} &\equiv \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} \text{Loss}_{h,i}^{\text{DA}}, \end{aligned} \quad (2.3)$$

for the average energy, congestion and loss components of the LMP, the futures payoff can further be decomposed based on (2.1) as

$$\text{Futures payoff}_{i,\mathcal{H}\mathcal{T}} = E_{\mathcal{H}\mathcal{T}} + L_{i,\mathcal{H}\mathcal{T}} + C_{i,\mathcal{H}\mathcal{T}}. \quad (2.4)$$

Denote by $F_{t,i,\mathcal{H}\mathcal{T}}$ the day- t closing price of the futures of a type- \mathcal{T} contract for zone i . The combination of on-peak and off-peak futures contracts weighted with their respective number of hours results in a monthly futures contract covering all hours in a given month.⁷ Thus, the combined futures price must be the following weighted average of the on-peak

⁵Due to the short time horizons considered for the futures contracts, the time value of the money is not considered here because it is immaterial in such a context.

⁶Due to the alternative reporting convention in the NYISO, the congestion portion of the payoff is given by $C_{i,\mathcal{H}} \equiv \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} (-\text{Congestion}_{h,i}^{\text{DA}})$ for that market.

⁷The number of on-peak and off-peak hours per month along with the NERC holidays can be found at <https://www.energygps.com/HomeTools/PowerCalendar>

and off-peak futures price:

$$F_{t,i,\mathcal{H}^{\text{all}}} = \frac{|\mathcal{H}^{\text{on}}|}{|\mathcal{H}^{\text{all}}|} F_{t,i,\mathcal{H}^{\text{on}}} + \frac{|\mathcal{H}^{\text{off}}|}{|\mathcal{H}^{\text{all}}|} F_{t,i,\mathcal{H}^{\text{off}}}. \quad (2.5)$$

Since the payoff of some of the FTRs cover all hours in a month, the combined off-peak and on-peak futures contracts are also considered in this study. In what follows, the *all*, *on* or *off* superscripts are dropped to make the notation more parsimonious.

2.2.3 Financial Transmission Rights

An FTR is a contract paying the day-ahead congestion cost spread, i.e. the difference between two points of reference⁸ cumulated over all hours of the associated duration. Such a duration can be 1, 6, 12 or 24 month, depending on the market. The holder of an FTR receives the congestion rent which is defined as the difference between the DA congestion cost at the sink (SI) node and the DA congestion cost at the source (SO) node.⁹ The congestion¹⁰ rent payment per contract for hour h is represented by

$$\text{Congestion rent payment}_{h,SO,SI} = \text{Congestion}_{h,SI}^{\text{DA}} - \text{Congestion}_{h,SO}^{\text{DA}}. \quad (2.6)$$

If the DA congestion is higher at the SI than at the SO, the FTR holder receives the difference for hour h . Contrarily, if the DA congestion is lower at the SI than at the SO, the FTR holder will have to pay the difference for that particular hour. The total rent is the sum of rent payments for all hours in the duration of the contract. FTRs are useful for hedging against price fluctuations between two zones, particularly when a participant expects congestion. By securing compensation based on price differences, FTRs help ensure predictable pricing and protect against unexpected price spikes caused by congestion.

An FTR's payoff is defined as the sum of congestion rent payments over its contract

⁸These can be between two zones, two nodes or a node and a zone.

⁹In the NYISO, the source and sink nodes are called the Point of Injection (POI) and the Point of Withdrawal (POW), respectively.

¹⁰For the NYISO market, because the congestion component has a reversed sign, we have $\text{Congestion rent payment}_{h,POI,POW} = (-\text{Congestion}_{h,POW}^{\text{DA}}) - (-\text{Congestion}_{h,POI}^{\text{DA}})$.

duration, where \mathcal{H} includes all hours of the specified duration. Therefore, an FTR from SO i to SI j has the following average payoff per hour:

$$\text{FTR payoff}_{i,j,\mathcal{H}} = C_{j,\mathcal{H}} - C_{i,\mathcal{H}}. \quad (2.7)$$

The nodal payoff of zone j is represented by setting $i = \emptyset$, with $C_{\emptyset,\mathcal{H}} = 0$.¹¹ Since futures are available only for zonal nodes and have a monthly duration, monthly zonal FTR price data are considered for this study. The day- $t_{\mathcal{H}}$ FTR price with SO i and SI j covering period \mathcal{H} is denoted by $\text{FTR}_{t_{\mathcal{H}},i,j,\mathcal{H}}$, where $t_{\mathcal{H}}$ is the day on which auction results for FTR contracts covering period \mathcal{H} are published.

2.2.3.1 Auction process

Unlike futures contracts, the FTRs are not traded in an exchange, but rather, are allocated through an auction process specific to each market. We present in-depth explanations for the auction process of the NYISO, and provide additional details for those of PJM and ISO-NE.

In the NYISO, there are two types of auctions in the TCC market. The first is the Centralized TCC Auction where FTRs with durations of six months and more are awarded through multiple rounds. These auctions take place twice a year. The second type of auction is the Balance-of-Period Auction, a monthly auction where only monthly FTRs are awarded. Before September 2017, only the FTRs for the month following the auction were traded. FTRs can now be traded up to six months in advance in such auctions, depending on the time of year. In this study, the *nearest auctions* are defined as the auctions closest to the contract month. Only results for such auctions are presented in the main text, with some results for earlier auctions being displayed in the Online Appendix.

To participate in the auction, market participants must submit their bids during a specified period. Then, on the day the NYISO publishes the results, participants are informed of whether their bids were retained and how many FTRs have been allocated to them. The bidding period is thus known in advance and is typically from a Friday at 8 a.m. to the

¹¹For the NYISO, \emptyset is called the NYISO LBMP Reference node.

following Monday at 5 p.m. for a monthly auction. The NYISO requires several days to calculate the Market Clearing Price (MCP) for each possible FTR. The results are then published on the FTR Automation site no later than 9 p.m. on the auction posting day.

The MCP for FTR prices is determined by the market operator through an optimization process involving the bids placed by market participants, taking into account physical constraints on the power grid. Once the bids and offers are received, which must be by the end of the bidding period as established in the schedule found in Attachment D of the NYISO website, the market operator validates the bids and offers and performs an Optimal Power Flow (OPF) analysis. The goal of this optimization process is to maximize the benefit to the market participants, who are the buyers of FTRs, while respecting the physical constraints and capacity of the power grid.¹²

One example of price constraints applied during the optimization is additive consistency; the price for an FTR between zone A (SO) and zone B (SI) should be the same as the sum of prices for the FTR from A to C and for the FTR from C to B. Such constraints allow NYISO to report FTR MCPs for standalone nodes (referred to as nodal prices) rather than for pairs. The price of an FTR associated with a node pair can then simply be obtained by subtracting the nodal MCP of the SO from that of the SI, ensuring that the additive consistency constraint is respected.

As detailed in Ma et al., 2003, the FTRs from the PJM market are also allocated through an auction bidding process. The *Monthly FTR Auctions* allow market participants to bid on residual FTR capability from the *Long-term FTR Auction* and the *Annual FTR Auction*. In this single-round auction, market participants may also offer FTRs that they already own. PJM has its own calendar¹³ for the bidding period and the results posting day. FTR instruments that are offered can take the form of obligations (as in the NYISO) or options. Another difference is that FTR contracts are also available separately for the on-peak and off-peak hours, in addition to the 24-hour contracts being offered.¹⁴ The ISO-NE market also has an auction bidding process with two rounds for annual FTRs and a monthly auction

¹²Additional details can be found in NYISO, 2021.

¹³<https://www.pjm.com/markets-and-operations/fttr>

¹⁴The split of hours into the on-peak and off-peak groups is the same as for standard futures contracts.

based on its own schedule.¹⁵ The market only offers obligation FTRs for on-peak or off-peak hours.

2.2.4 Data

Three datasets are considered in this study: FTR prices and the corresponding allocated volume from the monthly auctions, daily closing prices of monthly futures and historical realized values of the DAM LMP with its decomposition (2.1).

FTR prices and the corresponding allocated volume data are publicly available and informs the market participants of the actual quantity of FTRs that have been awarded for each SO and SI combination in each auction.¹⁶ The standalone nodes data are also available for all nodes on the grid even if FTR contracts are not necessarily ultimately awarded for each possible pair of nodes.¹⁷ For all three power markets, both nodal price and awarded quantities data¹⁸ are available from the starting point of our study, January 2012.

Quandl, a data supplier that sources its data from ICE, provides the daily prices of zonal electricity futures. Several contracts are available, but only the monthly futures contracts (on-peak and off-peak) on the DAM are used herein. We restrict our sample to FTR auction posting days $t_{\mathcal{H}}$ for all months \mathcal{H} . For the remainder of this paper, we denote by \mathcal{Z}_t the set of zones for which futures price data are available at day t . This set changes over time as futures availability depends on the power market and zone. For NYISO, futures in zones $\{A, C, F, G, J\}$ have been available since January 3, 2012, futures in zone D since June 1, 2015, and futures in zones $\{B, E, I, K\}$ since June 12, 2017. For PJM, futures are not

¹⁵<https://www.iso-ne.com/markets-operations/markets/financial-transmission-rights/#related-documents>

¹⁶*Summary of Awards* for NYISO at http://tcc.nyiso.com/tcc/public/view_awards_summary.do, *Historical FTR Monthly Auction Results* for PJM at <https://www.pjm.com/markets-and-operations/fttr/historical-fttr-auction> and *FTR Auction Results* for ISO-NE at <https://www.iso-ne.com/isoexpress/web/reports/auctions/-/tree/auction-results-fttr>

¹⁷*Nodal Prices report* for NYISO at http://tcc.nyiso.com/tcc/public/view_nodal_prices.do, PJM at <https://www.pjm.com/markets-and-operations/fttr/historical-fttr-auction> and *FTR Auction Clearing Prices* for ISO-NE from <https://www.iso-ne.com/isoexpress/web/reports/auctions/-/tree/auction-results-fttr>.

¹⁸In PJM, from October 2022 on, the off-peak contract is traded as two separate contracts: DailyOffPeak and WkndOnPeak. The off-peak nodal prices are still provided, but the awarded FTR for the off-peak contract is no longer available after that auction month.

available for the 4 zones PEPCO, RECO, EKPC and OVEC. Futures data start in January 2015 for the rest of the zones, with the exception of DAY and DOM where they start in July 2017 and March 2017, respectively. For ISO-NE, daily closing prices of monthly on-peak and off-peak futures are available for all zones with the exception of Vermont (VT) and Connecticut (CT). Maine (ME), New Hampshire (NH), Northeast Massachusetts (NEMA), Southeast Massachusetts (SEMA) and Western/Central Massachusetts (WCMA) futures price data are available from January 2015, whereas the data start in January 2016 for Rhode Island (RI).

Historical DA LMP data, including its decomposition into the three components, are publicly available on each market's website.¹⁹ We retrieve the data from January 2012 for NYISO and PJM and August 2015 for ISO-NE.

To summarize, the span of our dataset extend from January 2012 to November 2022. NYISO data have 10 zones and start at three different dates. PJM data cover 17 out of 21 zones; the data for 15 of those zones start in January 2015. ISO-NE data cover 6 out of 8 zones and start in August 2015, with the exception of Rhode-Island, which start in January 2016. Lastly, we also retrieved the auction results posting day from each market's own FTR calendar, which makes it possible to match FTR prices to their analogous futures prices on the day results are posted.

2.3 Pricing the Electricity Derivatives

The main objective of this paper is to assess whether futures and FTR contracts are priced in a consistent manner in ISO-driven markets of North America. This section describes the pricing framework considered herein, with empirical tests of consistency being reported subsequently. The empirical analyses revolve around isolating implied prices for the loss component of the LMP through proper comparison of FTR and futures price differentials. As outlined in this section, if the pricing in the futures and FTR markets is consistent, the variability of implied prices should be smaller than that of realized losses. We verify

¹⁹NYISO: <http://mis.nyiso.com/public/>, PJM: http://dataminer2.pjm.com/feed/da_hrl_lmps/definition and ISO-NE: <https://www.iso-ne.com/isoexpress/web/reports/pricing/-/tree/lmps-da-hourly>.

empirically that this inequality does not hold in any of the three markets considered.

2.3.1 Contract pricing framework

The two financial derivatives products considered, namely the futures and the FTRs, are not priced in the same market. Rather, futures are traded in an exchange and the FTRs through a bidding auction process. To formulate the pricing problem mathematically, two pricing operators are introduced, one for the futures market and one for the FTR market. A pricing operator $\Pi_{t_{\mathcal{H}}}$ is a mapping from a risk source—a random variable describing the contract cash flow—to a time- $t_{\mathcal{H}}$ price. Thus, the pricing operator for the futures market and the FTR market are represented, respectively, by $\Pi_{t_{\mathcal{H}}}^{\text{Fut}}$ and $\Pi_{t_{\mathcal{H}}}^{\text{FTR}}$, and subsequent tests will seek to determine if both are identical. A mild and common assumption applied at this point is that both pricing operators are linear²⁰ and that, for instance, the price of a set of payoff components can be decomposed as the sum of prices associated with each component.

Under such an assumption and based on (2.4), for any zone i in $\mathcal{Z}_{t_{\mathcal{H}}}$, the time- $t_{\mathcal{H}}$ futures price associated with period \mathcal{H} can be decomposed as the following sum of the three components:

$$F_{t_{\mathcal{H}},i,\mathcal{H}} = \Pi_{t_{\mathcal{H}}}^{\text{Fut}}(\underbrace{E_{\mathcal{H}} + C_{i,\mathcal{H}} + L_{i,\mathcal{H}}}_{\text{Futures payoff}_{i,\mathcal{H}}}) = \Pi_{t_{\mathcal{H}}}^{\text{Fut}}(E_{\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}^{\text{Fut}}(C_{i,\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}^{\text{Fut}}(L_{i,\mathcal{H}}).$$

Moreover, for zones i and j in $\mathcal{Z}_{t_{\mathcal{H}}}$, the time- $t_{\mathcal{H}}$ price per hour of the FTR for period \mathcal{H} can be decomposed as

$$\text{FTR}_{t_{\mathcal{H}},i,j,\mathcal{H}} = \Pi_{t_{\mathcal{H}}}^{\text{FTR}}(\underbrace{C_{j,\mathcal{H}} - C_{i,\mathcal{H}}}_{\text{FTR payoff}_{i,j,\mathcal{H}}}) = \Pi_{t_{\mathcal{H}}}^{\text{FTR}}(C_{j,\mathcal{H}}) - \Pi_{t_{\mathcal{H}}}^{\text{FTR}}(C_{i,\mathcal{H}}), \quad (2.8)$$

with $\Pi_{t_{\mathcal{H}}}^{\text{FTR}}(C_{\emptyset,\mathcal{H}}) = \Pi_{t_{\mathcal{H}}}^{\text{FTR}}(0) = 0$.

²⁰For any two cash flows X, Y and real numbers a, b , $\Pi_t(aX + bY) = a\Pi_t(X) + b\Pi_t(Y)$. Intuitively, this implies for instance that the price of two contracts is twice that of one contract, regardless of the contract cash flow.

The objective is to assess whether futures prices adjust to market information conveyed by FTR prices, leading to consistency between the pricing of both contracts.

2.3.2 Price consistency tests using zonal FTR prices

The procedure testing for the consistency of pricing operators $\Pi_{t_{\mathcal{H}}}^{\text{Fut}}$ and $\Pi_{t_{\mathcal{H}}}^{\text{FTR}}$, i.e. the assumption that $\Pi_{t_{\mathcal{H}}}^{\text{Fut}} = \Pi_{t_{\mathcal{H}}}^{\text{FTR}}$, is now outlined.

Under the assumption that $\Pi_{t_{\mathcal{H}}}^{\text{Fut}} = \Pi_{t_{\mathcal{H}}}^{\text{FTR}} = \Pi_{t_{\mathcal{H}}}$ for some pricing operator $\Pi_{t_{\mathcal{H}}}$ representing a coherent pricing system between the FTR and futures markets, the price for the two products can be rewritten by combining (2.4) and (2.7), implying that

$$\begin{aligned} F_{t_{\mathcal{H}},i,\mathcal{H}} - \text{FTR}_{t_{\mathcal{H}},\emptyset,i,\mathcal{H}} &= \Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}(C_{i,\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \Pi_{t_{\mathcal{H}}}(C_{i,\mathcal{H}}) \\ &= \Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}), \quad \forall i \in \mathcal{Z}_{t_{\mathcal{H}}}. \end{aligned} \quad (2.9)$$

Elements on the left-hand side of (2.9), namely the futures and FTR prices, are observable variables. On the right-hand side are quantities that are not directly observable: the energy price $\Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}})$, which is the same for all zones, and the loss price $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}})$, which is different for each zone.

The price consistency testing approach is based on the isolation of the loss component. At each auction time $t_{\mathcal{H}}$, the linear system (2.9) for all zones i has $|\mathcal{Z}_{t_{\mathcal{H}}}|$ independent equations for $|\mathcal{Z}_{t_{\mathcal{H}}}| + 1$ unknowns. There are $|\mathcal{Z}_{t_{\mathcal{H}}}|$ different prices for the energy loss component $L_{i,\mathcal{H}}$, $i \in \mathcal{Z}_{t_{\mathcal{H}}}$, and one price for the energy marginal cost $E_{\mathcal{H}}$. However, it is possible to isolate the price of the loss component relative to its cross-sectional average price among all zones in $\mathcal{Z}_{t_{\mathcal{H}}}$ as described below.

Differences in the cross-sectional averages of futures and TCC prices are

$$\left(\overline{F}_{t_{\mathcal{H}},\cdot,\mathcal{H}} - \overline{\text{FTR}}_{t_{\mathcal{H}},\emptyset,\cdot,\mathcal{H}} \right) = \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{F_{t_{\mathcal{H}},i,\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}}|} - \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{\text{FTR}_{t_{\mathcal{H}},\emptyset,i,\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}}|} \quad (2.10)$$

$$= \Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}}) + \underbrace{\sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{\Pi(L_{i,\mathcal{H}})}{|\mathcal{Z}_{t_{\mathcal{H}}}|}}_{\overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}}.$$

This leads to

$$\begin{aligned} (F_{t_{\mathcal{H}},i,\mathcal{H}} - \text{FTR}_{t_{\mathcal{H}},\emptyset,i,\mathcal{H}}) - (\overline{F}_{t_{\mathcal{H}},i,\mathcal{H}} - \overline{\text{FTR}}_{t_{\mathcal{H}},\emptyset,i,\mathcal{H}}) \\ = (\Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}}) + \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}})) - (\Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}}) + \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}) \\ = \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}, \\ = \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}) \end{aligned} \quad (2.11)$$

with $\overline{L}_{\mathcal{H}} = \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{L_{i,\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}}|}$, where the last equality is obtained from the linearity assumption on the pricing operator Π_t . $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ is subsequently referred to as the excess loss.

Since the price of energy $\Pi_{t_{\mathcal{H}}}(E_{\mathcal{H}})$ does not appear on the right-hand side of Equation (2.11), it becomes possible to calculate $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}})$ for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ since there are now $|\mathcal{Z}_{t_{\mathcal{H}}}|$ linear equations with $|\mathcal{Z}_{t_{\mathcal{H}}}|$ unknown values. Comparing prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}})$ with realized value $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ makes it possible to draw conclusions about whether futures and FTR contracts are consistently priced.

In what follows, operators E and Var refer to unconditional expectation and variance, whereas $E_{t_{\mathcal{H}}}$ and $\text{Var}_{t_{\mathcal{H}}}$ refer to conditional expectation and variance with respect to the information available at time $t_{\mathcal{H}}$, namely the time of auction for contracts awarded on month \mathcal{H} . To test whether both sets of values, i.e. priced versus realized excess losses over the cross-section average, are aligned, we assume that market participants value loss risk as the expected value of its possible outcomes plus a loss risk premium, which is

$$\forall i \in \mathcal{Z}_{t_{\mathcal{H}}}, \quad \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) = E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}}] + \lambda_{i,t_{\mathcal{H}}}. \quad (2.12)$$

Such a premium is assumed to be a deterministic function of time that is constant over seasonal buckets. The assumption of a piecewise constant risk premium is also used in other works related to pricing in electricity markets, see for instance Longstaff and Wang,

2004.

We deem such an assumption to be reasonable since the loss component is typically not very volatile, at least not relative to the energy and congestion component of the LMP. To ensure the linearity of the pricing operator, we also make the assumption that the loss risk premium is also linear; for instance $\Pi_{t_{\mathcal{H}}}(aX + bY) = aE_{t_{\mathcal{H}}}[X] + bE_{t_{\mathcal{H}}}[Y] + a\lambda_{X,t_{\mathcal{H}}} + b\lambda_{Y,t_{\mathcal{H}}}$ for all $a, b \in \mathbb{R}$ and contingent claims X, Y , where $\lambda_{X,t_{\mathcal{H}}}$ and $\lambda_{Y,t_{\mathcal{H}}}$ are their respective risk premiums.

This leads to the following implications about the price of a zone's loss component relative to its cross-sectional average: for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$,

$$\begin{aligned} \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})} &= \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}|} \frac{E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}}] + \lambda_{i,t_{\mathcal{H}}}}{|\mathcal{Z}_{t_{\mathcal{H}}|} \\ &= E_{t_{\mathcal{H}}} \left[\sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}|} \frac{L_{i,\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}|} \right] + \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}|} \frac{\lambda_{i,t_{\mathcal{H}}}}{|\mathcal{Z}_{t_{\mathcal{H}}|} = E_{t_{\mathcal{H}}}[L_{\mathcal{H}}] + \bar{\lambda}_{t_{\mathcal{H}}}, \end{aligned}$$

with $\bar{\lambda}_{t_{\mathcal{H}}} = \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}|} \frac{\lambda_{i,t_{\mathcal{H}}}}{|\mathcal{Z}_{t_{\mathcal{H}}|}$, and thus

$$\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}) = \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})} = E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] + (\lambda_{i,t_{\mathcal{H}}} - \bar{\lambda}_{t_{\mathcal{H}}}). \quad (2.13)$$

As a consequence, the excess loss component prices should be less volatile than their realized values. Indeed, the law of total variance,²¹ together with Equation (2.13), imply that

$$\begin{aligned} \text{Var} [\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})] &= \text{Var} [E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] + (\lambda_{i,t_{\mathcal{H}}} - \bar{\lambda}_{t_{\mathcal{H}}})] \\ &= \text{Var} [E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] \\ &\leq \text{Var} [E_{t_{\mathcal{H}}}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] + E[\text{Var}_{t_{\mathcal{H}}}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] \\ &= \text{Var} [L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}], \end{aligned} \quad (2.14)$$

²¹The law of total variance stipulates that for two random variables X and Y , the conditional moments of X with respect to Y can be used to compute the unconditional variance of X : $\text{Var}[X] = \text{Var}[E[X|Y]] + E[\text{Var}[X|Y]]$.

that is, the unconditional variance of excess loss prices should be smaller than the unconditional variance of the excess losses.

The second assumption relates to the stationarity of the premium-augmented excess loss process $\{L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}} + (\lambda_{i,t_{\mathcal{H}}} - \bar{\lambda}_{t_{\mathcal{H}}})\}$ over the various monthly periods \mathcal{H} . In practice, since electricity markets are well known to exhibit seasonality patterns that can break stationarity, a milder assumption is considered. The assumption we make is that the premium-augmented excess loss process is stationary within each seasonality bucket. The following three seasonality buckets, which follow the conventional classification used for FTRs, are considered herein: winter months (December to March), summer months (June to September) and the remaining months called shoulder months (April, May, October and November). Such an assumption implies first that the excess loss process has a constant variance within each seasonality bucket: $\text{Var}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}]$ is either V^{sum} , V^{win} or V^{sho} depending on whether \mathcal{H} corresponds to hours of a summer, winter or shoulder month. Assessing whether Inequality (2.14) holds is the basis of the pricing consistency test applied in this section. For each seasonality bucket, $\text{Var}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}]$ can be estimated with the sample variance of excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$. Conversely, estimates of $\text{Var}[\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})]$ can be obtained through the sample variance of implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})$, the latter being obtained by applying Equation (2.11). The failure of Inequality (2.14) to hold implies incompatibilities between market prices for futures and FTR contracts, as the price variance is estimated by combining price data for these two product types.

Empirical results for such tests, i.e. estimates of both sides of Inequality (2.14), their difference and their ratio, are presented for each seasonal bucket (or for the aggregation of the three buckets) in tables 2.1 to 2.6. In these tables, FTR prices used to compute implied excess loss prices are obtained from the nearest auction corresponding to the underlying monthly period \mathcal{H} .

The fifth column in each panel of tables 2.1 to 2.6, i.e. column $(A) - (B)$, shows that the variance of implied excess loss prices $\text{Var}[\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})]$ is always greater than $\text{Var}[L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}]$ for all zones of the three markets and all seasonality buckets,²² a result that contradicts Inequality (2.14). The price variance is sometimes orders of magnitude

²²With the exception of the COMED zone in summer months for the PJM market.

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
A	42	21.25	7.46	13.78	2.85	A	44	3.43	0.76	2.67	4.50
B	20	6.96	2.25	4.71	3.10	B	22	3.39	0.23	3.16	14.88
C	42	18.20	1.12	17.08	16.32	C	44	1.84	0.13	1.71	14.03
D	28	9.63	0.55	9.08	17.48	D	30	13.30	0.78	12.52	17.01
E	20	4.26	0.10	4.16	42.29	E	22	6.62	0.04	6.58	179.33
F	42	9.56	1.15	8.41	8.30	F	44	1.22	0.06	1.17	21.54
G	42	5.73	1.48	4.25	3.88	G	44	1.91	0.20	1.70	9.32
I	20	8.27	0.80	7.47	10.33	I	22	3.57	0.22	3.35	16.43
J	42	15.25	1.59	13.65	9.58	J	44	2.12	0.25	1.87	8.55
K	20	23.29	1.45	21.85	16.09	K	22	29.98	0.51	29.47	59.03

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
A	44	2.09	0.98	1.11	2.14	A	130	9.59	3.43	6.16	2.79
B	23	9.54	0.68	8.86	13.99	B	65	7.21	1.14	6.06	6.30
C	44	1.22	0.31	0.91	3.94	C	130	7.26	0.56	6.70	12.86
D	31	10.31	2.46	7.85	4.19	D	89	10.86	1.42	9.44	7.63
E	23	6.46	0.11	6.35	58.15	E	65	7.01	0.10	6.92	73.31
F	44	3.11	0.17	2.93	17.79	F	130	6.11	0.56	5.55	10.83
G	44	3.90	0.36	3.54	10.78	G	130	3.99	0.72	3.27	5.53
I	23	6.75	0.99	5.76	6.85	I	65	5.98	0.73	5.26	8.24
J	44	4.83	0.70	4.13	6.91	J	130	7.56	0.90	6.65	8.37
K	23	36.70	1.64	35.06	22.35	K	65	29.78	1.32	28.46	22.55

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.1: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - NYISO 24-hour contract

Chapter 2. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	31	11.38	0.72	10.66	15.86	AECO	32	23.06	0.08	22.97	285.52
BGE	31	1.67	0.25	1.42	6.75	BGE	32	3.13	0.30	2.83	10.41
DPL	31	7.93	1.57	6.36	5.05	DPL	32	25.88	0.07	25.81	354.31
JCPL	31	2.68	0.92	1.76	2.93	JCPL	32	1.67	0.06	1.61	28.59
METED	31	3.36	0.34	3.02	9.93	METED	32	3.19	0.04	3.15	71.03
PECO	31	1.94	0.52	1.42	3.74	PECO	32	4.52	0.11	4.41	40.37
PPL	31	3.29	0.22	3.07	15.01	PPL	32	2.70	0.12	2.58	22.55
PENELEC	31	2.16	0.07	2.09	32.59	PENELEC	32	0.86	0.05	0.81	16.16
PSEG	31	3.19	1.00	2.19	3.20	PSEG	32	1.47	0.09	1.38	16.98
APS	31	1.86	0.11	1.75	17.56	APS	32	0.91	0.04	0.87	23.78
AEP	29	1.81	0.73	1.08	2.48	AEP	32	0.86	0.06	0.81	15.48
COMED	31	3.39	2.32	1.07	1.46	COMED	32	1.92	0.55	1.37	3.47
DAY	20	7.54	0.97	6.57	7.77	DAY	22	1.74	0.32	1.41	5.36
DOM	21	1.41	0.07	1.35	21.16	DOM	24	2.54	0.37	2.17	6.85
DUQ	31	5.70	1.24	4.47	4.61	DUQ	32	3.20	0.12	3.07	25.83
ATSI	31	4.74	0.56	4.18	8.47	ATSI	32	3.78	0.07	3.71	53.67
DEOK	31	7.75	1.79	5.96	4.33	DEOK	32	2.59	0.20	2.40	13.24

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	32	35.11	0.29	34.82	121.92	AECO	95	22.88	0.43	22.45	52.78
BGE	32	4.95	0.78	4.17	6.34	BGE	95	3.41	0.45	2.96	7.59
DPL	32	42.01	0.47	41.55	90.02	DPL	95	25.90	0.83	25.06	31.12
JCPL	32	4.81	0.15	4.66	32.97	JCPL	95	3.14	0.47	2.67	6.62
METED	32	6.25	0.09	6.16	68.63	METED	95	4.94	0.19	4.75	26.44
PECO	32	14.07	0.14	13.93	102.84	PECO	95	6.81	0.31	6.50	21.73
PPL	32	13.22	0.14	13.09	95.23	PPL	95	6.85	0.20	6.65	34.40
PENELEC	32	2.27	0.07	2.20	31.99	PENELEC	95	1.82	0.07	1.75	27.69
PSEG	32	12.77	0.14	12.63	89.57	PSEG	95	5.95	0.52	5.42	11.42
APS	32	1.18	0.08	1.10	14.75	APS	95	1.36	0.10	1.27	13.84
AEP	32	1.28	0.32	0.96	4.01	AEP	93	1.41	0.42	0.98	3.32
COMED	32	1.62	0.78	0.84	2.07	COMED	95	2.42	1.30	1.12	1.87
DAY	23	2.63	0.55	2.08	4.77	DAY	65	4.98	0.74	4.24	6.73
DOM	24	10.11	0.28	9.83	35.98	DOM	69	5.02	0.28	4.74	18.19
DUQ	32	2.53	0.61	1.92	4.15	DUQ	95	4.11	0.75	3.37	5.50
ATSI	32	3.37	0.18	3.19	18.41	ATSI	95	4.20	0.31	3.90	13.77
DEOK	32	2.99	0.39	2.60	7.73	DEOK	95	5.35	0.94	4.40	5.67

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.2: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - PJM on-peak contract

Chapter 2. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	31	3.70	0.72	2.98	5.16	AECO	32	5.61	0.08	5.53	69.48
BGE	31	2.78	0.25	2.53	11.19	BGE	32	3.28	0.30	2.98	10.90
DPL	31	4.42	1.57	2.85	2.82	DPL	32	6.69	0.07	6.61	91.53
JCPL	31	3.01	0.92	2.09	3.29	JCPL	32	3.15	0.06	3.09	53.80
METED	31	2.46	0.34	2.13	7.29	METED	32	1.95	0.04	1.90	43.31
PECO	31	1.42	0.52	0.90	2.74	PECO	32	0.99	0.11	0.87	8.80
PPL	31	2.04	0.22	1.82	9.33	PPL	32	2.42	0.12	2.30	20.24
PENELEC	31	2.68	0.07	2.61	40.48	PENELEC	32	1.53	0.05	1.47	28.70
PSEG	31	3.48	1.00	2.49	3.50	PSEG	32	1.67	0.09	1.58	19.27
APS	31	1.27	0.11	1.16	11.97	APS	32	0.73	0.04	0.70	19.28
AEP	29	2.86	0.73	2.13	3.91	AEP	32	0.31	0.06	0.26	5.60
COMED	31	3.76	2.32	1.44	1.62	COMED	32	1.47	0.55	0.92	2.66
DAY	20	10.54	0.97	9.57	10.85	DAY	22	0.89	0.32	0.56	2.73
DOM	21	1.74	0.07	1.68	26.13	DOM	24	1.83	0.37	1.46	4.93
DUQ	31	4.05	1.24	2.82	3.28	DUQ	32	1.65	0.12	1.53	13.35
ATSI	31	4.85	0.56	4.29	8.66	ATSI	32	1.42	0.07	1.35	20.10
DEOK	31	6.59	1.79	4.80	3.68	DEOK	32	1.36	0.20	1.16	6.95

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	32	3.20	0.29	2.91	11.11	AECO	95	4.16	0.43	3.72	9.58
BGE	32	8.78	0.78	8.00	11.25	BGE	95	5.02	0.45	4.57	11.18
DPL	32	12.59	0.47	12.12	26.98	DPL	95	8.27	0.83	7.44	9.94
JCPL	32	0.70	0.15	0.55	4.80	JCPL	95	3.47	0.47	3.00	7.32
METED	32	1.88	0.09	1.79	20.63	METED	95	2.12	0.19	1.93	11.33
PECO	32	5.55	0.14	5.41	40.57	PECO	95	2.91	0.31	2.59	9.27
PPL	32	1.29	0.14	1.15	9.31	PPL	95	2.23	0.20	2.03	11.20
PENELEC	32	0.71	0.07	0.64	10.00	PENELEC	95	1.61	0.07	1.54	24.53
PSEG	32	4.92	0.14	4.77	34.47	PSEG	95	4.58	0.52	4.06	8.79
APS	32	0.84	0.08	0.76	10.53	APS	95	0.97	0.10	0.87	9.80
AEP	32	0.44	0.32	0.11	1.36	AEP	93	1.42	0.42	0.99	3.35
COMED	32	0.63	0.78	-0.15	0.80	COMED	95	2.17	1.30	0.87	1.67
DAY	23	1.19	0.55	0.64	2.16	DAY	65	4.05	0.74	3.31	5.46
DOM	24	2.75	0.28	2.47	9.78	DOM	69	2.09	0.28	1.81	7.58
DUQ	32	1.55	0.61	0.94	2.54	DUQ	95	2.86	0.75	2.11	3.82
ATSI	32	1.49	0.18	1.30	8.11	ATSI	95	2.96	0.31	2.66	9.71
DEOK	32	1.28	0.39	0.89	3.30	DEOK	95	3.85	0.94	2.91	4.08

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.3: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - PJM off-peak contract

Chapter 2. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	31	6.10	0.72	5.39	8.51	AECO	32	10.43	0.08	10.35	129.21
BGE	31	1.44	0.25	1.19	5.81	BGE	32	2.37	0.30	2.07	7.86
DPL	31	4.08	1.57	2.51	2.60	DPL	32	12.72	0.07	12.65	174.14
JCPL	31	1.99	0.92	1.08	2.18	JCPL	32	1.33	0.06	1.27	22.78
METED	31	1.65	0.34	1.31	4.88	METED	32	1.98	0.04	1.93	43.96
PECO	31	1.30	0.52	0.79	2.51	PECO	32	1.35	0.11	1.23	12.01
PPL	31	1.87	0.22	1.65	8.52	PPL	32	2.21	0.12	2.09	18.51
PENELEC	31	2.33	0.07	2.26	35.23	PENELEC	32	0.98	0.05	0.93	18.45
PSEG	31	2.71	1.00	1.72	2.72	PSEG	32	1.04	0.09	0.95	11.96
APS	31	1.37	0.11	1.26	12.96	APS	32	0.73	0.04	0.69	19.23
AEP	29	2.00	0.73	1.27	2.74	AEP	32	0.43	0.06	0.38	7.81
COMED	31	2.96	2.32	0.63	1.27	COMED	32	1.09	0.55	0.54	1.97
DAY	20	8.47	0.97	7.50	8.72	DAY	22	0.81	0.32	0.49	2.50
DOM	21	1.27	0.07	1.21	19.05	DOM	24	1.28	0.37	0.91	3.44
DUQ	31	4.33	1.24	3.10	3.50	DUQ	32	1.94	0.12	1.82	15.67
ATSI	31	3.14	0.56	2.58	5.61	ATSI	32	2.12	0.07	2.05	30.08
DEOK	31	6.75	1.79	4.96	3.77	DEOK	32	1.75	0.20	1.56	8.95

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	32	12.42	0.29	12.13	43.13	AECO	95	9.49	0.43	9.06	21.89
BGE	32	4.58	0.78	3.80	5.87	BGE	95	2.86	0.45	2.42	6.38
DPL	32	21.25	0.47	20.78	45.53	DPL	95	13.20	0.83	12.37	15.86
JCPL	32	1.40	0.15	1.26	9.62	JCPL	95	2.06	0.47	1.58	4.34
METED	32	2.13	0.09	2.04	23.42	METED	95	2.16	0.19	1.98	11.58
PECO	32	4.08	0.14	3.94	29.82	PECO	95	2.36	0.31	2.05	7.54
PPL	32	4.48	0.14	4.34	32.25	PPL	95	3.20	0.20	3.00	16.05
PENELEC	32	0.98	0.07	0.91	13.77	PENELEC	95	1.40	0.07	1.34	21.43
PSEG	32	5.09	0.14	4.95	35.71	PSEG	95	3.57	0.52	3.05	6.85
APS	32	0.88	0.08	0.80	10.95	APS	95	1.03	0.10	0.93	10.41
AEP	32	0.55	0.32	0.23	1.71	AEP	93	1.15	0.42	0.73	2.71
COMED	32	0.80	0.78	0.02	1.03	COMED	95	1.79	1.30	0.49	1.38
DAY	23	1.40	0.55	0.85	2.54	DAY	65	3.78	0.74	3.04	5.11
DOM	24	4.70	0.28	4.42	16.74	DOM	69	2.48	0.28	2.20	8.99
DUQ	32	1.63	0.61	1.02	2.67	DUQ	95	3.02	0.75	2.27	4.03
ATSI	32	1.74	0.18	1.56	9.51	ATSI	95	2.62	0.31	2.31	8.58
DEOK	32	1.65	0.39	1.27	4.28	DEOK	95	4.22	0.94	3.28	4.47

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.4: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - PJM 24-hour contract

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	28	0.70	0.50	0.20	1.40	ME	30	0.29	0.14	0.15	2.04
NH	28	0.18	0.01	0.17	13.27	NH	30	0.18	0.04	0.13	4.08
RI	27	1.22	0.09	1.13	13.13	RI	28	0.17	0.08	0.09	2.01
SEMA	28	0.17	0.05	0.12	3.64	SEMA	30	0.10	0.03	0.07	2.97
WCMA	28	0.49	0.10	0.38	4.79	WCMA	30	0.32	0.02	0.30	14.06
NEMA	28	0.70	0.03	0.67	22.53	NEMA	30	0.18	0.02	0.16	8.84

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	30	0.55	0.18	0.36	2.97	ME	88	0.51	0.28	0.23	1.82
NH	30	0.20	0.03	0.17	7.22	NH	88	0.18	0.03	0.15	5.50
RI	28	0.22	0.04	0.19	6.25	RI	83	0.52	0.08	0.44	6.43
SEMA	30	0.35	0.04	0.31	8.87	SEMA	88	0.21	0.04	0.17	4.84
WCMA	30	0.34	0.02	0.32	16.45	WCMA	88	0.37	0.05	0.33	8.12
NEMA	30	0.14	0.03	0.11	4.65	NEMA	88	0.34	0.03	0.31	12.92

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.5: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - ISO-NE on-peak contract

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	28	1.11	0.50	0.61	2.22	ME	30	0.31	0.14	0.16	2.17
NH	28	0.19	0.01	0.18	14.10	NH	30	0.05	0.04	0.01	1.24
RI	27	0.54	0.09	0.45	5.80	RI	28	0.17	0.08	0.08	1.98
SEMA	28	0.26	0.05	0.21	5.66	SEMA	30	0.07	0.03	0.03	1.90
WCMA	28	1.20	0.10	1.10	11.83	WCMA	30	0.31	0.02	0.29	13.59
NEMA	28	0.38	0.03	0.35	12.23	NEMA	30	0.12	0.02	0.10	5.92

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	30	0.26	0.18	0.07	1.39	ME	88	0.55	0.28	0.27	1.97
NH	30	0.08	0.03	0.05	2.83	NH	88	0.11	0.03	0.07	3.16
RI	28	0.21	0.04	0.17	5.78	RI	83	0.30	0.08	0.22	3.68
SEMA	30	0.14	0.04	0.10	3.61	SEMA	88	0.15	0.04	0.11	3.49
WCMA	30	0.43	0.02	0.41	20.65	WCMA	88	0.64	0.05	0.59	13.93
NEMA	30	0.15	0.03	0.12	5.03	NEMA	88	0.23	0.03	0.21	8.82

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of nearest auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.6: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their nearest-auction-implied prices - ISO-NE off-peak contract

higher than that of realized excess losses as evidenced by the sixth column $(A)/(B)$. If our assumptions holds,²³ these numbers unequivocally highlight the fact that futures and FTR prices are not priced consistently over the corresponding two financial markets.

To better visualize such outcomes, figures 2.4 to 2.6 present time series of zonal excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ (green lines) and of their nearest auction price $\Pi_{t,\mathcal{H}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})$ (black lines) for the three power markets.²⁴ Each panel represents a different zone, and each time point of the time series corresponds to one contract month. Some black lines do not cover the full 2012–2021 span as some of the futures contracts were not historically available.

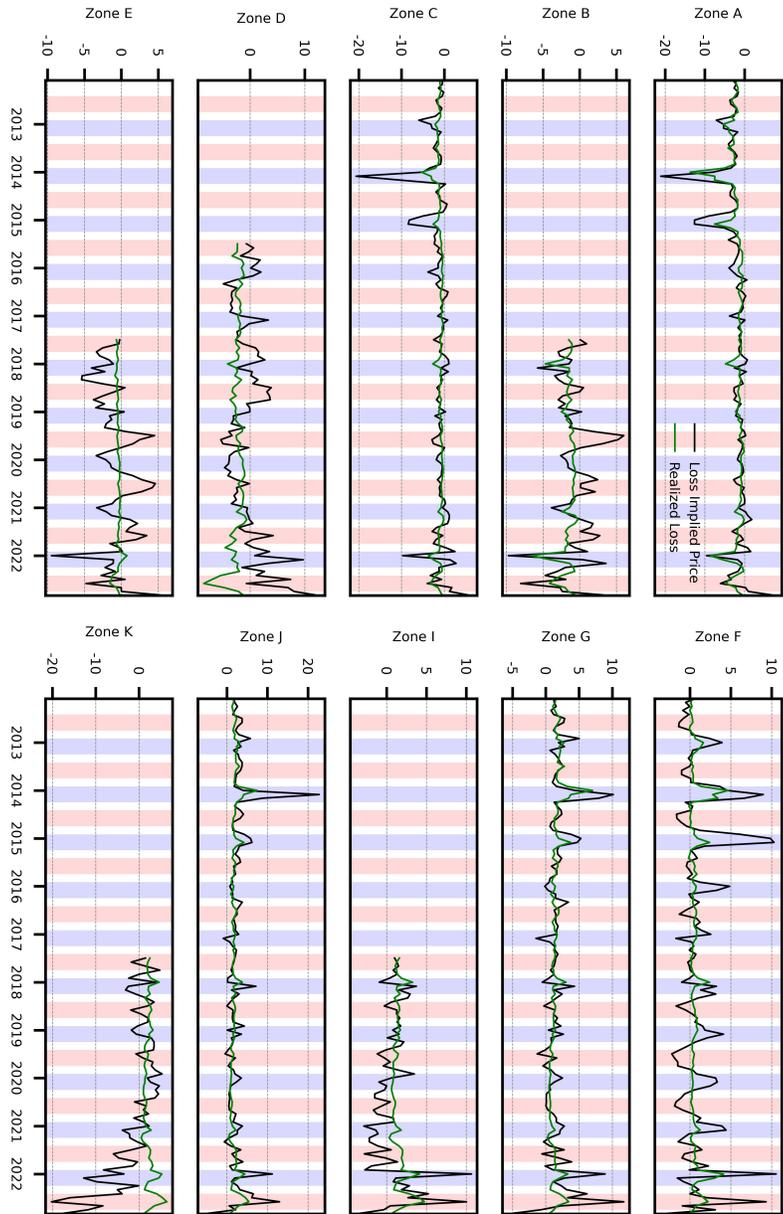
The implied excess loss prices are visually much more volatile than the realized excess losses, supporting the aforementioned evidence from tables 2.1 to 2.6. It is interesting to observe that for most of the zones, excess losses are, in some periods, priced at levels that realized values were never even close to reaching in the data sample. This confirms that a single pricing operator cannot consistently explain prices in both futures and FTR markets. Furthermore, the higher volatility for implied excess loss prices than for realized losses is pervasive throughout the entire duration of the sample; the misalignment between prices of futures and FTRs is therefore not due to a few standalone events; rather, it occurs systematically.

2.3.3 Price consistency tests using the most liquid FTR contracts

The results presented in the previous section leave no doubt as to the presence of price inconsistencies between the FTR and futures markets. However, it is important to assess whether such inconsistencies occur with liquid FTRs, or if they mostly apply to illiquid contracts. Indeed, the potential for acting on observed pricing inconsistencies depends on the ability to transact or place successful bids on identified contacts. The approach followed herein to answer this question is to perform variance tests similar to those outlined previously, but this time only on FTR contracts which were traded in the nearest auctions of the data sample. The results of the previous section are obtained from nodal FTR prices,

²³We relaxed the constant loss risk premium assumption by testing with a stochastic risk premium, the results remained unchanged. Details are in Appendix A.2.

²⁴For better readability, only the 24-hour contracts are presented, and separate on-peak and off-peak contracts results are omitted.



Comparison for each zone i of realized excess losses $L_{i,H} - \bar{L}_H$ (green lines) and their corresponding nearest-auction-implied price $\Pi_t(L_{i,H} - \bar{L}_H)$ (black lines) whose calculation is based on (2.11). Missing values correspond to months for which no futures contract is available for the corresponding zone. Blue areas are the winter months (December to March), white areas are the shoulder months (April, May, October and November) and red areas are the summer months (June to September).

Figure 2.4: Time series of excess losses $L_{i,H} - \bar{L}_H$ in $\$/MWh$ and their nearest-auction-implied price - NYISO

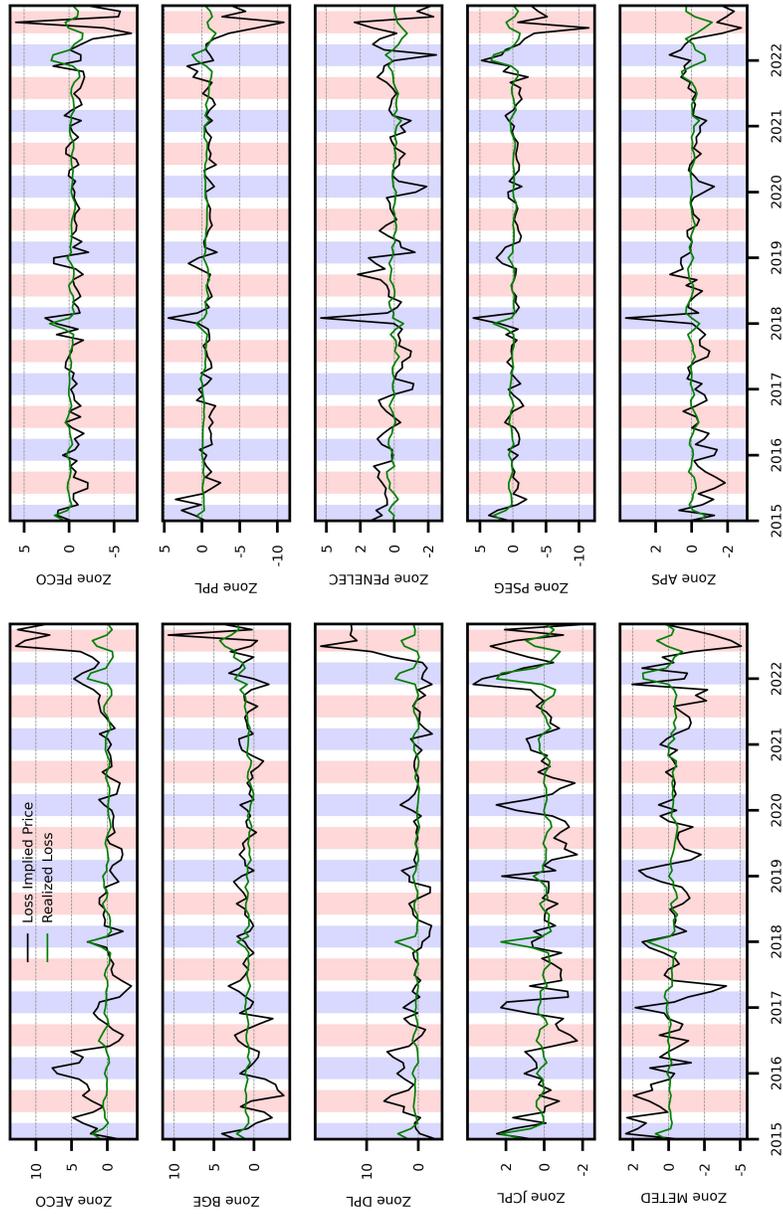
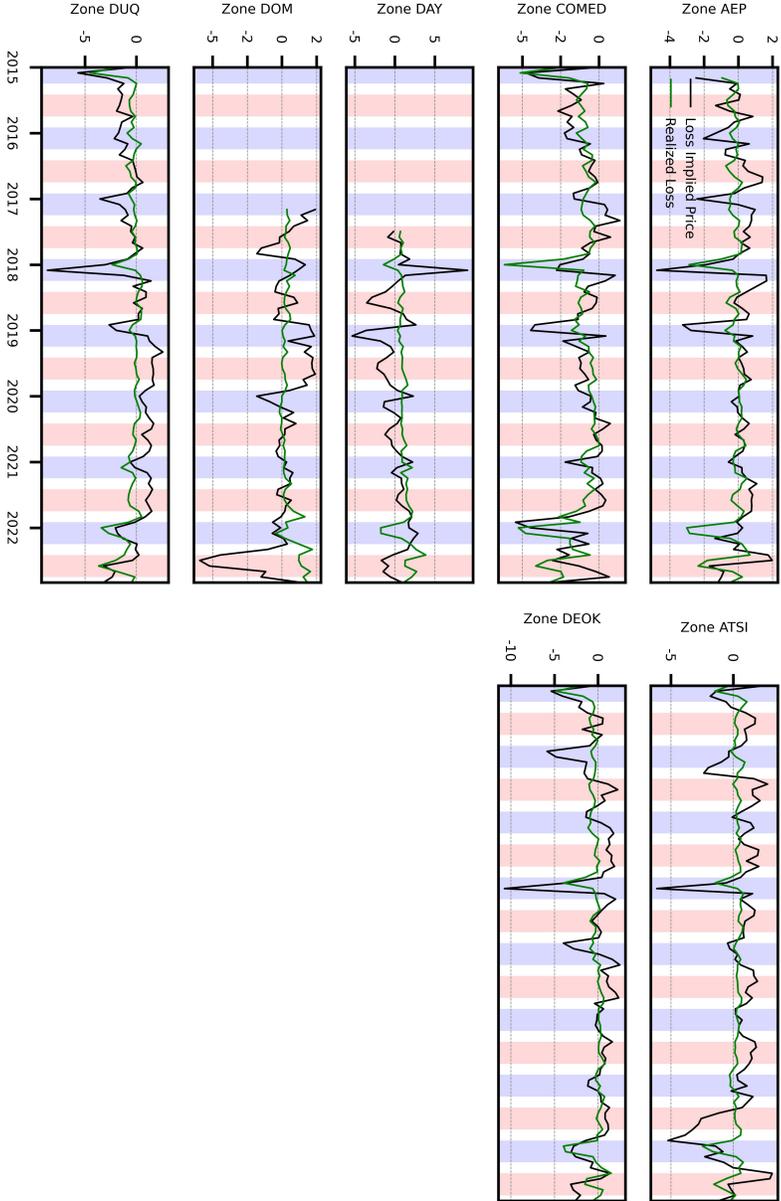
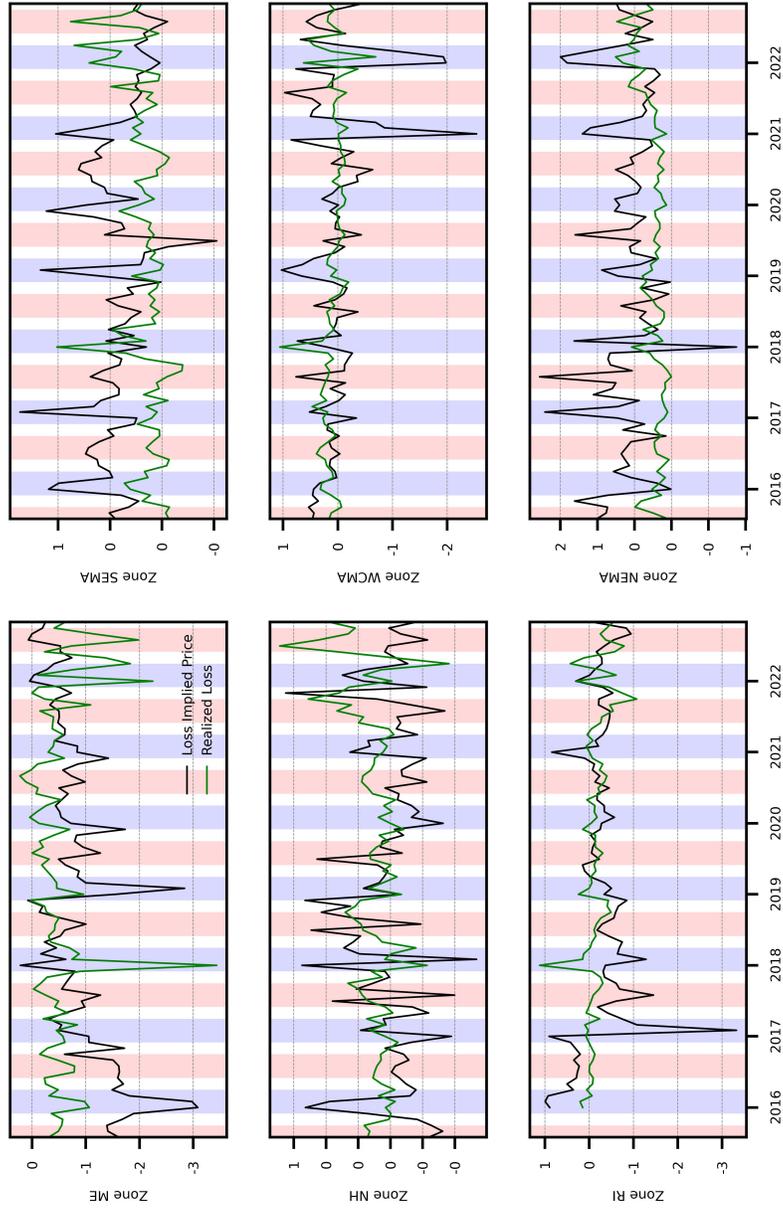


Figure 2.5: Time series of excess losses $L_{i,H} - \bar{L}_H$ in \$/MWh and their nearest-auction-implied price - PIM



Comparison for each zone i of realized excess losses $L_{i,\mathcal{H}} - \bar{L}_{\mathcal{H}}$ (green lines) and their corresponding nearest-auction-implied price $\Pi_t(L_{i,\mathcal{H}} - \bar{L}_{\mathcal{H}})$ (black lines) whose calculation is based on (2.11). Missing values correspond to months for which no futures contract is available for the corresponding zone. Blue areas are the winter months (December to March), white areas are the shoulder months (April, May, October and November) and red areas are the summer months (June to September).

Figure 2.5: Time series of excess losses $L_{i,\mathcal{H}} - \bar{L}_{\mathcal{H}}$ in \$/hour and their nearest-auction-implied price - PJM (cont.)



Comparison for each zone i of realized excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ (green lines) and their corresponding nearest-auction-implied price $\Pi_t(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}})$ (black lines) whose calculation is based on (2.11). Missing values correspond to months for which no futures contract is available for the corresponding zone. Blue areas are the winter months (December to March), white areas are the shoulder months (April, May, October and November) and red areas are the summer months (June to September).

Figure 2.6: Time series of excess losses $L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}$ in \$/MWh and their nearest-auction-implied price - ISO-NE

but nothing ensures that corresponding contracts are traded with sufficient liquidity.

Traded FTRs apply to zone pairs, i.e. with an SI j and an SO i . Based on Equation (2.9), the implied price differential between losses of two zones is given by

$$(F_{t_{\mathcal{H}},j,\mathcal{H}} - F_{t_{\mathcal{H}},i,\mathcal{H}}) - \text{FTR}_{t_{\mathcal{H}},i,j,\mathcal{H}} = \Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}}) - \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) = \Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}}). \quad (2.15)$$

Moreover,

$$\begin{aligned} \Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}}) &= \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) \\ &= \text{E}_{t_{\mathcal{H}}}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}] + (\lambda_{j,t_{\mathcal{H}}} - \lambda_{i,t_{\mathcal{H}}}). \end{aligned} \quad (2.16)$$

The energy component is naturally cancelled out by the differentiation between the two zones. Using assumptions analogous to those of the previous section, i.e. Equation (2.12) and the stationarity of premium-incremented losses over a given seasonality bucket, it is again possible to bound the variance of loss spread price, i.e. zone price differentials, by the variance of realized loss spreads:

$$\begin{aligned} \text{Var}[\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})] &= \text{Var}[\text{E}_{t_{\mathcal{H}}}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}] + (\lambda_{j,t_{\mathcal{H}}} - \lambda_{i,t_{\mathcal{H}}})] \\ &= \text{Var}[\text{E}_{t_{\mathcal{H}}}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}]] \\ &\leq \text{Var}[\text{E}_{t_{\mathcal{H}}}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}]] + \text{E}[\text{Var}_{t_{\mathcal{H}}}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}]] \\ &= \text{Var}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}]. \end{aligned} \quad (2.17)$$

Again, both sides of Inequality (2.17) can be estimated by computing the sample variance of either implied loss differential prices obtained from Equation (2.15) or realized loss spread, respectively. However this time, for a given zone pair i, j , data points are only included in the sample variance if the FTR has been traded in the corresponding months.

Tables 2.7 to 2.12 again report the estimate of $\text{Var}[\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})]$, the variance of implied price of loss spreads denoted (A), and the variance of realized loss differentials $\text{Var}[L_{j,\mathcal{H}} - L_{i,\mathcal{H}}]$, denoted (B), along with the difference and the ratio of the two former quantities. For each seasonality bucket, results are reported for the most liquid FTRs. These are defined as FTRs for which the corresponding SO/SI zone pair had the largest numbers

of nearest auctions with at least one unit allocated. The column labelled n in the table reports the number of nearest auctions in the sample for which such FTRs were allocated at least once.

Once again, the fifth column in each panel of tables 2.7 to 2.12, namely column (A) – (B), shows that estimates of the implied loss spread price variance are usually higher than estimates of the realized loss spread variance. Notably, in the NYISO market, this difference is positive for all ten zone pairs in the shoulder months and in the summer months. This relationship holds for most presented zone pairs, with the exception of pairs $A - C$ and $C - A$ in the winter months. In the PJM market, the difference is positive in all zones for the 24-hour, on-peak and off-peak FTR contracts. In the ISO-NE market, again, the difference is positive in the vast majority of zones for both contract types and in all seasonal buckets.

Panel a: Winter Months					Panel b: Shoulder Months						
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
C-A	40	1.73	3.02	-1.30	0.57	F-G	39	1.62	0.15	1.47	10.92
F-G	40	4.09	0.34	3.74	11.93	C-A	38	1.12	0.24	0.87	4.58
J-G	36	4.55	0.07	4.48	67.44	J-G	35	0.34	0.03	0.32	12.28
A-C	31	1.36	3.10	-1.74	0.44	A-C	34	0.73	0.18	0.55	4.04
G-J	30	0.62	0.05	0.57	12.48	G-F	33	1.36	0.19	1.17	7.21
G-F	28	3.49	0.14	3.36	25.22	G-C	25	2.59	0.35	2.24	7.40
A-G	23	20.25	6.98	13.27	2.90	A-G	22	13.94	1.36	12.58	10.24
G-A	23	68.11	21.94	46.17	3.10	G-J	22	0.38	0.02	0.37	21.13
C-G	20	29.14	3.32	25.82	8.77	I-G	21	1.53	0.02	1.51	73.52
G-C	20	64.86	3.25	61.61	19.97	G-A	20	2.15	1.45	0.70	1.49

Panel c: Summer Months					Panel d: All Months						
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
C-A	40	1.45	0.29	1.16	4.95	C-A	118	1.54	1.45	0.08	1.06
G-F	40	1.38	0.22	1.16	6.21	J-G	109	1.72	0.06	1.66	28.86
J-G	38	0.34	0.08	0.27	4.46	F-G	107	2.98	0.22	2.76	13.45
A-C	35	1.46	0.33	1.13	4.48	G-F	101	2.99	0.22	2.78	13.85
F-G	28	0.81	0.13	0.67	6.05	A-C	100	1.24	1.34	-0.11	0.92
A-G	26	3.72	0.91	2.81	4.09	A-G	71	13.07	3.45	9.63	3.79
G-C	24	12.03	1.36	10.68	8.87	G-J	69	0.49	0.05	0.44	9.19
G-A	24	3.68	1.15	2.53	3.19	G-C	69	24.35	1.71	22.64	14.27
B-A	21	6.33	0.05	6.28	121.92	G-A	67	29.88	10.19	19.69	2.93
I-G	20	0.95	0.12	0.83	7.98	C-G	58	16.88	1.86	15.02	9.09

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.7: Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - NYISO 24-hour contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$		$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$	
PSEG-JCPL	21	4.64	0.01	4.63	626.01	PSEG-JCPL	23	1.81	0.01	1.80	192.77
PSEG-PECO	20	3.19	0.13	3.06	24.61	PSEG-PPL	19	3.46	0.02	3.44	200.77
PPL-PECO	20	1.86	0.12	1.74	15.33	METED-PPL	17	3.87	0.07	3.80	56.41
METED-PECO	18	3.41	0.05	3.36	65.11	PSEG-PECO	17	0.81	0.01	0.80	57.05
DPL-PECO	15	5.15	0.30	4.85	17.28	METED-PECO	16	3.60	0.05	3.54	69.35
PSEG-PPL	15	3.56	0.54	3.02	6.64	PPL-PECO	15	3.51	0.02	3.49	195.57
JCPL-PECO	14	5.05	0.08	4.97	61.48	PECO-PPL	15	2.67	0.02	2.65	150.36
JCPL-PPL	14	11.57	0.47	11.10	24.66	PPL-METED	14	4.00	0.05	3.95	88.44
JCPL-PSEG	13	5.03	0.01	5.02	741.34	DPL-PECO	14	44.06	0.20	43.86	219.74
PECO-PPL	12	2.57	0.01	2.55	180.83	JCPL-PPL	14	2.23	0.03	2.20	86.51

Panel c: Summer Months						Panel d: All Months					
$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$		$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$	
AECO-PECO	26	64.71	0.12	64.60	549.46	PSEG-JCPL	63	15.69	0.01	15.68	2140.23
PSEG-PECO	24	7.61	0.02	7.59	473.91	PSEG-PECO	61	4.27	0.07	4.20	60.18
JCPL-PECO	21	30.53	0.02	30.51	1942.25	METED-PECO	52	3.51	0.05	3.46	70.49
METED-PPL	19	6.74	0.17	6.58	40.64	PPL-PECO	51	2.06	0.08	1.98	25.38
PSEG-JCPL	19	45.85	0.01	45.85	8883.34	METED-PPL	48	5.82	0.09	5.72	61.29
PPL-AECO	18	138.70	0.47	138.23	296.05	PSEG-PPL	46	3.00	0.23	2.77	13.18
METED-PECO	18	2.12	0.02	2.11	129.27	AECO-PECO	46	59.30	0.10	59.21	611.02
PPL-PECO	16	0.73	0.04	0.69	18.11	JCPL-PECO	46	16.44	0.04	16.40	366.31
JCPL-PPL	16	1.15	0.04	1.10	26.88	JCPL-PPL	44	4.60	0.19	4.41	24.19
DPL-PECO	14	18.73	0.09	18.65	219.86	DPL-PECO	43	21.39	0.20	21.19	107.59

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ is represented by (A) . The sample variance of the realized loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ is represented by (B) . The difference, column $(A) - (B)$, and the ratio, column $(A)/(B)$, are also provided in each panel.

Table 2.8: Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM on-peak contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	18	4.74	0.13	4.61	35.40	PPL-METED	16	1.93	0.06	1.87	31.25
METED-PECO	18	1.52	0.07	1.45	20.51	PSEG-JCPL	16	2.79	0.01	2.78	258.04
PECO-PPL	14	0.54	0.11	0.43	4.75	METED-PPL	14	2.13	0.08	2.05	25.45
JCPL-PPL	14	1.30	0.29	1.01	4.49	PPL-PECO	14	3.06	0.02	3.04	144.20
PPL-METED	13	2.98	0.02	2.96	120.50	PSEG-PECO	13	2.29	0.01	2.29	450.34
PSEG-PPL	13	3.70	0.59	3.11	6.26	AECO-PECO	13	9.03	0.00	9.03	2510.02
PPL-PECO	13	1.97	0.14	1.82	13.95	PECO-PPL	12	1.82	0.01	1.81	121.48
JCPL-AECO	13	1.42	0.05	1.37	31.30	PSEG-PPL	12	0.74	0.02	0.72	34.80
PSEG-JCPL	13	3.46	0.01	3.45	500.54	METED-PECO	10	2.88	0.05	2.83	54.49
JCPL-PECO	12	1.74	0.11	1.63	15.63	PECO-METED	10	0.75	0.07	0.69	11.15

Panel c: Summer Months						Panel d: All Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PPL-METED	23	3.34	0.11	3.23	31.14	PPL-METED	52	2.87	0.07	2.79	40.09
AECO-PECO	22	5.94	0.12	5.81	48.07	PSEG-PECO	50	3.97	0.07	3.90	54.36
PSEG-PECO	19	4.09	0.02	4.07	185.30	PPL-PECO	44	2.16	0.07	2.09	31.15
JCPL-PPL	18	1.13	0.07	1.06	16.56	AECO-PECO	43	6.57	0.10	6.47	65.82
PPL-PECO	17	1.07	0.04	1.04	29.75	PSEG-JCPL	43	3.61	0.01	3.60	496.86
METED-PPL	14	1.06	0.04	1.02	25.81	JCPL-PPL	42	1.38	0.14	1.24	9.57
PSEG-JCPL	14	2.89	0.00	2.88	670.60	METED-PECO	40	2.12	0.06	2.06	34.11
METED-PECO	12	2.03	0.02	2.01	133.60	PECO-PPL	38	0.95	0.06	0.89	15.13
PSEG-PPL	12	0.61	0.05	0.56	12.63	PSEG-PPL	37	2.03	0.25	1.78	8.17
PECO-PPL	12	0.55	0.03	0.51	16.97	METED-PPL	37	1.70	0.05	1.64	31.16

Each panel shows the results for one seasonally bucket: panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time t_H and underlying month H , implied loss prices $\Pi_{t_H}(L_{j,H} - L_{i,H})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_H}(L_{j,H} - L_{i,H})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,H} - L_{i,H}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.9: Sample variance of the loss spreads $L_{j,H} - L_{i,H}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM off-peak contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	14	3.26	0.09	3.17	37.24	PSEG-PECO	15	0.83	0.01	0.81	65.62
AECO-PECO	9	0.98	0.01	0.97	93.00	PPL-PECO	12	1.28	0.01	1.27	122.31
PSEG-JCPL	9	6.25	0.01	6.25	827.14	PSEG-PPL	9	0.86	0.02	0.84	50.65
JCPL-PPL	8	3.37	0.49	2.88	6.92	METED-PECO	8	1.09	0.07	1.03	16.41
PSEG-PPL	8	4.48	0.67	3.81	6.66	PSEG-JCPL	8	1.32	0.01	1.31	119.90
PPL-PECO	8	0.34	0.01	0.33	31.22	JCPL-PPL	8	0.77	0.02	0.75	31.37
METED-PECO	8	2.45	0.02	2.44	156.89	METED-PPL	7	2.64	0.11	2.53	24.91
METED-PPL	7	0.31	0.12	0.20	2.68	DPL-PECO	5	73.83	0.35	73.48	212.19
PECO-PPL	6	0.25	0.02	0.23	13.37	AECO-PECO	5	72.78	0.01	72.77	6034.80
JCPL-PECO	5	1.19	0.02	1.17	65.55	ATSI-DUQ	4	2.25	0.32	1.92	6.93

Panel c: Summer Months						Panel d: All Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	13	1.37	0.01	1.36	100.70	PSEG-PECO	42	1.89	0.04	1.85	43.55
PSEG-JCPL	9	4.18	0.00	4.18	7743.15	PPL-PECO	28	1.28	0.02	1.25	55.65
AECO-PECO	9	1.19	0.03	1.16	36.78	PSEG-JCPL	26	3.86	0.01	3.85	614.86
PPL-PECO	8	2.21	0.03	2.19	80.62	METED-PECO	24	1.49	0.05	1.44	29.81
METED-PECO	8	0.44	0.01	0.43	39.71	AECO-PECO	23	15.16	0.03	15.12	460.17
JCPL-PPL	7	0.59	0.08	0.51	7.51	PSEG-PPL	23	2.33	0.35	1.98	6.66
PSEG-PPL	6	0.25	0.03	0.22	8.24	JCPL-PPL	23	2.46	0.25	2.21	9.73
PECO-PPL	5	0.11	0.04	0.07	2.80	METED-PPL	16	1.36	0.10	1.26	13.27
ATSI-DUQ	5	1.91	0.38	1.53	5.00	PECO-PPL	15	0.58	0.03	0.56	22.53
PECO-AECO	5	23.93	0.06	23.87	413.05	PECO-AECO	11	13.65	0.07	13.59	208.40

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.10: Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - PJM 24-hour contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
NEMA-ME	11	1.47	0.61	0.86	2.40	NEMA-ME	12	0.18	0.04	0.14	4.06
RI-ME	5	1.45	0.09	1.36	15.66	NEMA-WCMA	10	0.95	0.03	0.93	37.97
NH-ME	3	0.29	0.07	0.22	4.11	NH-ME	5	0.20	0.05	0.15	3.86
SEMA-ME	3	0.68	1.78	-1.10	0.38	NEMA-SEMA	4	0.71	0.02	0.69	31.17
RI-SEMA	3	3.87	0.01	3.87	669.30	RI-ME	2	0.00	0.00	0.00	2.04
WCMA-NH	2	0.00	0.04	-0.04	0.04	SEMA-ME	2	0.00	0.04	-0.04	0.04
NEMA-SEMA	2	0.06	0.00	0.06	47.42	WCMA-RI	2	0.22	0.00	0.22	534.92

Panel c: Summer Months						Panel d: All Months					
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
NEMA-ME	14	0.27	0.06	0.21	4.49	NEMA-ME	37	0.71	0.23	0.48	3.09
NEMA-WCMA	8	0.33	0.02	0.31	19.24	NEMA-WCMA	20	1.64	0.07	1.57	24.37
NH-ME	5	0.15	0.02	0.13	6.40	NH-ME	13	0.23	0.06	0.18	4.12
NEMA-SEMA	5	1.31	0.01	1.30	177.17	NEMA-SEMA	11	1.00	0.01	0.99	69.35
SEMA-ME	3	0.00	0.03	-0.03	0.01	RI-ME	9	0.89	0.15	0.74	6.09
RI-ME	2	0.00	0.10	-0.10	0.00	SEMA-ME	8	0.20	0.75	-0.55	0.27
NH-RI	2	0.04	0.01	0.03	7.45	NH-SEMA	3	0.00	0.00	0.00	0.53

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{j\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{j\mathcal{H}}}(L_{j;\mathcal{H}} - L_{i;\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{j\mathcal{H}}}(L_{j;\mathcal{H}} - L_{i;\mathcal{H}})$ is represented by (A). The sample variance of the realized loss spreads $L_{j;\mathcal{H}} - L_{i;\mathcal{H}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table 2.11: Sample variance of the loss spreads $L_{j;\mathcal{H}} - L_{i;\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - ISO-NE on-peak contract

Panel a: Winter Months					Panel b: Shoulder Months						
$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$	$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$		
NEMA-ME	7	1.26	0.21	1.05	6.12	NEMA-ME	14	0.28	0.30	-0.02	0.93
RI-ME	6	1.17	0.28	0.89	4.21	NEMA-WCMA	9	0.60	0.04	0.56	16.13
NEMA-WCMA	4	0.92	0.12	0.80	7.56	RI-ME	6	0.53	0.34	0.20	1.58
NH-ME	3	4.93	0.06	4.87	80.42	NEMA-SEMA	4	0.09	0.02	0.07	5.10
WCMA-NH	3	0.11	0.05	0.06	2.15	NH-ME	3	0.02	0.03	-0.01	0.71
RI-SEMA	3	3.23	0.01	3.22	298.85	RI-NH	3	0.18	0.09	0.09	2.04
RI-NH	2	0.38	0.32	0.06	1.18	NH-SEMA	3	0.04	0.02	0.03	2.56

Panel c: Summer Months					Panel d: All Months						
$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$	$j - i$	(A)	(B)	$(A) - (B)$	$(A)/(B)$		
NEMA-ME	17	0.44	0.43	0.01	1.02	NEMA-ME	38	0.56	0.32	0.24	1.73
NEMA-WCMA	9	0.83	0.02	0.81	42.19	NEMA-WCMA	22	1.27	0.07	1.20	17.67
RI-ME	5	0.56	0.13	0.42	4.24	RI-ME	17	0.67	0.25	0.43	2.73
SEMA-ME	4	0.01	0.36	-0.35	0.03	NEMA-SEMA	7	0.18	0.03	0.15	6.97
NH-RI	2	0.08	0.01	0.07	15.14	NH-ME	6	2.40	0.08	2.32	30.61
NH-SEMA	2	0.01	0.00	0.01	21.46	SEMA-ME	6	0.01	0.23	-0.22	0.04
NEMA-NH	2	0.03	0.00	0.02	9.70	RI-NH	6	0.18	0.11	0.08	1.73

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the nearest auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of nearest auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ is represented by (A) . The sample variance of the realized loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ is represented by (B) . The difference, column $(A) - (B)$, and the ratio, column $(A)/(B)$, are also provided in each panel.

Table 2.12: Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their nearest-auction-implied prices for the most liquid FTRs - ISO-NE off-peak contract

2.4 Conclusion

An examination of historical futures and FTR price data for various zones of the NYISO, the ISO-NE and PJM reveals substantial pricing inconsistencies. By computing futures and FTR price differences over their cross-sectional averages, market-implied prices of the zonal LMPs' excess loss component can be isolated. Such implied loss prices are found to be incompatible with distributions of realized losses as loss-implied prices exhibit much higher variability than realized losses; this is incompatible with the loss prices being the expectation over all possible loss outcomes. Moreover, the loss price sometimes reaches levels never attained by realized losses, further confirming incompatibilities between the prices of futures and FTRs in the considered markets.

The misalignment between realized losses and their implied prices has several implications for the power market participants. The regulator could raise awareness about current existing pricing inconsistencies, which could lead to investors and specialized firms conducting arbitrage operations that would correct mispricing instances.²⁵ Another possibility that could be envisioned by the ISO would be to follow the suggestion of Rudkevich et al., 2005 and Sarkar and Khaparde, 2009, who propose the introduction of novel derivatives similar to FTRs, whose payoff would be linked to the loss part rather than to congestion. Indeed, the combination of an FTR and a futures contract on a given zone cannot be used to hedge standalone loss risk due to the inclusion of the energy component in the futures payoff. Furthermore, the pricing inconsistencies revealed in this study imply that such a combination might prove an inadequate loss hedge. The creation of loss-related financial derivatives could therefore alleviate this issue and potentially provide additional signals to improve market efficiency. Such contracts could facilitate arbitrage if misaligned prices persist, which would eventually lead to the market correcting pricing inconsistencies. Electricity futures would still remain relevant after such an inclusion due to futures being traded continuously, unlike auctions which occur sporadically. A key requirement for the inclusion of a novel loss-related derivative to work properly is sufficient appetite from market participants to trade in such contracts.

²⁵Arbitrage operations should account for the fact that the FTRs are auction-based and futures traded in an exchange as it can hurdle arbitrage opportunities.

A potential avenue for further work is the development of a pricing scheme for futures contracts that is compatible with observed FTR data (or conversely an FTR pricing scheme compatible with observed futures price data), which could be of interest for participants of such derivatives markets. To help with such an endeavour, statistical analysis on empirical data such as the PCA-based approach of Godin and Ibrahim, 2021, which demonstrates strong commonalities between congestion costs of the various zones, could be applied to provide insight on interrelationships between prices on all nodes of the grid. Such information can serve as a basis for the valuation of FTRs involving non-zonal nodes, by relating them to zonal FTRs.

References

- Adamson, S., & Englander, S. L. (2005). Efficiency of New York Transmission Congestion Contract auctions. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 59a–59a.
- Adamson, S., Noe, T., & Parker, G. (2010). Efficiency of Financial Transmission Rights markets in centrally coordinated periodic auctions. *Energy Economics*, 32(4), 771–778.
- Adamson, S., & Parker, G. (2013). Participation and efficiency in the New York Financial Transmission Rights markets. *Financial Transmission Rights: Analysis, Experiences and Prospects*, 289–303.
- Alsac, O., Bright, J., Brignone, S., Prais, M., Silva, C., Stott, B., & Vempati, N. (2004). The rights to fight price volatility. *IEEE Power and Energy Magazine*, 2(4), 47–57.
- Baltaduonis, R., Bonar, S., Carnes, J., & Mastrangelo, E. (2017). Risk and abnormal returns in markets for Congestion Revenue Rights. *Journal of Energy Markets*, 10(3).
- Bartholomew, E. S., Siddiqui, A. S., Marnay, C., & Oren, S. S. (2003). The New York Transmission Congestion Contract market: Is it truly working efficiently? *The Electricity Journal*, 16(9), 14–24.
- Celebi, M., Hajos, A., & Hanser, P. Q. (2010). Virtual bidding: The good, the bad and the ugly. *The Electricity Journal*, 23(5), 16–25.

- Deng, S.-J., Oren, S., & Meliopoulos, A. (2010). The inherent inefficiency of simultaneously feasible Financial Transmission Rights auctions. *Energy Economics*, 32(4), 779–785.
- Deng, S.-J., & Oren, S. S. (2006). Electricity derivatives and risk management. *Energy*, 31(6-7), 940–953.
- Godin, F., & Ibrahim, Z. (2021). An analysis of electricity congestion price patterns in North America. *Energy Economics*, 102, 105506.
- Hadsell, L., & Shawky, H. A. (2009). Efficiency and profit in the NYISO Transmission Congestion Contract market. *The Electricity Journal*, 22(9), 47–57.
- Hogan, W. (2002). Financial Transmission Right formulations [Mimeo, JFK School of Government, Harvard Electricity Policy Group, Harvard University].
- Hogan, W. W. (1992). Contract networks for electric power transmission. *Journal of Regulatory Economics*, 4(3), 211–242.
- Hogan, W., et al. (2013). Financial Transmission Rights, revenue adequacy and multi-settlement electricity markets. Available Harvard University web site: <http://www.hks.harvard.edu/whogan/HoganFTRRevAdequacy31813>.
- Kristiansen, T. (2005). Markets for Financial Transmission Rights. *Energy Studies Review*, 13(1).
- Kumar, A., Srivastava, S., & Singh, S. (2005). Congestion management in competitive power market: A bibliographical survey. *Electric Power Systems Research*, 76(1), 153–164.
- Ledgerwood, S. D., & Pfeifenberger, J. P. (2013). Using virtual bids to manipulate the value of Financial Transmission Rights. *The Electricity Journal*, 26(9), 9–25.
- Leslie, G. W. (2021). Who benefits from ratepayer-funded auctions of Transmission Congestion Contracts? Evidence from New York. *Energy Economics*, 93, 105025.

- Longstaff, F. A., & Wang, A. W. (2004). Electricity forward prices: A high-frequency empirical analysis. *The journal of finance*, 59(4), 1877–1900.
- Lyons, K., Fraser, H., & Parmesano, H. (2000). An Introduction to Financial Transmission Rights. *The Electricity Journal*, 13(10), 31–37.
- Ma, X., Sun, D. I., Rosenwald, G. W., & Ott, A. L. (2003). Advanced Financial Transmission Rights in the PJM market. *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No. 03CH37491)*, 2, 1031–1038.
- Mount, T. D., & Ju, J. (2014). An econometric framework for evaluating the efficiency of a market for Transmission Congestion Contracts. *Energy Economics*, 46, 176–185.
- NYISO. (2021). Transmission Congestion Contracts Manual.
- Olmstead, D. E. (2018). Ontario’s auction market for Financial Transmission Rights: An analysis of its efficiency. *The Energy Journal*, 39(1).
- Opgrand, J., Preckel, P. V., Gotham, D. J., & Liu, A. L. (2022). Price formation in auctions for Financial Transmission Rights. *The Energy Journal*, 43(3).
- Prete, C. L., Guo, N., & Shanbhag, U. V. (2018). Virtual bidding and Financial Transmission Rights: An equilibrium model for cross-product manipulation in electricity markets. *IEEE Transactions on Power Systems*, 34(2), 953–967.
- Rudkevich, A., Hausman, E., Tabors, R., Bagnall, J., & Kopel, C. (2005). Loss hedging rights: A final piece in the LMP puzzle. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 60c–60c.
- Sarkar, V., & Khaparde, S. A. (2008). A comprehensive assessment of the evolution of Financial Transmission Rights. *IEEE Transactions on Power Systems*, 23(4), 1783–1795.
- Sarkar, V., & Khaparde, S. A. (2009). Introduction to loss-hedging Financial Transmission Rights. *IEEE Transactions on Power Systems*, 24(2), 621–630.

- Siddiqui, A. S., Bartholomew, E. S., Marnay, C., & Oren, S. S. (2005). Efficiency of the New York independent system operator market for Transmission Congestion Contracts. *Managerial Finance*, 31(6), 1–45.
- Singla, A., Singh, K., & Yadav, V. K. (2014). Transmission congestion management in deregulated environment: A bibliographical survey. *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, 1–10.
- Zhang, N. (2009). Market performance and bidders' bidding behavior in the New York Transmission Congestion Contract market. *Energy Economics*, 31(1), 61–68.

Chapter 3

Computationally-Efficient Variance Filtering in Multidimensional Affine Models

Abstract¹

We design a precise analytical approximation to filter a multi-factor variance model in a multi-asset framework. Building on the informativeness of intraday statistics, such as the realized covariance, we propose a highly tractable approach to filtering the first four conditional moments of the latent state. Our filter is robust to nonlinearity and non-Gaussian measurement and state equations while remaining analytical.

¹Joint work with Christian Dorion and Geneviève Gauthier. Dorion and Gauthier are affiliated with HEC Montréal.

3.1 Introduction

Commonality in asset volatilities is extensively documented and goes well beyond stocks on a given market sharing some systematic component. Commonality is observed across several markets (Aït-Sahalia and Xiu, 2016; Andersen et al., 2013), and across idiosyncratic volatilities within a given market (Herskovic et al., 2016). We develop a filter that allows for multiple latent variance factors within a multi-asset framework, addressing a significant challenge in financial modeling. Our precise analytical filter approximation circumvents the curse of dimensionality inherent in existing techniques, handles nonlinear frameworks and filter the first four conditional moment of the latent variance components.

In single-asset analyses, stochastic volatility models with multiple components are standard. Methods like the auxiliary particle filter (APF) and Markov chain Monte Carlo (MCMC) integrate latent states and can theoretically extend to multi-asset frameworks. However, the curse of dimensionality hinders their practical application to multiple assets with shared volatility components. In this paper, we build upon existing literature of analytical filtering to offer a practical and fast solution to multivariate filtering problems.

To identify latent factors driving common volatility movements in multiple assets, we exploit intraday data. Our methodology relies on two types of observable variables: the daily forward log-return on each asset and the corresponding realized variances and covariances aggregated from 5-minute returns (Barndorff-Nielsen and Shephard, 2002, 2004). Realized covariance provides valuable insights into the relationships between asset volatilities, complementing the information captured by realized variances within each asset.

A popular application of realized variance is in the powerful Heterogeneous Autoregressive (HAR) model of Corsi, 2009. When it comes to forecasting daily realized volatilities, the HAR is both straightforward to implement and difficult to outperform. The HARQ model (Bollerslev et al., 2016) improves on the HAR by allowing for a time-varying AR parameter that depends linearly on an estimate of realized quarticity.² Bollerslev et al., 2018 use the HARQ model to predict not only variances, but also a full covariance matrix.

²Other extensions exist, such as the HAR-Kalman (HARK) model of Buccheri and Corsi, 2021, which uses a Kalman filter to diminish the effect of measurement errors in parameter estimation, an issue addressed by Bollerslev et al., 2016.

As the focus of these studies is on predicting (co)variances, our analysis complements them by proposing a filtering methodology that 1) disentangles the common variance factors, 2) incorporates the asymptotic behavior of measurement errors, and 3) offers a precise analytical approximation of the latent state conditional moments. A reliable filter can prove crucial for parameter estimation based on conditional moments,³ forecasting volatility⁴ and derivative pricing. Filtering is also essential for understanding the latent factors driving the underlying covariances. These factors should ideally be rooted in economic theory to ensure that they are economically sound.⁵

Our filter builds on the extensive literature of term structure and extends it to the stock market using realized variables. Specifically, when the model is in the affine class, it allows us to have conditional moments with respect to the model filtration that are affine in the latent variable. By incorporating the measurement error of the realized variables, we extend the filter's use beyond affine classes, as it is not linear in the latent state.

We calculate the first four conditional moments of observables and latent variables with respect to the model filtration. Using the law of total cumulants permits a seamless transition from model filtration to observed filtration. We propose an update to the latent variable estimator based on Kalman filter foundations without using classical Kalman equations. This includes calculating the expectation, variance, third and fourth moments of the latent variable, providing comprehensive estimators for the first four cumulants, including co-moments.

Additionally, we propose a correction for conditional moments of measurement error

³Conditional moments are useful tools to estimate parameters by GMM. Bollerslev and Zhou, 2002 show how to use the conditional moments of the realized variance to estimate the parameters through a GMM procedure. However, as they do not filter the latent state, their methodology becomes rapidly cumbersome in multivariate. Filtering the latent state could permit extending their work in multivariate settings more effectively.

⁴Buccheri and Corsi, 2021 shows that their HARK model estimates parameters better, especially the one-lag autoregressive which is a problem pointed out in Bollerslev et al., 2016, and leads to smaller MSE than both the HARQ and HAR models in MC study. They also present other models such as the SHAR and SHARK.

⁵Filtering these factors provide a much better understanding of the comovements between assets, with future potential applications such as options pricing. Indeed, understanding which risk factors contribute to the current high (low) covariance could also better inform the next courses of action; a practical example is the impact of stock market volatility in commodities discussed in Christoffersen et al., 2019.

at lower frequencies, as they do not reach their asymptotic distribution at these frequencies. This leads to an unbiased filter for the first four moments of the latent state. Monte Carlo simulations demonstrate our filter's unbiasedness and precision across various scenarios, validating it as a robust and accurate methodology.

In summary, we contribute to the existing literature in four ways. First, our analytical approach allows for rapid execution while handling increasing dimensionality, and despite being an approximation, the filter remains remarkably precise. Second, it accommodates nonlinear and non-Gaussian frameworks, which can be further extended to other financial applications. Third, we address the realized variance measurement error in intraday frequency, proposing a correction to account for its not-yet-attained asymptotic behavior. Lastly, our methodology provides not only the first two conditional moments of the latent variables but also extends to the fourth cumulant, including all co-moments of the latent variables, providing a comprehensive characterization of the instantaneous variance components.

The paper is organized as follows. After the literature review, Section 3.2 presents the continuous model, its discretization, and the observable variables. Section 3.3 derives the filter. Section 3.4 presents the Monte Carlo simulation study, and Section 3.5 concludes the paper.

Literature review

The Kalman Filter (KF), introduced by Kalman, 1960, has been extensively applied in finance due to its effectiveness in providing optimal linear state estimates. However, financial models are often non-linear with conditional non-normal innovations posing significant challenges for the standard KF.

Popular extensions of the KF include the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF). The EKF uses a first-order Taylor expansion to linearize the state and measurement equations.⁶ While this approach mitigates some non-linearities, it can introduce approximation errors, although Li and Zhao, 2006 derived a bias correction

⁶See Cortazar et al., 2017, Li and Zhao, 2006, and Trolle and Schwartz, 2009 for applications in finance.

term for their quadratic model. The UKF, introduced by Wan and Van Der Merwe, 2000 and Julier et al., 2000, utilizes a deterministic sampling technique to better capture the mean and covariance estimates of non-linear functions.

Christoffersen et al., 2014 and Javaheri et al., 2003 compare the performance of the standard KF, its non-linear extensions, and the Particle Filter (PF) on financial problems, highlighting the limitations of the standard KF in handling non-linearities. Christoffersen et al., 2014 findings suggest that the UKF outperforms the KF in handling non-linearities while being less computationally intensive compared to the PF. In this paper, we introduce an analytical filter that can effectively manage non-linearities without the need for linearization, sigma points, or particles.

Given that our variance process is based on the Heston model, there is a connection to the vast literature on the term structure of interest rates. For instance, Duan and Simonato, 1999 address the non-linearity in the state equation, which follows a CIR process, by deriving a quasi-optimal KF that does not require any Taylor expansion while remaining analytical. This approach is further explored in the works of Jong, 2000, Chen and Scott, 2003, Duffee and Stanton, 2012, Monfort et al., 2017⁷ among others. However, these methods assume linear measurement equations, distinguishing them from the approach presented in this paper where we generalized the non-linearity of both the measurement and state equation.

Feunou and Okou, 2018 apply a modified KF to the stock market (S&P 500) options. They construct risk-neutral cumulants using options data which have a linear relationship in the latent state and the latent state variance also depending on the latent state. This work is extended by Brignone et al., 2023 and Feunou et al., 2020.

Li and Zhao, 2006 study a multifactor quadratic term structure models. They use EKF as an estimation methodology and propose a bias correction for the linearisation. Monfort et al., 2015 presents a Quadratic Kalman Filter (QKF) where the filter covers the case when the measurement equation is quadratic in the latent state but with a linear state-space equation. They shows that it results in a better fit than the EKF and UKF. They create an augmented latent state vector, which includes the squared latent state to linearize the

⁷Refer to Appendix A.5.1 for a thorough discussion on the subject.

measurement equation. They also discuss potential curse of dimensionality problems, as the filtered vector is of size $N(N + 1)$. The filter proposes a practical solution to quadratic measurement error which is applied in Dubecq et al., 2016 and Roussellet, 2023.

The subtlety in our framework is that we do not need all the conditional cumulants of the state and measurement equations with respect to the model filtration (where the latent state is observed) to be strictly linear in the latent state. We show how to pass from the model filtration to the observed filtration using the law of total expectation, variance and cumulance. As we have a measurement error regarding the realized variance, the errors conditional variance are quadratic in the latent state, which is in continuity with the model presented in Monfort et al., 2015. Furthermore, we derived the higher moments of the latent state.

Even though the KF and its extentions are a popular choice, other methodologies have been studied, such as in Bates, 2006, 2012, 2019. Approximate Maximum Likelihood (AML) is introduced for parameter estimation and latent state filtering, achieving through the recursive update of the characteristic function. However, AML is less ideal for scenarios involving multiple data sources due to its reliance on numerical integration, thus suffering from the curse of dimensionality. Additionally, other filters such as particle filters, as demonstrated in Bégín et al., 2020, can filter the latent factors even with variance jumps. However, PF faces challenges related to dimensionality and computational intensity as the number of latent states increases. The likelihood function is not smooth with respect to the parameters, making particle filters less suitable in a multivariate context, especially when accommodating a large number of latent states.

Alizadeh et al., 2002 study the log range as a volatility proxy instead of the log absolute returns or others proxies, which is closer to a Gaussian distribution. They use the KF to retrieve information and acknowledge that if the prediction errors are not Gaussian, the projections of the KF are not the conditional expectation, which is a challenge we had to deal with in our framework. They estimate parameters through Gaussian maximum likelihood estimation. However, the further from the Gaussian distribution the errors are, the methodology is less effective. They also found that the two factor model performs better than one factor model on empirical study on exchange rates. Calzolari et al., 2021 study

log RV series in high dimensions, including commonality in their dynamics. They propose a dynamic factor model (DFM) that does not suffer from the curse of dimensionality and captures long persistence. However, working with log RV series was not an option in our framework, as we wanted to leverage the knowledge of the realized covariance to filter latent states. Since intraday statistics can be negative, this makes the log transformation unsuitable.

3.2 The Model

We consider the prices of N_S assets, influenced by N_V common volatility factors following a Heston, 1993 dynamics,

$$dS_{i,t} = \alpha_{i,t} S_{i,t} dt \tag{3.1}$$

$$+ S_t \left(\sum_{k=1}^{N_V} c_{i,k} \sqrt{V_{k,t}} \left(\rho_k dW_{k,t} + \sqrt{1 - \rho_k^2} dB_{k,t} \right) \right), \quad i \in \{1, 2, \dots, N_S\},$$

$$dV_{k,t} = \kappa_k (\theta_k - V_{k,t}) dt + \sigma_k \sqrt{V_{k,t}} dW_{k,t}, \quad k \in \{1, 2, \dots, N_V\}, \tag{3.2}$$

where $\{W_{k,t}, B_{k,t}\}_{k=1}^{N_V}$ are independent \mathbb{P} -Brownian motions. The continuous-time model filtration $(\mathcal{F}_t^C)_{t \geq 0}$ is generated by

$$\mathcal{F}_t^C = \sigma \left(\{W_{k,s}, B_{k,s}\}_{k=1}^{N_V} \mid s \leq t \right).$$

The ρ_k parameters control the correlation between the variance and price shocks, introducing a multivariate analogue to the so-called leverage effect.

To stay in the affine class, we assume that the drift is linear in the state variables. Namely

$$\alpha_{i,t} = r_t - q_{i,t} + \sum_{k=1}^{N_V} c_{ik} \lambda_k V_{k,t}, \tag{3.3}$$

where r_t is the risk-free rate, $q_{i,t}$ is the dividend rate, and λ_k are the prices of risk the associated with the volatility factors. Consequently, the time- t log-forward price for a

contract on asset i expiring at time T evolves according to⁸

$$d \log F_{i,t} = \sum_{k=1}^{N_V} \left(c_{ik} \lambda_k - \frac{c_{ik}^2}{2} \right) V_{k,t} dt + c_{ik} \sqrt{V_{k,t}} d \left(\rho_k W_{k,t} + \sqrt{1 - \rho_k^2} B_{k,t} \right). \quad (3.4)$$

3.2.1 Discretization

The aim is to use intraday data to filter variances at the daily frequency. We consider the daily forward log-variation. The log-forward price's daily quadratic covariation is given by

$$\begin{aligned} QCV_{i,j,t,t+1} &= \langle \log F_{i,\cdot,T}, \log F_{j,\cdot,T} \rangle_{(t+1)\Delta} - \langle \log F_{i,\cdot,T}, \log F_{j,\cdot,T} \rangle_{t\Delta} \\ &= \sum_{k=1}^{N_V} c_{i,k} c_{j,k} \mathcal{V}_{k,t,t+1}, \end{aligned} \quad (3.5)$$

where $\Delta = \frac{1}{252}$, and the daily integrated variances are

$$\mathcal{V}_{k,t,t+1} = \int_{t\Delta}^{(t+1)\Delta} V_{k,s} ds, \quad k \in \{1, 2, \dots, N_V\}.$$

The product $c_{i,k} c_{j,k}$ parameterize the level and sign of the covariances.

As usual, filtering the latent variables involves state and measurement equations. Whereas these equations, in their matrix form, are valid for any number of assets and factors, we illustrate them below for the nested case involving two assets ($N_S = 2$) and three variance factors ($N_V = 3$). We refer to this example as the 2×3 economy.

The state equation

Each daily integrated variance satisfies

$$\mathcal{V}_{k,t,t+1} = \theta_k \Delta + (V_{k,t\Delta} - \theta_k) \frac{(1 - e^{-\kappa_k \Delta})}{\kappa_k} + \frac{\sigma_k}{\kappa_k} (\mathcal{W}_{k,t,t+1} - \mathcal{Z}_{k,t,t+1}), \quad (3.6)$$

⁸This standard result is derived in Online Appendix D.1.

where

$$\begin{aligned}\mathcal{W}_{k,t,t+1} &= \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{k,s}} dW_{k,s}, & \mathcal{B}_{k,t,t+1} &= \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{k,s}} dB_{k,s}, \\ \mathcal{Z}_{k,t,t+1} &= \int_{t\Delta}^{(t+1)\Delta} e^{-\kappa_k((t+1)\Delta-s)} \sqrt{V_{k,s}} dW_{k,s}.\end{aligned}$$

The matrix form of the stochastic integrals are

$$\mathbf{W}_{t,t+1} = \begin{bmatrix} \mathcal{W}_{1,t,t+1} \\ \mathcal{W}_{2,t,t+1} \\ \mathcal{W}_{3,t,t+1} \end{bmatrix}, \quad \mathbf{Z}_{t,t+1} = \begin{bmatrix} \mathcal{Z}_{1,t,t+1} \\ \mathcal{Z}_{2,t,t+1} \\ \mathcal{Z}_{3,t,t+1} \end{bmatrix}, \quad \mathbf{B}_{t,t+1} = \begin{bmatrix} \mathcal{B}_{1,t,t+1} \\ \mathcal{B}_{2,t,t+1} \\ \mathcal{B}_{3,t,t+1} \end{bmatrix}, \quad (3.7)$$

where $\mathbf{W}_{t,t+1}$, $\mathbf{Z}_{t,t+1}$ and $\mathbf{B}_{t,t+1}$ are non-linear functions of the state variables $V_{k,t\Delta}$. However, the conditional moments of the stochastic integrals are linear in the state variables.⁹

The matrix form of the integrated variance is

$$\begin{aligned}\mathbf{V}_{t,t+1} &= \mathbf{G} + \mathbf{H}\mathbf{V}_{t\Delta} + \mathbf{J}\mathbf{W}_{t,t+1} + \mathbf{U}\mathbf{Z}_{t,t+1}, & (3.8) \\ \xrightarrow{2 \times 3} \begin{bmatrix} \mathcal{V}_{1,t,t+1} \\ \mathcal{V}_{2,t,t+1} \\ \mathcal{V}_{3,t,t+1} \end{bmatrix} &= \begin{bmatrix} \theta_1\Delta - \frac{(1-e^{-\kappa_1\Delta})}{\kappa_1}\theta_1 \\ \theta_2\Delta - \frac{(1-e^{-\kappa_2\Delta})}{\kappa_2}\theta_2 \\ \theta_3\Delta - \frac{(1-e^{-\kappa_3\Delta})}{\kappa_3}\theta_3 \end{bmatrix} + \begin{bmatrix} \frac{(1-e^{-\kappa_1\Delta})}{\kappa_1} & 0 & 0 \\ 0 & \frac{(1-e^{-\kappa_2\Delta})}{\kappa_2} & 0 \\ 0 & 0 & \frac{(1-e^{-\kappa_3\Delta})}{\kappa_3} \end{bmatrix} \begin{bmatrix} V_{1,t\Delta} \\ V_{2,t\Delta} \\ V_{3,t\Delta} \end{bmatrix} \\ &+ \begin{bmatrix} \frac{\sigma_1}{\kappa_1} & 0 & 0 \\ 0 & \frac{\sigma_2}{\kappa_2} & 0 \\ 0 & 0 & \frac{\sigma_3}{\kappa_3} \end{bmatrix} \begin{bmatrix} \mathcal{W}_{1,t,t+1} \\ \mathcal{W}_{2,t,t+1} \\ \mathcal{W}_{3,t,t+1} \end{bmatrix} + \begin{bmatrix} -\frac{\sigma_1}{\kappa_1} & 0 & 0 \\ 0 & -\frac{\sigma_2}{\kappa_2} & 0 \\ 0 & 0 & -\frac{\sigma_3}{\kappa_3} \end{bmatrix} \begin{bmatrix} \mathcal{Z}_{1,t,t+1} \\ \mathcal{Z}_{2,t,t+1} \\ \mathcal{Z}_{3,t,t+1} \end{bmatrix}\end{aligned}$$

Integrating Equation (3.2) over a day and replacing $\mathcal{V}_{k,t,t+1}$ with Equation (3.6), the instantaneous variance satisfies

$$\begin{aligned}V_{k,(t+1)\Delta} &= V_{k,t\Delta} + \theta_k\kappa_k\Delta - \kappa_k\mathcal{V}_{k,t,t+1} + \sigma_k\mathcal{W}_{k,t,t+1} \\ &= V_{k,t\Delta}e^{-\kappa_k\Delta} + (1 - e^{-\kappa_k\Delta})\theta_k + \sigma_k\mathcal{Z}_{k,t,t+1}, & (3.9)\end{aligned}$$

⁹See Appendix B.2 and Online Appendix D.6 for details.

which leads to the **state equation**

$$\begin{aligned}
 \mathbf{V}_{(t+1)\Delta} &= \mathbf{F} + \mathbf{KV}_{t\Delta} + \mathbf{SZ}_{t,t+1}, & (3.10) \\
 \stackrel{2 \times 3}{\implies} \begin{bmatrix} V_{1,(t+1)\Delta} \\ V_{2,(t+1)\Delta} \\ V_{3,(t+1)\Delta} \end{bmatrix} &= \begin{bmatrix} (1 - e^{-\kappa_1\Delta})\theta_1 \\ (1 - e^{-\kappa_2\Delta})\theta_2 \\ (1 - e^{-\kappa_3\Delta})\theta_3 \end{bmatrix} + \begin{bmatrix} e^{-\kappa_1\Delta} & 0 & 0 \\ 0 & e^{-\kappa_2\Delta} & 0 \\ 0 & 0 & e^{-\kappa_3\Delta} \end{bmatrix} \begin{bmatrix} V_{1,t\Delta} \\ V_{2,t\Delta} \\ V_{3,t\Delta} \end{bmatrix} \\
 &+ \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \begin{bmatrix} \mathcal{Z}_{1,t,t+1} \\ \mathcal{Z}_{2,t,t+1} \\ \mathcal{Z}_{3,t,t+1} \end{bmatrix}.
 \end{aligned}$$

Whereas Equation (3.10) is reminiscent of the state transition in a Kalman filter, it differs along two important dimensions: 1) the error term is a nonlinear function of the latent state because $\sqrt{V_{t\Delta}}$ appears in the stochastic integral $\mathcal{Z}_{t,t+1}$ and 2) the noise component $\mathcal{Z}_{t,t+1}$ is not Gaussian because the weak solution of the SDE in Equation (3.2) implies that $V_{k,t\Delta}$ is distributed according to a non central χ^2 .

The measurement equation

From Equation (3.4), the first set of observable variables are the daily forward log-variation

$$\begin{aligned}
 R_{i,t,t+1} &= \log F_{i,(t+1)\Delta} - \log F_{i,t\Delta}, \quad i \in \{1, 2\}, & (3.11) \\
 &= \sum_{k=1}^{N_V} \left(c_{ik} \lambda_k - \frac{c_{ik}^2}{2} \right) \mathcal{V}_{k,t,t+1} + \sum_{k=1}^{N_V} c_{ik} \left(\rho_k \mathcal{W}_{k,t,t+1} + \sqrt{1 - \rho_k^2} \mathcal{B}_{k,t,t+1} \right).
 \end{aligned}$$

The realized variances and covariances associated with assets i and j are computed from N intraday intervals from the forward log prices of maturity T :

$$RCV_{i,j,t,t+1} = \sum_{n=1}^N \left(\log \frac{F_{i,t\Delta + \frac{n\Delta}{N}}}{F_{i,t\Delta + \frac{(n-1)\Delta}{N}}} \right) \left(\log \frac{F_{j,t\Delta + \frac{n\Delta}{N}}}{F_{j,t\Delta + \frac{(n-1)\Delta}{N}}} \right). \quad (3.12)$$

Equation (3.12) asymptotically converges to the weighted sum of the integrated variances.

The link between the realized covariance measures and the variance components is

$$RCV_{i,j,t,t+1} = \underbrace{\sum_{k=1}^{N_V} c_{i,k} c_{j,k} \mathcal{V}_{k,t,t+1}}_{QCV_{i,j,t,t+1}} + \xi_{i,j,t,t+1}, \quad (3.13)$$

where the asymptotic theory of Barndorff-Nielsen and Shephard, 2002 and Barndorff-Nielsen and Shephard, 2004 of the measurement error $\xi_{i,j,t,t+1}$ are detailed in Appendix B.1.1.

The matrix form of the observable variables is

$$\begin{aligned} \mathbf{O}_{(t+1)\Delta} &= \mathbf{A} \mathbf{V}_{t,t+1} + \mathbf{C} \mathbf{W}_{t,t+1} + \mathbf{D} \mathbf{B}_{t,t+1} + \mathbf{Q} \boldsymbol{\xi}_{t,t+1}, \\ \xrightarrow{2 \times 3} \begin{bmatrix} R_{1,t,t+1} \\ R_{2,t,t+1} \\ RCV_{1,2,t,t+1} \\ RCV_{2,2,t,t+1} \\ RCV_{1,2,t,t+1} \end{bmatrix} &= \begin{bmatrix} c_{11}\lambda_1 - \frac{c_{11}^2}{2} & c_{12}\lambda_2 - \frac{c_{12}^2}{2} & c_{13}\lambda_3 - \frac{c_{13}^2}{2} \\ c_{21}\lambda_1 - \frac{c_{21}^2}{2} & c_{22}\lambda_2 - \frac{c_{22}^2}{2} & c_{23}\lambda_3 - \frac{c_{23}^2}{2} \\ c_{1,1}^2 & c_{1,2}^2 & c_{1,3}^2 \\ c_{2,1}^2 & c_{2,2}^2 & c_{2,3}^2 \\ c_{1,1}c_{2,1} & c_{1,2}c_{2,2} & c_{1,3}c_{2,3} \end{bmatrix} \begin{bmatrix} \mathcal{V}_{1,t,t+1} \\ \mathcal{V}_{2,t,t+1} \\ \mathcal{V}_{3,t,t+1} \end{bmatrix} \\ &+ \begin{bmatrix} c_{11}\sqrt{(1-\rho_1^2)} & c_{12}\sqrt{(1-\rho_2^2)} & c_{13}\sqrt{(1-\rho_3^2)} \\ c_{21}\sqrt{(1-\rho_1^2)} & c_{22}\sqrt{(1-\rho_2^2)} & c_{23}\sqrt{(1-\rho_3^2)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{B}_{1,t,t+1} \\ \mathcal{B}_{2,t,t+1} \\ \mathcal{B}_{3,t,t+1} \end{bmatrix} \\ &+ \begin{bmatrix} c_{11}\rho_1 & c_{12}\rho_2 & c_{13}\rho_3 \\ c_{21}\rho_1 & c_{22}\rho_2 & c_{23}\rho_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{W}_{1,t,t+1} \\ \mathcal{W}_{2,t,t+1} \\ \mathcal{W}_{3,t,t+1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_{1,1,t,t+1} \\ \xi_{2,2,t,t+1} \\ \xi_{1,2,t,t+1} \end{bmatrix}. \end{aligned}$$

Replacing the integrated variance with Equation (3.8),

$$\begin{aligned} \mathbf{O}_{(t+1)\Delta} &= \mathbf{A}\mathbf{G} + \mathbf{A}\mathbf{H}\mathbf{V}_{t\Delta} + \mathbf{A}\mathbf{U}\mathbf{Z}_{t,t+1} \\ &+ (\mathbf{A}\mathbf{J} + \mathbf{C})\mathbf{W}_{t,t+1} + \mathbf{D}\mathbf{B}_{t,t+1} + \mathbf{Q}\boldsymbol{\xi}_{t,t+1}. \end{aligned} \quad (3.14)$$

The **state Equation** (3.10) establishes a connection between the updated latent state $\mathbf{V}_{(t+1)\Delta}$ and the model constituents $\mathbf{V}_{t\Delta}$ and $\mathbf{Z}_{t,t+1}$. The **measurement Equation** (3.14) links the observable variables $\mathbf{O}_{(t+1)\Delta}$ and the model components $\mathbf{V}_{t\Delta}$, $\mathbf{W}_{t,t+1}$, $\mathbf{Z}_{t,t+1}$, and $\mathbf{B}_{t,t+1}$

along with the error measurement $\xi_{t,t+1}$. Neither the measurement errors and their conditional moments are linear in the latent state and from a Gaussian distribution, posing an additional challenge to our filter.

3.3 The Filter

3.3.1 An overview

Although our filter is an analytical approximation, it is extremely precise and does not require numerical integrals, particles, sigma points, or linearization. It can handle high dimensionality in both latent and observable variables without facing the curse of dimensionality. Furthermore, we propose updates for the first four conditional moments of the latent state, which offers a proper characterization of the variance components.

The stochastic integrals and measurement error induce a non-linear relationship with the latent state for both the measurement and state equations. Therefore, we build upon the work of Duan and Simonato, 1999 among others to extend their modified filter to a more general class of models and present an analytical approximation to the multivariate stochastic volatility filtering problem.

The discrete model filtration $\{\mathcal{F}_{t\Delta}\}_{t=0}^T$ associated to Equations (3.10) and (3.14) is generated by the stochastic integrals:

$$\mathcal{F}_{t\Delta} = \sigma\{\mathbf{W}_{s-1,s}, \mathbf{Z}_{s-1,s}, \mathbf{B}_{s-1,s}, \xi_{s-1,s}\}_{s=1}^t. \quad (3.15)$$

The econometrician filtration $\{\mathcal{G}_{t\Delta}\}_{t=0}^T$ is defined by the observable variables where $\mathcal{G}_{t\Delta} = \sigma\{\mathbf{O}_s\}_{s=1}^t$. Because the instantaneous variances are unobservable variables, the econometrician filtration \mathcal{G} is coarser than the discrete model filtration \mathcal{F} .

We do not require that the conditional moments of $\mathbf{O}_{(t+1)\Delta}$ and $\mathbf{V}_{(t+1)\Delta}$ with respect to the model filtration $\mathcal{F}_{t\Delta}$, $\mathbf{O}_{(t+1)\Delta}|\mathcal{F}_{t\Delta}$ and $\mathbf{V}_{(t+1)\Delta}|\mathcal{F}_{t\Delta}$, to have a strictly linear relationship with $\mathbf{V}_{t\Delta}$ (see Subsection 3.3.2).¹⁰

¹⁰More precisely, the conditional expectation should be linear in the latent state, the conditional variance

This is a generalization of the current literature as we include non-linearity in the measurement equation¹¹ with the stochastic integrals and the measurement error. We do not need to augment the latent state as in Monfort et al., 2015 to deal with the quadratic relationship, which is a potential issue as their augmented latent state dimensionality can grow rapidly.

Furthermore, we formalize the use of the law of total cumulance to transition from the model filtration $\mathcal{F}_{t\Delta}$ to the observed filtration $\mathcal{G}_{t\Delta}$ (see Subsection 3.3.3), allowing us to accurately express the conditional moments of the measurement $\mathbf{O}_{(t+1)\Delta}|\mathcal{G}_{t\Delta}$ and state equations $\mathbf{V}_{(t+1)\Delta}|\mathcal{G}_{t\Delta}$ in terms of the observed conditional moments of the latent state $\mathbf{V}_{t\Delta}|\mathcal{G}_{t\Delta}$.

With this in hand, we built (see Subsection 3.3.4) our update equations $\mathbf{V}_{(t+1)\Delta}|\mathcal{G}_{(t+1)\Delta}$ upon the fundamentals of the Kalman Filter, enabling us to reach a precise analytical approximation despite the nonlinearities and non-Gaussian framework while providing up to the first four conditional cumulants of the updated latent state $\mathbf{V}_{(t+1)\Delta}$.

3.3.2 Conditional moments of $\mathbf{V}_{(t+1)\Delta}$ and $\mathbf{O}_{(t+1)\Delta}$ with respect to $\mathcal{F}_{t\Delta}$

As mentioned in Appendix B.2, $\mathcal{W}_{t,t+1}$, $\mathcal{Z}_{t,t+1}$ and $\mathcal{B}_{t,t+1}$ are centered at zero and have a conditional covariance matrix that is a linear function of the instantaneous variances. That is

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \mathcal{W}_{t,t+1} \\ \mathcal{B}_{t,t+1} \\ \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \begin{bmatrix} \mathbf{P}_0^{\mathcal{W}\mathcal{W}} + \mathbf{P}_1^{\mathcal{W}\mathcal{W}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}) & \mathbf{0} & \mathbf{P}_0^{\mathcal{W}\mathcal{Z}} + \mathbf{P}_1^{\mathcal{W}\mathcal{Z}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}) \\ \mathbf{0} & \mathbf{P}_0^{\mathcal{B}\mathcal{B}} + \mathbf{P}_1^{\mathcal{B}\mathcal{B}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}) & \mathbf{0} \\ \mathbf{P}_0^{\mathcal{W}\mathcal{Z}} + \mathbf{P}_1^{\mathcal{W}\mathcal{Z}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}) & \mathbf{0} & \mathbf{P}_0^{\mathcal{Z}\mathcal{Z}} + \mathbf{P}_1^{\mathcal{Z}\mathcal{Z}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}) \end{bmatrix}, \end{aligned} \quad (3.16)$$

where \odot is an element-wise matrix multiplication and $\mathbf{1}_{1 \times 3} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. Since the Brownian motions are independent under the model filtration, each matrices \mathbf{P}_0 and \mathbf{P}_1

should be at most quadratic in the latent state and the conditional third central moment cubic in the latent state.

¹¹The CIR process as the latent state is covered in the term structure literature.

in Equation (3.16) are diagonal as derived in Equation (B.10).

As detailed in Appendix B.1, the measurement errors on the realized variances and covariances are assumed to be uncorrelated with the model components $\mathbf{Z}_{t,t+1}$, $\mathbf{W}_{t,t+1}$, and $\mathbf{B}_{t,t+1}$. Furthermore, since $\mathbb{E}[\boldsymbol{\xi}_{t,t+1} | \mathcal{F}_{t\Delta}]$ is zero, the second moment $\text{Var}[\boldsymbol{\xi}_{t,t+1} | \mathcal{F}_{t\Delta}]$ is

$$\begin{aligned} & \text{Cov}[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{k,\tilde{k}}^{i,\tilde{i},j,\tilde{j}} (\mathcal{A}_{k,\tilde{k}} + \mathcal{B}_{k,\tilde{k}} V_{k,t\Delta} V_{\tilde{k},t\Delta} + \mathcal{D}_{k,\tilde{k}} V_{k,t\Delta} + \mathcal{E}_{k,\tilde{k}} V_{\tilde{k},t\Delta}), \end{aligned} \quad (3.17)$$

where the details about the notation are provided in are in Equation (B.6).

Having established the fundamental building blocks, we can now utilize our state and measurement equations to derive the two conditional moments for $\mathbf{V}_{(t+1)\Delta}$ and $\mathbf{O}_{(t+1)\Delta}$ with respect to the model filtration. Based on Equation (3.14), the conditional expectation is

$$\mathbb{E}[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \mathbf{A}\mathbf{G} + \mathbf{A}\mathbf{H} \mathbf{V}_{t\Delta}, \quad (3.18)$$

and, using Equations (3.14) and (3.16), the conditional variance is

$$\begin{aligned} \text{Var}[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \mathbf{A}\mathbf{U} \text{Var}[\mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{U})^\top + \mathbf{D} \text{Var}[\mathbf{B}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbf{D}^\top \\ &+ (\mathbf{A}\mathbf{J} + \mathbf{C}) \text{Var}[\mathbf{W}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{J} + \mathbf{C})^\top \\ &+ (\mathbf{A}\mathbf{J} + \mathbf{C}) \text{Cov}[\mathbf{W}_{t,t+1}, \mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{U})^\top \\ &+ \mathbf{A}\mathbf{U} \text{Cov}[\mathbf{Z}_{t,t+1}, \mathbf{W}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{J} + \mathbf{C})^\top + \mathbf{Q} \text{Var}[\boldsymbol{\xi}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbf{Q}^\top. \end{aligned} \quad (3.19)$$

Based on Equation (3.16),

$$\begin{aligned} \text{Var}[\mathbf{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \mathbf{P}_0^{\mathcal{W}\mathcal{W}} + \mathbf{P}_1^{\mathcal{W}\mathcal{W}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}), \\ \text{Var}[\mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \mathbf{P}_0^{\mathcal{Z}\mathcal{Z}} + \mathbf{P}_1^{\mathcal{Z}\mathcal{Z}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}), \\ \text{Var}[\mathbf{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \mathbf{P}_0^{\mathcal{B}\mathcal{B}} + \mathbf{P}_1^{\mathcal{B}\mathcal{B}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}), \end{aligned}$$

$$\text{Cov}[\mathbf{W}_{t,t+1}, \mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = \mathbf{P}_0^{\mathcal{WZ}} + \mathbf{P}_1^{\mathcal{WZ}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}).$$

From Equation (3.10), the conditional moments of the latent variables are

$$\mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \mathbf{F} + \mathbf{K}\mathbf{V}_{t\Delta}, \quad (3.20)$$

$$\text{Var}[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \mathbf{S}\text{Var}[\mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}]\mathbf{S}^\top = \mathbf{S}(\mathbf{P}_0^{\mathcal{ZZ}} + \mathbf{P}_1^{\mathcal{ZZ}} \odot (\mathbf{V}_{t\Delta} \mathbf{1}_{1 \times 3}))\mathbf{S}^\top. \quad (3.21)$$

Using Equations (3.14) and (3.10), the conditional covariance matrix between the observable variables and the latent variables is

$$\begin{aligned} & \text{Cov}[\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] \\ &= \mathbf{S}\text{Var}[\mathbf{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{U})^\top + \mathbf{S}\text{Cov}[\mathbf{Z}_{t,t+1}, \mathbf{W}_{t,t+1} | \mathcal{F}_{t\Delta}] (\mathbf{A}\mathbf{J} + \mathbf{C})^\top. \end{aligned} \quad (3.22)$$

3.3.3 Conditional moments of $\mathbf{V}_{(t+1)\Delta}$ and $\mathbf{O}_{(t+1)\Delta}$ with respect to $\mathcal{G}_{t\Delta}$

With the help of the law of total expectation, law of total variance,¹² and Equations (3.18) and (3.19), the first two conditional moments of the observables variables are

$$\begin{aligned} \mathbb{E}[\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] &= \mathbb{E}[\mathbb{E}[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}], \\ \text{Var}[\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] &= \underbrace{\text{Var}[\mathbb{E}[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{= (\mathbf{A}\mathbf{H}) \text{Var}[\mathbf{V}_{t\Delta} | \mathcal{G}_{t\Delta}] (\mathbf{A}\mathbf{H})^\top} + \underbrace{\mathbb{E}[\text{Var}[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{\text{Expected value of Equation (3.19)}}. \end{aligned}$$

The second term measures the variability of $\mathbf{W}_{t,t+1}$, $\mathbf{Z}_{t,t+1}$, $\mathbf{B}_{t,t+1}$ and $\boldsymbol{\xi}_{t,t+1}$. From Equations (3.20) and (3.21), the conditional moments of the latent variables under the observed filtration are

$$\mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] = \mathbb{E}[\mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}],$$

¹²The law of total cumulance is a generalization of the law of total expectation and the law of total variance. It can be used to derive higher-order cumulants with respect to the observed filtration based on the cumulants with respect to the model filtration. For technical details, see https://en.wikipedia.org/wiki/Law_of_total_cumulance or Brillinger, 1969.

$$\text{Var} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] = \underbrace{\text{Var}[E[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{= \mathbf{K} \text{Var}[\mathbf{V}_{t\Delta} | \mathcal{G}_{t\Delta}] \mathbf{K}^\top} + \underbrace{E[\text{Var}[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{\text{Expected value of Equation (3.21)}}.$$

The second term measures the variability of $\mathcal{Z}_{t,t+1}$. Combining the law of total covariance and Equations (3.18), (3.20) and (3.22), the covariance matrix between the observable and the latent variables is

$$\begin{aligned} \text{Cov}[\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] &= \\ \underbrace{\text{Cov}[E[\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}], E[\mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{= \mathbf{K} \text{Var}[\mathbf{V}_{t\Delta} | \mathcal{G}_{t\Delta}] (\mathbf{A}\mathbf{H})^\top} &+ \underbrace{E[\text{Cov}[\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]}_{\text{Expected value of Equation (3.22)}}. \end{aligned}$$

The second term measures the variability of $\mathcal{W}_{t,t+1}$ and $\mathcal{Z}_{t,t+1}$.

3.3.4 Updating the latent variable $\mathbf{V}_{(t+1)\Delta}$ with respect to $\mathcal{G}_{(t+1)\Delta}$

To finalize the construction of the filter, we need an updated projection of the latent variables given the availability of the new set of observations $\mathbf{O}_{(t+1)\Delta}$. This involves adjusting the first two moments of $\mathbf{V}_{(t+1)\Delta}$ to accommodate the newly acquired information.

This process entails computing $E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]$ and $\text{Var}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]$, building upon the preceding expectation and variance represented by $E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]$ and $\text{Var}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]$ respectively.

The a posteriori approximation of the latent variables is obtained from its a priori prediction to which a correction term that account for the difference between the actual realization of the observable vectors and its a priori prediction is added:

$$\underbrace{E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]}_{\text{a posteriori estimation of the latent variables}} \simeq \underbrace{E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]}_{\text{a priori prediction of the latent variables}} + \underbrace{\Sigma_{t\Delta} (\mathbf{O}_{(t+1)\Delta} - E[\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}])}_{\epsilon_{t+1|t}^{\mathbf{O}}}, \quad (3.23)$$

where $E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]$ is our point estimate of our latent variables.¹³

¹³In cases where $E[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]$ has negative values, we distinguish here between the starting point in the filter update process and the variances used for application purposes. Therefore, we do as in Feunou et al., 2020 (p.493) and take the maximum between the right hand side of (3.23) and zero for application

As our measurement and state equation are neither conditional Gaussian nor linear, we are unable to analytically calculate the true conditional moment with respect to $\mathcal{G}_{t\Delta}$ as in the Kalman Filter. Indeed, as explained in Duan and Simonato, 1999, since the noise terms are not Gaussian, the latent variable update may differ from the true conditional mean and variance.¹⁴ We do not provide a theoretical background to quantify the discrepancy from each side of Equation (3.23). However, we show empirically in Section 3.4 that the filter is unbiased and precise.

The matrix $\Sigma_{t\Delta}$ of Equation (3.23) is constructed by minimizing the conditional variance of the latent variable estimator with respect to $\mathcal{G}_{t\Delta}$. Because

$$\underbrace{\mathbf{V}_{(t+1)\Delta} - \mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]}_{\text{filtering error: } \epsilon_{t+1|t+1}^{\mathbf{V}}} \simeq \underbrace{\mathbf{V}_{(t+1)\Delta} - \mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]}_{\epsilon_{t+1|t}^{\mathbf{V}}} - \Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}}, \quad (3.24)$$

the variance matrix of the estimator with respect to the information set $\mathcal{G}_{t\Delta}$ is

$$\begin{aligned} & \mathbb{E} \left[\left(\mathbf{V}_{(t+1)\Delta} - \mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}] \right) \left(\mathbf{V}_{(t+1)\Delta} - \mathbb{E}[\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}] \right)^\top \middle| \mathcal{G}_{t\Delta} \right] \\ & \simeq \mathbb{E} \left[\left(\epsilon_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right) \left(\epsilon_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right)^\top \middle| \mathcal{G}_{t\Delta} \right] \\ & = \mathbb{E} \left[\left(\epsilon_{t+1|t}^{\mathbf{V}} \right) \left(\epsilon_{t+1|t}^{\mathbf{V}} \right)^\top \middle| \mathcal{G}_{t\Delta} \right] - \mathbb{E} \left[\left(\epsilon_{t+1|t}^{\mathbf{V}} \right) \left(\Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right)^\top \middle| \mathcal{G}_{t\Delta} \right] \\ & \quad - \mathbb{E} \left[\left(\Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right) \left(\epsilon_{t+1|t}^{\mathbf{V}} \right)^\top \middle| \mathcal{G}_{t\Delta} \right] + \mathbb{E} \left[\left(\Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right) \left(\Sigma_{t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right)^\top \middle| \mathcal{G}_{t\Delta} \right] \\ & = \text{Var} \left[\mathbf{V}_{(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] + \Sigma_{t\Delta} \text{Var} \left[\mathbf{O}_{(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \Sigma_{t\Delta}^\top \\ & \quad - \text{Cov} \left[\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \Sigma_{t\Delta}^\top - \Sigma_{t\Delta} \text{Cov} \left[\mathbf{O}_{(t+1)\Delta}, \mathbf{V}_{(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right]. \end{aligned} \quad (3.25)$$

The posteriori covariance matrix of the right-hand side of Equation (3.25) is minimized

purposes. However, we let the negative value in the filter as the next starting point as it is observed with uncertainty.

¹⁴It is also discussed in Monfort et al., 2017 in section A.5.1, Fasano et al., 2013, Alizadeh et al., 2002 Feunou et al., 2020 in footnote 6. The resulting Kalman Filter is referred in Duan and Simonato, 1999 as the Quasi-optimal Kalman Filter. A detailed discussion on the impact of the Gaussian and linear assumptions is provided in Online Appendix D.4.

(see the Appendix B.3 for details) at

$$\Sigma_{t\Delta} = \text{Cov} [\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] (\text{Var} [\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}])^{-1}. \quad (3.26)$$

Replacing (3.26) in Equation (3.25) leads¹⁵ to

$$\begin{aligned} & \mathbb{E} \left[(\mathbf{V}_{(t+1)\Delta} - \mathbb{E} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]) (\mathbf{V}_{(t+1)\Delta} - \mathbb{E} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}])^\top \middle| \mathcal{G}_{t\Delta} \right] \\ & \simeq \text{Var} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] - \Sigma_{t\Delta} \text{Cov} [\mathbf{O}_{(t+1)\Delta}, \mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]. \end{aligned} \quad (3.27)$$

While we have minimized the squared error of our estimator akin to the standard Kalman Filter, we do not have the advantageous properties of a Gaussian linear framework,¹⁶ where Equation (3.27) and $\text{Var} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]$ would be equivalent.¹⁷

Furthermore, we also derive the third and fourth cumulants of our latent variables update, where calculations are provided in Appendix B.4. We propose the following updated values

$$\text{Cum}^{(3)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \quad (3.28)$$

$$\simeq \mathbb{E} [\text{Cum}^3 [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] | \mathcal{G}_{t\Delta}],$$

$$\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \quad (3.29)$$

$$\simeq \mathbb{E} [\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] | \mathcal{G}_{t\Delta}].$$

This permits us to better track the higher moments of all the model components, as under a Gaussian framework, all these components would be zero.

¹⁵We update the Cholesky decomposition of Equation (3.23) to ensure the variance matrix stays semi-definite-positive using the function cholupdate of Matlab. See Christoffersen et al., 2014 section 3.2 for details.

¹⁶For a comprehensive discussion on this matter, refer to Online Appendix D.4.

¹⁷To diminish the distance between the true conditional moments on the left hand side of Equation (3.23) and the filtered moments on the right hand side of (3.23), we introduce later an additional update step in Online Appendix D.8.

3.4 Simulation Study

We conduct a simulation study to assess the performance of the filter. The simulation procedure is detailed in Section 3.4.1. We simulate $T = 4000$ days with the realized measures computed at a frequency¹⁸ of 5 minutes (78 intraday steps). We explain how we compute our filtering error in Section 3.4.2 and finally present our results in Section 3.4.3.

3.4.1 Simulation procedure

This section presents a comprehensive framework for simulating the general model described in Section 3.2. The simulation process involves three different time frequencies and we simulate independent trajectories $\ell \in \{1, \dots, L = 1000\}$.

The instantaneous variance and return processes are simulated with $M = 3900$ intraday time steps, from which we calculate the integrated variances. We then aggregate the high-frequency returns at an intraday frequency $N \in \{78, 390\}$, which effectively represents the intraday returns at a five and one minutes frequencies respectively. Finally, these intraday returns are utilized to compute daily returns, realized variance and covariances which are our observable variables. We also aggregate the integrated variance at a daily frequency, which permits us to isolate the error measurement made on the realized variances and covariances.

3.4.1.1 High-frequency

The instantaneous variance $V_{k,t}^{(\ell,M)}$ is simulated at a high frequency using the Euler approximation with $M = 3900$ intraday steps with

$$\begin{aligned}
 V_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)} &\cong V_{k,t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} + \kappa_k \left(\theta_k - V_{t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} \right) \frac{\Delta}{M} \\
 &\quad + \sigma_k \sqrt{V_{k,t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} \frac{\Delta}{M}} w_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}.
 \end{aligned} \tag{3.30}$$

¹⁸Results at a minute frequency (390 intraday steps) are presented in Online Appendix D.10.

$\{w_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}\}_{m=1}^M$ is a sequence of independent standard normal random variables.¹⁹ The integrated variance is approximated by the trapezoidal rule with

$$\mathcal{V}_{k,t\Delta+\frac{(m-1)\Delta}{M},t\Delta+\frac{m\Delta}{M}}^{(\ell,M)} \approx \frac{\Delta}{M} \frac{1}{2} \left(V_{k,t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} + V_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)} \right).$$

The log returns on forward prices are approximated with

$$\begin{aligned} \log \frac{F_{i,t\Delta+\frac{m\Delta}{M},T}^{(\ell)}}{F_{i,t\Delta+\frac{(m-1)\Delta}{M},T}^{(\ell)}} &\cong \sum_{k=1}^{N_V} \left(c_{i,k} \lambda_k - \frac{c_{i,k}^2}{2} \right) \mathcal{V}_{k,t\Delta+\frac{(m-1)\Delta}{M},t\Delta+\frac{m\Delta}{M}}^{(\ell,M)} \\ &+ \sum_{k=1}^{N_V} c_{i,k} \rho_k \sqrt{V_{k,t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} \frac{\Delta}{M}} w_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)} + c_{i,k} \sqrt{1 - \rho_k} \sqrt{V_{k,t\Delta+\frac{(m-1)\Delta}{M}}^{(\ell,M)} \frac{\Delta}{M}} b_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}, \end{aligned} \quad (3.31)$$

with $b_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}$ as a standard normal random variables and independent from $w_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}$.

3.4.1.2 Intraday frequency

Mimicking what is done in practice, we aggregate the log prices based on a five-minute or a one-minute frequency, which is represented by N intraday steps return,

$$R_{i,t+\frac{(n-1)\Delta}{N},t+\frac{n\Delta}{N}} = \log \frac{F_{i,t\Delta+\frac{n\Delta}{N}}^{(\ell)}}{F_{i,t\Delta+\frac{(n-1)\Delta}{N}}^{(\ell)}} = \sum_{s=1}^{M/N} \log \frac{F_{i,t\Delta+\frac{(n-1)\Delta}{N}+\frac{s\Delta}{M}}^{(\ell)}}{F_{i,t\Delta+\frac{(n-1)\Delta}{N}+\frac{(s-1)\Delta}{M}}^{(\ell)}}, \quad n \in \{1, \dots, N\}.$$

3.4.1.3 Daily frequency

Our first daily observable is the return, which is calculated as

$$R_{i,t,t+1}^{(\ell)} = \sum_{n=1}^N \log \frac{F_{i,t\Delta+\frac{n\Delta}{N}}^{(\ell)}}{F_{i,t\Delta+\frac{(n-1)\Delta}{N}}^{(\ell)}}.$$

¹⁹To ensure the non negativity of $V_{k,t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}$, if it is below 0, we replace the value with 0.

The realized variances and covariances are

$$RCV_{i,j,t,t+1}^{(\ell,N)} = \sum_{n=1}^N R_{i,t+\frac{(n-1)}{N},t+\frac{n}{N}}^{(\ell)} R_{j,t+\frac{(n-1)}{N},t+\frac{n}{N}}^{(\ell)},$$

The daily quadratic covariation is

$$QCV_{i,j,t,t+1}^{(\ell,M)} = \sum_{m=1}^M c_{i,k} c_{j,k} \mathcal{V}_{k,t\Delta+\frac{(m-1)\Delta}{M},t\Delta+\frac{m\Delta}{M}}^{(\ell,M)}.$$

The error measurement is a consequence of Equation (3.13)

$$\xi_{i,j,t,t+1}^{(\ell,N,M)} = RCV_{i,j,t,t+1}^{(\ell,N)} - QCV_{i,j,t,t+1}^{(\ell,M)}. \quad (3.32)$$

The superscripts N , M and ℓ are dropped down for readability.

3.4.2 Filtering error

Table 3.1 presents the cases we are studying, which use different combinations of three variances. The parameters are inspired by Christoffersen et al., 2009, specifically from Table 3 (Panel A for $V_{1,t\Delta}$ and Panel B for $V_{2,t\Delta}$ and $V_{3,t\Delta}$). Specifically, $V_{1,t\Delta}$ has a moderate speed of mean-reversion ($\kappa_1 = 2.00$), moderate vol-of-vol ($\sigma_1 = 0.15$), and a strong negative leverage effect ($\rho_1 = -0.85$); $V_{2,t\Delta}$ exhibits a high speed of mean-reversion ($\kappa_2 = 4.00$), high vol-of-vol ($\sigma_2 = 0.30$), and an even stronger negative leverage effect ($\rho_2 = -0.95$); while $V_{3,t\Delta}$ shows a low speed of mean-reversion ($\kappa_3 = 0.20$), low vol-of-vol ($\sigma_3 = 0.12$), and a moderate negative leverage effect ($\rho_3 = -0.50$). We first study all three variances individually and then test different combinations. The risk premium parameters are set so the approximate annual return is 6%, that is $e^{\sum_k c_{i,k} \theta_k \lambda_k} - 1 \approx 6\%$.

The filter outputs for each time step (t) and each trajectory (ℓ) are the first four conditional cumulants in the multidimensional space of the latent state $\mathbf{V}_{t\Delta}^{(\ell)}$, that is

$$\mathbb{E} \left[\mathbf{V}_{t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right], \text{Var} \left[\mathbf{V}_{t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right], \text{Cum}^{(3)} \left[\mathbf{V}_{t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right] \text{ and } \text{Cum}^{(4)} \left[\mathbf{V}_{t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right].$$

Therefore, we have $T \times L \times N_V \times 4$ conditional moments where $L = 1000$ denotes the number of trajectories, and $T = 4000$ the number of days in the simulated time series.

For $k \in \{1, \dots, N_V\}$, the filtering error is

$$\varepsilon_{k,t}^{(\ell,1)} = \underbrace{V_{k,t\Delta}^{(\ell)}}_{\text{simulated}} - \underbrace{\mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right]}_{\text{filter output}} = \epsilon_{k,t|t}^{\mathbf{v}}. \quad (3.33)$$

Because the filter output is not only a point estimate, but a characterisation of the posterior distribution, we can assess the filtering error on higher order²⁰ centered moments:

$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \underbrace{\mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \middle| \mathcal{G}_{t\Delta} \right]}_{\text{filter output}}, \quad \text{for } n \in \{2, 3, 4\}. \quad (3.34)$$

We do not present the co-moments between variance factors. However, the model provides all co-moments up to the fourth cumulant, which is a strong tool for future applications of the filter. Indeed, we not only compute the uncertainty on the variance components but also how these behave together in higher dimensions.

Given the complexity involving time, variance, moment, and simulation axes, we model the filtering error using panel analysis. Specifically, the filtering error can be represented as follows:

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}. \quad (3.35)$$

Here, $z_{k,t}^{(\ell,n)}$ represents the noises allowing for a random effect, heteroscedasticity and autocorrelation.²¹ The $\vartheta_k^{(n)}$, $k \in \{1, \dots, N_V\}$, $n \in \{1, 2, 3, 4\}$ are the intercepts of the model. The term $b_{k,t}^{(\ell,n)}$ quantifies the dependence of the current filtering error on past steps, representing an autoregressive component.²² Fortunately, as detailed in Appendix B.6, $b_{k,t}^{(\ell,n)}$ is

²⁰Additional details on the filtering error derivation are provided in the Online Appendix D.12.

²¹The filtering error can also depend on errors from other variances (in a multivariate case) and lower moment when $n \geq 2$. Therefore, in our panel analysis, we account for residual autocorrelation and set the lag to $\lfloor 4(\frac{T}{100})^{2/9} \rfloor = 9$.

²²Empirical evidence of the importance of $b_{k,t}^{(\ell,n)}$ is showed in Figure D.1 where the filtering error exhibits

a known quantity, which is

$$b_{k,t}^{(\ell,n)} = \left(e^{-\kappa_k \Delta} - [\Sigma_{t\Delta} \mathbf{A} \mathbf{H}]_{k,k} \right)^n, \quad (3.36)$$

where $e^{-\kappa_k \Delta}$ is constant over time, while $[\Sigma_{t\Delta} \mathbf{A} \mathbf{H}]_{k,k}$ varies with time. Not only is the autoregressive parameter $b_{k,t}^{(\ell,n)}$ known, but it is also time-varying. Indeed, having a closed-form expression for the autoregressive parameter is an advantage, as we can quantify how much our filtered moments are affected by the last step's error.

In each panel of the Tables 3.2, 3.3 and 3.4, we present summary statistics such as the time and cross-section averages

$$\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{t,\ell}^{(k,n)}, \quad \bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{t,\ell}^{(k,n)}.$$

The cross-sectional averages of the filtering error quantify the overall filtering error without considering the autoregressive component where the significance of this term is highlighted by the column $\bar{b}^{(k,n)}$. We normalize $\bar{\varepsilon}^{(k,n)}$ by θ_k^n to present results as percentages of the long-term variances. Additionally, we scale this column by multiplying by 10,000, thereby displaying results in basis points (bp).

We aim to show that our filter is precise and unbiased up to the first four conditional moments of the latent variable. Therefore, our objective is to estimate $\vartheta_k^{(n)}$, the intercept of Model (3.36), and test whether it is statistically different from zero as it indicates whether the filtered moments are biased or not. Appendix B.7 details how we compute our statistical test where

$$H_0 : \vartheta_k^{(n)} = 0; \quad H_1 : \vartheta_k^{(n)} \neq 0. \quad (3.37)$$

high sample autocorrelation.

We estimate with panel analysis the vector $\hat{\vartheta}$ of intercepts of Model (3.36) as

$$\hat{\vartheta}_k^{(n)} = \frac{1}{TL} \sum_{t=2}^T \sum_{\ell=1}^L \left(\varepsilon_{k,t}^{(\ell,n)} - b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} \right),$$

along with its covariance matrix $\hat{\Omega}$ (see Appendix B.7 for details). The individual test for variance (k) and moment (n) is

$$\frac{\hat{\vartheta}_k^{(n)}}{\sqrt{\hat{\Omega}_{k,k}^{(n,n)}}} \sim \mathcal{N}(0, 1). \quad (3.38)$$

The joint test is: $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$ where

$$\hat{\vartheta}^\top \hat{\Omega}^{-1} \hat{\vartheta} \sim \chi_{4N_V}.$$

Therefore, Tables 3.2, 3.3, and 3.4 present the statistics on $\vartheta_k^{(n)}$, including its value (column $\vartheta_k^{(n)}$), the t-test of (3.38) (column t-stat), and the p-value (column p-value). We normalize $\vartheta_k^{(n)}$ by θ_k^n and scale it to present results in basis points (bp). The joint test is presented at the end of each panels.

3.4.3 The experiment

Tables 3.2, 3.3 and 3.4 present the filter's performance for three simulated cases: 1) 1 asset with one variance factor, 2) two assets and two variance factors and 3) two assets and three variance factors. The main contribution of the paper resides in the case where the number of latent variables is higher than the number of assets, represented by the last case.

In Equation (3.32), the measurement error is defined as the difference between the realized variance and the quadratic variance. We show empirically in the Appendix B.5 that the measurement error does not possess the sample moments of a Gaussian random variable (Assumption 3.4.1) when the number of intraday time steps N is not large enough.²³

²³We obtain the same conclusions with one-minute frequency $N = 390$.

Assumption 3.4.1. *If the number N of intraday steps is large enough, then the error measurement is normally distributed with*

$$\xi_{t,t+1} | \mathcal{F}_{t\Delta} \sim \mathcal{N}(\mathbf{0}, \text{Cov}[\xi_{t,t+1} | \mathcal{F}_{t\Delta}]),$$

where $\text{Cov}[\xi_{t,t+1} | \mathcal{F}_{t\Delta}]$ is defined in Equation (B.4) and $\xi_{t,t+1}$ is independent from the stochastic integrals $\mathcal{W}_{t,t+1}$, $\mathcal{B}_{t,t+1}$ and $\mathcal{Z}_{t,t+1}$.

The leftmost columns of each table assume that computing RV at the 5-minute frequency ($N = 82$) results in the measurement error

$$\xi_{i,j,t,t+1}^{(N)} = RCV_{i,j,t,t+1}^{(N)} - QCV_{i,j,t,t+1}$$

reaching its asymptotic behaviour, as described in Assumption 3.4.1. However, we demonstrate empirically that this assumption does not hold. We present these results for comparison with our corrected approach, detailed in Assumption 3.4.2 and shown in the rightmost columns.²⁴

On the other hand, the rightmost columns do not assume that $\xi_{i,j,t,t+1}^{(N)}$ reached its asymptotic Gaussian distribution.

We adjust the measurement error's third cumulants²⁵ by applying a Assumption detailed in Appendix B.5. This Assumption, at an intraday frequency step n , applies the Euler approximation to the stochastic integrals $\mathcal{Z}_{k,t+\frac{n-1}{N},t+\frac{n}{N}}$, $\mathcal{W}_{k,t+\frac{n-1}{N},t+\frac{n}{N}}$ and $\mathcal{B}_{k,t+\frac{n-1}{N},t+\frac{n}{N}}$, that is

Assumption 3.4.2. *With N , the number of intraday steps used to compute the realized variance and covariance, we assume that*

$$\mathcal{B}_{k,t+\frac{n-1}{N},t+\frac{n}{N}} = \int_{(t+\frac{n-1}{N})\Delta}^{(t+\frac{n}{N})\Delta} \sqrt{V_{k,s}} dB_{k,s} \simeq \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}}} \frac{\Delta}{N} b_{k,t\Delta+\frac{n\Delta}{N}},$$

²⁴Online Appendix D.10.3 presents results where we simulate the realized variance as the QCV plus an independently simulated error that follows the asymptotic theory.

²⁵We do not correct for the fourth cumulant.

$$\begin{aligned}
 \mathcal{W}_{k,t+\frac{n-1}{N},t+\frac{n}{N}} &= \int_{(t+\frac{n-1}{N})\Delta}^{(t+\frac{n}{N})\Delta} \sqrt{V_{k,s}} dW_{k,s} \simeq \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}}} \frac{\Delta}{N} w_{k,t\Delta+\frac{n\Delta}{N}}, \\
 \mathcal{Z}_{k,t+\frac{n-1}{N},t+\frac{n}{N}} &= \int_{(t+\frac{n-1}{N})\Delta}^{(t+\frac{n}{N})\Delta} e^{-\kappa_k((t+\frac{n}{N})\Delta-s)} \sqrt{V_{k,s}} dW_{k,s} \\
 &\simeq \underbrace{e^{-\kappa_k\Delta(1-\frac{n-1}{N})}}_{\approx 1-\kappa_k\Delta(1-\frac{n-1}{N})} \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}}} \frac{\Delta}{N} w_{k,t\Delta+\frac{n\Delta}{N}},
 \end{aligned}$$

with $w_{k,t\Delta+\frac{n\Delta}{N}}$ and $b_{k,t\Delta+\frac{n\Delta}{N}}$ iid standard normal random variables.

As N tends to infinity, Assumption 3.4.2 aligns with asymptotic theory.²⁶ However, at intraday frequencies, we use Assumption 3.4.2 to compute the distributional properties of measurement errors.

Consequently, as detailed in Appendix B.5, the third cumulants related to the measurement errors are not null and are a function of

$$\mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} \tilde{V}_{\hat{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] \text{ and } \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} \tilde{V}_{\hat{k},s} \hat{V}_{\hat{k},s} ds \middle| \mathcal{F}_{t\Delta} \right],$$

implying respectively a **quadratic** ($V_{k,t\Delta} \tilde{V}_{\hat{k},t\Delta}$) and **cubic** ($V_{k,t\Delta} \tilde{V}_{\hat{k},t\Delta} \hat{V}_{\hat{k},t\Delta}$) relationship with the latent state with respect to the model filtration.

The main conclusion of our simulation study is that the filter is unbiased (we do not reject H_0) when the approximation (Assumption 3.4.2) is applied to the realized variance measurement error. We are able to filter the first four conditional moments of the latent state with precision, even though our framework is non-linear and does not follow the Gaussian distribution.

One asset and one variance

Table 3.2 examines cases with one asset and one variance (two observable variables) individually for the three variances from Table 3.1.

²⁶Corollary D.7.8 shows that Equation (B.7) still holds under Assumption 3.4.2

For Panels A and C, the null hypothesis is not rejected in any instance, whether in individual tests or in the joint test, which follows a chi-square distribution with 4 degrees of freedom (corresponding to 4 moments). In the leftmost columns of Panel B, there is bias in the individual test for the third moment of the latent variable. Here, the t-stat is -2.08 with a p-value below 5%, and the joint test value is 9.48 with a p-value of 5.02%, very close to rejecting the null hypothesis, concluding in a bias in the third central conditional moment of the latent state when we assume a Gaussian distribution for the measurement error.

Correcting for the Gaussian Assumption improves the global test on the rightmost columns significantly, with the joint test statistic dropping from 9.48 to 1.70, and the p-value increasing from 5.02% to 79.04%.

Therefore, accounting for the approximate distribution of the measurement error permits to solve the issue on the leftmost columns of the table and obtain unbiased filtered moment in the case one asset and one variance up to the fourth conditional moment.

Two assets and two variances

In Table 3.3, the case with two assets and two variances is considered. Panel A covers the scenario without any common variance factor, while Panel B includes a common variance factor. In Table 3.3, each panel has five observable for two latent variables, increasing the filter's precision. Panels A and B encounter a similar issue as Panel B of Table 3.2. For both panels, the second variance's third moment have a p-value below 5% in the leftmost columns and global tests a p-value below 20%.

Again, correcting for the Gaussian assumption significantly increases the p-values and does not reject the null hypothesis for any moments, which indicates a substantial improvement in filter precision. These findings align with those from Table 3.2, particularly concerning Panel B.

Two assets and three variances

Table 3.4 presents the most interesting scenario as it has more latent variables (3) than assets (2) with five observable variables: two returns, two realized variances, and one realized

covariance. This table tests the filter's strength and capacity to estimate latent variables accurately as the number of latent variables increases. The filter succeeded to disentangle the latent variance components with the help of three realized measures.

In this table, the leftmost columns show issues with the third moment, with a p-value below 1%. For the rest, the t-statistics are below the rejection zone, but for the fourth centered moment second variance (2,4), the t-statistic is 8.63%. The global test is 30.58, rejected with a p-value of 0.23%. This indicates that as the model becomes more complex, the lack of correction for intraday frequencies affects the overall filter precision.

Applying the correction improves the global test from 30.58 to 9.76 with the p-value rising to 63.71%, along with the non-rejection of the null hypothesis for all individual tests. However, the individual test for the fourth moment remains at 5.66% due to the uncorrected fourth moment measurement error.

In Table 3.4, there is an increase in bias as the filtering error mean is higher than in previous tables. This increase is expected as adding more latent variables without increasing the number of observable variables reduces estimation precision. Despite this, the filter remains highly effective, performing well up to the fourth moment.

In summary,²⁷ the methodology provides a comprehensive and unbiased characterization of latent variables up to the fourth moment for multivariate stochastic variance model. We also present a correction for the third central moment of the measurement error to address the bias introduced by the Gaussian Assumption in higher moments. This correction is proven to be effective in our simulation study.

²⁷Online Appendix D.10 provides results on the most complex case for higher intraday steps, $N = 390$, with similar conclusions and a reduction in bias due to increased precision. In Online Appendix D.11, a comparison of sample moments of the filter with simulated moments shows highly similar characteristics, affirming the filter's accuracy.

	$V_{1,t\Delta}$	$V_{2,t\Delta}$	$V_{3,t\Delta}$
<i>Variance process parameters</i>			
κ_k	2.00	4.00	0.20
θ_k	0.05	0.03	0.02
σ_k	0.15	0.30	0.12
ρ_k	-0.85	-0.95	-0.50
<i>1 Asset & 1 Variance - vol 1</i>			
$c_{i,k}$	1.00	0.00	0.00
λ_k	1.15		
<i>1 Asset & 1 Variance - vol 2</i>			
$c_{i,k}$	0.00	1.00	0.00
λ_k		1.95	
<i>1 Asset & 1 Variance - vol 3</i>			
$c_{i,k}$	0.00	0.00	1.00
λ_k			2.90
<i>2 Assets & 2 Variances - no covariation</i>			
$c_{i,k}$	1.00	0.00	0.00
$c_{j,k}$	0.00	1.00	0.00
λ_k	1.15	1.95	
<i>2 Assets & 2 Variances - with covariation</i>			
$c_{i,k}$	1.00	0.00	0.00
$c_{j,k}$	3/5	2/3	0.00
λ_k	1.15	1.20	
<i>2 Assets & 3 Variances</i>			
$c_{i,k}$	1.00	0.00	1.00
$c_{j,k}$	0.00	1.00	1.00
λ_k	0.75	1.25	1.10
$T = 4000, L = 1000, \Delta = 1/252, M = 3900,$ $N = 78/390$ and $\eta = \frac{\Delta}{N}$.			

The Monte Carlo simulation study follows the simulation procedure outlined in Appendix 3.4.1. As established in Equation (B.4), η is not considered as a parameter as it is a function of known quantities.

Table 3.1: Parameters for the simulation study

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	
Assuming that $\xi_{i,j,t,t+1} \sim \text{Gaussian}$						$\xi_{i,j,t,t+1} = RCV_{i,j,t,t+1} - QCV_{i,j,t,t+1}$					
Panel A: Variance 1											
(1,1)	0.98	-0.3563	-0.0178	-0.16	87.45	0.98	-0.3563	-0.0178	-0.16	87.45	
(1,2)	0.97	0.0462	-0.0006	-0.05	96.13	0.97	0.0462	-0.0006	-0.05	96.13	
(1,3)	0.95	-0.0534	-0.0030	-1.31	19.15	0.95	0.0092	0.0002	0.07	94.70	
(1,4)	0.93	0.0026	0.0001	0.12	90.54	0.93	0.0035	0.0001	0.23	81.85	
Joint	χ_4^2	3.21			52.34	χ_4^2	0.20			99.52	
Panel B: Variance 2											
(1,1)	0.96	-0.5784	-0.0210	-0.14	89.22	0.96	-0.5784	-0.0210	-0.14	89.22	
(1,2)	0.92	0.4223	0.0033	0.14	89.21	0.92	0.4223	0.0033	0.14	89.21	
(1,3)	0.89	-0.1764	-0.0174	-2.08	3.79*	0.89	-0.0163	-0.0019	-0.23	81.65	
(1,4)	0.85	0.0437	0.0029	0.81	41.91	0.85	0.0473	0.0034	0.93	35.27	
Joint	χ_4^2	9.48			5.02	χ_4^2	1.70			79.04	
Panel C: Variance 3											
(1,1)	0.92	0.9375	-0.0098	-0.05	96.27	0.92	0.9375	-0.0098	-0.05	96.27	
(1,2)	0.85	0.1358	-0.0028	-0.04	96.85	0.85	0.1358	-0.0028	-0.04	96.85	
(1,3)	0.79	0.0846	-0.0145	-0.25	80.61	0.79	0.7569	0.0420	0.71	47.67	
(1,4)	0.73	-0.2537	-0.0192	-0.31	75.51	0.73	-0.2163	-0.0154	-0.25	80.15	
Joint	χ_4^2	0.17			99.65	χ_4^2	1.28			86.40	

This Monte Carlo study simulates $L = 1000$ independent trajectories for the Model (3.1), using $M = 3900$ intraday steps over a period of $T = 4000$ days. The daily RCV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell)} = RCV_{i,j,t,t+1}^{(\ell)} - QCV_{i,j,t,t+1}^{(\ell)}$. Filtering errors,

$$\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right],$$

are computed daily for each variance component k and each path ℓ . The filtering errors' higher moments are

$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \mid \mathcal{G}_{t\Delta} \right], n \in \{2, 3, 4\}.$$

The average filtering error $\bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{t,\ell}^{(k,n)}$ is normalized by the long-term expected variance component θ_k raised to the appropriate power. This ratio, $\frac{\bar{\varepsilon}^{(k,n)}}{\theta_k^n}$, is expressed in basis points (bp). The panel regression

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}$$

(as described in Section 3.4.2) accounts for autocorrelation and potential biases through the constants $\vartheta_k^{(n)}$. We test the absence of bias ($H_0 : \vartheta_k^{(n)} = 0$) and report the associated t-statistic (t-stat) and p-value. Standard errors are estimated considering spatial correlation, using the Newey-West estimator with a lag $Q = \lfloor 4(\frac{T}{100})^{2/9} \rfloor = 9$. An asterisk (*) denotes rejection of the null hypothesis at 5% and (***) at 1%. We provide the joint chi-square test statistic and p-value for the null hypothesis $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$. The average autocorrelation $\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{t,\ell}^{(k,n)}$ displays the cross-sectional and time-average of the autocorrelation.

Table 3.2: Filtering error for case 1 asset and 1 variance

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^{(n)}}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^{(n)}}$ (bp)	t-stat	p-value (%)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^{(n)}}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^{(n)}}$ (bp)	t-stat	p-value (%)	
Assuming that $\xi_{i,j,t,t+1} \sim \text{Gaussian}$						$\xi_{i,j,t,t+1} = RCV_{i,j,t,t+1} - QCV_{i,j,t,t+1}$					
Panel A: Without covariation											
(1,1)	0.98	-0.2320	-0.0003	0.00	99.78	0.98	-0.2320	-0.0003	0.00	99.78	
(2,1)	0.96	-0.1893	0.0113	0.07	94.19	0.96	-0.1893	0.0113	0.07	94.19	
(1,2)	0.97	0.1194	0.0022	0.19	84.93	0.97	0.1194	0.0022	0.19	84.93	
(2,2)	0.92	0.1844	-0.0090	-0.38	70.39	0.92	0.1844	-0.0090	-0.38	70.39	
(1,3)	0.95	-0.0695	-0.0033	-1.49	13.50	0.95	-0.0067	-0.0002	-0.10	91.83	
(2,3)	0.89	-0.1777	-0.0163	-2.08	3.78*	0.89	-0.0178	-0.0009	-0.11	90.90	
(1,4)	0.93	0.0014	0.0000	0.00	99.62	0.93	0.0023	0.0001	0.12	90.49	
(2,4)	0.85	0.0011	-0.0013	-0.42	67.19	0.85	0.0047	-0.0008	-0.28	78.24	
Joint	χ_8^2	11.59			17.06	χ_8^2	0.24			100.00	
Panel B: With covariation											
(1,1)	0.98	-0.2347	-0.0003	0.00	99.76	0.98	-0.2347	-0.0003	0.00	99.76	
(2,1)	0.96	-0.3036	0.0081	0.05	95.85	0.96	-0.3036	0.0081	0.05	95.85	
(1,2)	0.97	0.1193	0.0022	0.19	84.96	0.97	0.1193	0.0022	0.19	84.96	
(2,2)	0.92	0.1818	-0.0085	-0.36	72.02	0.92	0.1818	-0.0085	-0.36	72.02	
(1,3)	0.95	-0.0696	-0.0033	-1.50	13.48	0.95	-0.0068	-0.0002	-0.10	91.76	
(2,3)	0.89	-0.1812	-0.0171	-2.18	2.91*	0.89	-0.0234	-0.0014	-0.18	85.97	
(1,4)	0.93	0.0014	0.0000	0.00	99.64	0.93	0.0023	0.0001	0.12	90.51	
(2,4)	0.85	0.0009	-0.0012	-0.42	67.41	0.85	0.0044	-0.0008	-0.27	78.69	
Joint	χ_8^2	12.22			14.16	χ_8^2	0.24			100.00	

See the notes of Table 3.2. Panel A displays results where the two assets do not have a common variance factor, whereas Panel B shows results for two assets with a common variance factor.

Table 3.3: Filtering error for case 2 assets and 2 variances

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	
Assuming that $\xi_{i,j,t,t+1} \sim \text{Gaussian}$						$\xi_{i,j,t,t+1} = RCV_{i,j,t,t+1} - QCV_{i,j,t,t+1}$					
(1,1)	0.98	-1.9635	-0.0298	-0.23	81.48	0.98	-1.9635	-0.0298	-0.23	81.48	
(2,1)	0.96	-3.6735	-0.1086	-0.44	66.24	0.96	-3.6735	-0.1086	-0.44	66.24	
(3,1)	0.98	2.6461	-0.0819	-0.32	74.75	0.98	2.6461	-0.0819	-0.32	74.75	
(1,2)	0.97	0.2641	0.0049	0.32	74.96	0.97	0.2641	0.0049	0.32	74.96	
(2,2)	0.91	1.5650	0.0361	0.66	51.18	0.91	1.5650	0.0361	0.66	51.18	
(3,2)	0.96	0.4076	-0.0174	-0.19	84.71	0.96	0.4076	-0.0174	-0.19	84.71	
(1,3)	0.95	-0.0765	-0.0040	-1.21	22.54	0.95	0.0092	0.0002	0.07	94.72	
(2,3)	0.87	-0.7556	-0.0842	-3.23	0.12**	0.87	-0.2297	-0.0234	-0.90	36.86	
(3,3)	0.93	-1.7991	-0.1108	-1.53	12.65	0.93	-0.9102	-0.0505	-0.70	48.57	
(1,4)	0.93	0.0160	0.0008	0.93	35.24	0.93	0.0173	0.0008	1.03	30.14	
(2,4)	0.84	0.2651	0.0226	1.72	8.63	0.84	0.2819	0.0252	1.91	5.66	
(3,4)	0.91	-0.6580	-0.0577	-0.74	45.79	0.91	-0.6021	-0.0529	-0.68	49.62	
Joint	χ_{12}^2	30.58			0.23*	χ_{12}^2	9.76			63.71	

This Monte Carlo study simulates $L = 1000$ independent trajectories for the Model (3.1), using $M = 3900$ intraday steps over a period of $T = 4000$ days. The daily RCV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell)} = RCV_{i,j,t,t+1}^{(\ell)} - QCV_{i,j,t,t+1}^{(\ell)}$. Filtering errors,

$$\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right],$$

are computed daily for each variance component k and each path ℓ . The filtering errors' higher moments are

$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \mid \mathcal{G}_{t\Delta} \right], n \in \{2, 3, 4\}.$$

The average filtering error $\bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{k,t}^{(\ell,n)}$ is normalized by the long-term expected variance component θ_k raised to the appropriate power. This ratio, $\frac{\bar{\varepsilon}^{(k,n)}}{\theta_k^n}$, is expressed in basis points (bp). The panel regression

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}$$

(as described in Section 3.4.2) accounts for autocorrelation and potential biases through the constants $\vartheta_k^{(n)}$. We test the absence of bias ($H_0 : \vartheta_k^{(n)} = 0$) and report the associated t-statistic (t-stat) and p-value. Standard errors are estimated considering spatial correlation, using the Newey-West estimator with a lag $Q = \lceil 4(\frac{T}{100})^{2/9} \rceil = 9$. An asterisk (*) denotes rejection of the null hypothesis at 5% and (***) at 1%. We provide the joint chi-square test statistic and p-value for the null hypothesis $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$. The average autocorrelation $\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{k,t}^{(\ell,n)}$ displays the cross-sectional and time-average of the autocorrelation.

Table 3.4: Filtering error for case 2 assets and 3 variances

3.5 Conclusion

Our research introduces a precise analytical filter for multi-factor variance estimation within a multi-asset framework. This filter effectively disentangles latent variance components, an important step in understanding the commonality of asset volatilities across different markets. By leveraging intraday statistics, such as realized variances and covariances, our approach provides a tractable solution that overcomes the curse of dimensionality, a common challenge in multi-asset analysis.

A key strength of our filter lies in its ability to manage nonlinear and non-Gaussian frameworks, significantly broadening its applicability beyond traditional affine models. A critical innovation in our approach is the incorporation of measurement errors of realized variance data, particularly at lower frequencies where asymptotic theory is not fully applicable. We propose a correction that ensures the filter remains unbiased across the first four conditional moments, resulting in a more comprehensive and accurate estimation of the latent state. The robustness and precision of our method are further validated through extensive Monte Carlo simulations, confirming its reliability across a range of scenarios. Additionally, our methodology provides estimations up to the fourth cumulant, including co-moments, offering a detailed characterization of volatility dynamics.

References

- Aït-Sahalia, Y., & Xiu, D. (2016). Increased Correlation among Asset Classes: Are Volatility or Jumps to Blame, or Both? *Journal of Econometrics*, *194*(2), 205–219.
- Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance*, *57*(3), 1047–1091.
- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2013). Chapter 17 - Financial Risk Measurement for Financial Risk Management. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.). Elsevier.
- Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *64*(2), 253–280.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric Analysis of Realized Covariation: High Frequency based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, *72*(3), 885–925.
- Bates, D. S. (2006). Maximum Likelihood Estimation of Latent Affine Processes. *Review of Financial Studies*, *19*(3), 909–965.
- Bates, D. S. (2012). US Stock Market Crash Risk, 1926–2010. *Journal of Financial Economics*, *105*(2), 229–259.
- Bates, D. S. (2019). How Crashes Develop: Intradaily Volatility and Crash Evolution. *Journal of Finance*, *74*(1), 193–238.

- Bégin, J.-F., Amaya, D., Gauthier, G., & Malette, M.-È. (2020). On the Estimation of Jump-Diffusion Models Using Intraday Data: A Filtering-Based Approach. *SIAM Journal on Financial Mathematics*, *11*(4), 1168–1208.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting. *Journal of Econometrics*, *192*(1), 1–18.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2018). Modeling and Forecasting (Un)reliable Realized Covariances for more Reliable Financial Decisions. *Journal of Econometrics*, *207*(1), 71–91.
- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, *109*(1), 33–65.
- Brignone, R., Gonzato, L., & Lütkebohmert, E. (2023). Efficient Quasi-Bayesian Estimation of Affine Option Pricing Models using Risk-Neutral Cumulants. *Journal of Banking and Finance*, *148*, 106745.
- Brillinger, D. R. (1969). The Calculation of Cumulants via Conditioning. *Annals of the Institute of Statistical Mathematics*, *21*(1), 215–218.
- Buccheri, G., & Corsi, F. (2021). HARK the SHARK: Realized Volatility Modeling with Measurement Errors and Nonlinear Dependencies. *Journal of Financial Econometrics*, *19*(4), 614–649.
- Calzolari, G., Halbleib, R., & Zagidullina, A. (2021). A Latent Factor Model for Forecasting Realized Variances. *Journal of Financial Econometrics*, *19*(5), 860–909.
- Chen, R.-R., & Scott, L. (2003). Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter model. *Journal of Real Estate Finance and Economics*, *27*, 143–172.
- Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman Filtering in Affine Term Structure Models. *Management Science*, *60*(9), 2248–2268.

- Christoffersen, P., Heston, S., & Jacobs, K. (2009). The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. *Management Science*, 55(12), 1914–1932.
- Christoffersen, P., Lunde, A., & Olesen, K. V. (2019). Factor Structure in Commodity Futures Return and Volatility. *Journal of Financial and Quantitative Analysis*, 54(3), 1083–1115.
- Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196.
- Cortazar, G., Lopez, M., & Naranjo, L. (2017). A Multifactor Stochastic Volatility Model of Commodity Prices. *Energy Economics*, 67, 182–201.
- Duan, J.-C., & Simonato, J.-G. (1999). Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting*, 13, 111–135.
- Dubecq, S., Monfort, A., Renne, J.-P., & Roussellet, G. (2016). Credit and Liquidity in Interbank Rates: A Quadratic Approach. *Journal of Banking and Finance*, 68, 29–46.
- Duffee, G. R., & Stanton, R. H. (2012). Estimation of Dynamic Term Structure Models. *The Quarterly Journal of Finance*, 2(02), 1250008.
- Fasano, A., Germani, A., & Monteriu, A. (2013). Reduced-Order Quadratic Kalman-like Filtering of Non-Gaussian Systems. *IEEE Transactions on Automatic Control*, 58(7), 1744–1757.
- Feunou, B., Aliouchkin, R. L., Tédongap, R., & Xu, L. (2020). The Term Structures of Expected Loss and Gain Uncertainty. *Journal of Financial Econometrics*, 18(3), 473–501.
- Feunou, B., & Okou, C. (2018). Risk-Neutral Moment-based Estimation of Affine Option Pricing Models. *Journal of Applied Econometrics*, 33(7), 1007–1025.

- Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2016). The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications. *Journal of Financial Economics*, 119(2), 249–283.
- Heston, S. L. (1993). A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2), 327–343.
- Javaheri, A., Lautier, D., & Galli, A. (2003). Filtering in Finance. *Wilmott*, 3, 67–83.
- Jong, F. d. (2000). Time Series and Cross-Section Information in Affine Term-Structure Models. *Journal of Business and Economic Statistics*, 18(3), 300–314.
- Julier, S., Uhlmann, J., & Durrant-Whyte, H. F. (2000). A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. *IEEE Transactions on Automatic Control*, 45(3), 477–482.
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems.
- Li, H., & Zhao, F. (2006). Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives. *Journal of Finance*, 61(1), 341–378.
- Monfort, A., Pegoraro, F., Renne, J.-P., & Roussellet, G. (2017). Staying at Zero with Affine Processes: An Application to Term Structure Modelling. *Journal of Econometrics*, 201(2), 348–366.
- Monfort, A., Renne, J.-P., & Roussellet, G. (2015). A Quadratic Kalman Filter. *Journal of Econometrics*, 187(1), 43–56.
- Roussellet, G. (2023). The Term Structure of Macroeconomic Risks at the Effective Lower Bound. *Journal of Econometrics*, 105383.
- Trolle, A. B., & Schwartz, E. S. (2009). A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives. *Review of Financial Studies*, 22(5), 2007–2057.

Wan, E., & Van Der Merwe, R. (2000). The Unscented Kalman Filter for Nonlinear Estimation. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 153–158.

Chapter 4

Enhancing Parameter Estimation in Stochastic Volatility Models Using Intraday Data: Generalized Method of Moments Comparative Analysis

Abstract¹

Estimating stochastic volatility model parameters remains a significant challenge, especially when observations are limited to proxies of integrated variances. This study introduces a refined approach, building upon the Generalized Method of Moments (GMM) framework proposed by Bollerslev and Zhou, 2002, with extensions to account for leverage effects and multi-factor models. Additionally, we compare two GMM procedures widely discussed in the literature: conditional moments (CM) and unconditional moments (UM). Our findings suggest that without a filtering mechanism, CM-based GMM is less robust than UM-based GMM, as the required recursions in the CM approach introduce considerable noise into the moment conditions.

¹Joint work with Christian Dorion and Geneviève Gauthier. Dorion and Gauthier are affiliated with HEC Montréal.

4.1 Introduction

State-of-the-art stochastic volatility models assume multiple variance components. One of the main challenges in using these frameworks is parameter estimation within the context of many latent factors using intraday data. The Generalized Method of Moments (GMM) is a widely adopted tool in the literature, known for its flexibility and robustness across various model specifications. This paper focuses on comparing two prominent GMM procedures, those based on conditional moments (CM) and unconditional moments (UM) for one and two variance factor models.

Realized variance refers to intraday statistics that converge to the quadratic variation of an asset as the number of intraday steps increases (Barndorff-Nielsen and Shephard, 2002), making it a useful observable for quantifying variance. Bollerslev and Zhou, 2002 popularized a GMM methodology based solely on this observable to estimate variance parameters. GMM utilizes the conditional moments of the sample data, in this case, the realized variance, and estimates parameters through an optimization process. As models become more complex, traditional methods such as quasi-likelihood can pose challenges, thus creating opportunities for alternative approaches like GMM.²

Our main benchmarks for the two procedures compared here are Bollerslev and Zhou, 2002 for the CM method and Todorov, 2009 for the UM method. We present a thorough Monte Carlo study across multiple scenarios and model specifications to assess the strengths and weaknesses of each approach.

Our analysis uncovers notable differences in the effectiveness of these methods. The CM-based approach stands out for its substantially lower autocorrelation in GMM residuals compared to the UM-based approach, marking a clear advantage of the CM method. Moreover, the use of instrumental variables within the CM framework enhances the method's power and potential for extensions while also leading to more accurate estimation results when the integrated variance is observed without noise.

However, the CM approach is not without limitations. In the absence of a filtered value

²Zhang and Wang, 2023 compare the GMM performance to the MCMC and QML methods with the Realized Stochastic Volatility (RSV) model, concluding that GMM is faster than MCMC and also more effective as the QML method is less accurate.

of the instantaneous variances, the need for performing recursions introduces additional noise, which compromises the method's effectiveness. This issue is particularly acute in multi-factor models, where the recursive techniques proposed by Bollerslev and Zhou, 2002 for the two-factor model are less effective relative to UM methods. Additionally, the CM approach becomes less performant than the UM methods when measurement error is disproportionately large relative to the realized variance or the persistence of the variance is very high.

On the other hand, while UM approaches are generally more robust, especially in the context of two variance factors, they offer less potential for extension compared to CM methods combined with a filter. The absence of instrumental variables in UM methods makes it more difficult and cumbersome to derive additional moments as the model becomes more complex, particularly in a multivariate framework.

In addition to comparing these GMM procedures, this paper proposes the introduction of additional moment conditions within the CM framework to estimate the measurement error parameter. Through a comprehensive analysis of the methodology presented in Bollerslev and Zhou, 2002, we address the issue of measurement error specification of the the realized variance. We demonstrate that their specification allows dependence between the moments used in GMM and the instrumental variables. To rectify this issue, we follow the asymptotic theory outlined in Barndorff-Nielsen and Shephard, 2002 and introduce an additional moment to improve overall parameter estimation when the measurement error parameter is estimated.

In summary, we contribute to the existing literature by comparing two GMM procedures discussed in the literature, highlighting the advantages and limitations of each method depending on the model specifications and various parameter scenarios in our simulation study. Finally, we improve the CM procedure by introducing a method to adequately incorporate the measurement error parameter.

The paper is organized as follows: After the literature review, Section 4.2 presents the model, and Section 4.3 details the simulation procedure. Sections 4.4 and 4.5 respectively present the GMM derivation and Monte Carlo study for the one and two variance factors.

Finally, Section 4.6 concludes.

Literature Review

Bollerslev and Zhou, 2002 propose a parameter estimation method through GMM for stochastic volatility models, using intraday data. Their empirical implementation nest multiple models including multiple latent stochastic volatility factors, leverage effect and jumps on the returns. They address the measurement error of the realized variance by adding a nuisance parameter. However, this approach assumes a constant variance for the measurement error, which contradicts the asymptotic theory discussed in Barndorff-Nielsen and Shephard, 2002. Furthermore, there is no Monte Carlo simulation study for the more complex frameworks. Bollerslev et al., 2011 introduce an improved methodology for estimating the volatility risk premium. Their approach builds upon the insights from the work of Barndorff-Nielsen and Shephard, 2002. In their methodology, Bollerslev et al., 2011 include option-implied volatility measures as part of the moment selection process. However, they do not address the RV measurement error issue, as their methodology relies solely on first-order conditional moments.

Corradi and Distaso, 2006 propose a GMM moment selection based on unconditional moments for a larger class of stochastic volatility models, which they apply on three securities, General Electric, Intel, and Microsoft, for the case of one asset and one variance factor. Furthermore, they provide the conditions where the measurement error can be ignored asymptotically, which is used Garcia et al., 2011 among others. Ghosh and Linton, 2023 discuss the double asymptotic framework for the theory of estimated parameters, where N represents the intra-period returns and T denotes the low-frequency number of periods in which realized variance (RV) is used. They demonstrate a bias correction method when the asymptotic conditions are not met and the realized variance is used as a proxy for integrated variance in the GMM procedure.

As options are also explored as a source of data, Garcia et al., 2011 propose an innovative estimation procedure that combines option prices and intraday data to estimate the parameters of stochastic volatility model using a GMM procedure. By incorporating

option prices, they effectively integrate a risk premium parameter into their estimations. Their methodology builds upon the foundational work of Bollerslev and Zhou, 2002 and enhances it by including higher order moments and option-related moments. However, their methodology does not include the measurement error associated to RV. Ishida et al., 2011 detail and compare the estimation methodologies presented in Bollerslev and Zhou, 2002, Garcia et al., 2011 and Corradi and Distaso, 2006. Their model includes a leverage effect, but excludes a drift component. They estimate the leverage parameter by incorporating VIX data and numerically compares it to Bollerslev and Zhou, 2002. While their paper does not address the measurement error parameter issue, they do delve into the implications of calculating intraday data exclusively during trading hours, which is an additional contribution.

Todorov, 2009 proposes a GMM approach based on unconditional moments using the realized variance measure within a general model that includes jumps. He also suggests an asymptotic refinement for the realized measures, based on asymptotic theory, which reduces the bias in their estimates, a refinement not present in Todorov, 2010. Todorov, 2009 demonstrates the performance of intraday measures and their asymptotic refinement in a detailed Monte Carlo simulation study. Furthermore, the paper estimates the parameters for the S&P 500 futures contract using both the two-factor affine jump diffusion model and the CARMA(2,1) jump-driven SV model, although without the corresponding Monte Carlo study. Todorov, 2011 extends this analysis to exchange rates. Both papers conclude that the CARMA model better fits the data. Additionally, the leverage effect is not directly addressed or estimated in the results presented in these papers. Todorov, 2010 also applies GMM to isolate the variance risk premium. Ewald and Zou, 2021 implements a similar GMM procedure based on multipower variations to estimate an extended model with jumps in both the variance and return processes using S&P 500 data. Finally, Todorov et al., 2011 employs realized Laplace transform instead of realized variance as the observable, implements a GMM, and runs several model specifications with jumps and multi-factor variances in both Monte Carlo studies and on S&P 500 data. However, the paper does not include a Monte Carlo study on two stochastic volatility models, which is discussed in this paper.

Other studies extended the work of Bollerslev and Zhou, 2002 such as Bregantini, 2013

who perform a Monte Carlo simulation study along with real data estimations from the S&P 500 data for the model with one variance factor. They include the jump case in their paper along with leverage effect and present all the possible cases for the one-variance factor model, which make this paper complete for the one variance case. They also comment on some difficulties with the two-variance factor model of Bollerslev and Zhou, 2002. However, the error measurement is not considered in this study. In Baum and Zerilli, 2016, the authors apply the method of Bollerslev and Zhou, 2002 to crude oil futures data and include a jump component using returns and realized variance measures. They discuss that the leverage effect parameter is not significant from zero in the case of crude oil, but jumps are find significance evidence for the presence of jumps. However, they do not directly address the measurement error in the realized variance measures. They extend their work in Baum et al., 2021, by including S&P 500 futures data and commodities such as natural gas and crude oil. In Brix and Lunde, 2015, the authors compared the GMM methods with prediction-based estimating functions. They also discussed how to handle measurement error by assuming iid noise in the log process and assessing its impact both parametrically and non-parametrically. The latter is done by using realized kernels instead of realized variance. However, their assumption of iid measurement error contradicts the asymptotic theory, which typically requires different error structure assumptions.

Not all GMM methodologies used realized variance as observables. Indeed, in Feunou and Tédongap, 2012, they implement a GMM procedure to estimate their stochastic volatility model with time-varying conditional skewness in a discrete framework. They use a set of conditions related to higher moments of returns, including lagged returns. They compare their discrete time model with three existing benchmarks with one and two variance factors for application on real-data. Also, in Jiang and Oomen, 2007, the authors developed an unbiased estimator of instantaneous variance using the conditional characteristic function. This approach led to an estimator based on several conditional cumulants, which are subsequently weighted to yield the final estimator. They then performed a GMM procedure to estimate the parameters. The authors extended their methodology to a two-variance case, using two different sampling frequencies to isolate each variance. Notably, they did not use realized variance measures in their study.

Bolko et al., 2023 present a GMM approach to estimate a log-normal stochastic volatility model driven by a fractional Brownian motion with an unrestricted Hurst exponent. They incorporate measurement error on RV by analyzing its theoretical moments, either through the Central Limit Theorem or its exact distribution while assuming no leverage effect or drift. Building on these assumptions, they apply GMM to the unconditional moments of the integrated variance using realized variance measures. While their variance specifications differ from ours, they demonstrate a proper characterization of measurement error. Moreover, they employ unconditional moments, as we propose for the case of two variance factors. However, their model does not account for leverage effects, drift components or multiple latent variance factors.

Chaussé and Xu, 2018 compare four different GMM estimators, such as a robust, regularized and a PCA version of the standard GMM, applied to the realized stochastic volatility (RSV) model proposed by Takahashi et al., 2009. They found that alternatives GMM procedures improves the quality of the GMM estimator. Chacko and Viceira, 2003 derive a spectral GMM to estimate continuous-time volatility stochastic models based on characteristic functions. Jiang and Knight, 2002 also studied GMM using the characteristic function of the asset returns. In Escobar, 2018, they estimate the multivariate Heston model with the continuum GMM with the characteristic function on the returns. In Li and Xiu, 2016, they propose a novel GMM approach where, in the first step, they compute a non-parametric estimate for the spot variance. In the second step, they apply the GMM, correcting for the bias introduced by the estimator.

4.2 Theoretical Background

4.2.1 Continuous time model

The asset price is affected by K common volatility factors following a Heston, 1993 model,

$$dS_t = \alpha_t S_t dt + S_t \left(\sum_{k=1}^K \sqrt{V_{k,t}} dZ_{k,t} \right), \quad (4.1)$$

$$dV_{k,t} = \kappa_k (\theta_k - V_{k,t}) dt + \sigma_k \sqrt{V_{k,t}} dW_{k,t}, \quad k \in \{1, 2, \dots, K\}, \quad (4.2)$$

where $Z_{k,t} = \rho_k W_{k,t} + \sqrt{1 - \rho_k^2} B_{k,t}$ and $\{W_{k,t}, B_{k,t}\}_{k=1}^K$ are independent \mathbb{P} -Brownian motions. The leverage effect is reflected in the correlation between the Brownian motions of the price and variance factors.

As shown in Online Appendix E.1, the drift is composed of the risk-free rate r , the dividend rate q and the risk premium:

$$\alpha_t = r - q + \sum_{k=1}^K \lambda_k V_{k,t}, \quad (4.3)$$

where λ_k is the price of risk.

The log-price quadratic variation is

$$\langle \log S \rangle_t = \sum_{k=1}^K \int_0^t V_{k,s} ds.$$

The model filtration $\{\mathcal{F}_t\}_{t \geq 0}$ is generated by the Brownian motions, that is $\mathcal{F}_t = \sigma\{W_{k,u}, B_{k,u} : 0 \leq u \leq t, k \in \{1, \dots, K\}\}$. In our numerical implementation, we consider the cases $K \in \{1, 2\}$.

4.2.2 Discrete time model

4.2.2.1 The returns

For $t \in \{0, 1, 2, \dots, T\}$, the daily return³ satisfies

$$\begin{aligned} R_{t,t+1} &= \log \frac{S_{(t+1)\Delta}}{S_{t\Delta}} \\ &= (r - q) \Delta + \sum_{k=1}^K \left(\lambda_k - \frac{1}{2} \right) \int_{t\Delta}^{(t+1)\Delta} V_{t,s} ds + \sum_{k=1}^K \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{k,t}} dZ_{k,t}. \end{aligned} \quad (4.4)$$

³ Δ is intentionally left unspecified at this point. For the single variance case, Bollerslev and Zhou, 2002 use $\Delta = 1$. However, we propose a simulation study that is more closely related to the stock market and therefore use $\Delta = 1/252$.

4.2.2.2 The realized variance

Let

$$\mathcal{V}_{t,t+1} = \sum_{k=1}^K \mathcal{V}_{k,t,t+1},$$

be the daily integrated variance with

$$\mathcal{V}_{k,t,t+1} = \int_{t\Delta}^{(t+1)\Delta} V_{k,s} ds, \quad k \in \{1, \dots, K\}.$$

As shown by Barndorff-Nielsen and Shephard, 2002 and Barndorff-Nielsen and Shephard, 2004b, the daily realized variance calculated with N intraday returns

$$RV_{t,t+1}^{(N)} = \sum_{n=1}^N \left(\log \frac{S_{t\Delta + \frac{n}{N}\Delta}}{S_{t\Delta + \frac{(n-1)}{N}\Delta}} \right)^2 \quad (4.5)$$

converges weakly to the integrated variance

$$\frac{\sqrt{\frac{N}{\Delta}} \left(RV_{t,t+1}^{(N)} - \mathcal{V}_{t,t+1} \right)}{\sqrt{2RQ_{t,t+1}^{(N)}}} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1), \quad (4.6)$$

where the realized quarticity is defined as

$$RQ_{t,t+1}^{(N)} = \frac{N}{3\Delta} \sum_{n=1}^N \left(\log \frac{S_{t\Delta + \frac{n}{N}\Delta}}{S_{t\Delta + \frac{(n-1)}{N}\Delta}} \right)^4 \xrightarrow[N \rightarrow \infty]{\mathcal{P}} \int_{t\Delta}^{(t+1)\Delta} \left(\sum_{k=1}^K V_{k,s} \right)^2 ds, \quad (4.7)$$

and $\xrightarrow[N \rightarrow \infty]{\mathcal{P}}$ refers to a convergence in probability. The weak convergence⁴ in Equation (4.6) suggests that the realized variance is a noisy signal of the integrated variance, that is

$$RV_{t,t+1}^{(N)} = \mathcal{V}_{t,t+1} + \xi_{t,t+1}^{(N)}, \quad (4.8)$$

⁴The asymptotic results of Equation (4.6) can be extended to the case where the model exhibit leverage effect as discussed in Barndorff-Nielsen and Shephard, 2004a. Meddahi, 2002 concludes that the effect of the drift and the leverage effect are negligible compared to the mean of the integrated variance.

where the noise term $\xi_{t,t+1}^{(N)}$ is asymptotically Gaussian with

$$\mathbb{E} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] = 0, \quad \text{Var} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] = 2 \frac{\Delta}{N} \mathbb{E} \left[RQ_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right], \quad (4.9)$$

and from Equation (4.7),

$$\mathbb{E} \left[RQ_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] \xrightarrow{N \rightarrow \infty} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \left(\sum_{k=1}^K V_{k,s} \right)^2 ds \middle| \mathcal{F}_{t\Delta} \right].$$

As detailed in Appendix C.1, the unconditional variance of the measurement error is

$$\begin{aligned} \text{Var} \left[\xi_{t,t+1}^{(N)} \right] &= \mathbb{E} \left[2 \frac{\Delta}{N} \mathbb{E} \left[RQ_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] \right] \\ &\xrightarrow{N \rightarrow \infty} 2 \frac{\Delta^2}{N} \left(\left(\sum_{k=1}^K \theta_k \right)^2 + \sum_{k=1}^K \frac{\sigma_k^2 \theta_k}{2\kappa_k} \right). \end{aligned} \quad (4.10)$$

We assume that $\xi_{t,t+1}^{(N)}$ is uncorrelated with the integrated variance $\mathcal{V}_{k,t,t+1}$, the instantaneous variance $V_{k,t,t+1}$, and the return $R_{t,t+1}$. In Appendix C.2, we provide empirical evidence from Monte Carlo simulations supporting these assumptions, as they are essential for the GMM moment derivations. Additionally, we empirically confirm the validity of Equation (4.9) in finite sample.

The observable filtration $\{\mathcal{G}_t\}_{t \in \mathbb{N}}$ represents the information available for the investor at time t and is defined as

$$\mathcal{G}_t = \sigma \{ R_{u,u+1}, RV_{u,u+1}^{(N)}, RQ_{u,u+1}^{(N)} \}_{u=0}^t,$$

where we acknowledge that we do not directly observe the integrated variance.

4.3 Simulation Procedure

This Section presents a comprehensive framework for simulating the general model described in Section 4.2. The simulation process involves three different time frequencies and we simulate independent trajectories $l \in \{1, \dots, S = 1000\}$ over T days.

The instantaneous variance and return processes are simulated with $M = 820$ intraday time steps, approximately every 30 seconds, from which we calculate the integrated variances (Section 4.3.1). We then aggregate the high-frequency returns at an intraday frequency $N = 82$, which effectively represents the intraday returns at a five-minute frequency (Section 4.3.2). Finally, these intraday returns are aggregated to compute daily returns, realized variance, and realized quarticity, which are our observable variables (Section 4.3.3). We also aggregate the integrated variance at a daily frequency, to obtain the measurement error on the realized variance.

4.3.1 High-frequency

The instantaneous variance $V_{k,t}$ is simulated at a high-frequency using the Euler approximation with $M = 820$ intraday steps

$$V_{k,t\Delta + \frac{m\Delta}{M}} \cong V_{k,t\Delta + \frac{(m-1)\Delta}{M}} + \kappa_k \left(\theta_k - V_{t\Delta + \frac{(m-1)\Delta}{M}} \right) \frac{\Delta}{M} + \sigma_k \sqrt{V_{k,t\Delta + \frac{(m-1)\Delta}{M}}} \frac{\Delta}{M} w_{k,t\Delta + \frac{m\Delta}{M}}. \quad (4.11)$$

$\{w_{k,t\Delta + \frac{m\Delta}{M}}\}_{m=1}^M$ is a sequence of independent standard normal random variables.⁵ The integrated variance⁶ is approximated by the trapezoidal rule with

$$\mathcal{V}_{k,t\Delta + \frac{(m-1)\Delta}{M}, t\Delta + \frac{m\Delta}{M}} \approx \frac{\Delta}{M} \frac{\left(V_{k,t\Delta + \frac{(m-1)\Delta}{M}} + V_{k,t\Delta + \frac{m\Delta}{M}} \right)}{2}.$$

⁵To ensure the non negativity of $V_{k,t\Delta + \frac{m\Delta}{M}}$, if it is below 0, we replace the value with 0.

⁶Bollerslev and Zhou, 2002 use the realized variance at the high-frequency to approximate the integrated variance series.

The log returns are approximated with

$$\begin{aligned} \log \frac{S_{t\Delta + \frac{m}{M}\Delta}}{S_{t\Delta + \frac{(m-1)}{M}\Delta}} &= (r - q) \frac{\Delta}{M} + \sum_{k=1}^K (\lambda_k - \frac{1}{2}) \mathcal{V}_{k, t\Delta + \frac{(m-1)\Delta}{M}, t\Delta + \frac{m\Delta}{M}} \\ &+ \sum_{k=1}^K \rho_k \sqrt{V_{k, t\Delta + \frac{(m-1)\Delta}{M}} \frac{\Delta}{M}} w_{k, t\Delta + \frac{m\Delta}{M}} + \sqrt{1 - \rho_k} \sqrt{V_{k, t\Delta + \frac{(m-1)\Delta}{M}} \frac{\Delta}{M}} b_{k, t\Delta + \frac{m\Delta}{M}}, \end{aligned} \quad (4.12)$$

with $b_{k, t\Delta + \frac{m\Delta}{M}}$ as a standard normal random variables and independent from $w_{k, t\Delta + \frac{m\Delta}{M}}$.

4.3.2 Intraday frequency

Mimicking what is done in practice, we aggregate the log returns based on a five-minute interval, which corresponds to $N = 82$ intraday steps,

$$\log \frac{S_{t\Delta + \frac{n}{N}\Delta}}{S_{t\Delta + \frac{(n-1)}{N}\Delta}} = \sum_{s=1}^{M/N} \log \frac{S_{t\Delta + \frac{n-1}{N}\Delta + \frac{s\Delta}{M}}}{S_{t\Delta + \frac{n-1}{N}\Delta + \frac{(s-1)\Delta}{M}}}, \quad n \in \{1, \dots, N\}.$$

4.3.3 Daily frequency

Our first daily observable is the return calculated as

$$R_{t, t+1}^{(N)} = \log \frac{S_{(t+1)\Delta}}{S_{t\Delta}}.$$

The realized variance⁷ and quarticity are

$$RV_{t, t+1}^{(N)} = \sum_{n=1}^N \left(\log \frac{S_{t\Delta + \frac{n}{N}\Delta}}{S_{t\Delta + \frac{(n-1)}{N}\Delta}} \right)^2 \quad \text{and} \quad RQ_{t, t+1}^{(N)} = \frac{N}{3\Delta} \sum_{n=1}^N \left(\log \frac{S_{t\Delta + \frac{n}{N}\Delta}}{S_{t\Delta + \frac{(n-1)}{N}\Delta}} \right)^4.$$

⁷We do not use the estimator proposed by Zhang et al., 2005 because there is no microstructure noise in simulated data. However, if any real data were to be used, an adjustment to the realized measure would be done to minimize the microstructure noise.

The daily integrated variance is

$$\mathcal{V}_{k,t,t+1}^{(M)} = \sum_{m=1}^M \mathcal{V}_{k,t\Delta + \frac{(m-1)\Delta}{M}, t\Delta + \frac{m\Delta}{M}},$$

where, from Equation (4.8), the measurement error is

$$\xi_{t,t+1}^{(N,M)} = RV_{t,t+1}^{(N)} - \mathcal{V}_{t,t+1}^{(M)}. \quad (4.13)$$

For the remainder of the paper, we omit the superscript M in $\xi_{t,t+1}^{(N,M)}$ and $\mathcal{V}_{t,t+1}^{(M)}$.

4.4 GMM Estimation with One Variance

We derive the moment conditions that will be used in the GMM. This methodology uses the information contained in these moment conditions to estimate the model parameters. By deriving these moment conditions, we can construct a vector based on the sample data, which will be used in the optimization process to estimate the parameters efficiently.

We present a comprehensive derivation of the two classes of GMM procedures discussed in the literature on stochastic variance models. The first class is based on conditional moments (CM) as seen in Baum and Zerilli, 2016; Baum et al., 2021; Bollerslev and Zhou, 2002; Bollerslev et al., 2011; Bregantini, 2013; Garcia et al., 2011; Ishida et al., 2011, among others. The second class is based on unconditional moments (UM) as exemplified in Bolko et al., 2023; Corradi and Distaso, 2006; Ewald and Zou, 2021; Feunou and Tédongap, 2012; Todorov, 2009 and others.

We also introduce our proposed methodology to handle the measurement error of the realized variance for CM approaches, as the Bollerslev and Zhou, 2002 approach fails to estimate the measurement error parameter of the realized variance. In addition, we derive models with and without leverage effects and drift.

4.4.1 GMM derivation

4.4.1.1 Without a drift and a leverage effect

The first specification is the base case of Bollerslev and Zhou, 2002. The log price dynamic has no drift, no leverage effect, and, in opposition to the asymptotic theory, the measurement error on the realized variance is assumed to have a constant standard deviation, that is

$$R_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{1,t}} dB_{1,t}, \quad RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \gamma \epsilon_{t,t+1}^{(N)}, \quad (4.14)$$

where $\epsilon_{t,t+1}^{(N)}$ is centered at zero and of unit variance. There are four parameters to estimate, that is $\{\theta_1, \kappa_1, \sigma_1, \gamma\}$.

Based on the unconditional variance (4.10), the asymptotic theory and Appendix C.1, the specification (4.14) suggests⁸ a theoretical value for γ^2 :

$$\gamma^2 \cong 2 \frac{\Delta^2}{N} \left(\left(\sum_{k=1}^K \theta_k \right)^2 + \sum_{k=1}^K \frac{\sigma_k^2 \theta_k}{2\kappa_k} \right), \quad K = 1. \quad (4.15)$$

If N is sufficiently large, γ^2 is a function of the other parameters. In practice, we can treat γ^2 as a parameter on its own and compare its estimated value to its theoretical value.

The second specification proposes a measurement error⁹ coherent with the asymptotic theory (4.7)-(4.8) and the conditional variance (4.9):

$$R_{t,t+1} = \int_t^{t+1} \sqrt{V_{1,t}} dB_{1,t}, \quad RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \sqrt{\underbrace{2 \frac{\Delta}{N} RQ_{t,t+1}^{(N)} \epsilon_{t,t+1}^{(N)}}_{\cong \eta^2}}. \quad (4.16)$$

where $\epsilon_{t,t+1}^{(N)}$ is centered and of variance one and the parameter η^2 is asymptotically known. In specification (4.16), the measurement error is heteroskedastic, but its unconditional vari-

⁸In Bollerslev and Zhou, 2002, the conditional variance is γ , not γ^2 . We adjust the notation to emphasize the positivity of γ^2 .

⁹This specification is also coherent with Bollerslev et al., 2016

ance is

$$\text{Var} [\xi_{t,t+1}] = \eta^2 \text{E} [\text{E} [RQ_{t,t+1} | \mathcal{F}_{t\Delta}]] = \gamma^2, \quad (4.17)$$

which is coherent with Equation (4.10).

We estimate the parameters using sample moments derived from the realized variance data. For the CM methods, the first group of conditional moments is an obvious choice for estimating the long-term variance and mean-reverting parameters, as these conditional moments relate to the conditional expected value of the realized variance.

Additionally, we introduce instrumental variables, allowing us to increase the number of conditions with the same theoretical moments, which is not feasible with UM moments. The instrumental variables exploit the fact that the residuals of the moment conditions are uncorrelated with them, enabling a different weighting of the sample data and thereby creating an additional condition.

For parameters related to the volatility of volatility and measurement error variance, we rely on information from the squared realized variance with instrumental variables.

For the UM approach, we use the unconditional moments of the entire sample. Consequently, the expected value of the realized variance can only be used once in the GMM for estimating the long-term volatility. The mean-reverting parameter must then be estimated using other moments. As a result, more theoretical moments are required under this method since there are no instrumental variables involved.

In Table 4.1 we present the set of moment conditions. Panel A outlines the moments used in Bollerslev and Zhou, 2002 referred as BZ along with our proposed selection¹⁰ denoted CM to successfully estimate the parameters of both specifications (4.14) and (4.16).

Panel B provides moment selections based on unconditional moments (UM), inspired from Todorov, 2009. He derives a general model including jumps and estimates parameters on the S&P data on two nested models, the two-factor affine jump model and the CARMA

¹⁰An extensive investigation of the procedures is provided in the Online Appendix E.2, only the final procedure is exhibited here. While all the procedures we tested were decent, the one presented here minimized parameter bias.

jump driven SV model. Because there is no jumps component in our model, we exclude moments related to jump components parameters while including the asymptotic refinement.¹¹ In particular, from the seven moment conditions associated with the case of one variance,¹² we omit the two conditions moments regarding the bipower variation, resulting in five moments.¹³ We use γ^2 even though the measurement error is heteroskedastic since, as showed in Equation (4.17), γ^2 is the unconditional variance of the measurement error. As proposed in Corradi and Distaso, 2006, we subtract the sample average¹⁴ to center our observable variables instead of the theoretical average.

CM moment selection

As we are estimating the measurement error parameter, the BZ procedure faces two significant challenges that we solve with our CM procedure.¹⁵

Firstly, the specification is inadequate as the noise distribution contradicts the asymptotic theory. A detailed discussion on this matter is provided in Appendix C.2, where we demonstrate that moment V) and VI) result in an undesirable dependency between the square measurement error and the realized variance under the specification (4.14). To tackle the first challenge, we employ specification (4.16) as it aligns better with the asymptotic theory and solve the dependency problem.

Secondly, moments IV), V) and VI) cannot accurately isolate the measurement error parameter (see Appendix C.5 for details). The main concern is that the moments selected by BZ do not allow the measurement error parameter to have a relative importance with respect to the moment condition¹⁶ and is drowned in the moment's residuals. We address this

¹¹The asymptotic refinement is how Todorov, 2009 includes the measurement error in its moment conditions.

¹²Corradi and Distaso, 2006 used a similar moment selection where $h \in \{1, 2\}$ for moment III).

¹³We tried to add the unconditional mean of the realized quarticity for the case of one variance as it is used for the two variance factors, it did not yield to an improvement of the parameter estimates.

¹⁴We have tried to subtract the theoretical average $\overline{RV}_{t,t+1} = RV_{t,t+1} - \theta\Delta$, which is based on parameters, instead of the sample average $\overline{RV}_{t,t+1} = RV_{t,t+1} - \frac{1}{T} \sum_t RV_{t,t+1}$. It yielded in similar results.

¹⁵Garcia et al., 2011 proposed to use the third conditional moment, which we do not cover here as we include the measurement error. Indeed, this inclusion complicates assumptions and computations for higher moments.

¹⁶The measurement error does not significantly impact moment condition IV) and the addition of IV instruments did not introduce notable bias, as illustrated in Figure E.2 in Online Appendix E.8. For complete

Panel A: Moment selection based on conditional moments		BZ	CM
I)	$RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta$	X	X
II)	$(RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t}$	X	X
III)	$(RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t}^2$	X	
IV)	$(RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J$	X	X
V)	$((RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J) \times RV_{t-1,t}$	X	X
VI)	$((RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J) \times RV_{t-1,t}^2$	X	
VII)	$RV_{t,t+1}RV_{t+1,t+2} - \tilde{H} (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - \tilde{I}RV_{t,t+1} - \tilde{J}$		X
VIII)	$(RV_{t,t+1}RV_{t+1,t+2} - \tilde{H} (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - \tilde{I}RV_{t,t+1} - \tilde{J}) \times RV_{t-1,t}$		X
Panel B: Moment selection based on unconditional moments		UM	
I)	$RV_{t,t+1} - \theta \Delta$	X	
II)	$\overline{RV}_{t,t+1}^2 - \gamma^2 - \frac{\theta \sigma^2}{2\kappa} \left(\frac{2}{\kappa} (\Delta - a) \right)$	X	
III)	$\overline{RV}_{t,t+1} \overline{RV}_{t+h,t+h+1} - \alpha^{h-1} a^2 \frac{\theta \sigma^2}{2\kappa}$		$h \in \{1, 3, 6\}$

where $\overline{RV}_{t,t+1} = RV_{t,t+1} - \frac{1}{T} \sum_t RV_{t,t+1}$ and the unconditional variance of the measurement error is $\text{Var} \left[\xi_{t,t+1}^{(N)} \right] = \gamma^2$.

For Panel A, BZ stands for Bollerslev and Zhou, 2002 and uses specification (4.14) where $\eta^2 RQ_{t,t+1}$ is replaced by γ^2 . CM corresponds to our proposed conditional moment selection and specification (4.16); the measurement error variance is approximated by $\eta^2 RQ_{t,t+1}^{(N)}$. The moments are derived in Appendix C.3.1. For Panel B, we set $\text{Var} \left[\xi_{t,t+1}^{(N)} \right]$ equal to γ^2 as suggested in Equation (4.15) and moments are derived in Appendix C.4. Using a Taylor expansion, $\alpha \approx 1 - \kappa \Delta$, $\beta \approx \kappa \theta \Delta$, $H \approx 1 - 2\kappa \Delta$, $\tilde{H} \approx 1 - \kappa \Delta$, $I \approx (\sigma^2 + 2\kappa \theta) \Delta^2$, $\tilde{I} \approx \left(\frac{\sigma^2}{2} + 2\kappa \theta \right) \Delta^2$, $J \approx \left(\frac{\sigma^2 \theta \kappa}{3} + \kappa^2 \theta^2 \right) \Delta^4$ and $\tilde{J} \approx - \left(\frac{\sigma^2 \kappa \theta}{6} + \kappa^2 \theta^2 \right) \Delta^4$.

Table 4.1: GMM moment selection

second challenge by including moments VII) and VIII) in our moment selection allowing us to precisely estimate the measurement error parameter.

As detailed Appendix C.3.1, the construction of the moment VII) is based on

$$\mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = \mathcal{A}^* V_t^2 + \mathcal{B}^* V_t + \mathcal{C}^* \quad (4.18)$$

$$\mathbb{E} [\mathcal{V}_{t,t+1} \mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] = \mathcal{A} V_t^2 + \mathcal{B} V_t + \mathcal{C}, \quad (4.19)$$

where $\mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}]$ is related to $\mathbb{E} \left[RV_{t,t+1}^2 - \left(\xi_{t,t+1}^{(N)} \right)^2 \middle| \mathcal{F}_{t\Delta} \right]$ (Corollary E.6.4) and is affected by the measurement error. When we condition on the observed filtration $\mathcal{G}_{t\Delta}$, the lagged expected value of the realized variance is not affected by the measurement error (Corollary E.6.6) :

$$\begin{aligned} \mathbb{E} [RV_{t,t+1} RV_{t+1,t+2} | \mathcal{G}_{t\Delta}] &= \mathbb{E} [\mathbb{E} [RV_{t,t+1} RV_{t+1,t+2} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \mathbb{E} [\mathbb{E} [\mathcal{V}_{t,t+1} \mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]. \end{aligned}$$

By combining Equations (4.18) and (4.19), we obtain

$$\begin{aligned} \mathbb{E} [\mathcal{V}_{t,t+1} \mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] &= \frac{\mathcal{A}}{\mathcal{A}^*} \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \left(\mathcal{B} - \mathcal{B}^* \frac{\mathcal{A}}{\mathcal{A}^*} \right) V_t + \left(\mathcal{C} - \mathcal{C}^* \frac{\mathcal{A}}{\mathcal{A}^*} \right). \\ &= \tilde{H} \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \tilde{I} \mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \tilde{J}. \end{aligned}$$

Application to GMM

The GMM estimation method minimises the distance between theoretical moments and their empirical correspondents. Because BZ works with conditional moments, we define at each time step a vector $\mathbf{g}_{t,\theta}$ that, according to the model, should have a conditional expectation of zero. We add instrumental variables that provide additional valid conditions in the GMM framework, allowing for more accurate and consistent estimation of model parameters by leveraging the fact that the instruments are correlated with the observable variables but uncorrelated with the error term arising from the moment condition. If θ is

discussion on the subject see Appendix C.5.

the vector of parameters, the vector of moment conditions is

$$\mathbf{g}_{t,\theta} = \begin{bmatrix} RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta \\ (RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t} \\ (RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t}^2 \\ (RV_{t+1,t+2}^2 - \gamma^2) - H(RV_{t,t+1}^2 - \gamma^2) - IRV_{t,t+1} - J \\ ((RV_{t+1,t+2}^2 - \gamma^2) - H(RV_{t,t+1}^2 - \gamma^2) - IRV_{t,t+1} - J) \times RV_{t-1,t} \\ ((RV_{t+1,t+2}^2 - \gamma^2) - H(RV_{t,t+1}^2 - \gamma^2) - IRV_{t,t+1} - J) \times RV_{t-1,t}^2 \end{bmatrix},$$

where the moment selection is augmented with instrumental variables specifically for CM approaches.

As another example, the UM approach should have a null average over all the time steps, which we refer as unconditional moment. The vector of GMM moment restrictions, in that case, is

$$\mathbf{g}_{t,\theta} = \begin{bmatrix} RV_{t+1,t+2} - \theta \Delta \\ \overline{RV}_{t,t+1}^2 - \gamma^2 - \frac{\theta \sigma^2}{2\kappa} \left(\frac{2}{\kappa} (\Delta - a) \right) \\ \overline{RV}_{t+1,t+2} \overline{RV}_{t,t+1} - a^2 \frac{\theta \sigma^2}{2\kappa} \\ \overline{RV}_{t+3,t+4} \overline{RV}_{t,t+1} - \alpha^2 a^2 \frac{\theta \sigma^2}{2\kappa} \\ \overline{RV}_{t+6,t+7} \overline{RV}_{t,t+1} - \alpha^5 a^2 \frac{\theta \sigma^2}{2\kappa} \end{bmatrix},$$

where, unfortunately, we cannot use instrumental variables in UM approaches.

The average of the i^{th} moment condition over T observations is given by $\bar{g}_{i,\theta} = \frac{1}{T} \sum_{t=1}^T g_{t,i,\theta}$ and all the moments conditions form a vector $\bar{\mathbf{g}}_{\theta}$. The GMM estimator minimizes the following objective function

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \bar{\mathbf{g}}_{\theta}^{\top} \mathbf{W}_{\theta} \bar{\mathbf{g}}_{\theta},$$

where \mathbf{W}_{θ} is the weighting matrix. All details are provided in Appendix C.6. The weighting matrix plays a central role in the GMM procedure, as it optimally combines moment conditions by assigning different weights to each moment, with higher variance moments receiving less weight than more precise ones. Additionally, our choice for the weighting

matrix accounts for the dependence between moment conditions and autocorrelation. As commonly done, we do a 2-step procedure using Newey and West, 1987 for the weighing matrix where the lag is discussed in Online Appendix E.8. Indeed, this Online Appendix also discusses that the GMM moment conditions under UM approaches exhibit higher autocorrelation than under CM approaches, increasing the impact of the choice of the weighting matrix in the estimations.

4.4.1.2 With a drift and leverage effect

We now consider a model that contains a drift in the return dynamics and a leverage effect. This case was studied in Bollerslev and Zhou, 2002, Ishida et al., 2011 Garcia et al., 2011, Bregantini, 2013 and Baum et al., 2021 among others. The drift component includes the risk-free return, a risk premium, and the convexity correction. The proposed model is

$$\begin{aligned}
 R_{t,t+1} &= r\Delta + \left(\lambda - \frac{1}{2}\right) \mathcal{V}_{1,t,t+1} + \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{1,t\Delta}} dZ_{1,t} & (4.20) \\
 RV_{t,t+1}^{(N)} &= \mathcal{V}_{1,t,t+1} + \sqrt{\eta^2 RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}.
 \end{aligned}$$

We adapt the moment selections Table 4.1 to incorporate the two additional parameters, namely the leverage effect and the risk premium. The proposed additional moments are presented in Table 4.2.

In Table 4.2, we use moments related¹⁷ to the lagged value of the realized variance, that is, $R_{t,t+1}RV_{t+h,t+1+h}$, which is coherent with the moment selection in Bollerslev and Zhou, 2002; Bregantini, 2013; Garcia et al., 2011.

¹⁷We tested the conditional moment $R_{t,t+1}RV_{t,t+1} - \widehat{H}(RV_{t,t+1}^2 - \xi_{t,t+1}) - \widehat{I}RV_{t,t+1} - \widehat{J}$ and the unconditional moment $\overline{R}_{t,t+1}\overline{RV}_{t,t+1} - \left(\lambda - \frac{1}{2}\right) \left(a^2 + f(\Delta, \kappa)\right) \frac{\sigma^2\theta}{2\kappa} + \rho \frac{\sigma}{\kappa} \theta (\Delta - a)$, which led to similar results. All details are provided in Appendix C.3 and Appendix C.4 respectively. Results are available upon request.

 Panel A: Conditional moments

$$\begin{aligned}
 \text{IX)} \quad & R_{t,t+1} - \left(\lambda - \frac{1}{2}\right) RV_{t,t+1} - r\Delta \\
 \text{X)} \quad & \left(R_{t,t+1} - \left(\lambda - \frac{1}{2}\right) RV_{t,t+1} - r\Delta\right) \times RV_{t-1,t} \\
 \text{XI)} \quad & \left(R_{t,t+1} - r\Delta - \left(\lambda - \frac{1}{2}\right) RV_{t,t+1}\right) RV_{t+1,t+2} - \bar{I}RV_{t,t+1} - \bar{J} \\
 \text{XII)} \quad & \left(\left(R_{t,t+1} - r\Delta - \left(\lambda - \frac{1}{2}\right) RV_{t,t+1}\right) RV_{t+1,t+2} - \bar{I}RV_{t,t+1} - \bar{J}\right) \times RV_{t-1,t}
 \end{aligned}$$

 Panel B: Unconditional moments

$$\begin{aligned}
 \text{IV)} \quad & R_{t,t+1} - \left(\lambda - \frac{1}{2}\right) \theta\Delta - r\Delta \\
 \text{V)} \quad & \bar{R}_{t,t+1} \bar{R}V_{t+h,t+h+1} - \alpha^{h-1} a^2 \left(\left(\lambda - \frac{1}{2}\right) \frac{\theta\sigma^2}{2\kappa} + \rho\sigma\theta \right), \text{ For h, see Table 4.1}
 \end{aligned}$$

where $\bar{R}V_{t,t+1} = RV_{t,t+1} - \frac{1}{T} \sum_t RV_{t,t+1}$ and $\bar{R}_{t,t+1} = R_{t,t+1} - \frac{1}{T} \sum_t R_{t,t+1}$

Table 4.2: Details for Panels A and B are provided in Appendix C.3 and Appendix C.4 respectively.

Table 4.2: Additional moments for specification (4.20)

4.4.2 Simulation study

We conduct a Monte Carlo simulation study using two groups of parameters presented in Table 4.3. Panel A includes the first group with the three parameter sets proposed in BZ, which are widely used in the literature. Additionally, we introduce a second group of parameters that are more closely related to stock markets. The simulation scheme uses $\Delta = 1/252$. This table also includes the drift and leverage component, which is excluded in the drift-free specification. The choices for ρ and λ are discussed in the Section 4.4.2.2.

For Panel A, we annualized the parameters¹⁸ from BZ, with the initial parameters presented in parentheses; details on the rescaling are provided in Appendix C.6.3. The second group of parameters are inspired by Christoffersen et al., 2009, specifically from their Table 3 (Panel A for $V_{1,t\Delta}$ and Panel B for $V_{2,t\Delta}$ and $V_{3,t\Delta}$). In the Panel B of Table 4.3, the components have respectively a moderate, high and low speed of mean-reversion, vol-of-vol and leverage effect.

¹⁸The parameters in BZ are in percentage.

Parameters	$V_{k,t\Delta}$	$V_{k,t\Delta}$	$V_{k,t\Delta}$
<i>Panel A. Parameters related to BZ</i>			
Scenario	A	B	C
κ_k	7.56 (0.03)	25.2 (0.10)	25.2 (0.10)
θ_k	0.0063 (0.25)	0.0063 (0.25)	0.0063 (0.25)
σ_k	0.252 (0.10)	0.252 (0.10)	0.504 (0.20)
ρ_k	-0.50	-0.50	-0.50
<i>*There is no drift for those scenarios</i>			
$e^{-\kappa_k\Delta}$	0.9704	0.9048	0.9048
$\frac{\text{Var}[RV_{t,t+1}]}{\text{Var}[\mathcal{V}_{t,t+1}]}$	1.0610	1.1463	1.0549
γ^2	2.5407e-11 (0.0025)	1.8293e-11 (0.0018)	2.7439e-11 (0.0027)
η^2	9.6787e-05 (0.0244)	9.6787e-05 (0.0244)	9.6787e-05 (0.0244)
<i>Panel B. Stock Market</i>			
Scenario	D	E	F
κ_k	2.00	4.00	0.20
θ_k	0.05	0.03	0.02
σ_k	0.15	0.30	0.12
ρ_k	-0.85	-0.95	-0.50
λ_k	1.15	1.95	2.90
$e^{-\kappa_k\Delta}$	0.9920	0.9842	0.9992
$\frac{\text{Var}[RV_{t,t+1}]}{\text{Var}[\mathcal{V}_{t,t+1}]}$	1.2412	1.0894	1.0379
γ^2	1.0682e-09	4.7529e-10	4.3016e-10
η^2	9.6787e-05	9.6787e-05	9.6787e-05

Table 4.3: The Monte Carlo simulation study follows the simulation procedure outlined in Section 4.3 with $\Delta = 1/252$. In Panel A, we multiply κ by 252, θ by $252/100^2$ and σ by $252/100$. Initial parameters from BZ are in parenthesis.

Table 4.3: Parameters for the simulation study

4.4.2.1 Without drift and leverage effect

The model from Section 4.2 is set with $K = 1$ along with a null drift from Equation (4.1) and the leverage parameter set to $\rho_1 = 0$. We simulate $S = 1000$ independent trajectories of length of $T = 4000$ days.

For the BZ and CM selections of Panel A of Table 4.1, our procedure (CM) offers two significant contributions compared to BZ. Firstly, it provides a coherent modeling of the measurement error process by using specification (4.16) instead of specification (4.14). Furthermore, by incorporating additional moment conditions VII) and VIII), it improves the estimations for the measurement error parameter, as evidenced by the lower RMSE for the vol-of-vol parameter σ of Figure 4.1. The GMM estimation detailed results based on BZ parameters are provided in Appendix C.7, as it serves as a common benchmark in the literature.¹⁹

We test the methodology with another set of parameters inspired by stock markets. We first look at the first set of volatility parameters in Table 4.3, Scenario D. As demonstrated in Figure 4.2, when the integrated variance is observed, the CM moment selection results in a more precise volatility of volatility parameter than the UM procedure. However, when observing the realized variance, as shown in Figure 4.3, the CM moment selection experiences significant challenges for all parameters, even when the measurement error is fixed to its theoretical value.

This issue arises only with this variance. We believe it is because the proportion of the measurement error is much higher with this variance compared to the other variances. Indeed, if we examine the theoretical ratio of the variance of the measurement error over the integrated variance, we have

$$\frac{\text{Var} [RV_{t,t+1}]}{\text{Var} [\mathcal{V}_{t,t+1}]} = 1 + \frac{\text{Var} [\xi_{t,t+1}]}{\text{Var} [\mathcal{V}_{t,t+1}]} = 1 + \frac{2\frac{\Delta^2}{N} \left(\theta^2 + \frac{\sigma^2\theta}{2\kappa} \right)}{\frac{\theta\sigma^2}{2\kappa} \left(\frac{2}{\kappa} (\Delta - a) \right)}$$

¹⁹An extensive investigation of which CM procedures to use is provided in the Online Appendix E.2; however, only the final procedure is exhibited here. A detailed comparison of both conditional moment selection under specification (4.14) is provided in Online Appendix E.3. We conclude that adding additional moments allows for adequate estimation of the measurement error parameter, which is not achievable under the BZ procedure alone.

$$\approx 1 + \frac{2}{N} \frac{\theta^2 + \frac{\sigma^2 \theta}{2\kappa}}{\frac{\sigma^2 \theta}{2\kappa}} = 1 + \frac{2}{N} \left(1 + \frac{2\kappa \theta}{\sigma^2} \right),$$

since $\frac{2}{\kappa}(\Delta - a) \approx \Delta^2$. Table 4.3 shows the ratio for each parameter set, and the first variance in Panel B exhibits the highest ratio. Furthermore, running the Scenario A with $\theta = 0.02$ instead decreases the ratio to 1.11 compared to 1.24 and solves the estimation problem for the CM procedure. Therefore, we argue that the UM moment selection is more robust to the presence of a high measurement error.

A potential solution for the CM moment selection would be to derive a moment completely independent of the measurement error parameter and discard its estimation, such as the recursion $RV_{t+1,t+2}RV_{t,t+1}$ on its lagged values instead of the squared realized variance as it is impacted by the measurement error.

The second set of volatility parameters in Table 4.3 Panel B performs well in the GMM estimations based on CM and UM moment selection as displayed in Table 4.4. However, similar to the first set of parameters in Panel A of Table 4.3, the bias and RMSE for the mean-reverting parameter is smaller with CM, though there is slightly higher RMSE for the volatility of volatility under CM compared to UM.

The third set of volatility parameters in Table 4.3, Panel B, performs poorly for both UM and CM, even when using the true integrated variance as the realized variance. We attribute this to the high persistence of 0.9992. Indeed, when we replace the κ value with 1 instead of 0.20, the UM procedure yields satisfactory results, as demonstrated in Figure 4.5. However, when the measurement parameter is estimated, the CM procedure encounters more difficulties compared to UM. Additionally, when the measurement parameter is fixed, UM is more stable. Therefore, we observe that under conditions of high persistence, UM outperforms CM procedures.

In summary, UM methodologies seem more consistent and robust to parameter changes across the three stock market scenarios. While CM methods are generally more powerful due to their use of instrumental variables, when it is not possible to filter the conditional moment of the latent instantaneous variance, the errors introduced by the recursions can be more detrimental than beneficial in some cases, which may explain why UM appears to

be a more stable methodology.

4.4.2.2 With drift and leverage effect

The model from Section 4.2 is set with $K = 1$ along with a drift from Equation (4.1) and the non-null leverage parameter ρ_1 . For Panel A of Table 4.3, we are using the same parameter sets as in Section 4.4.2.1, but with a leverage parameter²⁰ set to $\rho = -0.5$ as in Garcia et al., 2011 Table 2. Additionally, Ishida et al., 2011 and Bregantini, 2013 also use the same leverage parameter for their simulation study.²¹ Furthermore, we exclude any drift component so the results are comparable with the literature, results are presented in Appendix C.7. The main conclusion is that we do as good or better performance compared to Garcia et al., 2011 and Bregantini, 2013. Furthermore, CM procedures performs better than UM procedures when there is no drift in the model, both with and without measurement error on the realized variance.

For Panel B of Table 4.3, we use use a median, high and lower negative leverage effect respectively. The Scenario F mean-reverting parameter is set to $\kappa = 1$. The risk premium parameters for stock markets scenarios are set so the approximate annual return of $e^{\sum \theta_{ann} \times \lambda} - 1 \approx 6\%$ for the second panel along with an annual risk-free rate to 2%.

When we include a drift and a leverage effect in the stock market scenarios, we arrive at the same conclusions as in the case without the leverage effect. First, when there is a high measurement error (Scenario D), the CM performs poorly, particularly for the leverage parameter (Table C.5 Panel A.2 to B.3), where the estimates fall outside the bounds of a correlation parameter. However, when there is no measurement error (Table C.5 Panel A.1 and B.1), the opposite is true with the leverage parameter estimated with great precision under the CM procedure.

Secondly, both CM and UM perform adequately for Scenario E (Table C.6), with CM performing slightly better. However, the leverage parameter falls outside its bounds espe-

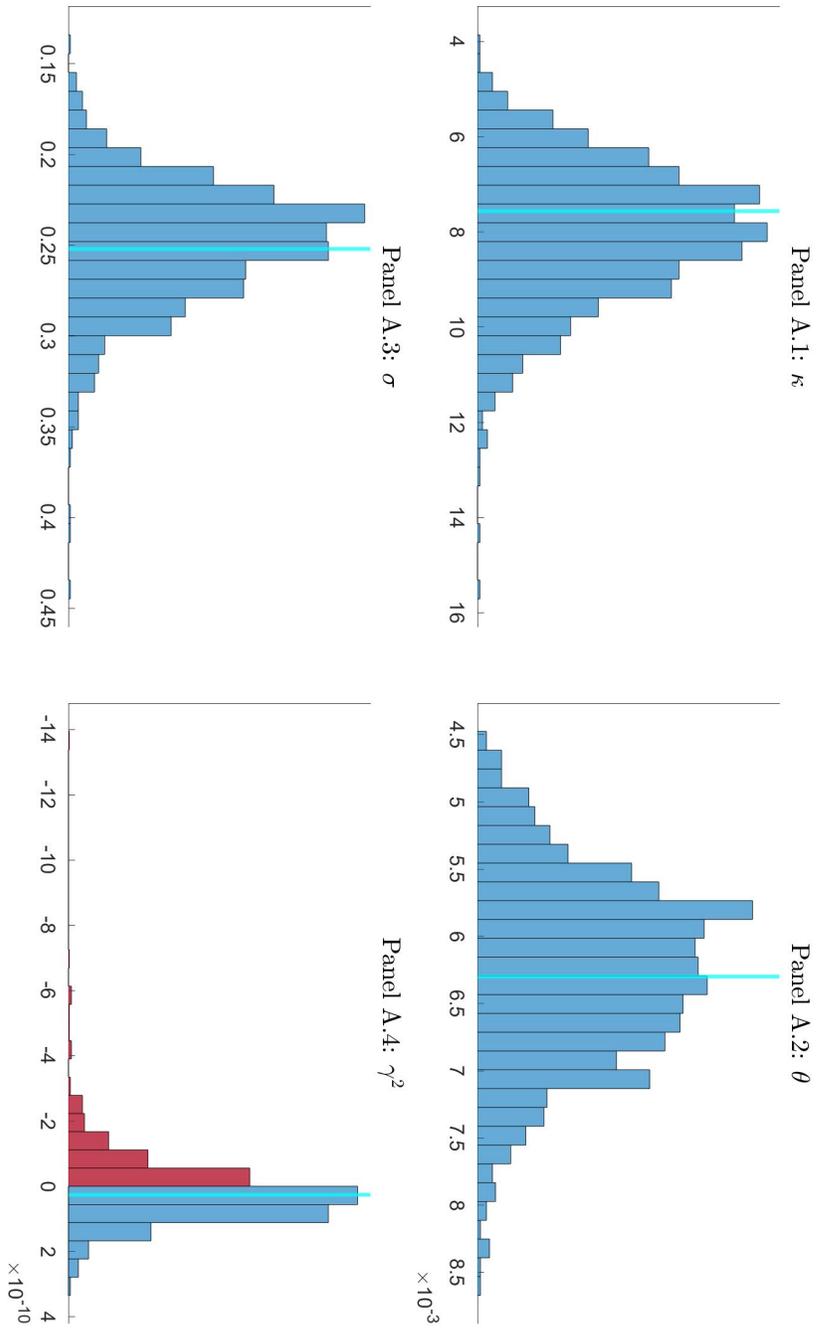
²⁰Baum et al., 2021 estimate the leverage parameter for the S&P500 of -0.44 before the financial crisis and -0.351 after. Bregantini, 2013 estimated a parameter of -0.46 for the S&P500 from 97-2011. Ishida et al., 2011 estimated a parameter of -0.57, motivating our parameter choice for our simulation study.

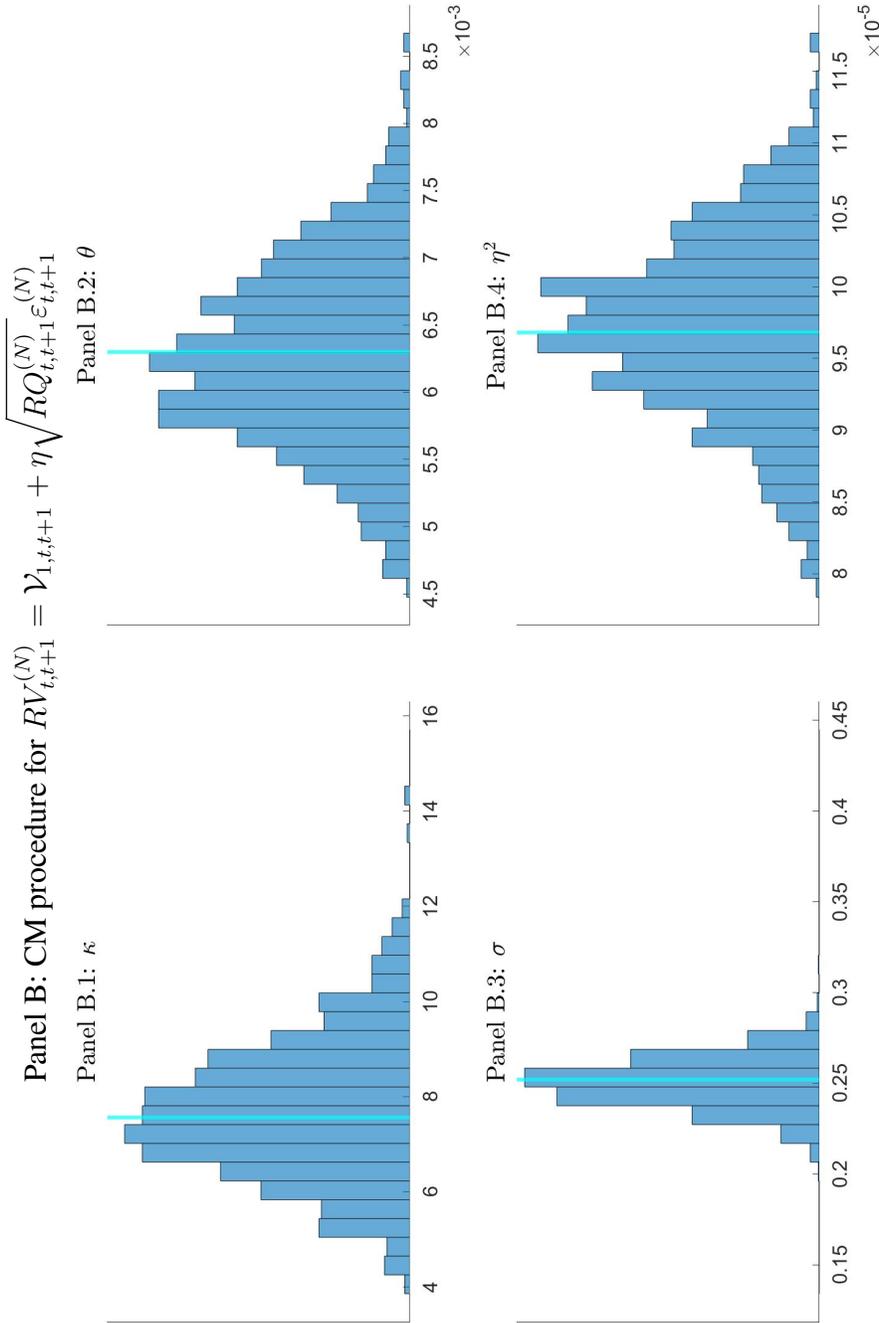
²¹The paper shows simulation results in Table 3. However, it does not address measurement error.

cially for a leverage parameter of -0.95 , which is very close to the boundary where issues can arise.

Lastly, in Scenario F (Table 4.5), without any measurement error, CM performs better. With measurement error fixed, UM proves have lower RMSE on θ, σ and ρ , but higher bias than CM for almost all parameters. However, we believe the UM to be a more robust procedure for estimating the parameters. Indeed, when the measurement parameter is estimated, the CM procedures encounter difficulties. This reinforces the conclusions of the last section, which suggest that UM procedures are more robust.

Panel A: BZ procedure for $RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \gamma \epsilon_{t,t+1}^{(N)}$

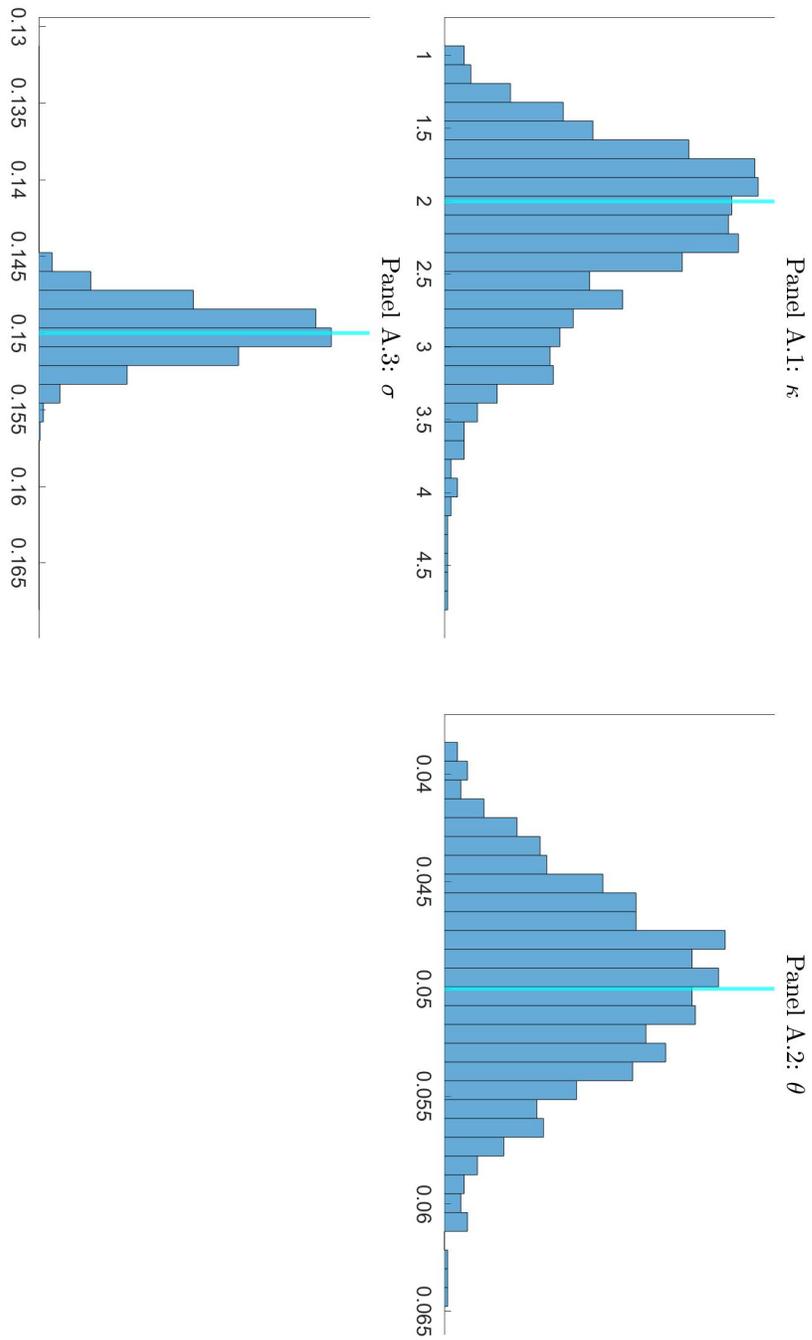


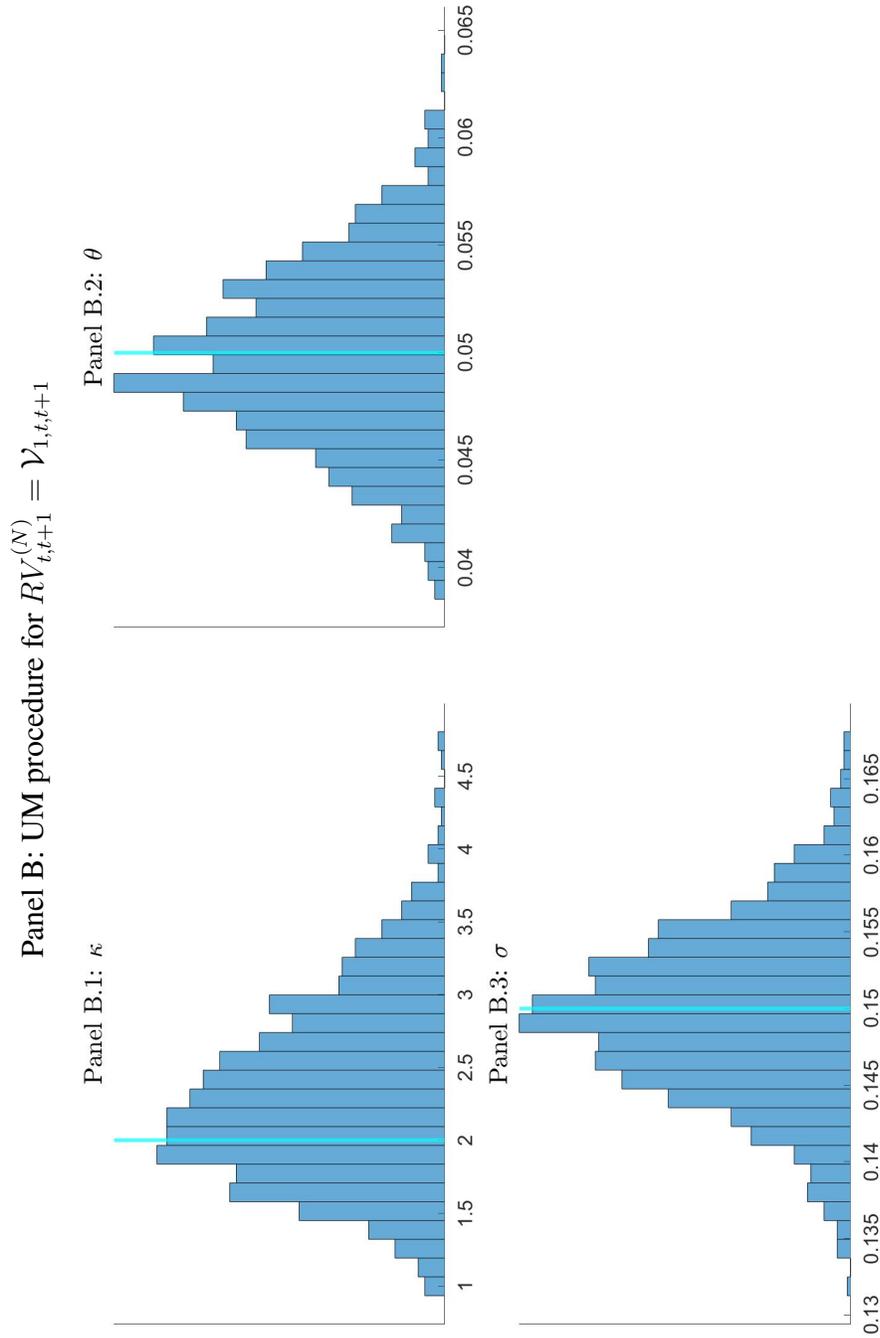


1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1/252$ are generated with parameter of scenario A $\{\kappa = 7.56, \theta = 0.0063, \sigma = 0.252\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. It uses the moments of column BZ (Panels A.1 to A.4) and CM (Panels B.1 to B.4) from Table 4.1 according to specification (4.14) and specification (4.16) respectively. The true parameters and the theoretical values of the error measurement parameter (γ^2 or η^2) and presented by the hard line. We run a 2-step GMM procedure using Newey-West covariance matrix estimator with a Bartlett-kernel with a lag of 5, details are provided in Appendix C.6.

Figure 4.1: GMM results comparison - Scenario A

Panel A: CM procedure for $RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1}$



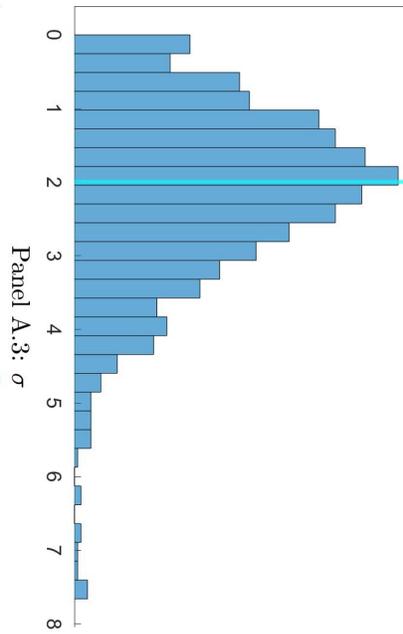


1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1/252$ are generated with parameter of scenario D $\{\kappa = 2, \theta = 0.05, \sigma = 0.15\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. It uses the moments of column CM (Panels A.1 to A.3) and UM (Panels B.1 to B.3) from Table 4.1. We run a 2-step GMM procedure using Newey-West covariance matrix estimator with a Bartlett-kernel with a lag of 5, details are provided in Appendix C.6.

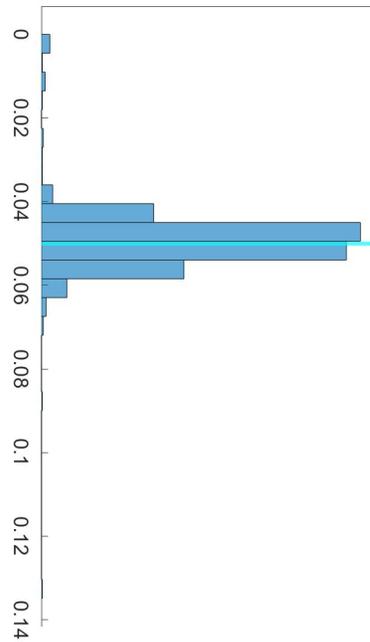
Figure 4.2: GMM results without error measurement - Scenario D

$$\text{Panel A: CM procedure for } RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \eta \sqrt{RQ_{t,t+1}^{(N)}} \varepsilon_{t,t+1}^{(N)}$$

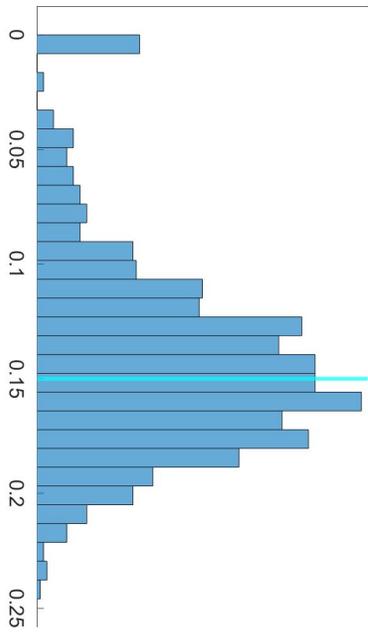
Panel A.1: κ

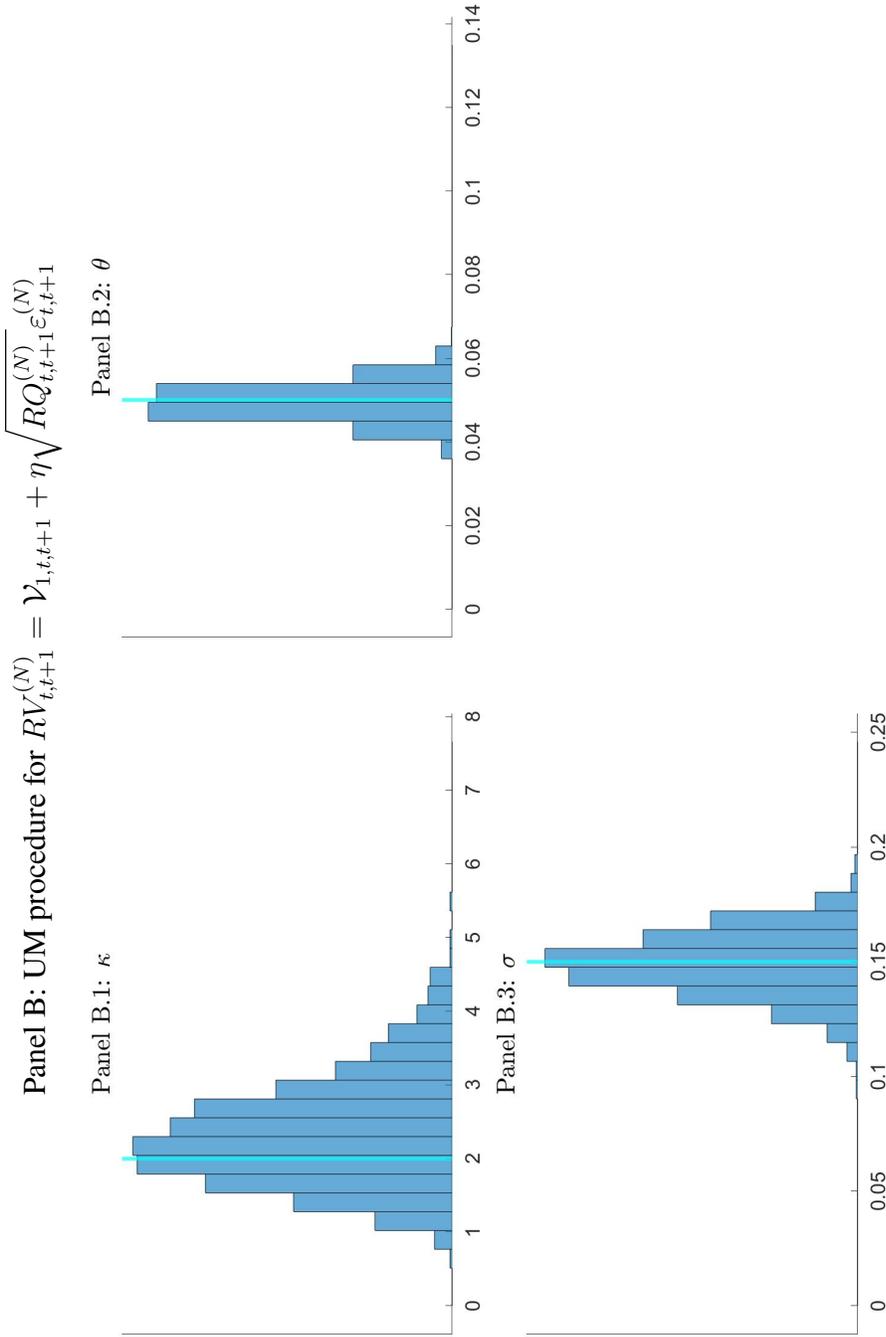


Panel A.2: θ



Panel A.3: σ





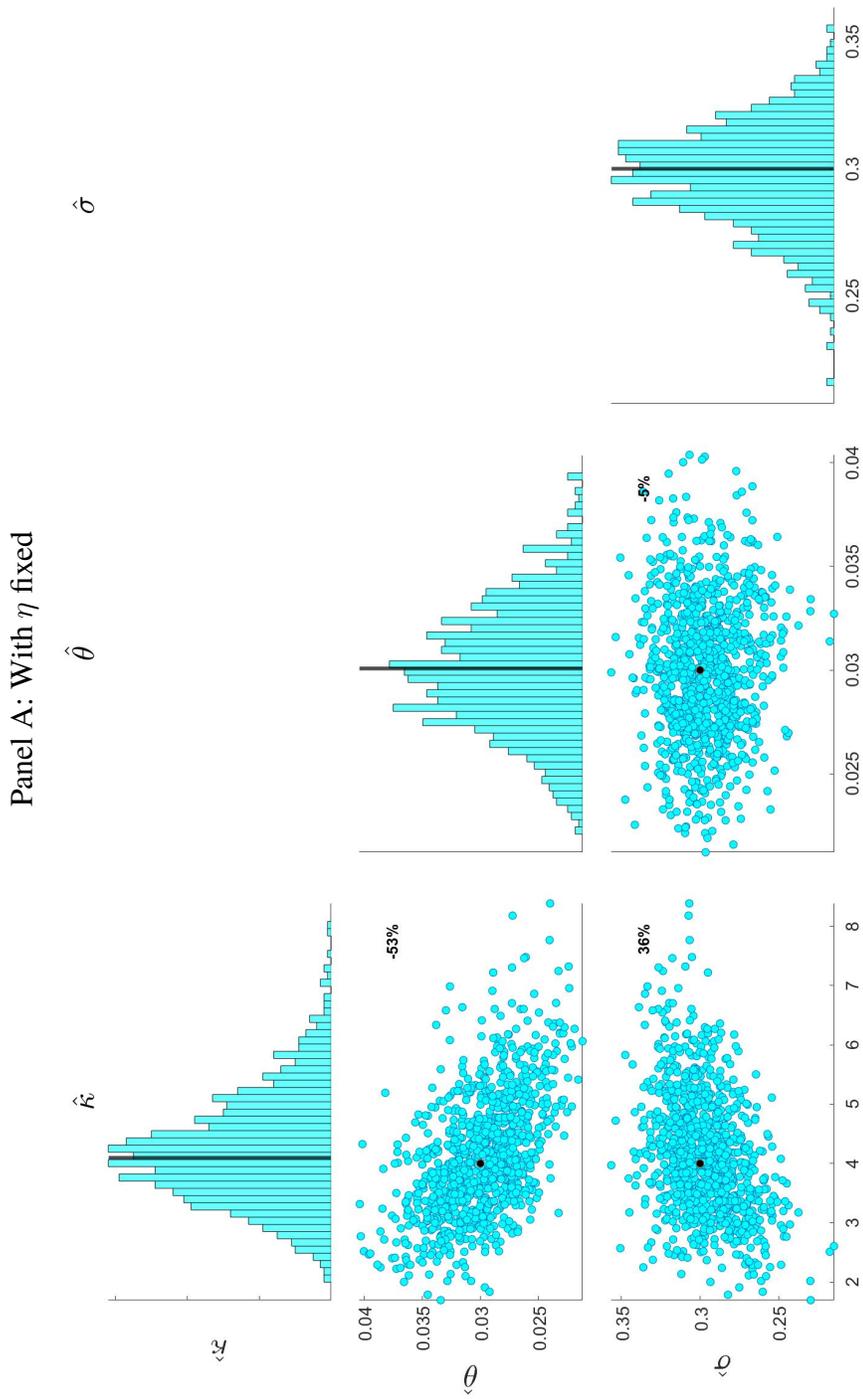
1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1/252$ are generated with parameter of scenario D $\{\kappa = 2, \theta = 0.05, \sigma = 0.15\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. It uses the moments of column CM (Panels A.1 to A.3) and UM (Panels B.1 to B.3) from Table 4.1. We run a 2-step GMM procedure using Newey-West covariance matrix estimator with a Bartlett-kernel with a lag of 5, details are provided in Appendix C.6.

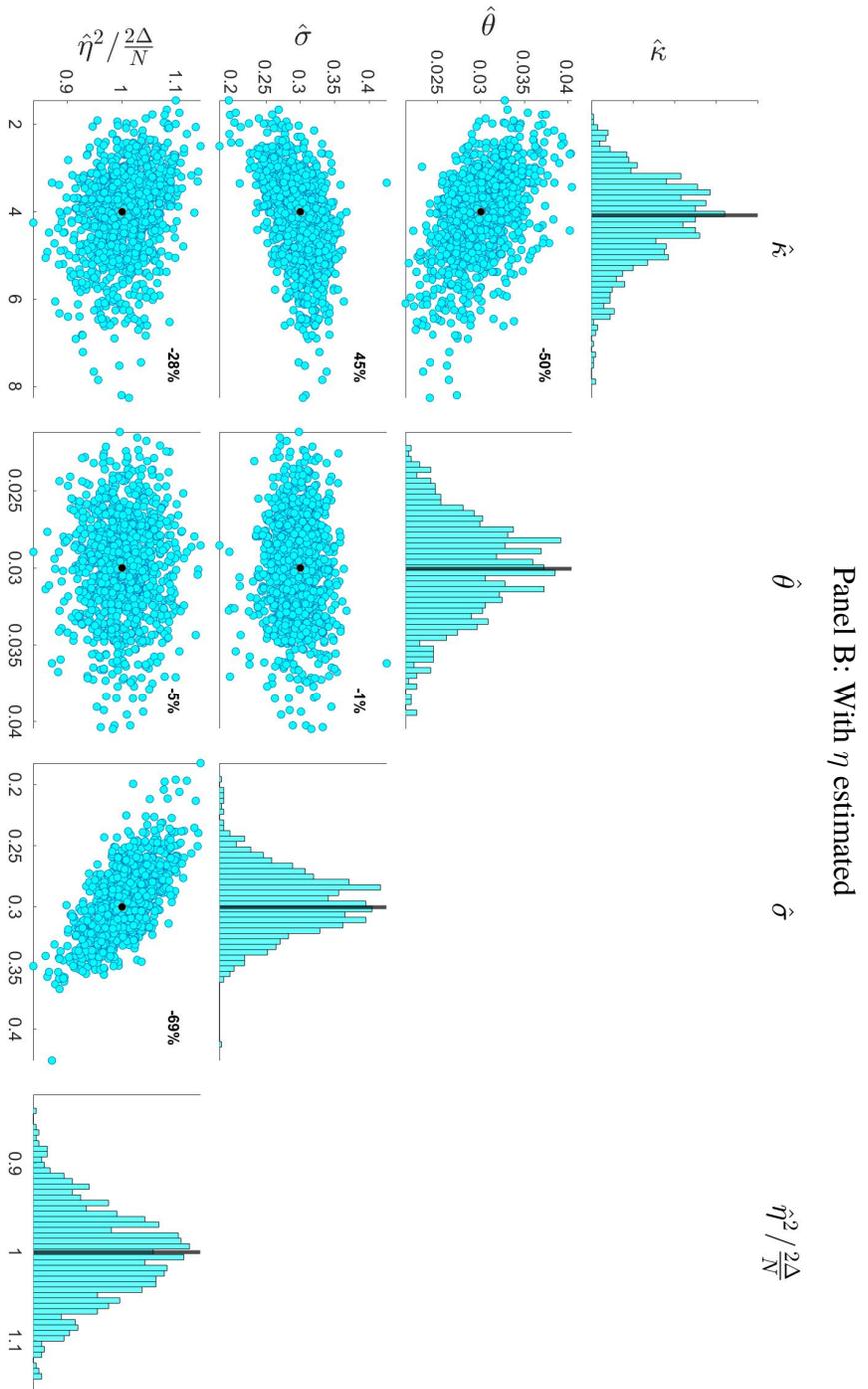
Figure 4.3: GMM results with error measurement - Scenario D

	True value	Mean	Median	RMSE	CR
Panel A.1: BZ with γ fixed and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.2859	4.2125	1.0578	0.9370
$10^2 \times \theta = 3.0000$		2.9886	2.9762	0.3359	0.9490
$\sigma = 0.3000$		0.2957	0.2962	0.0224	0.9370
Panel B.1: CM with η fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.1505	4.0500	1.0267	0.9460
$10^2 \times \theta = 3.0000$		2.9963	2.9867	0.3362	0.9520
$\sigma = 0.3000$		0.2968	0.2981	0.0206	0.9430
Panel C.1: UM with γ fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.5932	4.4878	1.2096	0.9040
$10^2 \times \theta = 3.0000$		2.9348	2.9274	0.3308	0.9460
$\sigma = 0.3000$		0.2994	0.2993	0.0138	0.9500
Panel A.2: BZ with γ estimated and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.2994	4.2233	1.0929	0.9440
$10^2 \times \theta = 3.0000$		2.9907	2.9868	0.3360	0.9460
$\sigma = 0.3000$		0.2892	0.2901	0.0667	0.9500
$10^{10} \times \gamma^2 = 4.7529$		3.3826	5.9998	23.0157	0.9580
Panel B.2: CM with η estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.1215	4.0284	1.0578	0.9440
$10^2 \times \theta = 3.0000$		2.9984	2.9873	0.3367	0.9480
$\sigma = 0.3000$		0.2953	0.2970	0.0288	0.9540
$10^4 \times \eta^2 = 0.9679$		0.9705	0.9707	0.0472	0.9490
Panel C.2: UM with γ estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$					
$\kappa = 4.0000$		4.5139	4.4144	1.1751	0.9120
$10^2 \times \theta = 3.0000$		2.9650	2.9574	0.3289	0.9480
$\sigma = 0.3000$		0.2997	0.2998	0.0140	0.9550
$10^{10} \times \gamma^2 = 4.7529$		4.6071	4.5001	1.0759	0.9580

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1/252$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

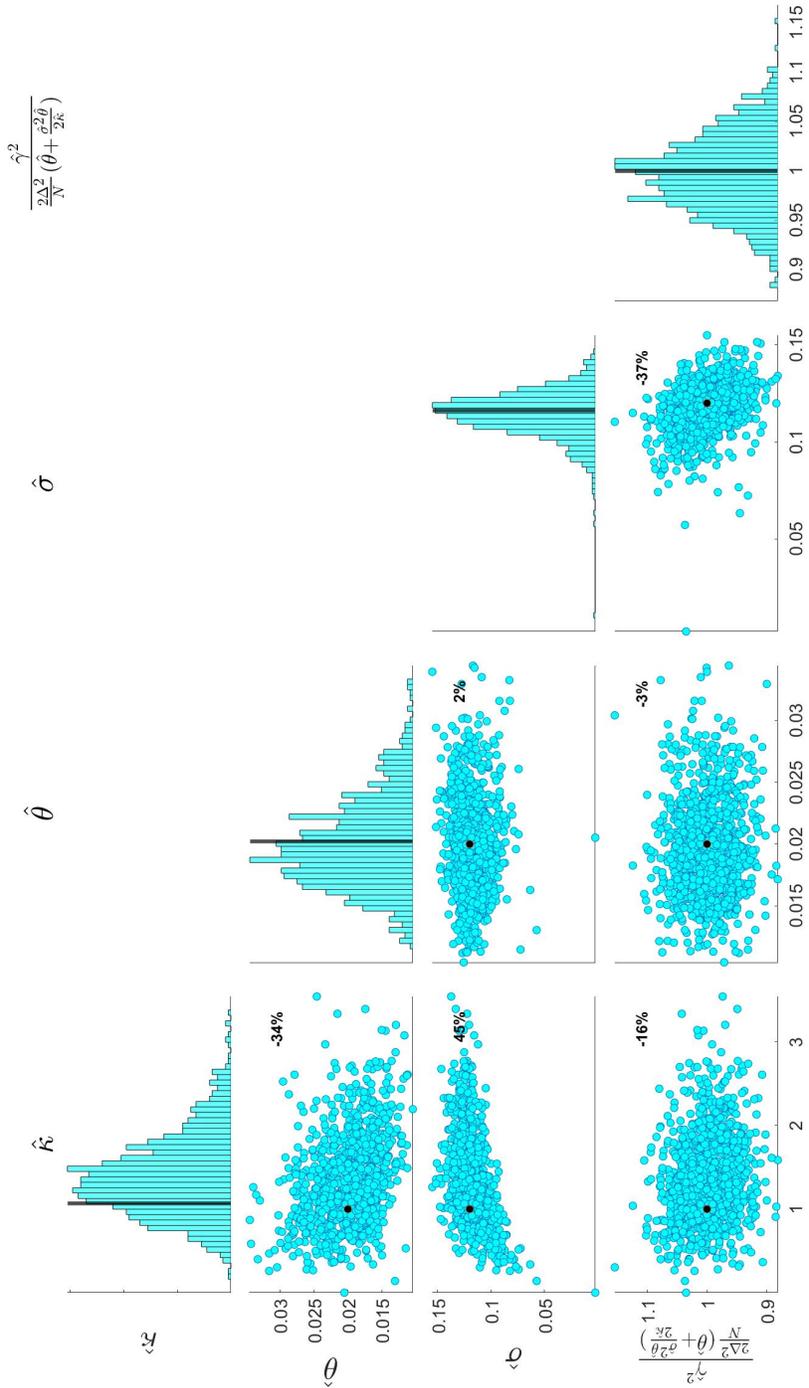
Table 4.4: Comparison results for parameter set E for specifications (4.14)-(4.16)





1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1/252$ are generated with parameter of scenario E $\{\kappa = 4, \theta = 0.03, \sigma = 0.30\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Each diagonal panel presents the histogram of the point estimate parameters based on the GMM optimization described in Appendix C.6. The other panels show scatter plots between two parameters. The black line or dark dot represents the theoretical value of the parameters. The number in the right corner of each panel is the correlation between the parameter estimates.

Figure 4.4: Results for specification (4.16) using CM moment selection and parameter of scenario E



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1/252$ are generated with parameter of scenario F $\{\kappa = 4, \theta = 0.03, \sigma = 0.30\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Each diagonal panel presents the histogram of the point estimate parameters based on the GMM optimization described in Appendix C.6. The other panels show scatter plots between two parameters. The black line or dark dot represents the theoretical value of the parameters. The number in the right corner of each panel is the correlation between the parameter estimates.

Figure 4.5: Results for specification (4.16) using UM moment selection and parameter of scenario F

True value	Mean	Median	RMSE	CR
Panel A.1: <i>CM estimations results with no measurement error</i>				
$\kappa = 1.0000$	1.1715	1.1182	0.4501	0.9150
$10^2 \times \theta = 2.0000$	1.9933	1.9409	0.4287	0.9570
$\sigma = 0.1200$	0.1199	0.1199	0.0014	0.9410
$\rho = -0.5000$	-0.4980	-0.4980	0.0227	0.9540
$\lambda = 2.9000$	3.1440	3.0319	1.8448	0.9500
Panel B.1: <i>UM estimations results with no measurement error</i>				
$\kappa = 1.0000$	1.3928	1.3211	0.6384	0.8670
$10^2 \times \theta = 2.0000$	1.9647	1.9126	0.4073	0.9590
$\sigma = 0.1200$	0.1191	0.1191	0.0053	0.9450
$\rho = -0.5000$	-0.4606	-0.4494	0.1763	0.9440
$\lambda = 2.9000$	3.2126	3.1252	1.8715	0.9460
Panel A.2: <i>CM estimations results with η^2 fixed</i>				
$\kappa = 1.0000$	1.1580	1.1034	0.5854	0.9331
$10^2 \times \theta = 2.0000$	2.0516	1.9652	0.6557	0.9777
$\sigma = 0.1200$	0.1156	0.1180	0.0248	0.9422
$\rho = -0.5000$	-0.5420	-0.5053	0.2493	0.9726
$\lambda = 2.9000$	3.1320	3.0145	1.8458	0.9513
Panel B.2: <i>UM estimations results with γ^2 fixed</i>				
$\kappa = 1.0000$	1.4227	1.3882	0.6970	0.8820
$10^2 \times \theta = 2.0000$	1.9215	1.8766	0.4055	0.9640
$\sigma = 0.1200$	0.1174	0.1182	0.0129	0.9410
$\rho = -0.5000$	-0.4703	-0.4481	0.2002	0.9590
$\lambda = 2.9000$	3.3053	3.2145	1.9270	0.9520

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1/252$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table 4.5: Comparison results for parameter set F for specification (4.20)

4.5 GMM Estimation with Two Variances

4.5.1 GMM derivation

We extend the specification (4.16) by adding an additional variance factor:

$$R_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{1,t}} dB_{1,t} + \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_{2,t}} dB_{2,t}, \quad (4.21)$$

$$RV_{t,t+1}^{(N)} \approx \mathcal{V}_{1,t,t+1} + \eta \sqrt{RQ_{t,t+1}^{(N)}} \varepsilon_{t,t+1}^{(N)},$$

where $\varepsilon_{t,t+1}^{(N)}$ is a standardized random variable. As for the case with one variance, we assume that the error measurement is not correlated²² with the Brownian motions B_1 and B_2 .

The parameter estimation of a two-variance stochastic model is explored in Bollerslev and Zhou, 2002, where the authors present a conditional moment approach. Additionally, Todorov et al., 2011, Todorov, 2009, and Todorov, 2011 propose a UM approach to estimate a two stochastic volatility models, among other more sophisticated models that also account for jumps.

Table 4.6 presents the moments used for the GMM optimization for approaches BZ and UM, inspired by Todorov, 2009.

In Appendix C.5.2, we conduct a comprehensive examination of issues related to BZ moment selection, presented in Panel A of Table 4.6. Particularly for moment II), we show that the recursions proposed by BZ for the two-variance case introduce a significant amount of noise, which is detrimental to the signal used for parameter estimation. For that reason, we do not present any additional CM procedures as the computation become rapidly cumbersome and unpractical.

In Panel B of Table 4.6, we report the moment selection from Todorov, 2009 without the two additional moments related to the separate the jump components from the continuous

²²A thorough discussion of the one variance factor model is provided in Appendix C.2.

Panel A: BZ moment selection

$$\begin{aligned} \text{I)} & \left((1 - e^{-\kappa_1 \Delta L})^2 (1 - e^{-\kappa_2 \Delta L})^2 (RV_{t+5,t+6} - (\theta_1 + \theta_2)\Delta) \right) \times \mathcal{Z}_t \\ \text{II)} & \left(\begin{array}{c} (1 - e^{-\kappa_1 \Delta L}) (1 - e^{-\kappa_2 \Delta L}) (1 - e^{-2\kappa_1 \Delta L}) \\ (1 - e^{-2\kappa_2 \Delta L}) (1 - e^{-\kappa_1 \Delta L} e^{-\kappa_2 \Delta L}) (RV_{t+5,t+6}^2 - \gamma^2) \end{array} - f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2) \right) \times \mathcal{Z}_t \end{aligned}$$

where L is the lag operator and \mathcal{Z}_t represents the instrumental variables: $1, RV_{t-1,t}, RV_{t-7,t-6}, RV_{t-1,t}^2$ and $RV_{t-7,t-6}^2$

Panel B: UM moment selection

$$\begin{aligned} \text{I)} & \overline{RV}_{t,t+1} \\ \text{II)} & RQ_{t,t+1} - (\theta_1 + \theta_2)^2 \Delta - \left(\frac{\theta_1 \sigma_1^2}{2\kappa_1} + \frac{\theta_2 \sigma_2^2}{2\kappa_2} \right) \Delta \\ \text{III)} & \overline{RV}_{t,t+1}^2 - \gamma^2 - \frac{\theta_1 \sigma_1^2}{2\kappa_1} \left(\frac{2}{\kappa_1} (\Delta - a_1) \right) - \frac{\theta_2 \sigma_2^2}{2\kappa_2} \left(\frac{2}{\kappa_2} (\Delta - a_2) \right) \\ \text{IV)} & \overline{RV}_{t,t+1} \overline{RV}_{t+h,t+h+1} - \alpha_1^{h-1} a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} - \alpha_2^{h-1} a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2}, \quad h \in \{1, 3, 6\} \\ \text{V)} & \frac{1}{10} \sum_{l=h+10}^{(h+1)10} \left(\overline{RV}_{t,t+1} \overline{RV}_{t+l,t+l+1} - \alpha_1^{l-1} a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} - \alpha_2^{l-1} a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2} \right), \quad h \in \{1, 2, 3\} \end{aligned}$$

where $\overline{RV}_{t,t+1} = RV_{t,t+1} - \frac{1}{T} \sum_t RV_{t,t+1}$

The details of the moments of Panel A and $f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2)$ are derived in Appendix C.3.2. Panel B and C are detailed in Appendix C.4.

Table 4.6: GMM moments selection for specification (4.21)

components. The paper also reparameterizes the model, as it does not estimate the two long-term variances separately.²³ Therefore, the parameter set is $\kappa_1, \kappa_2, \theta_1 + \theta_2, \frac{\theta_1 \sigma_1^2}{2\kappa_1}$ and $\frac{\theta_2 \sigma_2^2}{2\kappa_2}$.

In panel B, moment I) isolates the sum of the long-term variances $(\theta_1 + \theta_2)$, while moment II) deals with the sum of the unconditional variances $\left(\frac{\theta_1 \sigma_1^2}{2\kappa_1} + \frac{\theta_2 \sigma_2^2}{2\kappa_2} \right)$. Moment III) stands out as the only moment where the measurement error parameter comes into play. Finally, moments IV) and V) permit to have multiple conditions to disentangle the mean-reverting parameters and the long-term variances without being affected by the error measurement parameter.

4.5.2 Simulation study

The model from Section 4.2 is set with $K = 2$ along with a null drift from Equation (4.1) and leverage parameters set to zero.

We simulate $S = 1000$ independent trajectories of $T = 7500$ days. We did increase the number of time steps since two-variance factor case is more complex to pin down.

²³Todorov et al., 2011 estimated the two long term variances by using realized Laplace Transform.

Our attempt to replicate the results using the BZ-2FSV (Panel A of Table 4.6) procedure, even without introducing any measurement error, did not yield conclusive results;²⁴ a discussion of various potential issues is provided in Appendix C.5.2. In summary, the recursion shown in Panel A requires numerous lags of the realized variance, which amplifies the noise surrounding the signal. Additionally, the constant in moment II) includes both the long-term variances and the volatility of volatility parameters, making it challenging to disentangle these parameters.

However, we simulated a case, namely Scenario G, with the parameters of Table 6 of Todorov et al., 2011. The parameters are $\kappa = \{0.0186, 1.3776\}$, $\theta = \{0.6100, 0.4434\}$, $\sigma = \{0.1504, 1.1052\}$.²⁵ The results obtained using the UM procedures are presented in Figure 4.6. Although there are two cases of high correlation above 70% between the estimates, it does not appear to be any linear dependency between the parameter estimates. Furthermore, the black dots in the off-diagonal scatter representing the theoretical value are centered for almost all parameters, with the exception of κ_1 , which was the case with the one variance model.

We also tested all the possible pair combinations of Panel B of Table 4.3 without any success. Our intuition is that the persistence of each variances are too close.

To further our point, we run Scenario G several times while decreasing the value of the mean-reverting parameter of the second variance. We adjust the value of the vol-of-vol parameter to have the same long term variance (and also respect the Feller condition), that is $\frac{\sigma_2^2 \theta_2}{2\kappa_2} = 0.1965$. We try $\kappa_2 \in \{0.90, 0.60, 0.30, 0.10\}$.

Problems in the parameter estimation start to show when $\kappa_2 = 0.10$, which is equivalent to a persistence of 0.90, coupled with $\kappa_1 = 0.0186$ (persistence of 0.9816). For Panel B of Table 4.3, the persistence are respectively 0.9920, 0.9842 and 0.9992 (0.9960 if $\kappa = 1$), which are much higher than the one tested with Scenario G. We argue that for the UM methods to be effective, the two variances must have different levels of persistence.

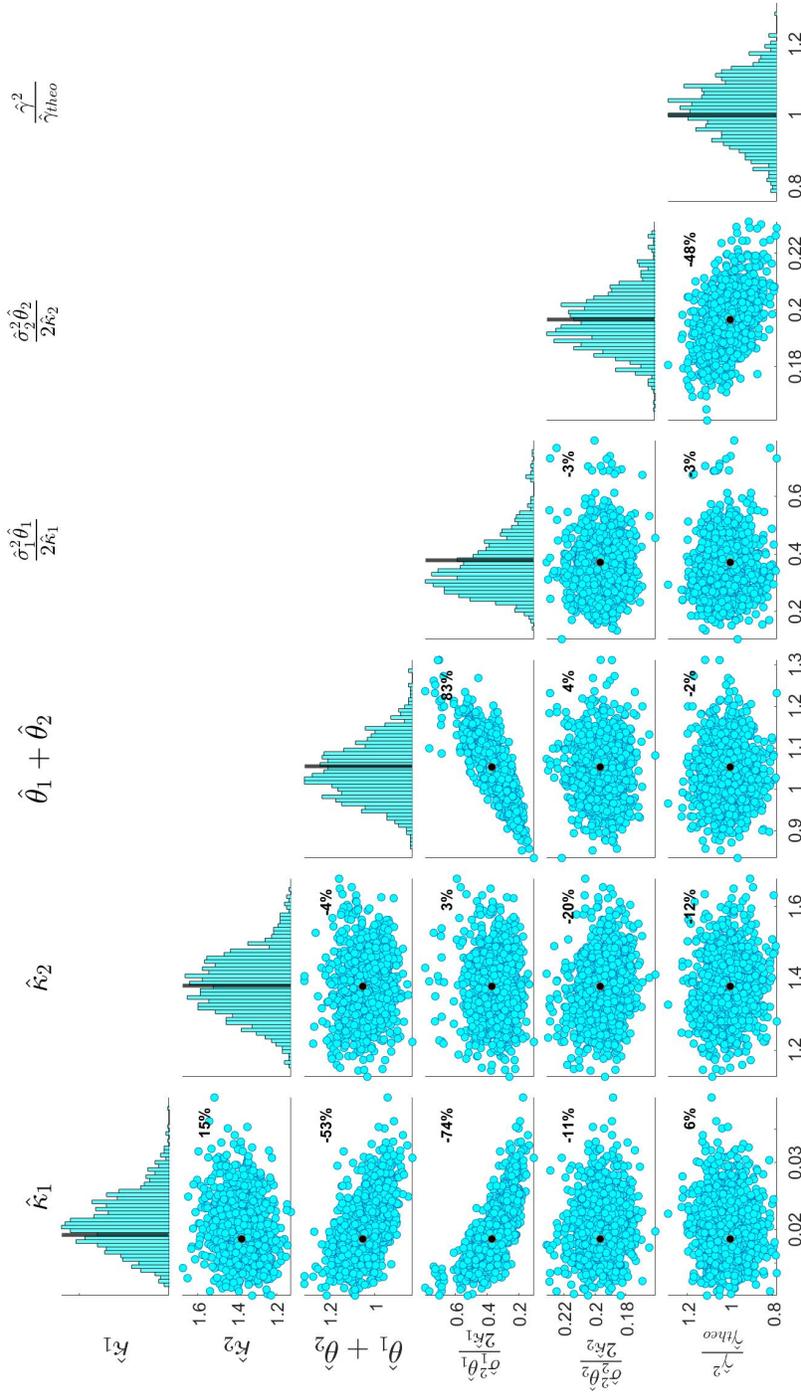
In summary, estimating a two-factor model with GMM remains challenging, as perfor-

²⁴Footnote 5 of Bregantini, 2013 also discuss issues with the two variances procedure of BZ.

²⁵The parameters are daily and in percentage.

mance is highly dependent on the nature of the variances. While the results were satisfactory for a single variance, the two-variance model does not guarantee success, particularly when using CM methods. Additionally, expanding these methods to a multi-asset context could quickly become computationally burdensome, especially with UM methods, due to the finite number of moments that can be calculated without using instrumental variables. Indeed, computing higher order moments could pose a challenge for the practicability of the method.

The inclusion of a filter for instantaneous variances, where the recursion of the CM methods could be simplified (reducing noise), could enhance the potential of CM methods for multi-factor extensions. As detailed in Panel A of Table 4.6, the second moment of $RV_{t,t+1}^2$ requires five lag factors to complete the recursion. With an appropriate filter, we could compute directly the expected value of $RV_{t,t+1}^2$ conditional on the observed filtration, thereby reducing the need to reference past data and minimizing noise in the moments.



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario F $\{\kappa = 4, \theta = 0.03, \sigma = 0.30\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Each diagonal panel presents the histogram of the point estimate parameters based on the GMM optimization described in Appendix C.6. The other panels show scatter plots between two parameters. The black line or dark dot represents the theoretical value of the parameters. The number in the right corner of each panel is the correlation between the parameter estimates.

Figure 4.6: Results for specification (4.21) using UM moment selection and parameter of scenario G

4.6 Conclusion

In conclusion, this paper presents a detailed comparison of two commonly used GMM procedures: conditional moments (CM) and unconditional moments (UM), in the context of one and two factor stochastic volatility models. Our analysis highlights the strengths and limitations of each approach, providing valuable insights into their suitability under different model specifications and parameter scenarios.

The CM-based approach stands out for its use of instrumental variables, which minimizes the derivation of additional moment conditions. This feature allows the CM method to extract more information from the same moment condition, particularly when the measurement error is minimal, resulting in highly precise parameter estimates. Furthermore, the CM method is relatively easy to implement as less moment are necessary, making it a practical choice in many applications. However, the absence of a filtering mechanism in the CM approach means it must rely on recursive techniques, which can introduce additional noise. This limitation becomes particularly challenging in models with two variance factors, and the complexity increases rapidly in multi-asset contexts.

On the other hand, the UM approach demonstrates greater robustness, especially in two-factor models. It also performs well even in the presence of high measurement error and high persistence in the variance. However, the UM method requires the derivation of more moments and lacks the benefit of instrumental variables, which can make the process more cumbersome as models become more complex.

References

- Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(2), 253–280.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004a). A Feasible Central Limit Theory for Realised Volatility Under Leverage.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004b). Econometric Analysis of Realized Covariation: High Frequency based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, 72(3), 885–925.
- Baum, C. F., & Zerilli, P. (2016). Jumps and Stochastic Volatility in Crude Oil Futures Prices using Conditional Moments of Integrated Volatility. *Energy Economics*, 53, 175–181.
- Baum, C. F., Zerilli, P., & Chen, L. (2021). Stochastic Volatility, Jumps and Leverage in Energy and Stock Markets: Evidence from High Frequency Data. *Energy Economics*, 93, 104481.
- Bolko, A. E., Christensen, K., Pakkanen, M. S., & Veliyev, B. (2023). A GMM Approach to Estimate the Roughness of Stochastic Volatility. *Journal of Econometrics*, 235(2), 745–778.
- Bollerslev, T., Gibson, M., & Zhou, H. (2011). Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. *Journal of Econometrics*, 160(1), 235–245.

- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting. *Journal of Econometrics*, 192(1), 1–18.
- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, 109(1), 33–65.
- Bregantini, D. (2013). Moment-based Estimation of Stochastic Volatility. *Journal of Banking and Finance*, 37(12), 4755–4764.
- Brix, A. F., & Lunde, A. (2015). Prediction-based Estimating Functions for Stochastic Volatility Models with Noisy Data: Comparison with a GMM Alternative. *ASTA Advances in Statistical Analysis*, 99, 433–465.
- Chacko, G., & Viceira, L. M. (2003). Spectral GMM Estimation of Continuous-Time Processes. *Journal of Econometrics*, 116(1-2), 259–292.
- Chaussé, P., & Xu, D. (2018). GMM Estimation of a Realized Stochastic Volatility Model: A Monte Carlo Study. *Econometric Reviews*, 37(7), 719–743.
- Christoffersen, P., Heston, S., & Jacobs, K. (2009). The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. *Management Science*, 55(12), 1914–1932.
- Corradi, V., & Distaso, W. (2006). Semi-Parametric Comparison of Stochastic Volatility Models using Realized Measures. *Review of Economic Studies*, 73(3), 635–667.
- Escobar, M. (2018). A Stochastic Volatility Factor Model of Heston Type. Statistical Properties and Estimation. *Stochastics*, 90(2), 172–199.
- Ewald, C., & Zou, Y. (2021). Stochastic Volatility: A Tale of Co-Jumps, Non-Normality, GMM and High Frequency Data. *Journal of Empirical Finance*, 64, 37–52.
- Feunou, B., & Tédongap, R. (2012). A Stochastic Volatility Model with Conditional Skewness. *Journal of Business and Economic Statistics*, 30(4), 576–591.

- Garcia, R., Lewis, M.-A., Pastorello, S., & Renault, É. (2011). Estimation of Objective and Risk-Neutral Distributions based on Moments of Integrated Volatility. *Journal of Econometrics*, *160*(1), 22–32.
- Ghosh, A., & Linton, O. (2023). Estimation with Mixed Data Frequencies: A Bias-Correction Approach. *Journal of Empirical Finance*, *74*, 101412.
- Heston, S. L. (1993). A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, *6*(2), 327–343.
- Ishida, I., McAleer, M., & Oya, K. (2011). Estimating the Leverage Parameter of Continuous-Time Stochastic Volatility Models Using High Frequency S&P 500 and VIX. *Managerial Finance*, *37*(11), 1048–1067.
- Jiang, G. J., & Knight, J. L. (2002). Estimation of Continuous-Time Processes via the Empirical Characteristic Function. *Journal of Business and Economic Statistics*, *20*(2), 198–212.
- Jiang, G. J., & Oomen, R. C. (2007). Estimating Latent Variables and Jump Diffusion Models using High-Frequency Data. *Journal of Financial Econometrics*, *5*(1), 1–30.
- Li, J., & Xiu, D. (2016). Generalized Method of Integrated Moments for High-Frequency Data. *Econometrica*, *84*(4), 1613–1633.
- Meddahi, N. (2002). A Theoretical Comparison between Integrated and Realized Volatility. *Journal of Applied Econometrics*, *17*(5), 479–508.
- Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*(3), 703–708.
- Takahashi, M., Omori, Y., & Watanabe, T. (2009). Estimating Stochastic Volatility Models using Daily Returns and Realized Volatility Simultaneously. *Computational Statistics and Data Analysis*, *53*(6), 2404–2426.

- Todorov, V. (2009). Estimation of Continuous-Time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics*, 148(2), 131–148.
- Todorov, V. (2010). Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, 23(1), 345–383.
- Todorov, V. (2011). Econometric Analysis of Jump-Driven Stochastic Volatility Models. *Journal of Econometrics*, 160(1), 12–21.
- Todorov, V., Tauchen, G., & Gryniv, I. (2011). Realized Laplace Transforms for Estimation of Jump Fiffusive Volatility Models. *Journal of Econometrics*, 164(2), 367–381.
- Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100(472), 1394–1411.
- Zhang, L., & Wang, L. (2023). Generalized Method of Moments Estimation of Realized Stochastic Volatility Model. *Journal of Risk and Financial Management*, 16(8), 377.

Chapter 5

Concluding Remarks

This thesis presents a filtering and estimation methods for multivariate stochastic volatility models along with an essay on congestion in power markets.

The first essay concentrates on the relationship between futures contracts and Financial Transmission Rights in electricity markets, focusing on their respective payoffs. While FTRs only cover the congestion component of electricity prices, futures contracts encompass the entire price, including energy and loss components. By comparing these instruments and accounting for pricing differentials across zones, we compute implied prices for the loss component. Our empirical study found significant pricing inconsistencies, with the variance of implied prices exceeding that of realized values in nearly all cases, indicating a disconnect in the pricing system between the futures and FTR markets.

The second essay presents a precise analytical filter for a multi-factor variance model within a multi-asset framework. Using intraday statistics, such as realized variances and covariances, our filter for instantaneous variances effectively manages dimensionality and addresses nonlinear, non-Gaussian data. We also tackle measurement errors in realized variance and incorporate corrections to ensure an unbiased estimation of the latent state's filtered moments. Extensive Monte Carlo simulations confirm the robustness and accuracy of our method, which provides detailed insights for the first four conditional cumulant and co-moments of the latent state.

The final essay provides a thorough analysis of two GMM procedures: conditional moments (CM) and unconditional moments (UM) in the context of stochastic volatility models and realized variance. The CM method stands out by leveraging instrumental variables, which enhances precision and simplifies implementation, making it advantageous for scenarios with minimal measurement error. However, it can suffer from additional noise due to its reliance on recursive techniques, especially in complex multi-factor models. Thus, having a filter for the latent state could help extend the methodology by combining the contributions of the last two essays of this thesis. In contrast, the UM approach is notably robust, performing well even with high measurement error and persistence, and is particularly effective in two variance factor models. Despite its advantages, UM methods require more moment derivations compared to CM methods, which can complicate their application in more complex models. This comparison underscores the strengths and trade-offs of each approach, providing valuable insights for selecting the appropriate method based on model complexity and data characteristics.

Bibliography

- Adamson, S., & Englander, S. L. (2005). Efficiency of New York Transmission Congestion Contract auctions. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 59a–59a.
- Adamson, S., Noe, T., & Parker, G. (2010). Efficiency of Financial Transmission Rights markets in centrally coordinated periodic auctions. *Energy Economics*, 32(4), 771–778.
- Adamson, S., & Parker, G. (2013). Participation and efficiency in the New York Financial Transmission Rights markets. *Financial Transmission Rights: Analysis, Experiences and Prospects*, 289–303.
- Aït-Sahalia, Y., & Xiu, D. (2016). Increased Correlation among Asset Classes: Are Volatility or Jumps to Blame, or Both? *Journal of Econometrics*, 194(2), 205–219.
- Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance*, 57(3), 1047–1091.
- Alsac, O., Bright, J., Brignone, S., Prais, M., Silva, C., Stott, B., & Vempati, N. (2004). The rights to fight price volatility. *IEEE Power and Energy Magazine*, 2(4), 47–57.
- Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2013). Chapter 17 - Financial Risk Measurement for Financial Risk Management. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.). Elsevier.

- Andersen, T. G., & Sørensen, B. E. (1996). GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *Journal of Business and Economic Statistics*, 14(3), 328–352.
- Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3), 817–858.
- Baltaduonis, R., Bonar, S., Carnes, J., & Mastrangelo, E. (2017). Risk and abnormal returns in markets for Congestion Revenue Rights. *Journal of Energy Markets*, 10(3).
- Baltagi, B. H. (2021). *Econometric Analysis of Panel Data (6th ed)*. Springer.
- Barndorff-Nielsen, O. E., & Shephard, N. (2002a). Estimating Quadratic Variation Using Realized Variance. *Journal of Applied Econometrics*, 17(5), 457–477.
- Barndorff-Nielsen, O. E., & Shephard, N. (2002b). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(2), 253–280.
- Barndorff-Nielsen, O. E., & Shephard, N. (2003). Realized Power Variation and Stochastic Volatility Models. *Bernoulli*, 9(2), 243–265.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004a). A Feasible Central Limit Theory for Realised Volatility Under Leverage.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004b). Econometric Analysis of Realized Covariation: High Frequency based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, 72(3), 885–925.
- Bartholomew, E. S., Siddiqui, A. S., Marnay, C., & Oren, S. S. (2003). The New York Transmission Congestion Contract market: Is it truly working efficiently? *The Electricity Journal*, 16(9), 14–24.
- Bates, D. S. (2006). Maximum Likelihood Estimation of Latent Affine Processes. *Review of Financial Studies*, 19(3), 909–965.

- Bates, D. S. (2012). US Stock Market Crash Risk, 1926–2010. *Journal of Financial Economics*, 105(2), 229–259.
- Bates, D. S. (2019). How Crashes Develop: Intradaily Volatility and Crash Evolution. *Journal of Finance*, 74(1), 193–238.
- Baum, C. F., & Zerilli, P. (2016). Jumps and Stochastic Volatility in Crude Oil Futures Prices using Conditional Moments of Integrated Volatility. *Energy Economics*, 53, 175–181.
- Baum, C. F., Zerilli, P., & Chen, L. (2021). Stochastic Volatility, Jumps and Leverage in Energy and Stock Markets: Evidence from High Frequency Data. *Energy Economics*, 93, 104481.
- Bégin, J.-F., Amaya, D., Gauthier, G., & Malette, M.-È. (2020). On the Estimation of Jump-Diffusion Models Using Intraday Data: A Filtering-Based Approach. *SIAM Journal on Financial Mathematics*, 11(4), 1168–1208.
- Bolko, A. E., Christensen, K., Pakkanen, M. S., & Veliyev, B. (2023). A GMM Approach to Estimate the Roughness of Stochastic Volatility. *Journal of Econometrics*, 235(2), 745–778.
- Bollerslev, T., Gibson, M., & Zhou, H. (2011). Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. *Journal of Econometrics*, 160(1), 235–245.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting. *Journal of Econometrics*, 192(1), 1–18.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2018). Modeling and Forecasting (Un)reliable Realized Covariances for more Reliable Financial Decisions. *Journal of Econometrics*, 207(1), 71–91.

- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, 109(1), 33–65.
- Bregantini, D. (2013). Moment-based Estimation of Stochastic Volatility. *Journal of Banking and Finance*, 37(12), 4755–4764.
- Brignone, R., Gonzato, L., & Lütkebohmert, E. (2023). Efficient Quasi-Bayesian Estimation of Affine Option Pricing Models using Risk-Neutral Cumulants. *Journal of Banking and Finance*, 148, 106745.
- Brillinger, D. R. (1969). The Calculation of Cumulants via Conditioning. *Annals of the Institute of Statistical Mathematics*, 21(1), 215–218.
- Brix, A. F., & Lunde, A. (2015). Prediction-based Estimating Functions for Stochastic Volatility Models with Noisy Data: Comparison with a GMM Alternative. *AStA Advances in Statistical Analysis*, 99, 433–465.
- Buccheri, G., & Corsi, F. (2021). HARK the SHARK: Realized Volatility Modeling with Measurement Errors and Nonlinear Dependencies. *Journal of Financial Econometrics*, 19(4), 614–649.
- Calzolari, G., Halbleib, R., & Zagidullina, A. (2021). A Latent Factor Model for Forecasting Realized Variances. *Journal of Financial Econometrics*, 19(5), 860–909.
- Celebi, M., Hajos, A., & Hanser, P. Q. (2010). Virtual bidding: The good, the bad and the ugly. *The Electricity Journal*, 23(5), 16–25.
- Chacko, G., & Viceira, L. M. (2003). Spectral GMM Estimation of Continuous-Time Processes. *Journal of Econometrics*, 116(1-2), 259–292.
- Chaussé, P., & Xu, D. (2018). GMM Estimation of a Realized Stochastic Volatility Model: A Monte Carlo Study. *Econometric Reviews*, 37(7), 719–743.
- Chen, H., Cheng, Y., Liu, Y., & Tang, K. (2023). Teaching Economics to the Machines. Available at SSRN 4642167

- Chen, R.-R., & Scott, L. (2003). Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter model. *Journal of Real Estate Finance and Economics*, 27, 143–172.
- Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear Kalman Filtering in Affine Term Structure Models. *Management Science*, 60(9), 2248–2268.
- Christoffersen, P., Feunou, B., Jacobs, K., & Meddahi, N. (2014). The Economic Value of Realized Volatility: Using High-Frequency Returns for Option Valuation. *Journal of Financial and Quantitative Analysis*, 49(3), 663–697.
- Christoffersen, P., Feunou, B., & Jeon, Y. (2015). Option Valuation with Observable Volatility and Jump Dynamics. *Journal of Banking and Finance*, 61, S101–S120.
- Christoffersen, P., Heston, S., & Jacobs, K. (2009). The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. *Management Science*, 55(12), 1914–1932.
- Christoffersen, P., Lunde, A., & Olesen, K. V. (2019). Factor Structure in Commodity Futures Return and Volatility. *Journal of Financial and Quantitative Analysis*, 54(3), 1083–1115.
- Corradi, V., & Distaso, W. (2006). Semi-Parametric Comparison of Stochastic Volatility Models using Realized Measures. *Review of Economic Studies*, 73(3), 635–667.
- Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2), 174–196.
- Cortazar, G., Lopez, M., & Naranjo, L. (2017). A Multifactor Stochastic Volatility Model of Commodity Prices. *Energy Economics*, 67, 182–201.
- Deng, S.-J., Oren, S., & Meliopoulos, A. (2010). The inherent inefficiency of simultaneously feasible Financial Transmission Rights auctions. *Energy Economics*, 32(4), 779–785.

- Deng, S.-J., & Oren, S. S. (2006). Electricity derivatives and risk management. *Energy*, 31(6-7), 940–953.
- Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4), 549–560.
- Duan, J.-C., & Simonato, J.-G. (1999). Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting*, 13, 111–135.
- Dubecq, S., Monfort, A., Renne, J.-P., & Roussellet, G. (2016). Credit and Liquidity in Interbank Rates: A Quadratic Approach. *Journal of Banking and Finance*, 68, 29–46.
- Duffee, G. R., & Stanton, R. H. (2012). Estimation of Dynamic Term Structure Models. *The Quarterly Journal of Finance*, 2(02), 1250008.
- Durham, G., Geweke, J., & Ghosh, P. (2015). A Comment on Christoffersen, Jacobs, and Ornathanalai (2012), “Dynamic Jump Intensities and Risk Premiums: Evidence from S&P 500 Returns and Options”. *Journal of Financial Economics*, 115(1), 210–214.
- Escobar, M. (2018). A Stochastic Volatility Factor Model of Heston Type. Statistical Properties and Estimation. *Stochastics*, 90(2), 172–199.
- Ewald, C., & Zou, Y. (2021). Stochastic Volatility: A Tale of Co-Jumps, Non-Normality, GMM and High Frequency Data. *Journal of Empirical Finance*, 64, 37–52.
- Fasano, A., Germani, A., & Monteriu, A. (2013). Reduced-Order Quadratic Kalman-like Filtering of Non-Gaussian Systems. *IEEE Transactions on Automatic Control*, 58(7), 1744–1757.
- Feunou, B., Aliouchkin, R. L., Tédongap, R., & Xu, L. (2020). The Term Structures of Expected Loss and Gain Uncertainty. *Journal of Financial Econometrics*, 18(3), 473–501.

- Feunou, B., & Okou, C. (2018). Risk-Neutral Moment-based Estimation of Affine Option Pricing Models. *Journal of Applied Econometrics*, 33(7), 1007–1025.
- Feunou, B., & Tédongap, R. (2012). A Stochastic Volatility Model with Conditional Skewness. *Journal of Business and Economic Statistics*, 30(4), 576–591.
- Fiebig, D. G. (2003). *A Companion to Theoretical Econometrics*. John Wiley & Sons, Ltd.
- Garcia, R., Lewis, M.-A., Pastorello, S., & Renault, É. (2011). Estimation of Objective and Risk-Neutral Distributions based on Moments of Integrated Volatility. *Journal of Econometrics*, 160(1), 22–32.
- Ghosh, A., & Linton, O. (2023). Estimation with Mixed Data Frequencies: A Bias-Correction Approach. *Journal of Empirical Finance*, 74, 101412.
- Godin, F., & Ibrahim, Z. (2021). An analysis of electricity congestion price patterns in North America. *Energy Economics*, 102, 105506.
- Greene, W. (2017). *Econometric Analysis (8th ed)*. Pearson.
- Hadsell, L., & Shawky, H. A. (2009). Efficiency and profit in the NYISO Transmission Congestion Contract market. *The Electricity Journal*, 22(9), 47–57.
- Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.
- Hansen, P. R., & Lunde, A. (2014). Estimating the Persistence and the Autocorrelation Function of a Time Series that is Measured with Error. *Econometric Theory*, 30(1), 60–93.
- Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2016). The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications. *Journal of Financial Economics*, 119(2), 249–283.
- Heston, S. L. (1993). A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2), 327–343.

- Hogan, W. (2002). Financial Transmission Right formulations [Mimeo, JFK School of Government, Harvard Electricity Policy Group, Harvard University].
- Hogan, W. W. (1992). Contract networks for electric power transmission. *Journal of Regulatory Economics*, 4(3), 211–242.
- Hogan, W., et al. (2013). Financial Transmission Rights, revenue adequacy and multi-settlement electricity markets. Available Harvard University web site: <http://www.hks.harvard.edu/fs/whogan/HoganFTRRevAdequacy31813>.
- Ishida, I., McAleer, M., & Oya, K. (2011). Estimating the Leverage Parameter of Continuous-Time Stochastic Volatility Models Using High Frequency S&P 500 and VIX. *Managerial Finance*, 37(11), 1048–1067.
- Javaheri, A., Lautier, D., & Galli, A. (2003). Filtering in Finance. *Wilmott*, 3, 67–83.
- Jiang, G. J., & Knight, J. L. (2002). Estimation of Continuous-Time Processes via the Empirical Characteristic Function. *Journal of Business and Economic Statistics*, 20(2), 198–212.
- Jiang, G. J., & Oomen, R. C. (2007). Estimating Latent Variables and Jump Diffusion Models using High-Frequency Data. *Journal of Financial Econometrics*, 5(1), 1–30.
- Joanes, D. N., & Gill, C. A. (1998). Comparing Measures of Sample Skewness and Kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 183–189.
- Jong, F. d. (2000). Time Series and Cross-Section Information in Affine Term-Structure Models. *Journal of Business and Economic Statistics*, 18(3), 300–314.
- Julier, S., Uhlmann, J., & Durrant-Whyte, H. F. (2000). A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. *IEEE Transactions on Automatic Control*, 45(3), 477–482.
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems.

- Kristiansen, T. (2005). Markets for Financial Transmission Rights. *Energy Studies Review*, 13(1).
- Kumar, A., Srivastava, S., & Singh, S. (2005). Congestion management in competitive power market: A bibliographical survey. *Electric Power Systems Research*, 76(1), 153–164.
- Ledgerwood, S. D., & Pfeifenberger, J. P. (2013). Using virtual bids to manipulate the value of Financial Transmission Rights. *The Electricity Journal*, 26(9), 9–25.
- Leslie, G. W. (2021). Who benefits from ratepayer-funded auctions of Transmission Congestion Contracts? Evidence from New York. *Energy Economics*, 93, 105025.
- Li, H., & Zhao, F. (2006). Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives. *Journal of Finance*, 61(1), 341–378.
- Li, J., & Xiu, D. (2016). Generalized Method of Integrated Moments for High-Frequency Data. *Econometrica*, 84(4), 1613–1633.
- Longstaff, F. A., & Wang, A. W. (2004). Electricity forward prices: A high-frequency empirical analysis. *The journal of finance*, 59(4), 1877–1900.
- Lyons, K., Fraser, H., & Parmesano, H. (2000). An Introduction to Financial Transmission Rights. *The Electricity Journal*, 13(10), 31–37.
- Ma, X., Sun, D. I., Rosenwald, G. W., & Ott, A. L. (2003). Advanced Financial Transmission Rights in the PJM market. *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No. 03CH37491)*, 2, 1031–1038.
- Maddala, G. S. (1971). The Use of Variance Components Models in Pooling Cross Section and Time Series Data. *Econometrica: Journal of the Econometric Society*, 341–358.
- Meddahi, N. (2002). A Theoretical Comparison between Integrated and Realized Volatility. *Journal of Applied Econometrics*, 17(5), 479–508.

- Meddahi, N. (2003). ARMA Representation of Integrated and Realized Variances. *Econometrics Journal*, 6(2), 335–356.
- Monfort, A., Pegoraro, F., Renne, J.-P., & Roussellet, G. (2017). Staying at Zero with Affine Processes: An Application to Term Structure Modelling. *Journal of Econometrics*, 201(2), 348–366.
- Monfort, A., Renne, J.-P., & Roussellet, G. (2015). A Quadratic Kalman Filter. *Journal of Econometrics*, 187(1), 43–56.
- Mount, T. D., & Ju, J. (2014). An econometric framework for evaluating the efficiency of a market for Transmission Congestion Contracts. *Energy Economics*, 46, 176–185.
- Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708.
- NYISO. (2021). Transmission Congestion Contracts Manual.
- Olmstead, D. E. (2018). Ontario’s auction market for Financial Transmission Rights: An analysis of its efficiency. *The Energy Journal*, 39(1).
- Opgrand, J., Preckel, P. V., Gotham, D. J., & Liu, A. L. (2022). Price formation in auctions for Financial Transmission Rights. *The Energy Journal*, 43(3).
- Prete, C. L., Guo, N., & Shanbhag, U. V. (2018). Virtual bidding and Financial Transmission Rights: An equilibrium model for cross-product manipulation in electricity markets. *IEEE Transactions on Power Systems*, 34(2), 953–967.
- Roussellet, G. (2023). The Term Structure of Macroeconomic Risks at the Effective Lower Bound. *Journal of Econometrics*, 105383.
- Rudkevich, A., Hausman, E., Tabors, R., Bagnall, J., & Kopel, C. (2005). Loss hedging rights: A final piece in the LMP puzzle. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 60c–60c.

- Sarkar, V., & Khaparde, S. A. (2008). A comprehensive assessment of the evolution of Financial Transmission Rights. *IEEE Transactions on Power Systems*, 23(4), 1783–1795.
- Sarkar, V., & Khaparde, S. A. (2009). Introduction to loss-hedging Financial Transmission Rights. *IEEE Transactions on Power Systems*, 24(2), 621–630.
- Siddiqui, A. S., Bartholomew, E. S., Marnay, C., & Oren, S. S. (2005). Efficiency of the New York independent system operator market for Transmission Congestion Contracts. *Managerial Finance*, 31(6), 1–45.
- Singla, A., Singh, K., & Yadav, V. K. (2014). Transmission congestion management in deregulated environment: A bibliographical survey. *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, 1–10.
- Takahashi, M., Omori, Y., & Watanabe, T. (2009). Estimating Stochastic Volatility Models using Daily Returns and Realized Volatility Simultaneously. *Computational Statistics and Data Analysis*, 53(6), 2404–2426.
- Todorov, V. (2009). Estimation of Continuous-Time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics*, 148(2), 131–148.
- Todorov, V. (2010). Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies*, 23(1), 345–383.
- Todorov, V. (2011). Econometric Analysis of Jump-Driven Stochastic Volatility Models. *Journal of Econometrics*, 160(1), 12–21.
- Todorov, V., Tauchen, G., & Gryniv, I. (2011). Realized Laplace Transforms for Estimation of Jump Diffusive Volatility Models. *Journal of Econometrics*, 164(2), 367–381.
- Trolle, A. B., & Schwartz, E. S. (2009). A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives. *Review of Financial Studies*, 22(5), 2007–2057.

- Uhlmann, J., & Julier, S. J. (2022). Gaussianity and the Kalman Filter: A Simple Yet Complicated Relationship. *Journal de Ciencia e Ingeniería*, 14(1), 21–26.
- Vogelsang, T. J. (2012). Heteroskedasticity, Autocorrelation, and Spatial Correlation Robust Inference in Linear Panel Models with Fixed-Effects. *Journal of Econometrics*, 166(2), 303–319.
- Wan, E., & Van Der Merwe, R. (2000). The Unscented Kalman Filter for Nonlinear Estimation. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 153–158.
- Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American statistical Association*, 57(298), 348–368.
- Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100(472), 1394–1411.
- Zhang, L., & Wang, L. (2023). Generalized Method of Moments Estimation of Realized Stochastic Volatility Model. *Journal of Risk and Financial Management*, 16(8), 377.
- Zhang, N. (2009). Market performance and bidders' bidding behavior in the New York Transmission Congestion Contract market. *Energy Economics*, 31(1), 61–68.

Chapter A

Appendices of Pricing Inconsistency between the Futures and Financial Transmission Right Markets in North America

A.1 Supplementary Material

Results from the main paper are obtained by examining data from nearest auctions, which are the closest auctions to the corresponding month \mathcal{H} . However, in all three markets, FTRs can be traded up to several months before. This appendix examines whether similar conclusions can be reached when studying FTRs prices from earlier auctions.

In this appendix, data from all earlier auctions are grouped regardless of the time between the auction and the contract month. Thus, results presented for earlier auctions incorporate all earlier auctions simultaneously, i.e. excluding the nearest auction.

A.1.1 Price consistency tests using nodal FTR prices

From a qualitative perspective, the results for earlier auctions are similar to those presented in Section 2.3.2 for nearest auctions. When examining the variances in Tables A.1 to A.6, the conclusions are identical to those made for nearest auctions, namely that Inequality (2.14) does not hold generally, confirming the presence of pricing misalignment between futures and FTR contracts. For almost all zones in the NYISO and in PJM, all seasonal

buckets and all contract types, the inequality does not hold. Exception are the COMED zone for off-peak and 24-hour contracts during the summer months, or the DEOK off-peak contracts for winter months. For the ISO-NE, the inequality does not hold in the majority of the zones.

A.1.2 Price consistency tests using the most liquid FTR contracts

To alleviate potential illiquidity effects in analyses, this section considers the FTRs that were the most traded instead of looking at FTR nodal prices. The results pertaining to price and realized loss variance in earlier auction results are provided in Tables A.7 to A.10, which can be compared with the analogous Tables 2.7 to 2.10 for nearest auctions. Again, Inequality (2.14) is systematically violated for PJM and the NYISO, reinforcing the main findings of this paper. Results for ISO-NE are not presented since there is insufficient liquidity.

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
A	50	3.55	3.51	0.04	1.01	A	51	1.17	0.35	0.81	3.32
B	50	5.16	1.98	3.18	2.61	B	51	3.47	0.21	3.26	16.53
C	50	5.04	0.71	4.34	7.13	C	51	1.75	0.11	1.64	15.90
D	50	9.36	0.59	8.76	15.75	D	51	8.10	0.87	7.23	9.35
E	50	2.92	0.09	2.83	31.77	E	51	3.73	0.05	3.68	79.65
F	50	4.18	0.93	3.24	4.47	F	51	2.76	0.03	2.73	85.60
G	50	2.68	0.52	2.16	5.16	G	51	1.40	0.14	1.25	9.65
I	50	5.26	0.72	4.54	7.27	I	51	4.15	0.24	3.91	17.63
J	50	4.07	0.84	3.23	4.87	J	51	1.72	0.35	1.37	4.97
K	50	22.87	1.38	21.49	16.55	K	51	13.30	0.53	12.78	25.32

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
A	50	1.48	1.13	0.35	1.31	A	151	2.10	1.83	0.26	1.14
B	50	6.71	0.74	5.96	9.04	B	151	5.98	1.09	4.89	5.48
C	50	1.20	0.51	0.70	2.38	C	151	2.67	0.49	2.18	5.45
D	50	11.10	3.50	7.60	3.17	D	151	9.40	1.85	7.54	5.07
E	50	5.43	0.11	5.32	47.63	E	151	6.10	0.10	6.01	63.48
F	50	5.25	0.24	5.00	21.54	F	151	4.48	0.46	4.03	9.80
G	50	4.57	0.56	4.02	8.23	G	151	2.99	0.43	2.56	6.90
I	50	5.83	1.09	4.74	5.37	I	151	5.04	0.74	4.30	6.79
J	50	5.30	1.36	3.94	3.91	J	151	3.92	0.92	2.99	4.24
K	50	57.81	1.81	56.00	31.87	K	151	32.58	1.35	31.23	24.06

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.1: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - NYISO 24-hour contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	106	3.62	0.89	2.73	4.06	AECO	113	14.35	0.08	14.27	179.01
BGE	106	3.80	0.28	3.51	13.39	BGE	113	2.65	0.47	2.18	5.65
DPL	106	3.11	1.77	1.34	1.76	DPL	113	12.81	0.08	12.73	160.35
JCPL	106	1.70	1.10	0.59	1.54	JCPL	113	2.28	0.07	2.21	34.92
METED	106	1.80	0.49	1.30	3.64	METED	113	1.51	0.04	1.46	35.96
PECO	106	0.97	0.64	0.33	1.52	PECO	113	2.05	0.15	1.89	13.36
PPL	106	1.22	0.34	0.89	3.62	PPL	113	1.35	0.14	1.21	9.80
PENELEC	106	1.23	0.08	1.15	14.54	PENELEC	113	0.70	0.07	0.64	10.68
PSEG	106	2.91	1.21	1.70	2.41	PSEG	113	1.85	0.10	1.76	19.03
APS	106	1.38	0.12	1.26	11.39	APS	113	1.47	0.05	1.42	28.79
AEP	100	1.56	1.03	0.53	1.51	AEP	112	0.59	0.07	0.52	8.93
COMED	106	3.57	2.58	0.99	1.38	COMED	113	1.83	0.67	1.15	2.71
DAY	84	4.65	1.35	3.29	3.43	DAY	93	1.57	0.36	1.22	4.42
DOM	84	1.62	0.06	1.56	25.76	DOM	96	3.30	0.50	2.81	6.67
DUQ	106	3.75	1.22	2.54	3.08	DUQ	113	2.15	0.16	1.99	13.15
ATSI	106	4.90	0.71	4.19	6.86	ATSI	113	2.25	0.08	2.17	27.87
DEOK	106	4.12	2.11	2.01	1.96	DEOK	113	1.75	0.20	1.55	8.58

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	43	24.85	0.30	24.55	83.63	AECO	262	11.66	0.61	11.05	19.13
BGE	43	3.80	0.93	2.86	4.06	BGE	262	3.44	0.47	2.97	7.35
DPL	43	40.87	0.53	40.34	76.67	DPL	262	13.99	1.13	12.86	12.39
JCPL	43	4.54	0.13	4.41	33.87	JCPL	262	2.87	0.72	2.14	3.96
METED	43	5.18	0.09	5.09	56.68	METED	262	2.85	0.29	2.55	9.67
PECO	43	8.15	0.11	8.04	72.62	PECO	262	2.84	0.48	2.35	5.87
PPL	43	3.79	0.11	3.68	35.90	PPL	262	1.95	0.28	1.67	6.94
PENELEC	43	1.12	0.05	1.07	24.27	PENELEC	262	0.99	0.07	0.92	13.56
PSEG	43	10.30	0.12	10.18	84.44	PSEG	262	4.27	0.81	3.46	5.25
APS	43	1.54	0.10	1.44	15.35	APS	262	1.63	0.12	1.51	13.06
AEP	43	1.62	0.34	1.27	4.72	AEP	255	1.71	0.65	1.06	2.62
COMED	43	1.56	0.95	0.61	1.64	COMED	262	3.07	1.61	1.45	1.90
DAY	31	2.97	0.35	2.62	8.44	DAY	208	4.59	1.05	3.55	4.39
DOM	33	9.26	0.38	8.88	24.39	DOM	213	3.87	0.42	3.45	9.28
DUQ	43	2.56	0.79	1.77	3.23	DUQ	262	4.08	0.85	3.22	4.78
ATSI	43	2.62	0.24	2.38	10.88	ATSI	262	4.66	0.45	4.20	10.25
DEOK	43	2.43	0.34	2.09	7.06	DEOK	262	4.26	1.27	2.99	3.35

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.2: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM on-peak contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	106	1.83	0.89	0.94	2.05	AECO	113	3.95	0.08	3.87	49.25
BGE	106	3.08	0.28	2.80	10.86	BGE	113	3.33	0.47	2.86	7.10
DPL	106	2.57	1.77	0.80	1.46	DPL	113	5.04	0.08	4.96	63.13
JCPL	106	1.26	1.10	0.15	1.14	JCPL	113	1.87	0.07	1.80	28.69
METED	106	2.89	0.49	2.40	5.86	METED	113	1.04	0.04	1.00	24.94
PECO	106	1.76	0.64	1.12	2.76	PECO	113	0.87	0.15	0.72	5.68
PPL	106	1.95	0.34	1.61	5.78	PPL	113	1.45	0.14	1.31	10.47
PENELEC	106	0.88	0.08	0.80	10.40	PENELEC	113	0.59	0.07	0.52	8.95
PSEG	106	2.76	1.21	1.55	2.28	PSEG	113	1.12	0.10	1.02	11.48
APS	106	0.48	0.12	0.36	3.96	APS	113	0.61	0.05	0.55	11.86
AEP	100	1.68	1.03	0.65	1.62	AEP	112	0.36	0.07	0.30	5.53
COMED	106	3.33	2.58	0.74	1.29	COMED	113	0.94	0.67	0.27	1.39
DAY	84	3.90	1.35	2.54	2.88	DAY	93	0.80	0.36	0.45	2.26
DOM	84	0.96	0.06	0.90	15.27	DOM	96	0.97	0.50	0.47	1.95
DUQ	106	2.44	1.22	1.23	2.01	DUQ	113	1.36	0.16	1.19	8.28
ATSI	106	3.88	0.71	3.16	5.42	ATSI	113	1.01	0.08	0.93	12.53
DEOK	106	2.09	2.11	-0.02	0.99	DEOK	113	1.19	0.20	0.99	5.83

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	43	2.58	0.30	2.29	8.69	AECO	262	2.97	0.61	2.36	4.87
BGE	43	4.85	0.93	3.92	5.19	BGE	262	3.49	0.47	3.02	7.45
DPL	43	9.86	0.53	9.32	18.49	DPL	262	5.03	1.13	3.90	4.46
JCPL	43	3.68	0.13	3.54	27.43	JCPL	262	3.06	0.72	2.34	4.23
METED	43	2.05	0.09	1.96	22.40	METED	262	2.16	0.29	1.87	7.34
PECO	43	3.97	0.11	3.86	35.36	PECO	262	2.48	0.48	2.00	5.13
PPL	43	1.82	0.11	1.72	17.27	PPL	262	2.20	0.28	1.92	7.83
PENELEC	43	0.30	0.05	0.25	6.45	PENELEC	262	0.66	0.07	0.59	9.04
PSEG	43	4.69	0.12	4.57	38.47	PSEG	262	3.88	0.81	3.07	4.77
APS	43	0.79	0.10	0.69	7.83	APS	262	0.64	0.12	0.51	5.12
AEP	43	0.68	0.34	0.33	1.97	AEP	255	1.56	0.65	0.91	2.39
COMED	43	0.84	0.95	-0.11	0.89	COMED	262	2.61	1.61	1.00	1.62
DAY	31	0.92	0.35	0.56	2.60	DAY	208	2.32	1.05	1.28	2.22
DOM	33	3.08	0.38	2.70	8.12	DOM	213	1.34	0.42	0.93	3.23
DUQ	43	1.42	0.79	0.63	1.79	DUQ	262	2.64	0.85	1.79	3.10
ATSI	43	1.70	0.24	1.46	7.07	ATSI	262	3.44	0.45	2.99	7.58
DEOK	43	0.80	0.34	0.46	2.33	DEOK	262	2.26	1.27	0.99	1.78

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.3: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM off-peak contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	106	2.16	0.89	1.27	2.42	AECO	113	7.16	0.08	7.08	89.28
BGE	106	2.95	0.28	2.66	10.38	BGE	113	2.23	0.47	1.76	4.75
DPL	106	2.33	1.77	0.57	1.32	DPL	113	7.73	0.08	7.65	96.71
JCPL	106	1.18	1.10	0.07	1.06	JCPL	113	1.13	0.07	1.06	17.26
METED	106	1.89	0.49	1.40	3.84	METED	113	0.94	0.04	0.90	22.39
PECO	106	1.03	0.64	0.39	1.61	PECO	113	0.85	0.15	0.69	5.52
PPL	106	1.19	0.34	0.85	3.52	PPL	113	1.16	0.14	1.02	8.39
PENELEC	106	0.96	0.08	0.87	11.25	PENELEC	113	0.55	0.07	0.48	8.34
PSEG	106	2.16	1.21	0.95	1.79	PSEG	113	0.96	0.10	0.86	9.82
APS	106	0.75	0.12	0.63	6.18	APS	113	0.85	0.05	0.80	16.62
AEP	100	1.46	1.03	0.42	1.41	AEP	112	0.38	0.07	0.31	5.76
COMED	106	3.02	2.58	0.43	1.17	COMED	113	0.83	0.67	0.16	1.24
DAY	84	3.56	1.35	2.20	2.63	DAY	93	0.82	0.36	0.46	2.30
DOM	84	1.01	0.06	0.95	16.02	DOM	96	1.42	0.50	0.93	2.88
DUQ	106	2.66	1.22	1.44	2.19	DUQ	113	1.51	0.16	1.35	9.23
ATSI	106	3.82	0.71	3.11	5.35	ATSI	113	1.33	0.08	1.25	16.53
DEOK	106	2.64	2.11	0.54	1.25	DEOK	113	1.29	0.20	1.08	6.31

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO	43	9.54	0.30	9.24	32.10	AECO	262	5.53	0.61	4.93	9.08
BGE	43	2.38	0.93	1.44	2.55	BGE	262	2.59	0.47	2.12	5.52
DPL	43	20.99	0.53	20.45	39.37	DPL	262	8.05	1.13	6.92	7.13
JCPL	43	3.12	0.13	2.99	23.28	JCPL	262	2.27	0.72	1.54	3.13
METED	43	2.84	0.09	2.75	31.13	METED	262	2.00	0.29	1.71	6.81
PECO	43	2.27	0.11	2.16	20.21	PECO	262	1.64	0.48	1.16	3.39
PPL	43	1.89	0.11	1.79	17.95	PPL	262	1.67	0.28	1.39	5.95
PENELEC	43	0.42	0.05	0.37	9.03	PENELEC	262	0.69	0.07	0.62	9.49
PSEG	43	5.71	0.12	5.59	46.83	PSEG	262	3.25	0.81	2.44	4.00
APS	43	0.88	0.10	0.78	8.72	APS	262	0.92	0.12	0.79	7.35
AEP	43	0.75	0.34	0.41	2.20	AEP	255	1.47	0.65	0.82	2.25
COMED	43	0.82	0.95	-0.13	0.87	COMED	262	2.37	1.61	0.76	1.47
DAY	31	1.34	0.35	0.99	3.80	DAY	208	2.59	1.05	1.54	2.48
DOM	33	4.69	0.38	4.31	12.34	DOM	213	1.78	0.42	1.36	4.27
DUQ	43	1.53	0.79	0.74	1.94	DUQ	262	2.97	0.85	2.12	3.49
ATSI	43	1.41	0.24	1.17	5.84	ATSI	262	3.56	0.45	3.11	7.84
DEOK	43	0.97	0.34	0.63	2.83	DEOK	262	2.84	1.27	1.57	2.24

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.4: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - PJM 24-hour contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	33	0.29	0.07	0.22	4.28	ME	79	0.16	0.19	-0.03	0.83
NH	33	0.10	0.02	0.08	5.86	NH	79	0.39	0.06	0.32	5.97
RI	33	0.07	0.04	0.02	1.57	RI	79	0.10	0.14	-0.05	0.68
SEMA	33	0.12	0.01	0.11	21.28	SEMA	79	0.13	0.03	0.10	4.82
WCMA	33	0.61	0.04	0.57	15.06	WCMA	79	0.27	0.03	0.24	10.06
NEMA	33	0.19	0.01	0.18	15.64	NEMA	79	0.25	0.02	0.22	10.66

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	78	0.08	0.41	-0.34	0.18	ME	190	0.17	0.26	-0.09	0.64
NH	78	0.04	0.04	0.01	1.20	NH	190	0.22	0.05	0.17	4.56
RI	78	0.07	0.03	0.05	2.91	RI	190	0.09	0.08	0.01	1.11
SEMA	78	0.18	0.08	0.11	2.44	SEMA	190	0.15	0.04	0.10	3.34
WCMA	78	0.22	0.02	0.20	13.17	WCMA	190	0.33	0.03	0.30	10.87
NEMA	78	0.05	0.04	0.01	1.26	NEMA	190	0.16	0.03	0.13	5.79

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.5: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - ISO-NE on-peak contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months						Panel b: Shoulder Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	33	0.28	0.07	0.21	4.08	ME	79	0.06	0.19	-0.14	0.29
NH	33	0.12	0.02	0.11	7.02	NH	79	0.05	0.06	-0.02	0.75
RI	33	0.07	0.04	0.03	1.62	RI	79	0.05	0.14	-0.09	0.34
SEMA	33	0.28	0.01	0.27	48.89	SEMA	79	0.22	0.03	0.19	8.06
WCMA	33	1.34	0.04	1.30	33.12	WCMA	79	0.44	0.03	0.42	16.59
NEMA	33	0.29	0.01	0.27	23.27	NEMA	79	0.14	0.02	0.12	6.11

Panel c: Summer Months						Panel d: All Months					
Zone	n	(A)	(B)	(A) - (B)	(A)/(B)	Zone	n	(A)	(B)	(A) - (B)	(A)/(B)
ME	78	0.07	0.41	-0.34	0.17	ME	190	0.11	0.26	-0.15	0.41
NH	78	0.08	0.04	0.04	2.21	NH	190	0.07	0.05	0.02	1.51
RI	78	0.11	0.03	0.09	4.32	RI	190	0.09	0.08	0.01	1.07
SEMA	78	0.29	0.08	0.21	3.84	SEMA	190	0.25	0.04	0.21	5.71
WCMA	78	1.09	0.02	1.07	65.67	WCMA	190	0.87	0.03	0.84	28.55
NEMA	78	0.29	0.04	0.25	8.03	NEMA	190	0.23	0.03	0.20	8.32

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied excess loss prices $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ are calculated for all zones $i \in \mathcal{Z}_{t_{\mathcal{H}}}$ via Equation (2.11). The second column exhibits the number of earlier auctions for which the data are available. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) - \overline{\Pi_{t_{\mathcal{H}}}(L_{\mathcal{H}})}$ is represented by (A). The sample variance of the realized excess loss values $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.6: Sample variance of the excess losses $L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}$ and of their earlier-auction-implied prices - ISO-NE off-peak contract

Appendix A. Pricing inconsistency between the futures and FTR markets

Panel a: Winter Months					Panel b: Shoulder Months						
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
A-C	40	1.19	0.96	0.23	1.24	A-C	42	1.18	0.10	1.08	11.41
F-G	29	1.83	0.20	1.63	9.25	I-G	33	1.73	0.02	1.72	102.72
G-J	28	0.50	0.03	0.47	14.43	J-G	30	0.19	0.04	0.15	4.43
G-A	25	11.83	2.77	9.06	4.27	G-F	26	0.64	0.12	0.51	5.25
G-F	25	1.18	0.05	1.13	24.18	F-G	24	1.54	0.09	1.45	16.35
J-G	24	0.46	0.06	0.41	8.33	G-C	23	7.21	0.42	6.80	17.29
E-C	24	2.02	1.65	0.38	1.23	C-A	21	1.15	0.12	1.03	9.89
I-G	23	3.03	0.01	3.02	375.61	A-G	21	1.18	0.79	0.39	1.49
G-I	23	1.90	0.02	1.88	87.97	G-A	20	0.83	0.53	0.29	1.55
D-C	22	0.83	0.16	0.67	5.12	G-I	19	2.00	0.04	1.96	55.38

Panel c: Summer Months					Panel d: All Months						
$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j-i$	n	(A)	(B)	(A) - (B)	(A)/(B)
A-C	44	1.31	0.17	1.14	7.67	A-C	126	1.38	0.47	0.92	2.96
I-G	36	0.90	0.13	0.77	6.86	I-G	92	1.89	0.07	1.81	25.64
G-C	33	7.51	1.49	6.02	5.03	G-F	79	0.95	0.08	0.87	11.82
G-F	28	0.81	0.06	0.75	13.68	J-G	76	0.37	0.11	0.26	3.36
G-J	24	0.22	0.04	0.18	5.64	F-G	68	1.45	0.16	1.30	9.17
J-G	22	0.48	0.23	0.25	2.05	G-C	67	7.05	1.04	6.01	6.77
A-G	21	2.30	0.72	1.58	3.20	G-J	67	0.33	0.05	0.28	6.82
C-G	21	14.87	3.73	11.14	3.99	C-A	61	1.67	0.60	1.07	2.78
C-A	19	1.78	0.18	1.60	10.14	G-A	58	8.45	2.07	6.38	4.09
D-C	17	7.95	0.22	7.74	36.71	C-G	56	12.82	2.96	9.86	4.33

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of earlier auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j,\mathcal{H}} - L_{i,\mathcal{H}})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.7: Sample variance of the loss spreads $L_{j,\mathcal{H}} - L_{i,\mathcal{H}}$ and of their earlier-auction-implied prices for the most liquid FTRs - NYISO 24-hour contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	54	1.69	0.12	1.57	14.64	PPL-PECO	51	0.95	0.03	0.92	37.60
PPL-PECO	48	1.10	0.14	0.96	8.03	PSEG-PECO	41	2.89	0.02	2.87	159.47
METED-PECO	43	3.46	0.04	3.43	95.23	METED-PPL	40	2.41	0.05	2.35	46.22
PSEG-JCPL	38	3.03	0.01	3.03	445.03	PSEG-JCPL	34	2.24	0.01	2.23	263.51
PSEG-PPL	28	3.67	0.28	3.39	13.23	AECO-PECO	34	14.75	0.03	14.72	582.02
JCPL-PPL	26	2.01	0.50	1.52	4.06	PSEG-PPL	28	0.62	0.01	0.61	41.84
JCPL-PSEG	25	2.36	0.01	2.35	284.16	PPL-AECO	27	33.47	0.02	33.45	1896.22
PECO-PPL	24	0.55	0.18	0.37	3.00	PECO-PPL	26	3.39	0.01	3.38	438.05
DPL-PECO	24	2.57	0.10	2.47	25.66	METED-PECO	25	2.07	0.07	2.00	29.01
JCPL-PECO	23	1.10	0.05	1.05	21.31	JCPL-PPL	23	1.54	0.02	1.52	69.88

Panel c: Summer Months						Panel d: All Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
AECO-PECO	25	52.52	0.08	52.44	632.06	PPL-PECO	115	1.10	0.11	0.99	10.07
PSEG-PECO	19	2.68	0.02	2.66	142.41	PSEG-PECO	114	2.31	0.08	2.22	27.97
PSEG-JCPL	18	34.27	0.01	34.26	3875.92	PSEG-JCPL	90	8.81	0.01	8.80	1148.14
PPL-PECO	16	1.23	0.08	1.15	15.43	AECO-PECO	80	22.46	0.05	22.41	416.89
PPL-AECO	14	78.89	0.47	78.42	167.51	METED-PECO	77	3.58	0.06	3.52	60.09
PECO-PPL	13	0.16	0.01	0.14	10.93	METED-PPL	75	2.64	0.08	2.56	35.15
METED-PPL	12	6.38	0.24	6.14	26.20	PSEG-PPL	63	1.92	0.17	1.75	11.43
PECO-PSEG	10	3.03	0.03	3.01	114.94	PECO-PPL	63	1.75	0.10	1.65	17.61
JCPL-PPL	9	0.57	0.02	0.54	24.20	PPL-AECO	62	35.13	0.29	34.84	122.03
METED-PECO	9	7.56	0.02	7.55	441.84	JCPL-PPL	58	1.59	0.35	1.24	4.57

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time t_H and underlying month H , implied loss prices $\Pi_{t_H}(L_{j,H} - L_{i,H})$ are calculated for the most liquid SO i and SL j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of earlier auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_H}(L_{j,H} - L_{i,H})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,H} - L_{i,H}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.8: Sample variance of the loss spreads $L_{j,H} - L_{i,H}$ and of their earlier-auction-implied prices for the most liquid FTRs - PJM on-peak contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	n	(A)	(B)	$(A) - (B)$	$(A)/(B)$	$j - i$	n	(A)	(B)	$(A) - (B)$	$(A)/(B)$
PSEG-PECO	54	2.50	0.12	2.38	20.93	PPL-PECO	46	2.59	0.03	2.57	100.96
PPL-METED	36	2.01	0.05	1.96	41.33	AECO-PECO	42	4.79	0.02	4.76	200.51
PSEG-PPL	35	2.15	0.21	1.94	10.43	METED-PPL	39	1.16	0.06	1.10	18.33
METED-PECO	34	1.85	0.04	1.81	46.30	PSEG-PECO	37	0.67	0.02	0.65	28.80
PPL-PECO	32	0.76	0.13	0.63	5.86	PECO-PPL	36	0.98	0.02	0.97	57.71
PECO-PPL	32	1.05	0.17	0.88	6.25	PPL-METED	35	1.64	0.09	1.55	19.07
METED-PPL	30	1.42	0.03	1.40	56.38	PPL-AECO	26	11.93	0.02	11.91	623.53
DPL-PECO	27	3.61	0.35	3.26	10.29	PSEG-JCPL	26	2.10	0.01	2.09	192.38
PSEG-JCPL	27	4.25	0.01	4.24	593.27	PPL-JCPL	26	0.90	0.02	0.88	41.39
PPL-JCPL	23	1.19	0.28	0.92	4.30	PSEG-PPL	24	0.26	0.02	0.24	15.11

Panel c: Summer Months						Panel d: All Months					
$j - i$	n	(A)	(B)	$(A) - (B)$	$(A)/(B)$	$j - i$	n	(A)	(B)	$(A) - (B)$	$(A)/(B)$
PSEG-PECO	21	0.93	0.02	0.90	37.07	PSEG-PECO	112	1.85	0.09	1.77	21.17
PPL-METED	17	0.59	0.02	0.58	39.31	PPL-PECO	94	2.07	0.09	1.98	23.06
PECO-PPL	16	0.25	0.02	0.23	11.44	PPL-METED	88	1.66	0.06	1.60	28.60
PPL-PECO	16	2.64	0.04	2.61	73.10	PECO-PPL	84	1.03	0.09	0.93	11.03
AECO-PECO	14	1.67	0.12	1.55	13.73	METED-PPL	81	1.39	0.04	1.34	33.68
METED-PPL	12	1.66	0.00	1.66	358.00	AECO-PECO	75	4.21	0.05	4.16	82.52
JCPL-PPL	11	0.46	0.20	0.26	2.31	PSEG-PPL	67	1.66	0.16	1.51	10.67
PPL-JCPL	10	1.71	0.06	1.64	26.71	METED-PECO	64	1.45	0.12	1.34	12.41
AEP-DAY	9	1.66	0.15	1.52	11.17	PSEG-JCPL	62	3.58	0.01	3.57	384.85
PPL-AECO	9	8.16	0.37	7.79	22.23	PPL-JCPL	59	1.40	0.16	1.24	8.63

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time $t_{\mathcal{H}}$ and underlying month \mathcal{H} , implied loss prices $\Pi_{t_{\mathcal{H}}}(L_{j;\mathcal{H}} - L_{i;\mathcal{H}})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of earlier auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_{\mathcal{H}}}(L_{j;\mathcal{H}} - L_{i;\mathcal{H}})$ is represented by (A) . The sample variance of the realized loss spreads $L_{j;\mathcal{H}} - L_{i;\mathcal{H}}$ is represented by (B) . The difference, column $(A) - (B)$, and the ratio, column $(A)/(B)$, are also provided in each panel.

Table A.9: Sample variance of the loss spreads $L_{j;\mathcal{H}} - L_{i;\mathcal{H}}$ and of their earlier-auction-implied prices for the most liquid FTRs - PJM off-peak contract

Panel a: Winter Months						Panel b: Shoulder Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	62	0.94	0.14	0.80	6.63	PPL-PECO	36	0.91	0.02	0.89	41.53
METED-PECO	42	2.26	0.04	2.22	60.87	PSEG-PECO	33	0.88	0.02	0.86	37.24
PECO-PPL	27	0.23	0.05	0.18	4.37	PECO-PPL	23	0.47	0.02	0.46	30.63
PPL-PECO	24	0.73	0.17	0.57	4.36	JCPL-PPL	20	3.61	0.03	3.58	139.21
JCPL-PPL	22	0.35	0.35	0.00	0.99	PSEG-PPL	19	0.08	0.02	0.05	3.15
PSEG-JCPL	21	1.86	0.01	1.85	208.45	METED-PECO	15	1.11	0.13	0.98	8.31
PSEG-PPL	21	1.86	0.18	1.67	10.09	PSEG-JCPL	14	1.51	0.00	1.51	525.60
JCPL-PECO	16	2.04	0.12	1.93	17.24	AECO-PECO	13	16.19	0.03	16.17	575.84
METED-PPL	13	2.18	0.09	2.09	24.42	PPL-JCPL	12	0.69	0.03	0.66	26.34
JCPL-PSEG	9	2.10	0.01	2.10	235.42	PPL-PSEG	11	2.33	0.02	2.31	94.54

Panel c: Summer Months						Panel d: All Months					
$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)	$j - i$	n	(A)	(B)	(A) - (B)	(A)/(B)
PSEG-PECO	14	0.21	0.01	0.19	13.96	PSEG-PECO	109	1.00	0.12	0.88	8.54
PPL-PECO	9	1.77	0.01	1.76	215.09	PPL-PECO	69	1.03	0.09	0.94	11.20
DPL-PSEG	8	41.40	0.63	40.77	65.62	METED-PECO	57	2.08	0.10	1.98	20.66
JCPL-PPL	8	0.08	0.04	0.03	1.86	PECO-PPL	56	0.53	0.06	0.47	9.56
PSEG-PPL	6	0.28	0.03	0.25	8.18	JCPL-PPL	50	1.60	0.22	1.38	7.33
ATSI-DUQ	6	3.35	0.24	3.11	13.80	PSEG-PPL	46	1.06	0.16	0.90	6.76
PECO-PPL	6	1.47	0.04	1.43	39.42	PSEG-JCPL	39	1.51	0.01	1.50	167.76
JCPL-PECO	6	3.50	0.04	3.46	88.65	JCPL-PECO	33	2.24	0.10	2.14	22.84
AECO-PECO	6	0.96	0.04	0.92	24.61	AECO-PECO	26	10.02	0.03	9.99	339.54
METED-PPL	5	0.22	0.00	0.22	228.08	METED-PPL	24	1.55	0.10	1.45	15.15

Each panel shows the results for one seasonality bucket: Panel a is for winter months (December to March); panel b is for shoulder months (April, May, October and November); panel c is for summer months (June to September); panel d covers all months of the year. For the earlier auction time t_H and underlying month H , implied loss prices $\Pi_{t_H}(L_{j,H} - L_{i,H})$ are calculated for the most liquid SO i and SI j zone combinations in each seasonality bucket via Equation (2.15). The second column exhibits the number of earlier auctions for which at least one unit of such a contract was traded. The sample variance of $\Pi_{t_H}(L_{j,H} - L_{i,H})$ is represented by (A). The sample variance of the realized loss spreads $L_{j,H} - L_{i,H}$ is represented by (B). The difference, column (A) - (B), and the ratio, column (A)/(B), are also provided in each panel.

Table A.10: Sample variance of the loss spreads $L_{j,H} - L_{i,H}$ and of their earlier-auction-implied prices for the most liquid FTRs - PJM 24-hour contract

A.2 Stochastic Risk Premium

We construct a model containing a stochastic risk premium for the loss component and use it as a robustness check to verify whether our conclusions are robust when dropping the piecewise deterministic risk premium assumption used in the paper.

As commonly used in derivative pricing, the price of $L_{i,\mathcal{H}}$ is its expected discounted value where the discount factor is the risk-free rate to which we add a risk premium proportional to its conditional variance,

$$\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) = \mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} \left[e^{-(r+\eta_i \text{Var}_{t_{\mathcal{H}}}^{\mathbb{P}}[L_{i,\mathcal{H}}])} L_{i,\mathcal{H}} \right].$$

Assuming the price of risk is the same for each node, that is $\eta_i = \eta$ for all $i \in \mathcal{Z}_{t_{\mathcal{H}}}$, the risk premium is node dependent and time-varying only through the conditional variance.

Letting

$$\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} = (e^{-(r+\eta \text{Var}_{t_{\mathcal{H}}}^{\mathbb{P}}[L_{i,\mathcal{H}}])} - 1)L_{i,\mathcal{H}},$$

the price of $L_{i,\mathcal{H}}$ is decomposed into two components:

$$\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}}) = \mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} [L_{i,\mathcal{H}}] + \mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} [\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}}].$$

With the cross-sectional averages defined as

$$\overline{L}_{\mathcal{H}} = \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{L_{i,\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}}|}, \quad \text{and} \quad \overline{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}} = \sum_{i=1}^{|\mathcal{Z}_{t_{\mathcal{H}}}|} \frac{\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}}}{|\mathcal{Z}_{t_{\mathcal{H}}}|},$$

the relation (2.14) becomes

$$\begin{aligned} \text{Var}^{\mathbb{P}} [\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}})] &= \text{Var}^{\mathbb{P}} [\mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} [(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}})] + \mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} [(\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}})]] \quad (\text{A.1}) \\ &\leq \text{Var}^{\mathbb{P}} [\mathbb{E}_{t_{\mathcal{H}}}^{\mathbb{P}} [(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}) + (\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}})]] \\ &+ \mathbb{E}^{\mathbb{P}} [\text{Var}_{t_{\mathcal{H}}}^{\mathbb{P}} [(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}) + (\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}})]] \\ &= \text{Var}^{\mathbb{P}} [(L_{i,\mathcal{H}} - \overline{L}_{\mathcal{H}}) + (\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}})] \end{aligned}$$

$$= \text{Var}^{\mathbb{P}} [L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] + \text{Var}^{\mathbb{P}} [\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}] + 2\text{Cov}^{\mathbb{P}} [L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}, \mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}] .$$

The left-hand side is measured through the sample variance of the monthly sample of loss prices $\{\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})\}_{t_{\mathcal{H}}}$ obtained from a portfolio of FTR and futures prices.

To compute the right hand-side of Inequality (A.1), we need to construct the monthly time series $\{\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}\}_{t_{\mathcal{H}}}$ which depends on the time-varying loss conditional variance. The latter is captured using a GARCH framework widely used in the derivative pricing literature. Because $L_{i,\mathcal{H}}$ is the aggregation of hourly losses across the whole month \mathcal{H} , then at the auction time $t_{\mathcal{H}}$ located sometime in the previous month $\mathcal{H} - 1$, the aggregated loss $L_{i,\mathcal{H}-1}$ is only partially observed. For that reason, the loss model autoregressive component depends on $L_{i,\mathcal{H}-2}$. The FTR and futures prices are also used as explanatory variables:

$$\begin{aligned} L_{i,\mathcal{H}} &= \phi_{0,i} + \phi_{1,i}L_{i,\mathcal{H}-2} + \phi_{2,i}\text{FTR}_{t_{\mathcal{H}},i,\mathcal{H}} + \phi_{3,i}F_{t_{\mathcal{H}},i,\mathcal{H}} + \sigma_{t_{\mathcal{H}},i}\epsilon_{i,\mathcal{H}}, \\ \sigma_{t_{\mathcal{H}},i}^2 &= \omega_i + \alpha_i\epsilon_{i,\mathcal{H}-1}^2 + \beta_i\sigma_{t_{\mathcal{H}-1},i}^2, \end{aligned} \quad (\text{A.2})$$

where $\epsilon_{i,\mathcal{H}}$ is a standard normal random variable and $\sigma_{t_{\mathcal{H}},i}^2$ stands for the loss conditional variance. The model is estimated by a maximum likelihood approach and we retrieve the monthly time series of conditional variance that becomes a proxy for $\text{Var}_{t_{\mathcal{H}}}^{\mathbb{P}} [L_{i,\mathcal{H}}]$. We approximate $\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} \approx \left(e^{-(r+\eta)\sigma_{t_{\mathcal{H}},i}^2} - 1 \right) L_{i,\mathcal{H}}$.

For any given value of η , the right hand-side of Inequality (A.1) is evaluated with the sample variances and covariance of the two monthly time series $\{L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}\}_{t_{\mathcal{H}}}$ and $\{\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}\}_{t_{\mathcal{H}}}$.

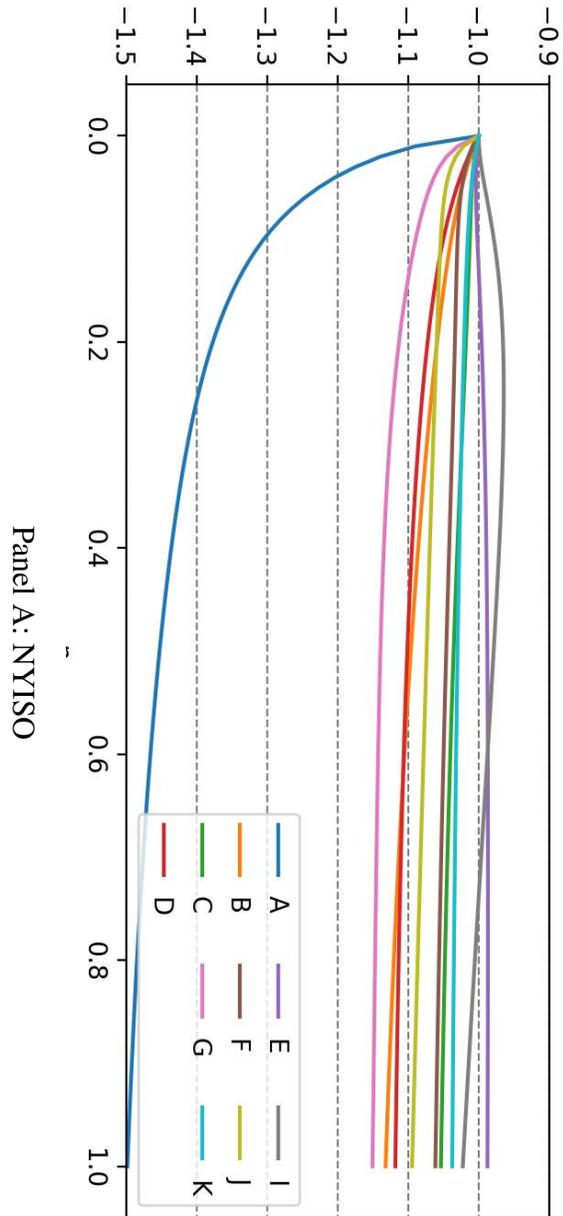
To verify if Inequality (A.1) is empirically verified, we compute

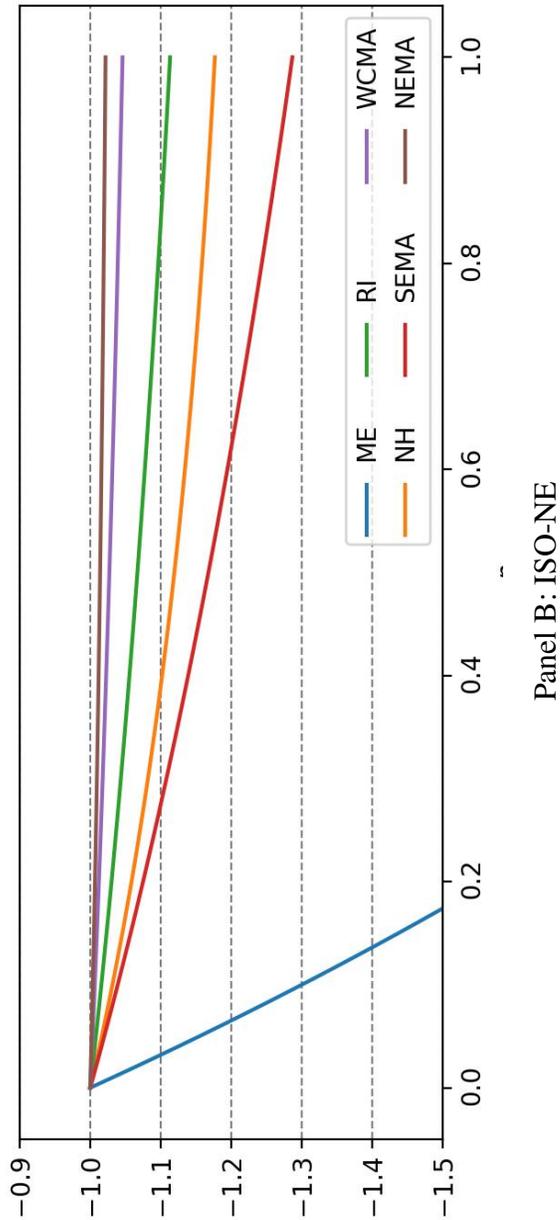
$$\begin{aligned} \mathcal{R}(\eta) &= \text{Var}^{\mathbb{P}} [L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}] - \text{Var}^{\mathbb{P}} [\Pi_{t_{\mathcal{H}}}(L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}})] \\ &\quad + \text{Var}^{\mathbb{P}} [\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}] + 2\text{Cov}^{\mathbb{P}} [L_{i,\mathcal{H}} - \overline{L_{\mathcal{H}}}, \mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \overline{\mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}}] , \end{aligned} \quad (\text{A.3})$$

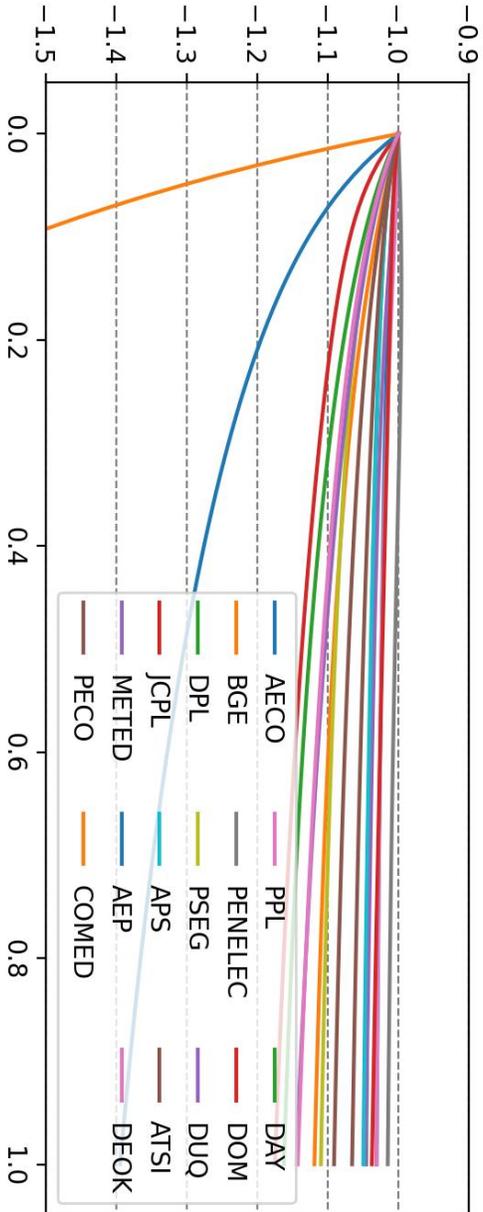
for many prices of risk η . Because the loss variances vary from one node to the other, we normalized Equation (A.3) by a positive constant to present all the curves on the same

figure, the important point being that Inequality (A.1) is violated if the normalized values $\frac{\mathcal{R}(\eta)}{|\mathcal{R}(0)|}$ are negative.

Figure A.1 shows that for a wide range of η , the curves always take negative values, which invalidates Inequality (A.1) and supports the main conclusion of the paper.







For each node and market, we evaluated the loss model (A.2) and retrieve its conditional variance to compute the excess price of risk contribution $\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}$ and its sample variance $\text{Var}^{\mathbb{P}}[\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \mathcal{L}_{t_{\mathcal{H}},\mathcal{H}}]$. For each zone i , we use the realized excess losses $L_{i,t_{\mathcal{H}}} - \bar{L}_{t_{\mathcal{H}}}$ to compute the variance $\text{Var}^{\mathbb{P}}[L_{i,t_{\mathcal{H}}} - \bar{L}_{t_{\mathcal{H}}}]$ and its sample covariance with $\mathcal{L}_{i,t_{\mathcal{H}},\mathcal{H}} - \bar{\mathcal{L}}_{t_{\mathcal{H}},\mathcal{H}}$. We also take its corresponding nearest-auction-
implied price $\Pi_t(L_{i,t_{\mathcal{H}}} - \bar{L}_{t_{\mathcal{H}}})$ to compute its sample variance $\text{Var}^{\mathbb{P}}[\Pi_{t_{\mathcal{H}}}(L_{i,t_{\mathcal{H}}} - \bar{L}_{t_{\mathcal{H}}})]$, permitting to evaluate quantity (A.3) that is graphed above for each zone. We display the normalized values $\frac{\mathcal{R}^{(i)}}{\mathcal{R}^{(0)}}$ for presentation purposes.

Panel C: PJM

Chapter B

Appendices of *Computationally-Efficient Variance Filtering in Multidimensional Affine Models*

B.1 Measurement Error

B.1.1 Asymptotic theory

Section 2.3 of Barndorff-Nielsen and Shephard, 2004b shows that the general asymptotic convergence of the multivariate realized covariance measures is

$$\sqrt{\frac{N}{\Delta}} \left(\begin{bmatrix} RCV_{i,j,t,t+1} - QCV_{i,j,t,t+1} \\ RCV_{\tilde{i},\tilde{j},t,t+1} - QCV_{\tilde{i},\tilde{j},t,t+1} \end{bmatrix} \right) \left(\int_{t\Delta}^{(t+1)\Delta} \Sigma_{i\tilde{i},j\tilde{j},s} ds \right)^{-1/2} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{1}), \quad (\text{B.1})$$

where¹

$$\Sigma_{ij\tilde{i}\tilde{j},s} = \begin{bmatrix} \Sigma_{ii,s}\Sigma_{jj,s} + \Sigma_{ij,s}^2 & \Sigma_{i\tilde{i},s}\Sigma_{j\tilde{j},s} + \Sigma_{ij,s}\Sigma_{\tilde{i}\tilde{j},s} \\ \Sigma_{i\tilde{i},s}\Sigma_{j\tilde{j},s} + \Sigma_{ij,s}\Sigma_{\tilde{i}\tilde{j},s} & \Sigma_{\tilde{i}\tilde{i},s}\Sigma_{\tilde{j}\tilde{j},s} + \Sigma_{\tilde{i}\tilde{j},s}^2 \end{bmatrix},$$

and

$$\Sigma_{ij,s} = \sum_{k=1}^{N_V} c_{i,k}c_{j,k}V_{k,s}. \quad (\text{B.2})$$

Furthermore, Barndorff-Nielsen and Shephard, 2004a shows that the leverage effect does not affect this asymptotic convergence.

The weak convergence in Equation (B.1) suggests that the realized covariance is a noisy signal of the integrated covariance, as described in Equation (3.13),

$$RCV_{i,j,t,t+1} = QCV_{i,j,t,t+1} + \xi_{i,j,t,t+1},$$

where the measurement error $\xi_{i,j,t,t+1}$ is asymptotically Gaussian with²

$$\mathbb{E} [\xi_{i,j,t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad (\text{B.3})$$

$$\text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] = \frac{\Delta}{N} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \Sigma_{i\tilde{i},s}\Sigma_{j\tilde{j},s} + \Sigma_{ij,s}\Sigma_{\tilde{i}\tilde{j},s} ds \middle| \mathcal{F}_{t\Delta} \right]. \quad (\text{B.4})$$

Based on the asymptotic theory,³ we set the variance of the measurement error to

¹With Itô's Isometry, the expected value of quadratic covariation is

$$\begin{aligned} \mathbb{E} [QCV_{i,j,t,t+1}] &= \mathbb{E} \left[\left(\sum_{k=1}^{N_V} \int_{t\Delta}^{(t+1)\Delta} c_{i,k} \sqrt{V_{k,s}} dW_{k,s} \right) \left(\sum_{k=1}^{N_V} \int_{t\Delta}^{(t+1)\Delta} c_{j,k} \sqrt{V_{k,s}} dW_{k,s} \right) \right] \\ &= \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \sum_{k=1}^{N_V} c_{i,k}c_{j,k}V_{k,s} ds \right] = \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \Sigma_{ij,s} ds \right], \end{aligned}$$

motivating the shorted notation $\Sigma_{ij,s}$.

²The information available at time $t\Delta$, $\mathcal{F}_{t\Delta}$, is defined in Equation (3.15).

³Additional discussion on the multivariate properties is provided in Online Appendix D.2.

$\eta \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \Sigma_{i\tilde{i},s} \Sigma_{j\tilde{j},s} + \Sigma_{i\tilde{j},s} \Sigma_{j\tilde{i},s} ds \middle| \mathcal{F}_{t\Delta} \right]$, where $\eta = \frac{\Delta}{N}$. However, in practice, the realized variance suffers from microstructure noise (Zhang et al., 2005), which impacts the variability of the intraday statistics. Therefore, estimating η can be useful to mitigate this effect and motivates our decision to not replace η by its asymptotic theoretical value.

The forthcoming sections provide a comprehensive derivation of the conditional moments of the error variance with respect to the model and observed filtration respectively.

B.1.2 Moments of $\xi_{t,t+1}$ with respect to the model filtration $\mathcal{F}_{t\Delta}$

Like most modeling approaches, the measurement error is assumed to be uncorrelated with the model components.

Assumption B.1.1. *The error measurement is uncorrelated with the stochastic integrals, that is*

$$\text{Cov} \left[\xi_{t,t+1}, \begin{bmatrix} \mathcal{W}_{t,t+1} \\ \mathcal{B}_{t,t+1} \\ \mathcal{Z}_{t,t+1} \end{bmatrix} \middle| \mathcal{F}_{t\Delta} \right] = \mathbf{0}.$$

A Monte Carlo experiment presented in Online Appendix D.10.2 shows that Assumption B.1.1 is empirically satisfied.

The covariance matrix follows the asymptotic theory as in Equation (B.4)

$$\text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] = \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \quad (\text{B.5})$$

where,

$$\mathcal{C}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} = c_{i,k} c_{\tilde{i},k} c_{j,\tilde{k}} c_{\tilde{j},\tilde{k}} + c_{i,k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\tilde{i},\tilde{k}}, \quad (\text{B.6})$$

and⁴

$$\mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] = \mathcal{A}_{k,\bar{k}} + \mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta}, \quad (\text{B.7})$$

where

$$\begin{aligned} \mathcal{A}_{k,\bar{k}} &= \frac{\kappa_{\bar{k}} \theta_{\bar{k}} + \mathbb{1}_{k=\bar{k}} \sigma_k \sigma_{\bar{k}}}{(\kappa_k + \kappa_{\bar{k}})} \theta_k \left(\Delta - \frac{1 - e^{-(\kappa_k + \kappa_{\bar{k}})\Delta}}{(\kappa_k + \kappa_{\bar{k}})} \right) - \theta_k \mathcal{D}_{k,\bar{k}} \\ &\quad + \frac{\kappa_k \theta_k}{(\kappa_k + \kappa_{\bar{k}})} \theta_{\bar{k}} \left(\Delta - \frac{1 - e^{-(\kappa_k + \kappa_{\bar{k}})\Delta}}{(\kappa_k + \kappa_{\bar{k}})} \right) - \theta_{\bar{k}} \mathcal{E}_{k,\bar{k}}, \\ \mathcal{B}_{k,\bar{k}} &= \frac{(1 - e^{-(\kappa_k + \kappa_{\bar{k}})\Delta})}{(\kappa_k + \kappa_{\bar{k}})}, \\ \mathcal{D}_{k,\bar{k}} &= \frac{\kappa_{\bar{k}} \theta_{\bar{k}} + \mathbb{1}_{k=\bar{k}} \sigma_k \sigma_{\bar{k}}}{(\kappa_k + \kappa_{\bar{k}})} \left(\frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} - e^{-\kappa_k \Delta} \frac{1 - e^{-\kappa_{\bar{k}} \Delta}}{\kappa_{\bar{k}}} \right), \\ \mathcal{E}_{k,\bar{k}} &= \frac{\kappa_k \theta_k}{(\kappa_k + \kappa_{\bar{k}})} \left(\frac{1 - e^{-\kappa_{\bar{k}} \Delta}}{\kappa_{\bar{k}}} - e^{-\kappa_{\bar{k}} \Delta} \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right). \end{aligned}$$

Therefore, the expected value of the integrated square variance is a quadratic function of the latent variable.

B.1.3 Moments of $\xi_{t,t+1}$ with respect to the observed filtration $\mathcal{G}_{t\Delta}$

Now that the conditional moments of $\xi_{i,j,t,t+1}$ with respect to the model filtration $\mathcal{F}_{t\Delta}$ are derived, we can compute conditional moments with respect to the observed filtration $\mathcal{G}_{t\Delta}$. For any $i, j \in \{1, 2, \dots, N_S\}$:

$$\begin{aligned} \text{Cov} [\xi_{i,j,t,t+1}, \mathbf{V}_{(t+1)\Delta} \middle| \mathcal{G}_{t\Delta}] &= \mathbb{E} \left[\underbrace{\text{Cov} [\xi_{i,j,t,t+1}, \mathbf{V}_{(t+1)\Delta} \middle| \mathcal{F}_{t\Delta}]}_{=0} \middle| \mathcal{G}_{t\Delta} \right] \\ &\quad + \text{Cov} \left[\underbrace{\mathbb{E} [\xi_{i,j,t,t+1} \middle| \mathcal{F}_{t\Delta}]}_{=0}, \mathbb{E} [\mathbf{V}_{(t+1)\Delta} \middle| \mathcal{F}_{t\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\ &= 0. \end{aligned}$$

⁴Additional details are available in the Online Appendix D.3.

When substituting $\mathbf{V}_{(t+1)\Delta}$ with $\mathcal{W}_{t,t+1}$, $\mathcal{B}_{t,t+1}$, or $\mathcal{V}_{t,t+1}$ in the covariance expression provided above, the same rationale applies.

The conditional covariance matrix $\text{Cov} [\xi_{t,t+1} | \mathcal{G}_{t\Delta}]$ is

$$\begin{aligned} \text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{G}_{t\Delta}] &= \text{Cov} \left[\underbrace{\text{E} [\xi_{i,j,t,t+1} | \mathcal{F}_{t\Delta}]}_{=0}, \underbrace{\text{E} [\xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}]}_{=0} \middle| \mathcal{G}_{t\Delta} \right] \\ &\quad + \text{E} [\text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}]. \end{aligned}$$

Using Equation (B.7), the expectation with respect to $\mathcal{G}_{t\Delta}$ of Equation (B.5) is

$$\begin{aligned} &\text{E} [\text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,\tilde{i},j,\tilde{j}} \text{E} \left[\text{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\ &= \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,\tilde{i},j,\tilde{j}} \text{E} [\mathcal{A}_{k,\tilde{k}} + \mathcal{B}_{k,\tilde{k}} V_{k,t\Delta} V_{\tilde{k},t\Delta} + \mathcal{D}_{k,\tilde{k}} V_{k,t\Delta} + \mathcal{E}_{k,\tilde{k}} V_{\tilde{k},t\Delta} | \mathcal{G}_{t\Delta}] \\ &= \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,\tilde{i},j,\tilde{j}} \left(\begin{array}{c} \mathcal{A}_{k,\tilde{k}} + \mathcal{D}_{k,\tilde{k}} \text{E} [V_{k,t\Delta} | \mathcal{G}_{t\Delta}] + \mathcal{E}_{k,\tilde{k}} \text{E} [V_{\tilde{k},t\Delta} | \mathcal{G}_{t\Delta}] \\ + \mathcal{B}_{k,\tilde{k}} (\text{Cov} [V_{k,t\Delta}, V_{\tilde{k},t\Delta} | \mathcal{G}_{t\Delta}] + \text{E} [V_{k,t\Delta} | \mathcal{G}_{t\Delta}] \text{E} [V_{\tilde{k},t\Delta} | \mathcal{G}_{t\Delta}]) \end{array} \right), \end{aligned}$$

which is a non linear function of the conditional moments of the latent variables.

B.2 First two Conditional Moments

This section summarizes the moments related to instantaneous variance, integrated variance and stochastic integrals that are used in the filter.⁵

The subscript k is omitted for readability. We denote $\mathcal{V}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} V_s ds$, $\mathcal{W}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dW_s$, $\mathcal{Z}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} e^{-\kappa((t+1)\Delta-s)} \sqrt{V_s} dW_s$ and $\mathcal{B}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dB_s$.

$$\text{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] = V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right),$$

⁵Proofs are available upon request.

$$\begin{aligned}
\text{Var} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \frac{\sigma^2}{\kappa^2} V_{t\Delta} \left(\frac{1 - e^{-2\kappa\Delta}}{\kappa} - 2e^{-\kappa\Delta} \Delta \right) \\
&\quad + \frac{\sigma^2}{\kappa^2} \theta \left(\Delta - \frac{1 - e^{-2\kappa\Delta}}{2\kappa} - 2\frac{1 - e^{-\kappa\Delta}}{\kappa} + 2e^{-\kappa\Delta} \Delta \right), \\
\text{E} [\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\
\text{Var} [\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\
\text{E} [\mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\
\text{Var} [\mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\
\text{E} [\mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\
\text{Var} [\mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} e^{-\kappa\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\frac{1 - e^{-2\kappa\Delta}}{2\kappa} - e^{-\kappa\Delta} \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\
\text{Cov} [\mathcal{W}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\
\text{Cov} [\mathcal{Z}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\
\text{Cov} [\mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} e^{-\kappa\Delta} \Delta + \theta \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right), \\
\text{Cov} [\mathcal{W}_{t,t+1}, \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \frac{\sigma}{\kappa} V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right) \\
&\quad + \frac{\sigma}{\kappa} \theta \left(\Delta - 2\frac{1 - e^{-\kappa\Delta}}{\kappa} + e^{-\kappa\Delta} \Delta \right), \\
\text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0.
\end{aligned}$$

The strong solution of Equation (3.2) is

$$V_{(t+1)\Delta} = V_{t\Delta} e^{-\kappa\Delta} + \theta (1 - e^{-\kappa\Delta}) + \sigma \mathcal{Z}_{t,t+1}, \quad (\text{B.8})$$

which leads to

$$\mathcal{V}_{t,t+1} = (V_{t\Delta} - \theta) \frac{(1 - e^{-\kappa\Delta})}{\kappa} + \theta \Delta + \frac{\sigma}{\kappa} (\mathcal{W}_{t,t+1} - \mathcal{Z}_{t,t+1}). \quad (\text{B.9})$$

The matrices defined in Equation (3.16) are diagonal where

$$\begin{aligned}
 \mathbf{P}_{1,k,k}^{\mathcal{W}\mathcal{W}} &= \mathbf{P}_{1,k,k}^{\mathcal{B}\mathcal{B}} = \left(\frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right), & \mathbf{P}_{0,k,k}^{\mathcal{W}\mathcal{W}} &= \mathbf{P}_{0,k,k}^{\mathcal{B}\mathcal{B}} = \theta_k \left(\Delta - \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right), \\
 \mathbf{P}_{1,k,k}^{\mathcal{Z}\mathcal{Z}} &= e^{-\kappa_k \Delta} \left(\frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right), & \mathbf{P}_{0,k,k}^{\mathcal{Z}\mathcal{Z}} &= \theta_k \left(\frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} - e^{-\kappa_k \Delta} \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right), \\
 \mathbf{P}_{1,k,k}^{\mathcal{W}\mathcal{Z}} &= e^{-\kappa_k \Delta} \Delta, & \mathbf{P}_{0,k,k}^{\mathcal{W}\mathcal{Z}} &= \theta_k \left(\frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} - e^{-\kappa_k \Delta} \Delta \right).
 \end{aligned} \tag{B.10}$$

B.3 Derivation of the Update Matrix

The goal is to choose $\Sigma_{t\Delta}$ such that it minimizes the conditional variance derived in Equation (3.25). We choose to minimize the trace of Equation (3.25), which corresponds to the sum of the prediction error variances. The first-order condition is

$$2\Sigma_{t\Delta} \text{Var} [\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] - 2\text{Cov} [\mathbf{O}_{(t+1)\Delta}, \mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]^\top = 0,$$

which implies that

$$\Sigma_{t\Delta} = \text{Cov} [\mathbf{V}_{(t+1)\Delta}, \mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] (\text{Var} [\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}])^{-1}.$$

The second derivative is

$$2\text{Var} [\mathbf{O}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}],$$

which is positive and confirms that the prediction error variance has been minimized.

B.4 Third and Fourth Conditional Cumulant

Here, we summarize the derivations of the third and fourth cumulants of the stochastic integrals, observable variables and latent variables. We begin by formally defining these cumulants and establishing their connection with central moments. Subsequently, we present

the third and fourth cumulants of the latent variance both before and after the update.

We employ cumulants instead of central moments due to the model's affine nature in the latent state. Consequently, the model's conditional cumulants with respect to the model filtration, excluding measurement errors, exhibit linearity in the latent state up to the fourth conditional cumulant. This linearity specifically applies to the fourth cumulant rather than the fourth central moment, simplifying our computations when working with cumulants.

B.4.1 Third cumulant

We define the third cumulant and its application to the law of total cumulance it in Section B.4.1.1. The latter is useful as it permits the transition from the model filtration to the observed filtration. In Section B.4.1.2, we show the results regarding the third central moment of the stochastic integrals $\mathcal{B}_{t,t+1}$, $\mathcal{W}_{t,t+1}$, and $\mathcal{Z}_{t,t+1}$. Finally, we apply these results to the model's latent variable and observations before the update in Section B.4.1.3 and present our update of the latent variable in Section B.4.1.4.

B.4.1.1 Definitions

Definition B.4.1. *For a random variable X_i , $i \in \{1, 2, 3\}$, the third cumulant is equivalent to the third central moment, that is:*

$$\text{Cum}^{(3)} [X_1, X_2, X_3] = \text{E} [(X_1 - E[X_1]) (X_2 - E[X_2]) (X_3 - E[X_3])].$$

Lemma B.4.2. *If $\mathcal{G}_{t\Delta} \subseteq \mathcal{F}_{t\Delta}$, with the random variables $X_{k,(t+1)\Delta}$, $k \in \{i, j, l\}$, the law of total cumulance applied to the third cumulant is*

$$\begin{aligned} & \text{Cum}^{(3)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{E} \left[\text{Cum}^{(3)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{i,(t+1)\Delta}, | \mathcal{F}_{t\Delta}], \text{E} [X_{j,(t+1)\Delta}, | \mathcal{F}_{t\Delta}], \text{E} [X_{l,(t+1)\Delta}, | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Var} [\text{E} [X_{i,(t+1)\Delta}, | \mathcal{F}_{t\Delta}], \text{Var} [X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Var} [\text{E} [X_{j,(t+1)\Delta}, | \mathcal{F}_{t\Delta}], \text{Var} [X_{i,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \end{aligned}$$

$$+ \text{Var} \left[\mathbb{E} \left[X_{l,(t+1)\Delta}, \mathcal{F}_{t\Delta} \right], \text{Var} \left[X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta} \right] | \mathcal{G}_{t\Delta} \right].$$

Proof: Direct application of the law of total cumulance (Brillinger, 1969). \square

B.4.1.2 Stochastic integrals

This subsection details how to calculate the third moments of stochastic integrals $\mathcal{B}_{t,t+1}$, $\mathcal{W}_{t,t+1}$ and $\mathcal{Z}_{t,t+1}$ in Corollaries B.4.3 and B.4.4. The subscript k is omitted for readability.

Corollary B.4.3.

$$\begin{aligned} \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds = 0, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds = 0, \\ \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds = 0, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds = 0, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds, \\ \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds, \\ \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 2e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &\quad + e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds, \\ \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 2e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \end{aligned}$$

$$+ e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s \Delta} ds.$$

Proof. Direct application of Lemma D.6.1 and

$$\mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}], \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] \text{ and } \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}]$$

are detailed in Corollary B.4.4. □

Corollary B.4.4.

$$\begin{aligned} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] &= \sigma \mathbb{E} [\mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] = \sigma e^{-2\kappa s \Delta} \int_t^s \mathbb{E} [V_{r\Delta} | \mathcal{F}_{t\Delta}] e^{2\kappa r \Delta} dr, \\ \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] &= \sigma \mathbb{E} [\mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] = \sigma e^{-\kappa s \Delta} \int_t^s \mathbb{E} [V_{r\Delta} | \mathcal{F}_{t\Delta}] e^{\kappa r \Delta} dr, \\ \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] &= 0. \end{aligned}$$

Proof. Direct application of Lemma D.6.1. □

B.4.1.3 Model components

We show here how the previous section is used to compute the model component's third central moment with respect to the model filtration. Indeed, the third central moment (third cumulant) of the observable variables without measurement error and the latent variables $\mathbf{V}_{(t+1)\Delta}$ with respect to the model filtration are a linear combination of the stochastic integrals $\mathcal{Z}_{t,t+1}$, $\mathcal{W}_{t,t+1}$ and $\mathcal{B}_{t,t+1}$ third cumulant.

Therefore, based on Equations (3.10) and (3.14), we have

$$\begin{aligned} &\text{Cum}^{(3)} [\mathbf{V}_{i,(t+1)\Delta}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \\ &= \sum_{k=1}^{N_V} \mathbf{S}_{i,k} \mathbf{S}_{j,k} \mathbf{S}_{l,k} \mathbb{E} [\mathcal{Z}_{k,t,t+1} \mathcal{Z}_{k,t,t+1} \mathcal{Z}_{k,t,t+1} | \mathcal{F}_{t\Delta}], \\ &\text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{F}_{t\Delta}] \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^{N_V} \mathbb{E} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right) \\ \left(\mathbf{A}\mathbf{U}_{j,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{j,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{j,k} \mathcal{B}_{k,t,t+1} \right) \\ \left(\mathbf{A}\mathbf{U}_{l,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{l,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{l,k} \mathcal{B}_{k,t,t+1} \right) \end{array} \middle| \mathcal{F}_{t\Delta} \right], \\
&\text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \\
&= \sum_{k=1}^{N_V} \mathbb{E} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right) \\ \left(\mathbf{A}\mathbf{U}_{j,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{j,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{j,k} \mathcal{B}_{k,t,t+1} \right) \\ (\mathbf{S}_{l,k} \mathcal{Z}_{k,t,t+1}) \end{array} \middle| \mathcal{F}_{t\Delta} \right], \\
&\text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \\
&= \sum_{k=1}^{N_V} \mathbb{E} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right) \\ (\mathbf{S}_{j,k} \mathcal{Z}_{k,t,t+1}) (\mathbf{S}_{l,k} \mathcal{Z}_{k,t,t+1}) \end{array} \middle| \mathcal{F}_{t\Delta} \right],
\end{aligned}$$

To pass from the model filtration to the observed filtration, the law of total cumulance (Lemma B.4.2) is used. The measurement error is then added depending on the assumption used (asymptotic theory, Assumption 3.4.1, or Gaussian approximation, Assumption 3.4.2).⁶

B.4.1.4 Latent variable update

Corollary B.4.5.

$$\begin{aligned}
&\mathbb{E} \left[\text{Cum}^{(3)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\
&= \mathbb{E} \left[\epsilon_{i,t+1|t+1}^{\mathbf{V}} \epsilon_{j,t+1|t+1}^{\mathbf{V}} \epsilon_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{t\Delta} \right]
\end{aligned}$$

Proof. Since $\mathcal{G}_{t\Delta} \subseteq \mathcal{G}_{(t+1)\Delta}$, applying Lemma B.4.2 to the filtering error $\epsilon_{t+1|t+1}^{\mathbf{V}}$ third central moment leads to

$$\begin{aligned}
&\text{Cum}^{(3)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{t\Delta}] \\
&= \mathbb{E} \left[\text{Cum}^{(3)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \middle| \mathcal{G}_{t\Delta} \right],
\end{aligned}$$

⁶See Appendix B.5 for higher cumulant of the measurement error under the model filtration.

as $E \left[\boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{V}}, \mathcal{G}_{(t+1)\Delta} \right] = 0$, therefore canceling all others terms in Lemma B.4.2.

$\text{Cum}^{(3)} \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{t\Delta} \right]$ can be rewritten as

$$\begin{aligned} & \text{Cum}^{(3)} \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{t\Delta} \right] \\ &= E \left[E \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{(t+1)\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\ &= E \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{t\Delta} \right]. \quad \square \end{aligned}$$

We set the latent variable updated third cumulant to its expected value with respect to $\mathcal{G}_{t\Delta}$, and with Corollary B.4.5, we have

$$\begin{aligned} & \text{Cum}^{(3)} \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{(t+1)\Delta} \right] \tag{B.11} \\ & \simeq E \left[\text{Cum}^{(3)} \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}}, \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{(t+1)\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\ &= E \left[\boldsymbol{\epsilon}_{i,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{j,t+1|t+1}^{\mathbf{V}} \boldsymbol{\epsilon}_{l,t+1|t+1}^{\mathbf{V}} \middle| \mathcal{G}_{t\Delta} \right] \\ &= E \left[\left(\boldsymbol{\epsilon}_{i,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{i,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{j,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{j,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{l,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{l,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \middle| \mathcal{G}_{t\Delta} \right], \end{aligned}$$

which is a linear combination of the third cumulant of $\mathbf{O}_{t+1\Delta}$ and $\mathbf{V}_{t+1\Delta}$ with respect to the information set $\mathcal{G}_{t\Delta}$ (see section B.4.1.3).

B.4.2 Fourth cumulant

We derive here the equivalent of the last section but for the fourth cumulant. We first define the fourth cumulant and its application to the law of total cumulance in Section B.4.2.1. In Section B.4.2.2, we show the results regarding the fourth cumulant of the stochastic integrals $\mathcal{B}_{t,t+1}$, $\mathcal{W}_{t,t+1}$, and $\mathcal{Z}_{t,t+1}$. Finally, we apply these results to the model's latent variable and observations before the update in Section B.4.2.3 and present our update of the latent variable in Section B.4.2.4.

B.4.2.1 Definitions

Definition B.4.6. For a random variable X_i , $i \in \{1, 2, 3, 4\}$, the fourth cumulant is

$$\begin{aligned} \text{Cum}^{(4)} [X_1, X_2, X_3, X_4] &= \text{E} [(X_1 - E[X_1]) (X_2 - E[X_2]) (X_3 - E[X_3]) (X_4 - E[X_4])] \\ &\quad - \text{Var} [X_1, X_2] \text{Var} [X_3, X_4] - \text{Var} [X_1, X_3] \text{Var} [X_2, X_4] \\ &\quad - \text{Var} [X_1, X_4] \text{Var} [X_2, X_3]. \end{aligned}$$

Lemma B.4.7. If $\mathcal{G}_{t\Delta} \subseteq \mathcal{F}_{t\Delta}$, with the random variables $X_{k,(t+1)\Delta}$, $k \in \{i, j, l, m\}$, the law of total cumulance applied to the fourth cumulant is

$$\begin{aligned} &\text{Cum}^{(4)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{E} \left[\text{Cum}^{(4)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cum}^{(4)} \left[\begin{array}{c} \text{E} [X_{i,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \\ \text{E} [X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \end{array} | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Var} [\text{Var} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Var} [\text{Var} [X_{i,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{j,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Var} [\text{Var} [X_{i,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Var} \left[\text{E} [X_{i,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Cum}^{(3)} [X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Var} \left[\text{E} [X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Cum}^{(3)} [X_{i,(t+1)\Delta}, X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Var} \left[\text{E} [X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Cum}^{(3)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Var} \left[\text{E} [X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Cum}^{(3)} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{i,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{l,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{i,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{j,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{i,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{j,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{i,(t+1)\Delta}, X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\text{E} [X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{i,(t+1)\Delta}, X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \end{aligned}$$

$$+ \text{Cum}^{(3)} \left[\mathbb{E} [X_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \mathbb{E} [X_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [X_{i,(t+1)\Delta}, X_{j,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta} \right].$$

Proof: Direct application of the law of total cumulance (Brillinger, 1969). \square

B.4.2.2 Stochastic integrals

We summarize the results detailed in Online Appendix D.6, which includes fifteen different combinations of the three stochastic integrals. First, six fourth cumulants are null. That is

$$\begin{aligned} \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0. \end{aligned}$$

Finally, the nine non-null fourth cumulants are:

$$\begin{aligned} \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \\ \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= 6\sigma \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{B}_{t,s}^2 | \mathcal{F}_{t\Delta}] ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \\ \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= \sigma \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s}^2 + \mathcal{B}_{t,s}^2) | \mathcal{F}_{t\Delta}] ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \\ \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= \sigma e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s}^2 + \mathcal{B}_{t,s}^2) e^{2\kappa s\Delta} | \mathcal{F}_{t\Delta}] ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \\ \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= \sigma e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s} \mathcal{W}_{t,s} e^{\kappa s\Delta} + \mathcal{B}_{t,s}^2 e^{\kappa s\Delta}) | \mathcal{F}_{t\Delta}] ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \\ \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= 6\sigma \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{W}_{t,s}^2 | \mathcal{F}_{t\Delta}] ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \\ \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= 3\sigma e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{W}_{t,s}^2) | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \\ \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= \sigma e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} \left[\mathcal{Z}_{t,s} \left(\begin{array}{c} \mathcal{W}_{t,s}^2 + \mathcal{Z}_{t,s}^2 \\ + 4\mathcal{W}_{t,s} \mathcal{Z}_{t,s} \end{array} \right) \middle| \mathcal{F}_{t\Delta} \right] e^{2\kappa s\Delta} ds, \end{aligned}$$

$$\begin{aligned} \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1}, \\ \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= 3\sigma e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} \left[\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s}^2) \middle| \mathcal{F}_{t\Delta} \right] e^{3\kappa s\Delta} ds, \\ \text{Cum}^{(4)} \left[\begin{array}{c} \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \\ \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] &= 6\sigma e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} \left[\mathcal{Z}_{t,s}^3 \middle| \mathcal{F}_{t\Delta} \right] e^{4\kappa s\Delta} ds, \end{aligned}$$

which are a function of the third cumulant of the stochastic integrals.

B.4.2.3 Model components

As in section B.4.1.3, we show here how the previous section is used to compute the model component's fourth cumulant with respect to the model filtration. Complete details on the derivation of these cumulants are provided in Online Appendix D.5.

Therefore, based on Equations (3.10) and (3.14), we have

$$\begin{aligned} &\text{Cum}^{(4)} \left[\mathbf{V}_{i,(t+1)\Delta}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \sum_{k=1}^{N_V} \mathbf{S}_{i,k} \mathbf{S}_{j,k} \mathbf{S}_{l,k} \mathbf{S}_{m,k} \text{Cum}^{(4)} \left[\mathcal{Z}_{k,t,t+1}, \mathcal{Z}_{k,t,t+1}, \mathcal{Z}_{k,t,t+1}, \mathcal{Z}_{k,t,t+1} \middle| \mathcal{F}_{t\Delta} \right], \\ &\text{Cum}^{(4)} \left[\begin{array}{c} \mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \\ \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{m,(t+1)\Delta} - \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \sum_{k=1}^{N_V} \text{Cum}^{(4)} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{j,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{j,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{j,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{l,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{l,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{l,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{m,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{m,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{m,k} \mathcal{B}_{k,t,t+1} \right) \end{array} \middle| \mathcal{F}_{t\Delta} \right], \\ &\text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \sum_{k=1}^{N_V} \text{Cum}^{(4)} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{j,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{j,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{j,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{l,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{l,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{l,k} \mathcal{B}_{k,t,t+1} \right), \\ (\mathbf{S}_{m,k} \mathcal{Z}_{k,t,t+1}) \end{array} \middle| \mathcal{F}_{t\Delta} \right], \\ &\text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right] \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=1}^{N_V} \text{Cum}^{(4)} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right), \\ \left(\mathbf{A}\mathbf{U}_{j,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{j,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{j,k} \mathcal{B}_{k,t,t+1} \right), \\ (\mathbf{S}_{l,k} \mathcal{Z}_{k,t,t+1}), (\mathbf{S}_{m,k} \mathcal{Z}_{k,t,t+1}) \end{array} \middle| \mathcal{F}_{t\Delta} \right], \\
 &\text{Cum}^{(4)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \\
 &= \sum_{k=1}^{N_V} \text{Cum}^{(4)} \left[\begin{array}{c} \left(\mathbf{A}\mathbf{U}_{i,k} \mathcal{Z}_{k,t,t+1} + (\mathbf{A}\mathbf{J} + \mathbf{C})_{i,k} \mathcal{W}_{k,t,t+1} + \mathbf{D}_{i,k} \mathcal{B}_{k,t,t+1} \right), \\ (\mathbf{S}_{j,k} \mathcal{Z}_{k,t,t+1}), (\mathbf{S}_{l,k} \mathcal{Z}_{k,t,t+1}), (\mathbf{S}_{m,k} \mathcal{Z}_{k,t,t+1}) \end{array} \middle| \mathcal{F}_{t\Delta} \right].
 \end{aligned}$$

B.4.2.4 Latent variable update

Corollary B.4.8.

$$\begin{aligned}
 &\mathbb{E} \left[\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{t\Delta}] \\
 &\quad - \text{Cov} [\text{Cov} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}], \text{Cov} [\epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] | \mathcal{G}_{t\Delta}] \\
 &\quad - \text{Cov} [\text{Cov} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}], \text{Cov} [\epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] | \mathcal{G}_{t\Delta}] \\
 &\quad - \text{Cov} [\text{Cov} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}], \text{Cov} [\epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] | \mathcal{G}_{t\Delta}].
 \end{aligned}$$

Proof. Since $\mathcal{G}_{t\Delta} \subseteq \mathcal{G}_{(t+1)\Delta}$, applying Lemma B.4.7 to the filtering error $\epsilon_{t+1|t+1}^{\mathbf{V}}$ fourth cumulants leads to the result as $\mathbb{E} [\epsilon_{t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] = 0$, therefore canceling all others terms in Lemma B.4.7. \square

We set the latent variable updated fourth cumulant to its expected value with respect to $\mathcal{G}_{t\Delta}$.

$$\begin{aligned}
 &\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \tag{B.12} \\
 &\simeq \mathbb{E} \left[\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Cum}^{(4)} \left[\begin{array}{c} \left(\epsilon_{i,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{i,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right), \left(\epsilon_{j,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{j,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right), \\ \left(\epsilon_{l,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{l,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right), \left(\epsilon_{m,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{m,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right],
 \end{aligned}$$

which is a linear combination of the second and fourth cumulant⁷ of $\mathbf{O}_{t+1\Delta}$ and fourth cumulant of $\mathbf{V}_{t+1\Delta}$ with respect to the information set $\mathcal{G}_{t\Delta}$ (see section B.4.2.3).

B.5 Adjustment for Intraday Frequency

We investigate two distinct assumptions regarding error measurement with respect to the model filtration. The first assumption entails a conditionally Gaussian error measurement, aligning with asymptotic theory (Assumption 3.4.1). The second assumption considers a conditional Gaussian distribution of the stochastic integrals at the intraday frequency (Assumption 3.4.2).

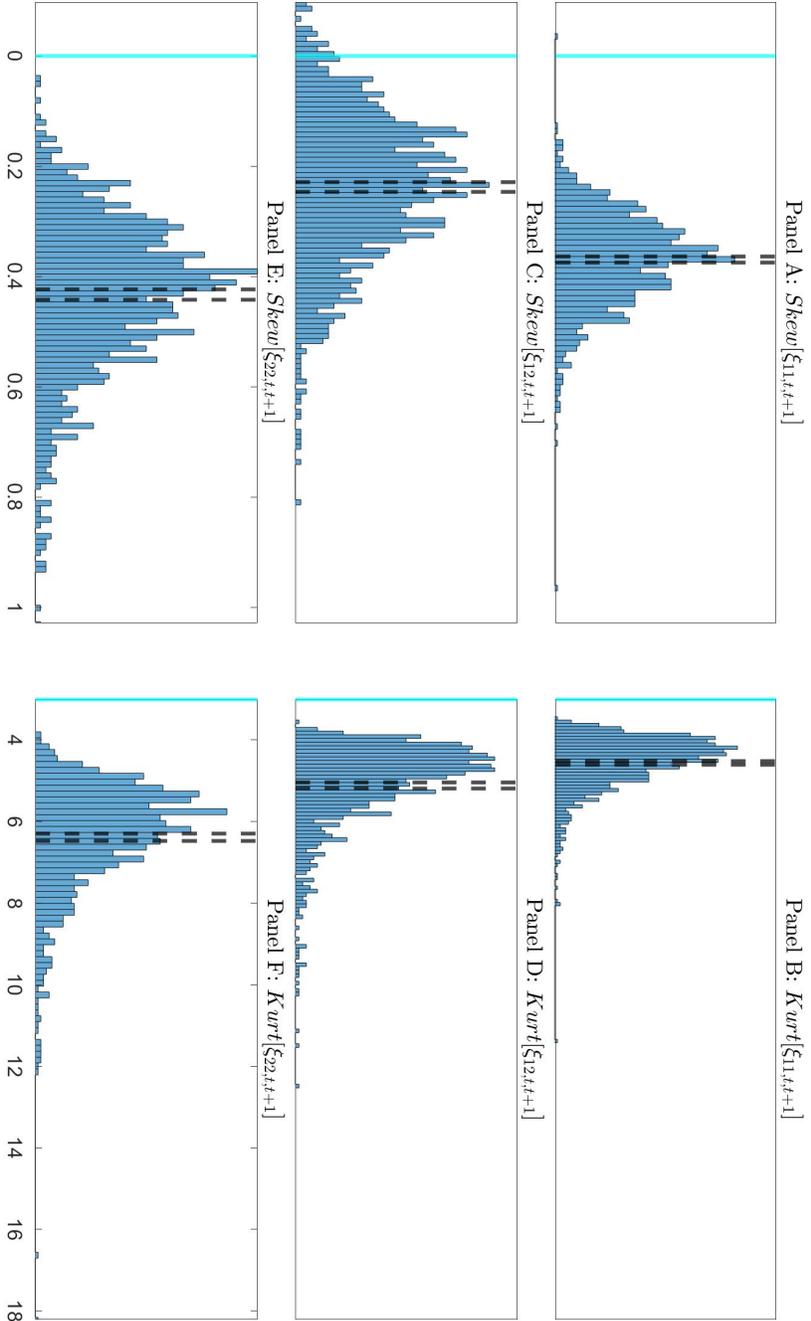
Under Assumption 3.4.1, the error measurement has the following properties:

Corollary B.5.1. *If Assumption 3.4.1 holds, for $\mathcal{X}_{(t+1)\Delta} \in \{\mathbf{O}_{(t+1)\Delta} - \mathbf{Q}\xi_{t,t+1}, \mathbf{V}_{(t+1)\Delta}\}$, then*

$$\begin{aligned} \text{Cum}^{(3)} [\xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\xi_{t,t+1}, \xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cum}^{(4)} [\xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0 \end{aligned}$$

Despite satisfying the first two conditional moments of the error measurement, the measurement error sample moments differ from what would be expected of a normally distributed random variable. Indeed, in Figure B.1, the sample moments of the simulated measurement error are not Gaussian, as the skewness and kurtosis differ from 0 and 3, respectively. Therefore, Assumption 3.4.1 is not valid if N is not large enough.

⁷If the additional step detailed in the Online Appendix D.8 is used, modifications are required to this results.



This Monte Carlo study generates $L = 1000$ independent trajectories with $M = 3900$ intraday time steps for $T = 4000$ days. The daily RV is aggregated at the 5 minutes frequency with initiate. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell,N)} = RCV_{i,j,t,t+1}^{(\ell,N)} - QCV_{i,j,t,t+1}^{(\ell)}$. The results are presented for case 2 assets and 3 variances. The vertical blue lines correspond to the skewness and kurtosis of a standard normal random variable. The black dashed lines are the 95%-confidence interval around the sample mean.

Figure B.1: Sample skewness and kurtosis of $\xi_{t,t+1}$

Consequently, while Assumption 3.4.1 holds asymptotically, it is invalid for smaller frequencies. Therefore, we propose a more flexible set of assumptions based on Euler approximation (see Assumption 3.4.2).

As N tends to infinity, Assumption 3.4.2 aligns with asymptotic theory.⁸ However, at intraday frequencies, we still utilize Assumption 3.4.2 to compute the distributional properties of measurement errors. Indeed, since the realized variance is the sum of N intraday squared returns, it converges to a Gaussian distribution as N tends to infinity. However, at intraday frequencies, it does not exhibit such properties. Therefore, if the third or fourth central moments are not employed in the filter or application, the additional computations under Assumption 3.4.2 are not useful.

We present the general results of the third central moment of the error measurement. The proofs are provided in the Online Appendix D.7.

$$\begin{aligned}
 & \text{Cum} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\
 & \approx \frac{\Delta^2}{N^2} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\hat{i},j,\tilde{j},\hat{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} V_{\hat{k},s} ds \mid \mathcal{F}_{t\Delta} \right], \\
 & \text{Cum} \left[R_{i,t,t+1}, R_{j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\
 & \approx \frac{\Delta}{N} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \mid \mathcal{F}_{t\Delta} \right], \\
 & \text{Cum} \left[V_{k,t+1}, R_{i,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\
 & \approx \frac{\Delta}{N} \sum_{\tilde{k}=1}^{N_V} \mathcal{H}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \mid \mathcal{F}_{t\Delta} \right], \\
 & \text{Cum} \left[V_{k,t+1}, V_{\hat{k},t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\
 & \approx \frac{\Delta}{N} \sum_{\tilde{k}=1}^{N_V} \mathcal{I}_{k,\tilde{k}}^{i,\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \mid \mathcal{F}_{t\Delta} \right], \\
 & \text{Cum} \left[V_{k,t+1}, QCV_{\hat{i},\hat{j},t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right]
 \end{aligned}$$

⁸Corollary D.7.8 shows that Equation (B.7) still holds under Assumption 3.4.2

$$\begin{aligned}
 &\approx \frac{\Delta}{N} \frac{\Delta}{2} \sum_{k=1}^{N_V} \mathcal{L}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\
 \text{Cum} [R_{i,t,t+1}, QCV_{i,\hat{j},t,t+1}, \xi_{i,\bar{j},t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\approx \frac{\Delta}{N} \frac{\Delta}{2} \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{M}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\
 \text{Cum} [QCV_{i,j,t,t+1}, QCV_{i,\hat{j},t,t+1}, \xi_{i,\bar{j},t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\approx \frac{\Delta}{N} \frac{\Delta^2}{3} \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{J}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\
 \text{Cum} [R_{i,t,t+1}, \xi_{i,\bar{j},t,t+1}, \xi_{i,\hat{j},t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\approx \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} (\sigma_k c_{i,k} \rho_k + \sigma_{\bar{k}} \rho_{\bar{k}} c_{i,\bar{k}}) \mathcal{C}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \frac{\Delta}{N} \frac{\Delta}{2} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\
 \text{Cum} [V_{k,(t+1)\Delta}, \xi_{i,\bar{j},t,t+1}, \xi_{i,\hat{j},t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\approx \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} (\sigma_k^2 + \sigma_{\bar{k}}^2) \mathcal{C}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \frac{\Delta}{N} \frac{\Delta}{2} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\
 \text{Cum} [QCV_{i,j,t,t+1}, \xi_{i,\bar{j},t,t+1}, \xi_{i,\hat{j},t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\approx \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} (\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\bar{k}}^2 c_{i,\bar{k}} c_{j,\bar{k}}) \mathcal{C}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{3} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right].
 \end{aligned}$$

where

$$\begin{aligned}
 \mathcal{S}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} &= c_{i,k} c_{i,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{j,\hat{k}} c_{j,\hat{k}} + c_{i,k} c_{i,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{j,\hat{k}} c_{j,\hat{k}} + c_{i,k} c_{j,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{i,\hat{k}} c_{j,\hat{k}} \\
 &\quad + c_{i,k} c_{j,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{i,\hat{k}} c_{j,\hat{k}} + c_{i,k} c_{i,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{j,\hat{k}} c_{j,\hat{k}} + c_{i,k} c_{i,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{j,\hat{k}} c_{j,\hat{k}} \\
 &\quad + c_{i,k} c_{j,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{i,\hat{k}} c_{j,\hat{k}} + c_{i,k} c_{j,k} c_{j,\bar{k}} c_{i,\bar{k}} c_{i,\hat{k}} c_{j,\hat{k}}, \\
 \mathcal{H}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} &= \sigma_k c_{i,k} c_{i,\bar{k}} c_{j,\bar{k}}^2 \rho_k, \\
 \mathcal{I}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} &= c_{i,\bar{k}}^2 c_{j,\bar{k}}^2 \sigma_k^2 \rho_k^2, \\
 \mathcal{L}_{k,\bar{k}}^{i,\hat{i},\bar{i},\hat{j}} &= \sigma_k^2 c_{i,k} c_{j,k} c_{i,\bar{k}} c_{j,\bar{k}}^2 \rho_k^2,
 \end{aligned}$$

$$\begin{aligned}\mathcal{M}_{k,\tilde{k}}^{i,\hat{i},j,\tilde{j}} &= \sigma_k c_{i,k} c_{i,k}^{\hat{i}} c_{j,k}^{\tilde{j}} c_{i,\tilde{k}}^{\hat{i}} c_{j,\tilde{k}}^{\tilde{j}} 2\rho_k, \\ \mathcal{J}_{k,\tilde{k}}^{i,\hat{i},j,\tilde{j}} &= \sigma_k^2 c_{i,k} c_{j,k} c_{i,k}^{\hat{i}} c_{j,k}^{\tilde{j}} c_{i,\tilde{k}}^{\hat{i}} c_{j,\tilde{k}}^{\tilde{j}} 2\rho_k^2.\end{aligned}$$

The third cumulant are a function of

$$\mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] \text{ and } \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right],$$

which implies a cubic and quadratic relationship with the latent state⁹ respectively.

For the fourth cumulant, we assume the following:

Assumption B.5.2. *The fourth cumulant of the measurement error is null and is independent of all other model components.*

Even though the kurtosis of the measurement error is different from 3, we saw in Section 3.4 that the bias for the fourth central moment is not statistically different from zero under Assumption 3.4.1, thus not justifying the complete derivation of the fourth cumulant of the measurement error.¹⁰

B.6 Autocorrelation of Filtering Errors

This Appendix assesses the influence of using the observed filtration in contrast to the model filtration in the filter derived in Section 3.3, resulting in an autocorrelation of the filtering error.

Based on Equation (3.10), we define the errors made on the variance with respect to the model and observed filtration, respectively, as follows:

$$\begin{aligned}\mathbf{e}_{t+1|t}^{\mathbf{V}} &= \mathbf{V}_{(t+1)\Delta} - \mathbb{E} [\mathbf{V}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \mathbf{S} \mathbf{Z}_{t,t+1} \quad (\text{Model Filtration}), \\ \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{V}} &= \mathbf{V}_{(t+1)\Delta} - \mathbb{E} [\mathbf{V}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] = \mathbf{K} \boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} + \mathbf{e}_{t+1|t}^{\mathbf{V}} \quad (\text{Observed Filtration}).\end{aligned}$$

⁹See Online Appendix D.3 for derivations.

¹⁰We leave that for future research.

Based on Equation (3.8) and (3.14), the errors made on the observation are defined as

$$\begin{aligned}
 \mathbf{e}_{t+1|t}^{\mathbf{O}} &= \mathbf{O}_{(t+1)\Delta} - \underbrace{(\mathbf{A}\mathbf{G} + \mathbf{A}\mathbf{H}\mathbf{V}_{t\Delta})}_{\mathbb{E}[\mathbf{O}_{(t+1)\Delta}|\mathcal{F}_{t\Delta}]} \quad (\text{Model Filtration}), \\
 &= \mathbf{A}(\mathbf{U}\mathbf{Z}_{t,t+1} + \mathbf{J}\mathbf{W}_{t,t+1}) + \mathbf{C}\mathbf{W}_{t,t+1} + \mathbf{D}\mathbf{B}_{t,t+1} + \mathbf{Q}\boldsymbol{\xi}_{t,t+1}, \\
 \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} &= \mathbf{O}_{(t+1)\Delta} - \mathbb{E}[\mathbf{O}_{(t+1)\Delta}|\mathcal{G}_{t\Delta}] = \mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} + \mathbf{e}_{t+1|t}^{\mathbf{O}} \quad (\text{Observed Filtration}).
 \end{aligned}$$

Therefore, $\boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{V}}$, the updated filtering error on our latent variables, can be expressed as follows:

$$\begin{aligned}
 \boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{V}} &\simeq \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta}\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \quad (\text{B.13}) \\
 &= \mathbf{e}_{t+1|t}^{\mathbf{V}} + \mathbf{K}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} - \Sigma_{t\Delta}(\mathbf{e}_{t+1|t}^{\mathbf{O}} + \mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}) \\
 &= \mathbf{e}_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{e}_{t+1|t}^{\mathbf{O}} + (\mathbf{K} - \Sigma_{t\Delta}\mathbf{A}\mathbf{H})\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}.
 \end{aligned}$$

Therefore, the filtering error $\boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{V}}$ is inherently linked to $\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}$.

We can further investigate the variance of the latent variable and disentangle the variance originating from the variance of $\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}$ and the variance implied by the stochastic integrals:

$$\begin{aligned}
 &\text{Var}[\boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{V}}|\mathcal{G}_{t\Delta}] \quad (\text{B.14}) \\
 &\simeq \mathbb{E}\left[\begin{pmatrix} \left(\mathbf{e}_{t+1|t}^{\mathbf{V}} + \mathbf{K}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} - \Sigma_{t\Delta}(\mathbf{e}_{t+1|t}^{\mathbf{O}} + \mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}})\right) \\ \left(\mathbf{e}_{t+1|t}^{\mathbf{V}} + \mathbf{K}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} - \Sigma_{t\Delta}(\mathbf{e}_{t+1|t}^{\mathbf{O}} + \mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}})\right)^{\top} \end{pmatrix} \middle| \mathcal{G}_{t\Delta}\right] \\
 &= \mathbb{E}\left[\begin{pmatrix} \left(\mathbf{e}_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{e}_{t+1|t}^{\mathbf{O}}\right) \left(\mathbf{e}_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{e}_{t+1|t}^{\mathbf{O}}\right)^{\top} \\ + \left(\mathbf{K}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}\right) \left(\mathbf{K}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{A}\mathbf{H}\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}\right)^{\top} \end{pmatrix} \middle| \mathcal{G}_{t\Delta}\right] \\
 &= \mathbb{E}\left[\underbrace{\text{Var}[\mathbf{e}_{t+1|t}^{\mathbf{V}} - \Sigma_{t\Delta}\mathbf{e}_{t+1|t}^{\mathbf{O}}|\mathcal{F}_{t\Delta}]}_{\text{Variance of the stochastic integrals}} \middle| \mathcal{G}_{t\Delta}\right] \\
 &\quad + \underbrace{(\mathbf{K} - \Sigma_{t\Delta}\mathbf{A}\mathbf{H})\text{Var}[\boldsymbol{\epsilon}_{t|t}^{\mathbf{V}}|\mathcal{G}_{t\Delta}]}_{\text{Variance due to last step uncertainty}}(\mathbf{K} - \Sigma_{t\Delta}\mathbf{A}\mathbf{H})^{\top},
 \end{aligned}$$

leading to autocorrelation in the variance of the filtering error.

Regarding the filtering errors studied in Section 3.4, the Equation (B.13) implies that $(\varepsilon_{k,t}^{(\ell,1)})^n$ is autocorrelated with $(\varepsilon_{k,t-1}^{(\ell,1)})^n$ for $n \in \{1, 2, 3, 4\}$. We then define $b_{k,t}^{(\ell,n)}$ as the autoregressive parameter with

$$b_{k,t}^{(\ell,n)} = (e^{-\kappa_k \Delta} - [\Sigma_{t\Delta} \mathbf{A} \mathbf{H}]_{k,k})^n. \quad (\text{B.15})$$

However, $b_{k,t}^{(\ell,n)}$ might not capture all autocorrelation in multivariate or higher-order moments, as the error term has a complex structure and can include other residual terms exhibiting autocorrelation.

Also, since, from Equation (B.14), the filtering error variance $\mathbb{E} \left[(\varepsilon_{k,t}^{(\ell,1)})^2 \middle| \mathcal{G}_{t\Delta} \right]$ has an autocorrelation parameter of $b_{k,t}^{(\ell,2)}$, resulting in an autocorrelation parameter for $\varepsilon_{k,t}^{(\ell,2)}$ of $b_{k,t}^{(\ell,2)}$. The same conclusions for central moment 3 and 4 can be obtained with Equations (3.28) and (3.29).

B.7 Panel Analysis

This Appendix describes the methodology employed to derive our confidence interval for the Monte Carlo simulation study outlined in Section 3.4.

B.7.1 Model

We have a panel model with a time axis (t), a path axis (l), and a third axis (m), which is a combination of the factor axis (k) and moment axis (n) with $m = nk$, $n \in 1, \dots, 4$ and $k \in 1, \dots, N_V$. Let

$$y_{l,m,t} = \underbrace{\vartheta_m}_{\text{fixed effect}} + \underbrace{u_{l,m}}_{\text{random effect}} + \underbrace{e_{l,m,t}}_{\text{noise}} = \vartheta_m + z_{l,m,t}, \quad (\text{B.16})$$

which is the equivalent to Equation (3.35) without any autoregressive components ($y_{l,m,t} = \varepsilon_{t,l}^{(k,n)}$). We have modified the notation to adapt it to panel analysis regression.

Given that Equation (3.35) does have an autoregressive components $b_{k,t}^{(\ell,n)}$ and it is **known**, not estimated, our panel analysis focuses then on the transformed variable:

$$y_{\ell,m,t} = \varepsilon_{k,t}^{(\ell,n)} - b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)}.$$

Our objective is to estimate the vector of fixed effects¹¹ $\hat{\boldsymbol{\vartheta}}$ ($M \times 1$) and its covariance matrix $\text{Cov}[\hat{\boldsymbol{\vartheta}}]$. The random effects $u_{l,m}$ captures the variability associated with a given path l . Because paths are simulated independently, there is no covariation in that axis. The dependence between bins m and m' is taken into account in the residuals term. Therefore,

$$\mathbb{E}[u_{l,m}] = 0 \quad \text{and} \quad \text{Cov}[u_{l,m}, u_{l',m'}] = \begin{cases} \sigma_m^2 & \text{if } l = l' \text{ and } m = m', \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.17})$$

The residuals $e_{l,m,t}$ have the following moments:

$$\mathbb{E}[e_{l,m,t}] = 0 \quad \text{and} \quad \text{Cov}[e_{l,m,t}, e_{l',m',t'}] = \begin{cases} \Gamma_{m,m'} & \text{if } l = l' \text{ and } t = t', \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.18})$$

The matrix $\boldsymbol{\Gamma} = [\Gamma_{m,m'}]_{m,m' \in \{1, \dots, M\}}$ is a $M \times M$ covariance matrix across bins. The structure in Equation (B.18) relies on a homoskedasticity and no serial correlation assumption, which we will soon relax. In addition, we define $z_{l,m,t} = u_{l,m} + e_{l,m,t}$ as the combined error: the random effect plus the residual, that is,

$$\begin{aligned} \mathbb{E}[z_{l,m,t}] &= 0 \\ \text{Cov}[z_{l,m,t}, z_{l',m',t'}] &= \begin{cases} \sigma_m^2 \mathbb{1}_{m=m'} + \Gamma_{m,m'} & \text{if } l = l' \text{ and } t = t', \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (\text{B.19})$$

¹¹Even though Vogelsang, 2012 discusses how to estimate parameters in panel analysis with fixed effects using Driscoll and Kraay, 1998 with a complex error structure, it does not apply to our case as they estimate the coefficient of the regressor, not the fixed effect parameters.

B.7.2 Covariance estimates

We have two distinct groups of variance parameters to estimate: the variance of the random effect σ_m^2 for $m = 1, \dots, M$, and the nuisance covariance matrix $\Gamma_{m,m'}$.

The first regression is Equation (B.16). Because $E[y_{l,m,t}] = \vartheta_m$,

$$\hat{\vartheta}_m = \bar{y}_m = \frac{1}{TL} \sum_{t=1}^T \sum_{l=1}^L y_{l,m,t} \quad (\text{B.20})$$

is a consistent estimator of ϑ_m . Define the residual $\hat{z}_{l,m,t} = y_{l,m,t} - \hat{\vartheta}_m$, and note that

$$\hat{\bar{z}}_m = \frac{1}{LT} \sum_l \sum_t \underbrace{y_{l,m,t} - \bar{y}_m}_{\hat{z}_{l,m,t}} = 0.$$

The sample covariance

$$\frac{1}{LT-1} \sum_l \sum_t \hat{z}_{l,m,t} \hat{z}_{l,m',t}, \quad (\text{B.21})$$

is a consistent estimator of

$$\text{Cov}[z_{l,m,t}, z_{l,m',t}] = \text{Cov}[u_{l,m}, u_{l,m'}] + \text{Cov}[e_{l,m,t}, e_{l,m',t}] = \sigma_m^2 \mathbb{1}_{m=m'} + \Gamma_{m,m'}.$$

The second regression is

$$y_{l,m,t} - \bar{y}_{l,m} = \vartheta_m + u_{l,m} + e_{l,m,t} - (\vartheta_m + u_{l,m} + \bar{e}_{l,m}) = e_{l,m,t} - \bar{e}_{l,m},$$

where $\bar{y}_{l,m} = \frac{1}{T} \sum_{t=1}^T y_{l,m,t}$ and $\bar{e}_{l,m} = \frac{1}{T} \sum_{t=1}^T e_{l,m,t}$. The sample covariance

$$\frac{1}{L(T-1)} \sum_l \sum_t (y_{l,m,t} - \bar{y}_{l,m})(y_{l,m',t} - \bar{y}_{l,m'}), \quad m, m' \in \{1, \dots, M\} \quad (\text{B.22})$$

is a consistent estimator of $\Gamma_{m,m'}$.

Therefore, the estimator of σ_m^2 is given by

$$\frac{\sum_l \sum_t (y_{l,m,t} - \bar{y}_m)(y_{l,m',t} - \bar{y}_{m'})}{LT - 1} - \frac{\sum_l \sum_t (y_{l,m,t} - \bar{y}_{l,m})(y_{l,m',t} - \bar{y}_{l,m'})}{L(T - 1)}. \quad (\text{B.23})$$

Even though other estimates are possible,¹² Equations (B.21) and (B.22) are used to estimate the variances. However, $\hat{\sigma}_m^2$ is not guaranteed to be positive. Therefore, we do as suggested by Greene, 2017 and Maddala, 1971: if a negative value arises, we set $\hat{\sigma}_m^2 = 0$ and use $\hat{z}_{l,m,t}$ in Equation (B.22) instead of $(y_{l,m,t} - \bar{y}_{l,m})$ for the bin m .

The covariance of $\hat{\vartheta}_m$ is then

$$\begin{aligned} \text{Cov} \left[\hat{\vartheta}_m, \hat{\vartheta}_{m'} \right] &= \text{Cov} \left[\frac{1}{LT} \sum_t \sum_l y_{l,m,t}, \frac{1}{LT} \sum_{t'} \sum_{l'} y_{l',m',t'} \right] \\ &= \frac{1}{(LT)^2} \text{Cov} \left[\sum_t \sum_l y_{l,m,t}, \sum_{t'} \sum_{l'} y_{l',m',t'} \right] \\ &= \frac{1}{(LT)^2} \sum_l \text{Cov} \left[\sum_t y_{l,m,t}, \sum_{t'} y_{l,m',t'} \right] \quad \text{Independence of trajectories} \\ &= \frac{1}{(LT)^2} \sum_l \left(T^2 \hat{\sigma}_m^2 \mathbb{1}_{m=m'} + T \hat{\Gamma}_{m,m'} \right) \\ &= \frac{1}{LT} \left(T \hat{\sigma}_m^2 \mathbb{1}_{m=m'} + \hat{\Gamma}_{m,m'} \right). \end{aligned}$$

B.7.3 Heteroskedastic errors with serial and spatial autocorrelation

The assumption of homoskedasticity is now relaxed for the residual $e_{l,m,t}$, as it includes noise with serial (t axis) and spatial (m axis) autocorrelation along with heteroskedasticity. Furthermore, the random effect also exhibits heteroskedasticity, as stated in Equation

¹²The homoskedastic case without spatial correlation is covered in Maddala, 1971. Baltagi, 2021 also proposes another way to isolate the variance parameters in section 5.1 and discuss the case with heteroskedasticity with no serial correlation. Greene, 2017 also proposed two methodologies. See <https://online.stat.psu.edu/stat503/lesson/14/14.1> as a reference for two-stage nested design model with either fixed or random effect.

(B.17). Therefore, $\mathbf{\Gamma}_{m,m'}$ is a $T \times T$ matrix where

$$\text{Cov}[e_{l,m,t}, e_{l',m',t'}] = \begin{cases} [\mathbf{\Gamma}_{m,m'}]_{t,t'} & \text{if } l = l', \\ 0 & \text{otherwise.} \end{cases}$$

We estimate $[\mathbf{\Gamma}_{m,m'}]_{t,t'}$ by $[\hat{\mathbf{\Gamma}}_{l,m,m'}]_{t,t'}$ for each trajectory l and we follow the Newey and West, 1987 estimator with Bartlett weights and bandwidth Q :

$$\begin{aligned} & [\hat{\mathbf{\Gamma}}_{l,m,m'}]_{t,t'} && \text{(B.24)} \\ & = \begin{cases} \frac{T}{T-1} \left(1 - \frac{|t-t'|}{Q+1}\right) (y_{l,m,t} - \bar{y}_{l,m})(y_{l,m',t'} - \bar{y}_{l,m'}) & \text{if } |t-t'| \leq Q, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The factor $\frac{T}{T-1}$ is to adjust for degrees of freedom as suggested in Andrews, 1991. The random effect parameter is estimated with

$$\hat{\sigma}_m^2 + \frac{\sum_l \sum_t (y_{l,m,t} - \bar{y}_{l,m})^2}{L(T-1)} = \frac{\sum_l \sum_t \hat{z}_{l,m,t}^2}{TL-1},$$

where $\hat{\sigma}_m^2$ is isolated as in Equation (B.23). If $\hat{\sigma}_m^2$ is negative, Equation (B.24) is calculated with $\hat{z}_{l,m,t}$ instead.

The covariance matrix of the vector $\hat{\boldsymbol{\vartheta}} = [\hat{\vartheta}_1, \dots, \hat{\vartheta}_M]$, denoted $\hat{\mathbf{\Omega}}$, is then

$$\begin{aligned} [\hat{\mathbf{\Omega}}]_{m,m'} &= \text{Cov} \left[\hat{\vartheta}_m, \hat{\vartheta}_{m'} \right] = \frac{1}{(LT)^2} \sum_l \text{Cov} \left[\sum_t y_{l,m,t}, \sum_{t'} y_{l,m',t'} \right] \\ &= \frac{1}{(LT)^2} \sum_l \left(T^2 \hat{\sigma}_m^2 \mathbb{1}_{m=m'} + \mathbf{i}_T \hat{\mathbf{\Gamma}}_{l,m,m'} \mathbf{i}_T^\top \right), \end{aligned}$$

where \mathbf{i}_T is a $T \times 1$ column vector of ones.

We might consider conducting individual estimations and statistical tests for each m bin. Nevertheless, due to interrelations among the M equations, Zellner, 1962 argue that it's more effective to estimate the entire system, incorporating spatial correlation.¹³ There-

¹³He elaborates on the estimation procedure in situations where there's no serial correlation and consistent

fore, a global test can be made on the vector $\hat{\boldsymbol{\vartheta}} = [\hat{\vartheta}_1, \dots, \hat{\vartheta}_M]$, where

$$\hat{\boldsymbol{\vartheta}}^\top \hat{\boldsymbol{\Omega}}^{-1} \hat{\boldsymbol{\vartheta}} \sim \chi_M.$$

We can also present individual m bin test based on

$$\frac{\hat{\vartheta}_m}{\sqrt{[\hat{\boldsymbol{\Omega}}]_{m,m}}} \sim \mathcal{N}(0, 1).$$

variance within each group; other explanations can be found in Fiebig, 2003.

References

- Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3), 817–858.
- Baltagi, B. H. (2021). *Econometric Analysis of Panel Data (6th ed)*. Springer.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004a). A Feasible Central Limit Theory for Realised Volatility Under Leverage.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004b). Econometric Analysis of Realized Covariation: High Frequency based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, 72(3), 885–925.
- Brillinger, D. R. (1969). The Calculation of Cumulants via Conditioning. *Annals of the Institute of Statistical Mathematics*, 21(1), 215–218.
- Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4), 549–560.
- Fiebig, D. G. (2003). *A Companion to Theoretical Econometrics*. John Wiley & Sons, Ltd.
- Greene, W. (2017). *Econometric Analysis (8th ed)*. Pearson.
- Maddala, G. S. (1971). The Use of Variance Components Models in Pooling Cross Section and Time Series Data. *Econometrica: Journal of the Econometric Society*, 341–358.
- Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708.

Vogelsang, T. J. (2012). Heteroskedasticity, Autocorrelation, and Spatial Correlation Robust Inference in Linear Panel Models with Fixed-Effects. *Journal of Econometrics*, 166(2), 303–319.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American statistical Association*, 57(298), 348–368.

Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data. *Journal of the American Statistical Association*, 100(472), 1394–1411.

Chapter C

Appendices of *Enhancing Parameter Estimation in Stochastic Volatility Models Using Intraday Data: Generalized Method of Moments Comparative Analysis*

C.1 Measurement Error Variance

The goal is to derive $E \left[\int_0^{t\Delta} \left(\sum_{k=1}^K V_{k,s} \right)^2 ds \right]$ (Corollary C.1.3), which is a key quantity for computing the conditional and unconditional measurement error variances in Equations (4.9) and (4.10).

Corollary C.1.1.

$$E \left[\int_0^{t\Delta} V_{k,s}^2 ds \right] = \left(\frac{\sigma_k^2 \theta_k}{2\kappa_k} + \theta_k^2 \right) t\Delta.$$

Proof. From Fubini's theorem

$$\begin{aligned} E \left[\int_0^{t\Delta} V_{k,s}^2 ds \right] &= \int_0^{t\Delta} E [V_{k,s}^2] ds = \int_0^{t\Delta} ((E[V_{k,s}])^2 + \text{Var}[V_{k,s}]) ds \\ &= \int_0^{t\Delta} \left(\frac{\sigma_k^2 \theta_k}{2\kappa_k} + \theta_k^2 \right) ds = \left(\frac{\sigma_k^2 \theta_k}{2\kappa_k} + \theta_k^2 \right) t\Delta. \quad \square \end{aligned}$$

Corollary C.1.2. Let $V_{k,t\Delta}$ and $V_{\bar{k},t\Delta}$ follow a CIR process with independent Brownian motions, then

$$\mathbb{E} \left[\int_0^{t\Delta} V_{k,s} V_{\bar{k},s} ds \right] = \theta_{\bar{k}} \theta_k t \Delta.$$

Proof. From Fubini's theorem,

$$\begin{aligned} \mathbb{E} \left[\int_0^{t\Delta} V_{k,s} V_{\bar{k},s} ds \right] &= \int_0^{t\Delta} \mathbb{E} [V_{k,s} V_{\bar{k},s}] ds = \int_0^{t\Delta} \mathbb{E} [V_{k,s}] \mathbb{E} [V_{\bar{k},s}] ds \\ &= \int_0^{t\Delta} \theta_{\bar{k}} \theta_k ds = \theta_{\bar{k}} \theta_k t \Delta. \quad \square \end{aligned}$$

Corollary C.1.3.

$$\mathbb{E} \left[\int_0^{t\Delta} \left(\sum_{k=1}^K V_{k,s} \right)^2 ds \right] = \left(\sum_{k=1}^K \theta_k \right)^2 t \Delta + \sum_{k=1}^K \frac{\sigma_k^2 \theta_k}{2\kappa_k} t \Delta.$$

Proof. See Corollaries C.1.1 and C.1.2. □

C.2 Measurement Error Empirical Property

This Appendix discusses the model assumptions regarding the measurement error, the assumptions used for the moment calculation for Appendix C.3.1 and the impact of model specification with the instrumental variables used in the GMM moment selection. We proceed with the Monte Carlo simulation detailed in Section 4.3 using the parameters¹ of Scenario A from Bollerslev and Zhou, 2002.

¹The results for Scenario B and C leads to the same conclusion. They are available upon request.

C.2.1 Model assumptions regarding the measurement error

For specifications (4.14) and (4.16), we assume a null conditional expectation and two specifications for the measurement error conditional variance. That is,

$$\mathbb{E} [\xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad \text{Var} [\xi_{t,t+1} | \mathcal{F}_{t\Delta}] = \gamma^2 \quad \text{and} \quad \text{Var} [\xi_{t,t+1} | \mathcal{F}_{t\Delta}] = \eta^2 \mathbb{E} [RQ_{t,t+1} | \mathcal{F}_{t\Delta}].$$

Since the measurement error is not simulated based on the models, but rather calculated with Equation (4.13), we must evaluate whether the assumptions for the measurement error are valid.

For each trajectory, we compute the averaged error $\frac{1}{T} \sum_{t=1}^T \xi_{t,t+1}$, the squared error under both specifications, $\frac{1}{T} \sum_{t=1}^T \xi_{t,t+1}^2 - \gamma^2$ and $\frac{1}{T} \sum_{t=1}^T \xi_{t,t+1}^2 - \eta^2 RQ_{t,t+1}$. The results are shown in Figure C.1. As anticipated by the asymptotic theory, the assumption that the measurement error is centered at zero (Panel A) appears to be reasonable. Panels B and C present the second moment of $\xi_{t,t+1}$ under both specifications. Panel B corresponds to specification (4.14) presented by BZ, while Panel C relates to specification (4.16), which is supported by the asymptotic theory. Even though both Panels B and C are centered around zero, specification (4.16)'s panel shows significantly less variation between the squared error and its conditional assumed variance, providing an initial indication that specification (4.16) is superior to specification (4.14).

C.2.2 Covariance with model components

To obtain the results of Appendix C.3.1, we rely on three assumptions detailed in the Online Appendix E.6.1, which are

$$\mathbb{E} [\mathcal{V}_{t,t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad \mathbb{E} [V_{t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0 \quad \text{and} \quad \mathbb{E} [R_{t,t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0. \quad (\text{C.1})$$

Furthermore, in Appendix C.4, we also use assumptions based on unconditional expectations:

$$\mathbb{E} [\mathcal{V}_{k,t,t+1} \xi_{t,t+1}] = 0, \quad \mathbb{E} [V_{k,(t+1)\Delta} \xi_{t,t+1}] = 0 \quad \text{and} \quad \mathbb{E} [R_{t,t+1} \xi_{t,t+1}] = 0. \quad (\text{C.2})$$

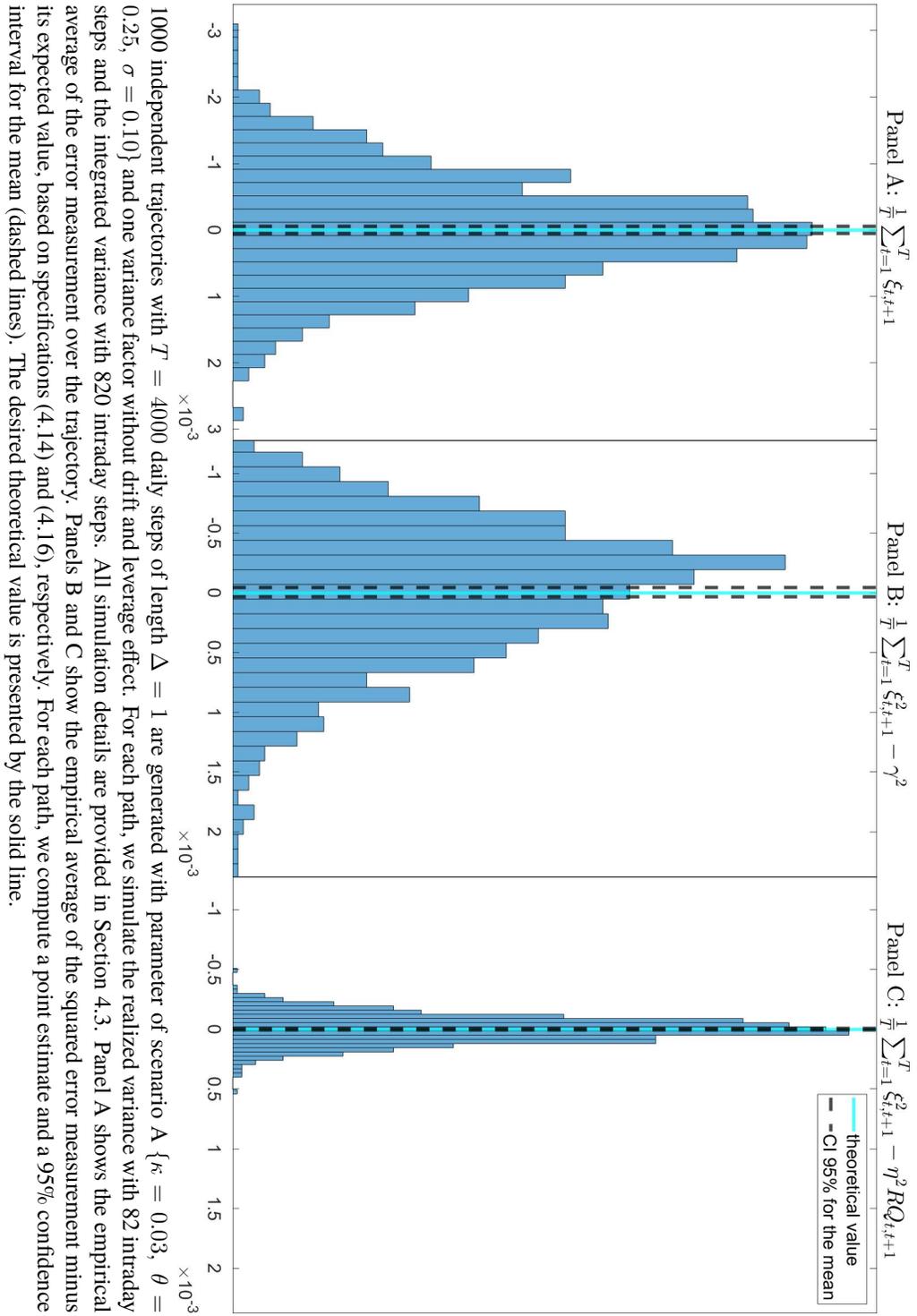


Figure C.1: Empirical distribution of $\xi_{t,t+1}$ with parameter of scenario A

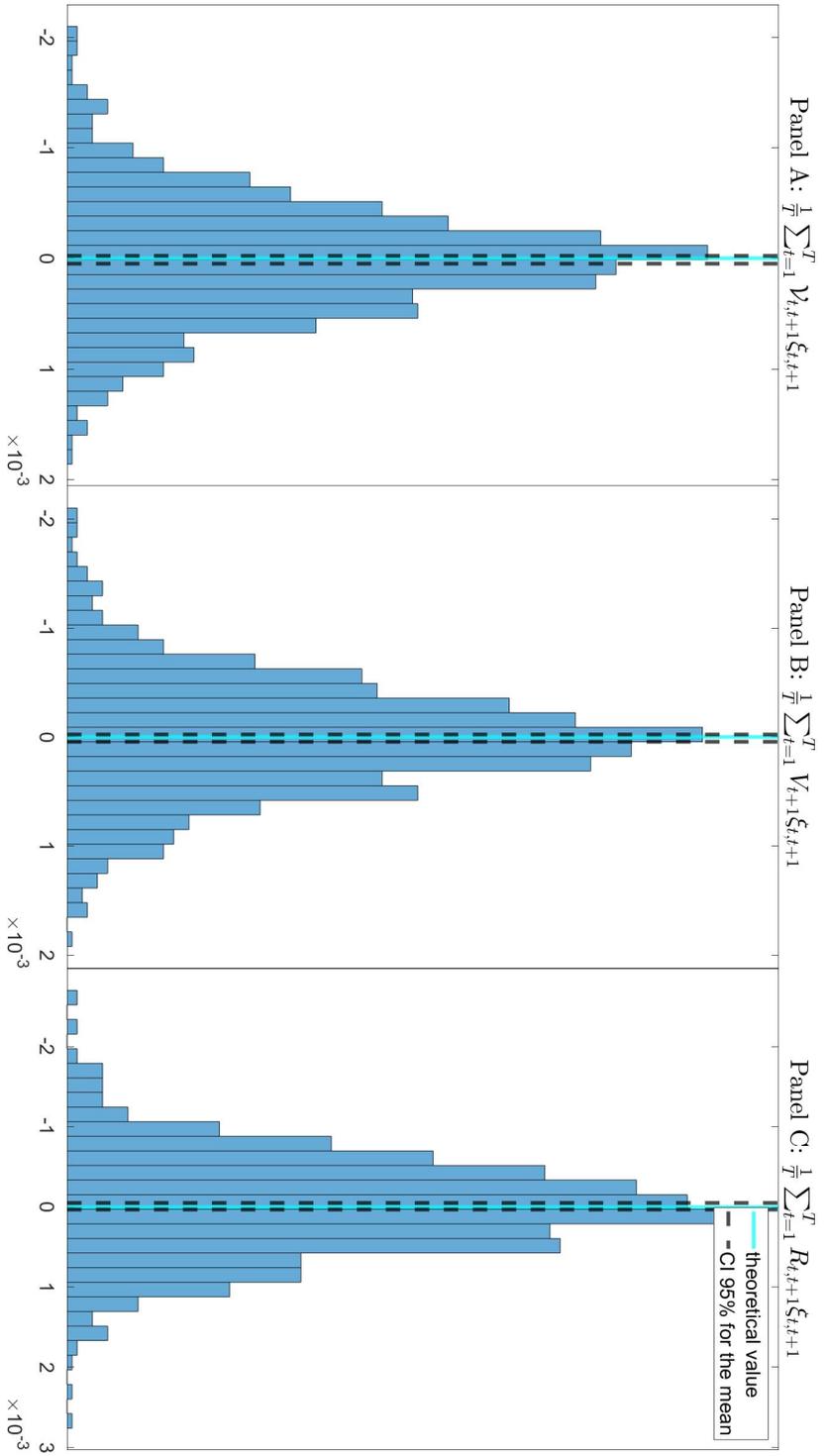
Figure C.2 shows numerical evidence that Assumptions (C.2) hold, as the sample average for all 1000 trajectories are centered at zero for all three panels. While we numerically justify Assumptions (C.2), it does not directly imply (C.1). However, based on the results of Figure C.2, we assume that Assumptions (C.1) are satisfied.

C.2.3 Interaction with instrumental variables

The GMM moment selection described in Sections 4.4 and 4.5 assumed that the measurement error $\xi_{t,t+1}$ is not correlated with the instrumental variables $RV_{t-1,t}$ and $RV_{t-1,t}^2$, thereafter $RV_{t-1,t}^p$. However, assuming that the variance of $\xi_{t,t+1}$ is constant over time can impact the orthogonality of $\xi_{t,t+1}^2$ with $RV_{t-1,t}^p$ via the discrepancy between the time-varying condition variance $\text{Var}[\xi_{t,t+1} | \mathcal{F}_{t\Delta}]$ and the constant γ^2 .

Panels A.1 and A.2 of Figure C.3 show numerically that the correlations between the measurement error $\xi_{t,t+1}$ and the instrumental variables $RV_{t-1,t}$ and $RV_{t-1,t}^2$ are centered at zero.

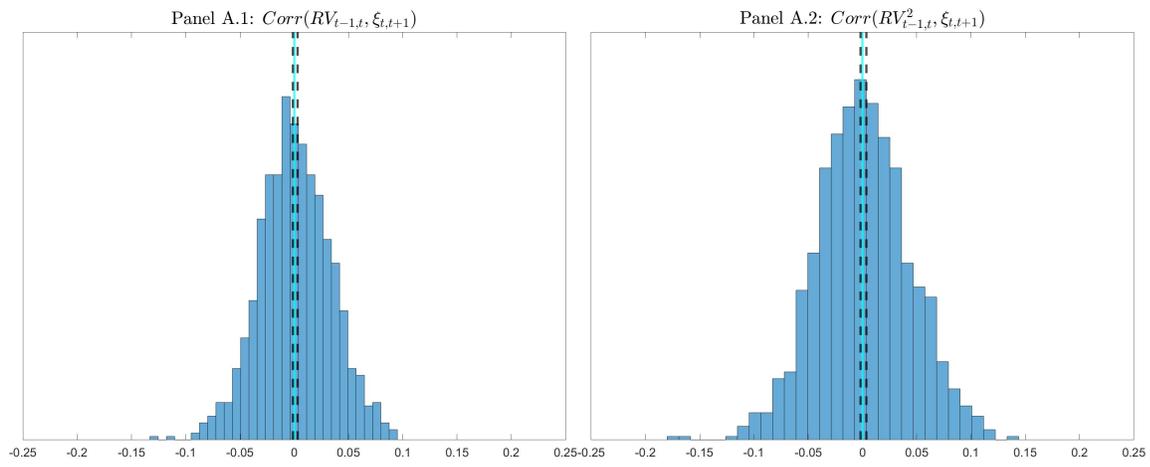
Panels B.1 and B.2 illustrate that when we don't acknowledge for the time-varying variance of the quadratic variation measurement error, then the residual $(\xi_{t,t+1}^2 - \gamma^2)$ is correlated with $RV_{t-1,t}$ and $RV_{t-1,t}^2$. However, when the conditional variance of $\xi_{t,t+1}$ is measured according to the asymptotic theory, the correlation between $(\xi_{t,t+1}^2 - \eta^2 RQ_{t,t+1})$ and the instrument variables are centered around zero (Figure C.3, Panels C.1 and C.2).



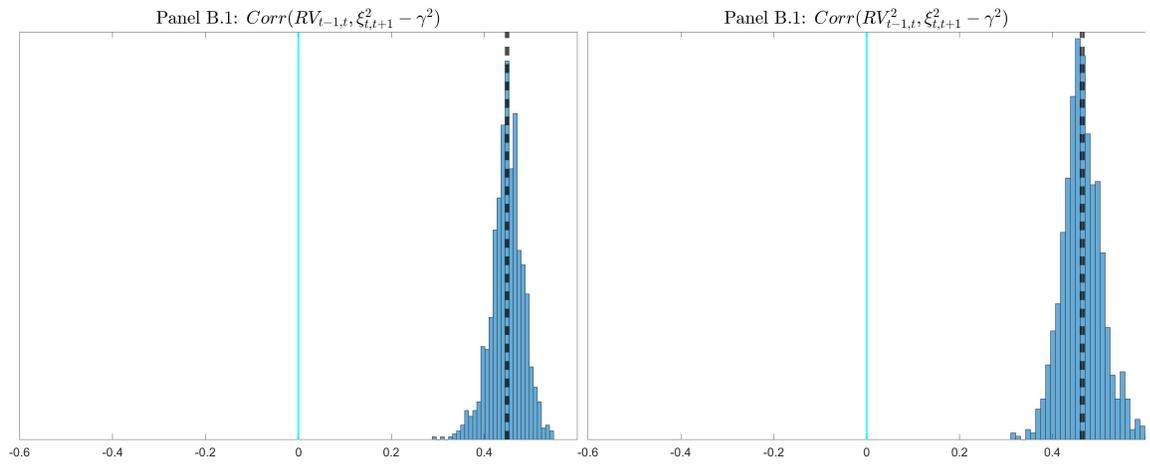
1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Panel A is the sample average of the error measurement multiplied by the integrated variance, Panel B is the sample average of the error measurement multiplied by the instantaneous variance and Panel C is the error measurement multiplied by the return. For each path, we compute a point estimate and a 95% confidence interval for the mean (dashed lines). The desired theoretical value is presented by the solid line.

Figure C.2: Sample correlation between $\xi_{t,t+1}$, the instantaneous, integrated variance and the returns with parameter of scenario A

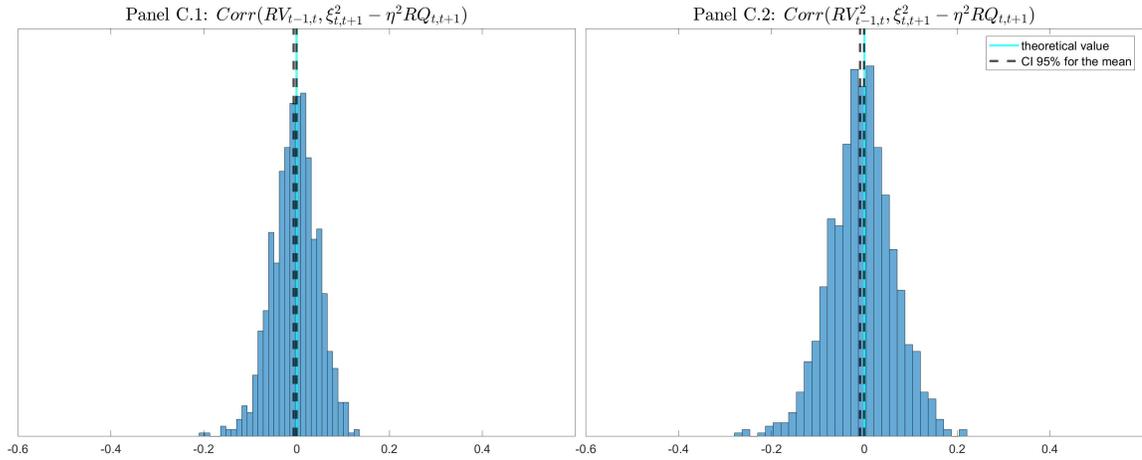
Panel A: measurement error and instrumental



Panel B: Squared measurement error and instrumental variable with specification (4.14)



Panel C: Squared measurement error and instrumental variable with specification (4.16)



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Panels A are the empirical correlation between measurement error $\xi_{t,t+1}$ and the instrumental variable $RV_{t-1,t}$ and $RV_{t-1,t}^2$ over each trajectory. Panels B are the empirical correlation between $\xi_{t,t+1}^2 - \gamma^2$ and the instrumental variable $RV_{t-1,t}$ and $RV_{t-1,t}^2$ over each trajectory. Panels C are the empirical correlation between $\xi_{t,t+1}^2 - \eta^2 RQ_{t,t+1}$ and the instrumental variable $RV_{t-1,t}$ and $RV_{t-1,t}^2$ over each trajectory. For each path, we compute a point estimate and a 95% confidence interval for the mean (dashed lines). The desired theoretical value is presented by the solid line.

Figure C.3: Correlation between $\xi_{t,t+1}$ and instrumental variables

C.3 Conditional Moment Derivation

C.3.1 One-variance factor

This Appendix summarizes the conditional moments related to the instantaneous variance, the integrated variance and the return. The subscript k is omitted for readability. We denote $\mathcal{V}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} V_s ds$, $\mathcal{W}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dW_s$ and $\mathcal{B}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dB_s$. Complete derivations are available in the Online Appendices E.4 and E.6.1.

From Corollaries E.4.3 and E.4.4, the first moment of the instantaneous variance and integrated variance are

$$\mathbb{E} [V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \underbrace{e^{-\kappa\Delta}}_{\alpha} V_{t\Delta} + \underbrace{\theta(1 - e^{-\kappa\Delta})}_{\beta} \approx (1 - \kappa\Delta) V_{t\Delta} + \kappa\theta\Delta, \quad (\text{C.3})$$

$$\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] = V_{t\Delta} \underbrace{\left(\frac{1 - e^{-\kappa\Delta}}{\kappa}\right)}_a + \underbrace{\theta\Delta - \theta \frac{(1 - e^{-\kappa\Delta})}{\kappa}}_b \approx \Delta V_{t\Delta} + \frac{1}{2} \kappa\theta\Delta^2. \quad (\text{C.4})$$

Their respective variances, detailed in Corollaries E.4.9 and E.4.10, are

$$\text{Var} [V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \underbrace{\frac{\sigma^2}{\kappa} (e^{-\kappa\Delta} - e^{-2\kappa\Delta}) V_{t\Delta}}_C + \underbrace{\frac{\sigma^2\theta}{2\kappa} (1 - e^{-\kappa\Delta})^2}_D \quad (\text{C.5})$$

$$\approx \sigma^2 \Delta V_{t\Delta} + \frac{\sigma^2 \kappa \theta}{2} \Delta^2,$$

$$\text{Var} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] = \underbrace{\frac{\sigma^2}{\kappa^2} \left(\frac{1 - e^{-2\kappa\Delta}}{\kappa} - 2e^{-\kappa\Delta} \Delta \right)}_A V_{t\Delta} \quad (\text{C.6})$$

$$+ \underbrace{\frac{\sigma^2\theta}{\kappa^2} \left(\Delta - \frac{1 - e^{-2\kappa\Delta}}{2\kappa} - 2 \frac{1 - e^{-\kappa\Delta}}{\kappa} + 2e^{-\kappa\Delta} \Delta \right)}_B$$

$$\approx \frac{1}{3} \sigma^2 \Delta^3 V_{t\Delta} + \frac{\sigma^2 \kappa \theta}{12} \Delta^4.$$

Finally, from Corollaries E.4.8, E.4.11 and E.4.12, the conditional covariances are

$$\text{Cov} [\mathcal{V}_{t,t+1}, V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \underbrace{\frac{\sigma^2}{\kappa} e^{-\kappa\Delta} \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right)}_F V_{t\Delta} \quad (\text{C.7})$$

$$+ \underbrace{\theta \frac{\sigma^2}{\kappa} \left(\frac{(1 - e^{-\kappa\Delta})^2}{2\kappa} - e^{-\kappa\Delta} \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right) \right)}_G$$

$$\approx \frac{\sigma^2}{2} \Delta^2 V_{t\Delta} + \frac{\sigma^2}{6} \kappa \theta \Delta^3,$$

$$\text{Cov} [\mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = \underbrace{\frac{\sigma}{\kappa} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right)}_{\bar{F}} V_{t\Delta} \quad (\text{C.8})$$

$$+ \underbrace{\frac{\sigma}{\kappa} \theta \left(\Delta - 2 \frac{1 - e^{-\kappa\Delta}}{\kappa} + e^{-\kappa\Delta} \Delta \right)}_{\tilde{G}}$$

$$\approx \frac{\sigma}{2} \Delta^2 V_{t\Delta} + \frac{\kappa \sigma \theta}{6} \Delta^3,$$

$$\text{Cov} [\mathcal{W}_{t,t+1}, V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \underbrace{\sigma e^{-\kappa\Delta} \Delta}_{\bar{F}} V_{t\Delta} + \underbrace{\sigma \theta \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right)}_{\bar{G}}. \quad (\text{C.9})$$

The recursive relationships between conditional moments are

$$\text{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] = \underbrace{e^{-\kappa\Delta}}_{\alpha} \text{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \underbrace{(1 - e^{-\kappa\Delta}) \theta \Delta}_{\beta \Delta} \quad (\text{C.10})$$

$$\approx (1 - \kappa \Delta) \text{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + (\kappa \theta \Delta^2),$$

$$\text{E} [\mathcal{V}_{t+1,t+2}^2 | \mathcal{F}_{t\Delta}] = \underbrace{\alpha^2}_{\bar{H}} \text{E} [\mathcal{V}_{t\Delta, (t+1)\Delta}^2 | \mathcal{F}_{t\Delta}] \quad (\text{C.11})$$

$$+ \underbrace{\frac{1}{a} (a^2 (C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A))}_{\bar{I}} \text{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]$$

$$\begin{aligned}
 & + \underbrace{-bI + (a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B))}_J \\
 & \approx (1 - 2\kappa\Delta) \mathbb{E} [\mathcal{V}_{t\Delta, (t+1)\Delta}^2 | \mathcal{F}_{t\Delta}] \\
 & + (\sigma^2 + 2\kappa\theta) \Delta^2 \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] + \left(\kappa^2\theta^2 + \frac{\sigma^2\kappa\theta}{3} \right) \Delta^4, \\
 \mathbb{E} [\mathcal{V}_{t+1, t+2} \mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] & = \underbrace{\alpha}_{\tilde{H}} \mathbb{E} [\mathcal{V}_{t\Delta, (t+1)\Delta}^2 | \mathcal{F}_{t\Delta}] \tag{C.12} \\
 & + \underbrace{\frac{1}{a}(a(F + b + ab + a\beta) - \alpha(2ab + A))}_{\tilde{I}} \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] \\
 & + \underbrace{-b\tilde{I} + (aG + b^2 + ab\beta - \alpha(B + b^2))}_{\tilde{J}} \\
 & \approx (1 - \kappa\Delta) \mathbb{E} [\mathcal{V}_{t, t+1}^2 | \mathcal{F}_{t\Delta}] \\
 & + \left(\frac{\sigma^2}{2} + 2\kappa\theta \right) \Delta^2 \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] - \left(\frac{\sigma^2}{6} \kappa\theta + \frac{\kappa^2\theta^2}{2} \right) \Delta^4, \\
 \mathbb{E} [R_{t, t+1} | \mathcal{F}_{t\Delta}] & = \left(\lambda - \frac{1}{2} \right) \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] + r\Delta, \\
 \mathbb{E} [R_{t, t+1} \mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] & = \underbrace{\left(\lambda - \frac{1}{2} \right)}_{\hat{H}} \mathbb{E} [\mathcal{V}_{t, t+1}^2 | \mathcal{F}_{t\Delta}] \\
 & + \underbrace{\left(\frac{\rho\tilde{F}}{a} + r\Delta \right)}_{\hat{J}} \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] + \underbrace{\left(\rho\tilde{G} - \rho\tilde{F} \frac{b}{a} \right)}_{\hat{I}}, \\
 \mathbb{E} [R_{t, t+1} \mathcal{V}_{t+1, t+2} | \mathcal{F}_{t\Delta}] & = \mathbb{E} \left[\left(r\Delta + \left(\lambda - \frac{1}{2} \right) \mathcal{V}_{t, t+1} \right) \mathcal{V}_{t+1, t+2} \middle| \mathcal{F}_{t\Delta} \right] \\
 & + \underbrace{\rho\tilde{F}}_{\tilde{I}} \mathbb{E} [\mathcal{V}_{t, t+1} | \mathcal{F}_{t\Delta}] + \underbrace{a\rho\tilde{G} - b\rho\tilde{F}}_{\tilde{J}}.
 \end{aligned}$$

When the integrated variance is replaced with the observable realized variance, we have:

$$\mathbb{E} [RV_{t+1, t+2} | \mathcal{G}_{t\Delta}] - \alpha \mathbb{E} [RV_{t, t+1} | \mathcal{G}_{t\Delta}] - \Delta\beta = 0,$$

$$\begin{aligned}
 \mathbb{E} [RV_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] - H (\mathbb{E} [RV_{t,t+1}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{G}_{t\Delta}]) \\
 - I \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] - J = 0, \\
 \mathbb{E} [RV_{t+1,t+2} RV_{t,t+1} | \mathcal{G}_{t\Delta}] - \tilde{H} (\mathbb{E} [RV_{t,t+1}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{G}_{t\Delta}]) \\
 - \tilde{I} \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] - \tilde{J} = 0.
 \end{aligned}$$

C.3.2 Two-variance factor

This Appendix summarizes the conditional moments related to the case with two variance factors, additional details are available in the Online Appendix E.6.2. We have

$$\begin{aligned}
 \alpha_1 &= e^{-\kappa_1 \Delta}, & \alpha_2 &= e^{-\kappa_2 \Delta}, \\
 \beta_1 &= (1 - \alpha_1)\theta_1, & \beta_2 &= (1 - \alpha_2)\theta_2, \\
 1 - A &= (1 - \alpha_1)(1 - \alpha_2) = 1 - (\alpha_1 + \alpha_2) + \alpha_1\alpha_2, \\
 B &= (1 - \alpha_1)\beta_2\Delta + (1 - \alpha_2)\beta_1\Delta = (1 - A)(\theta_1 + \theta_2)\Delta.
 \end{aligned}$$

From Lemma E.6.14, the recursive relationship for the integrated variance² is

$$(1 - \alpha_1 L)(1 - \alpha_2 L)\mathcal{V}_{t+k+2,t+k+3} = B + MA(2) = (1 - A)(\theta_1 + \theta_2)\Delta + MA(2),$$

where L is the lag operator. We can generalize with

$$\begin{aligned}
 (1 - \alpha_1 L)^q (1 - \alpha_2 L)^q \mathcal{V}_{t+k+2q,t+k+2q+1} &= (1 - \alpha_1)^{q-1} (1 - \alpha_2)^{q-1} B + MA(2q) \\
 &= (1 - A)^q (\theta_1 + \theta_2)\Delta + MA(2q).
 \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 (1 - \alpha_1 L)^q (1 - \alpha_2 L)^q \mathbb{E}[\mathcal{V}_{t+k+2q,t+k+2q+1} | \mathcal{G}_{t\Delta}] \\
 = (1 - \alpha_1)^q (1 - \alpha_2)^q (\theta_1 + \theta_2)\Delta,
 \end{aligned}$$

²See Equation B.4 of BZ or Proposition 3.3 and 3.4 of Meddahi, 2003.

where $\mathcal{V}_{t+k+2q,t+k+2q+1}$ can be replaced with the realized variance $RV_{t+k+2q,t+k+2q+1}$. To reproduce the results of Bollerslev and Zhou, 2002, we set $q = 2$ and $k = 1$, which implies that

$$\begin{aligned} & \mathcal{V}_{t+5,t+6} - 2(\alpha_1 + \alpha_2)\mathcal{V}_{t+4,t+5} + (\alpha_1^2 + \alpha_2^2 + 4\alpha_1\alpha_2)\mathcal{V}_{t+3,t+4} \\ & - 2(\alpha_1^2\alpha_2 + \alpha_2^2\alpha_1)\mathcal{V}_{t+2,t+3} + (\alpha_1^2\alpha_2^2)\mathcal{V}_{t+1,t+2} - (1 - \alpha_1)^2(1 - \alpha_2)^2(\theta_1 + \theta_2)\Delta, \end{aligned}$$

has a null expected value conditional to the model filtration at time t .

For the second moment, from Equation (B.8) of BZ, we have

$$\begin{aligned} & (1 - \alpha_1 L)(1 - \alpha_2 L)(1 - H_1 L)(1 - H_2 L)(1 - \alpha_1 \alpha_2 L) \mathcal{V}_{(t+5)\Delta, (t+6)\Delta}^2 \\ & = f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2) + MA(5). \end{aligned}$$

Therefore, we have

$$\begin{aligned} & \mathcal{V}_{t+5,t+6}^2 (1 - \alpha_1 \alpha_2 L) - \underbrace{(\alpha_1 + \alpha_2 + H_1 + H_2)}_{\phi_1} \mathcal{V}_{t+4,t+5}^2 (1 - \alpha_1 \alpha_2 L) \\ & + \underbrace{(\alpha_1 \alpha_2 + H_1 H_2 + (\alpha_1 + \alpha_2)(H_1 + H_2))}_{\phi_2} \mathcal{V}_{t+3,t+4}^2 (1 - \alpha_1 \alpha_2 L) \\ & - \underbrace{((\alpha_1 + \alpha_2)H_1 H_2 + \alpha_1 \alpha_2 (H_1 + H_2))}_{\phi_3} \mathcal{V}_{t+2,t+3}^2 (1 - \alpha_1 \alpha_2 L) \\ & + \underbrace{\alpha_1 \alpha_2 H_1 H_2}_{\phi_4} \mathcal{V}_{t+1,t+2}^2 (1 - \alpha_1 \alpha_2 L) - f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2), \end{aligned}$$

which can be further simplified to

$$\begin{aligned} & \mathcal{V}_{t+5,t+6}^2 - (\phi_1 + \alpha_1 \alpha_2) \mathcal{V}_{t+4,t+5}^2 + (\phi_2 + \alpha_1 \alpha_2 \phi_1) \mathcal{V}_{t+3,t+4}^2 - (\phi_3 + \alpha_1 \alpha_2 \phi_2) \mathcal{V}_{t+2,t+3}^2 \\ & + (\phi_4 + \alpha_1 \alpha_2 \phi_3) \mathcal{V}_{t+1,t+2}^2 - (\alpha_1 \alpha_2 \phi_4) \mathcal{V}_{t,t+1}^2 - f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2), \end{aligned}$$

which has a null expected value conditional to the model filtration at time t .

C.4 Unconditional Moment Derivation

This Appendix summarizes the unconditional moments used in the moment selection of Sections 4.4.1 and 4.5.1. Complete derivations are available in the Online Appendix E.7.

The first two unconditional moments of the realized variance are

$$\mathbb{E} [RV_{t,t+1}] = \sum_k \theta_k \Delta \text{ and } \text{Var} [RV_{t,t+1}] = \mathbb{E} [\xi_{t,t+1}^2] + \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} \left(\frac{2}{\kappa_k} (\Delta - a_k) \right).$$

The expected value of the realized quarticity is

$$\mathbb{E} [RQ_{t,t+1}] = \left(\sum_k \theta_k \right)^2 \Delta + \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} \Delta.$$

The unconditional moments related to the lagged values of the realized variance are useful since they are not affected by the measurement error. If we assume that $\mathbb{E} [\xi_{t,t+1} V_{t+1}] = 0$, then, for $h \geq 1$, the unconditional covariance between lagged realized variances is

$$\text{Cov} [RV_{t,t+1}, RV_{t+h,t+h+1}] = \sum_k \alpha_k^{h-1} a_k^2 \frac{\theta_k \sigma_k^2}{2\kappa_k},$$

implying the following recursion for a model with two variance factors ($K = 2$)

$$\begin{aligned} \text{Cov} [RV_{t,t+1}, RV_{t+1+2n,t+2+2n}] &= (\alpha_1^n + \alpha_2^n) \text{Cov} [RV_{t,t+1}, RV_{t+1+n,t+2+n}] \\ &\quad - (\alpha_1^n \alpha_2^n) \text{Cov} [RV_{t,t+1}, RV_{t+1,t+2}]. \end{aligned}$$

If we assume that $\mathbb{E} [R_{t,t+1} \xi_{t,t+1}] = 0$, then the unconditional covariance between the realized variance and returns is

$$\text{Cov} [RV_{t,t+1}, R_{t,t+1}] = \sum_k \left(\lambda_k - \frac{1}{2} \right) (a_k^2 + f(\Delta, \kappa_k)) \frac{\sigma_k^2 \theta_k}{2\kappa_k} + \rho_k \frac{\sigma_k}{\kappa_k} \theta_k (\Delta - a_k),$$

where $f(\Delta, \kappa_k) = \frac{2}{\kappa_k} \left(\Delta + \frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} - 2 \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right)$ is derived in Corollary E.7.4. The covariance between the lagged realized variance and the returns is for $h \geq 1$

$$\text{Cov} [RV_{t+h, t+h+1}, R_{t, t+1}] = \sum_k \alpha_k^{h-1} a_k^2 \left(\left(\lambda_k - \frac{1}{2} \right) \frac{\theta_k \sigma_k^2}{2\kappa_k} + \rho_k \sigma_k \theta_k \right).$$

C.5 Linear Regression Study

C.5.1 One variance

In this Appendix,³ we examine the issue regarding the moment selection of BZ as discussed in Table 4.1 of Section 4.4.1. The BZ moment selection alone does not fully isolate all the parameters in the GMM estimation when the error measurement parameter is estimated.

More precisely, the parameters κ and θ can be fully identified by using the moment I) with $\alpha = e^{-\Delta\kappa} \approx 1 - \kappa\Delta$ and $\beta = \theta(1 - e^{-\Delta\kappa})\Delta \approx \theta\kappa\Delta$ along with instrumental variables in moments II) and III). The remaining parameters, σ and γ^2 , appear in moments IV) to VI). While it appears that the parameter σ can be effectively identified based on the given moments, the parameter γ^2 has very little effect on the moment condition.

To demonstrate our intuition numerically, we write moments I) to III) with the following regressions and also estimate them using the Monte Carlo Simulation (see Section 4.4.2)

$$\begin{aligned} RV_{t+1, t+2} RV_{t-1, t}^p &= \alpha RV_{t, t+1} RV_{t-1, t}^p \\ &+ (1 - \alpha) \theta \Delta RV_{t-1, t}^p + \varepsilon_{t+1, t+2}^{(p)}, \quad p \in \{0, 1, 2\}. \end{aligned} \quad (\text{C.13})$$

In Figure C.4, we present the R^2 results for the regression (C.13) in red, in panels A and B, representing Models (4.14) and (4.16), respectively. The R^2 values are high and show that the error $\varepsilon_{t+1, t+2}^{(p)}$ does not dominate the right-hand side of regression (C.13).

Furthermore, we write moments IV) to VI) with the following regressions

$$RV_{t+1, t+2}^2 RV_{t-1, t}^p = \text{Var} \left[\xi_{t+1, t+2}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] - H \text{Var} \left[\xi_{t, t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] + H RV_{t, t+1}^2 RV_{t-1, t}^p$$

³We present the results for parameter set A, the Figures for parameter sets B and C are available upon request.

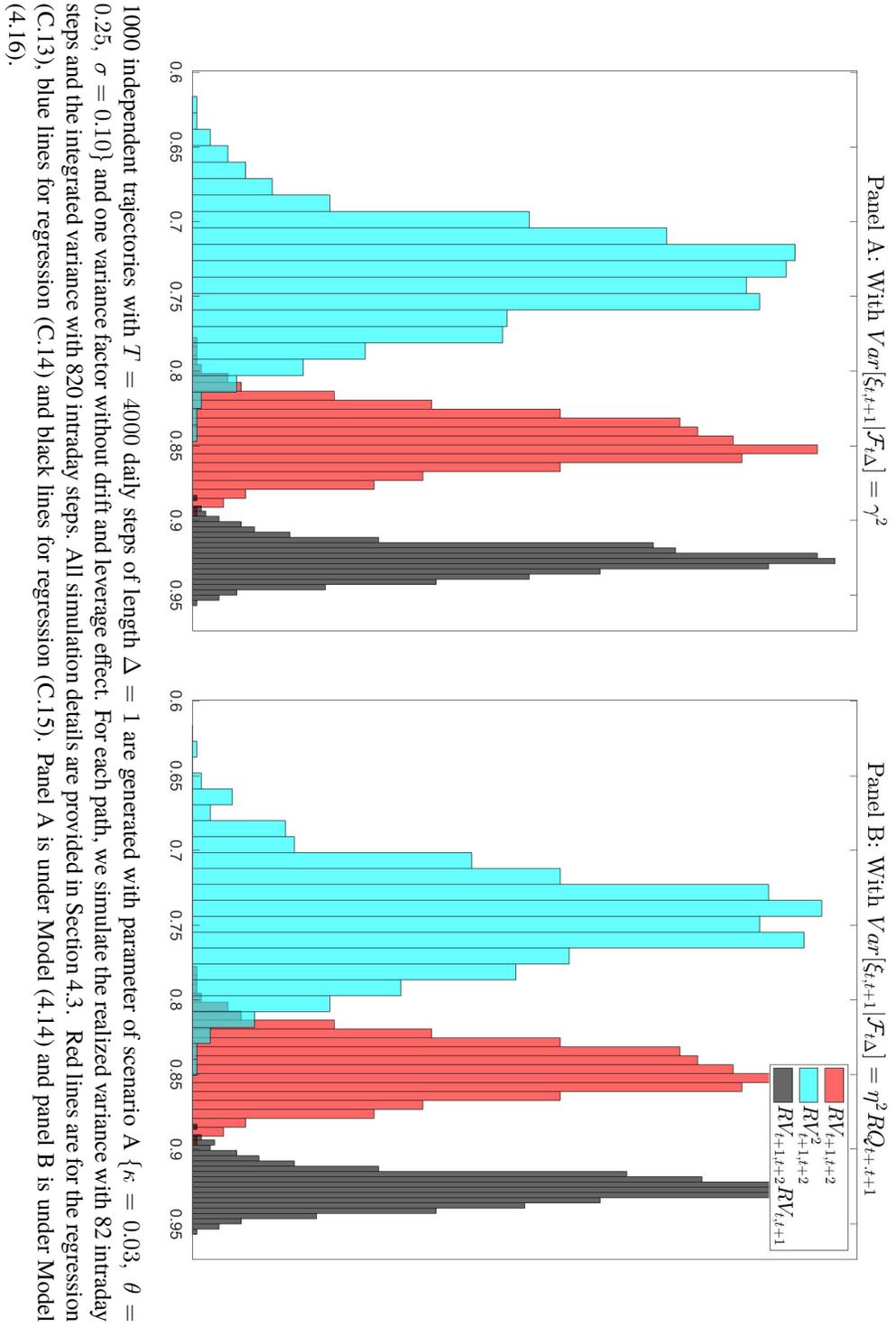


Figure C.4: Regressions (C.13) to (C.15) R^2 comparison

$$+ IRV_{t,t+1}RV_{t-1,t}^p + JRV_{t-1,t}^p + e_{t+1,t+2}^{(p)}, \quad p \in \{0, 1, 2\}. \quad (\text{C.14})$$

For Model (4.14) Regression (C.14) becomes

$$\begin{aligned} RV_{t+1,t+2}^2 RV_{t-1,t}^p &= (1 - H)\gamma^2 + HRV_{t,t+1}^2 RV_{t-1,t}^p \\ &+ IRV_{t,t+1}RV_{t-1,t}^p + JRV_{t-1,t}^p + e_{t+1,t+2}^{(p)}, \quad p \in \{0, 1, 2\}, \end{aligned}$$

and for Model (4.16), Regression (C.14) becomes

$$\begin{aligned} RV_{t+1,t+2}^2 RV_{t-1,t}^p &= \eta^2 RQ_{t+1,t+2} - H\eta^2 RQ_{t,t+1} + HRV_{t,t+1}^2 RV_{t-1,t}^p \\ &+ IRV_{t,t+1}RV_{t-1,t}^p + JRV_{t-1,t}^p + e_{t+1,t+2}^{(p)}, \quad p \in \{0, 1, 2\}. \end{aligned}$$

From Appendix C.3.1, we have the theoretical values of H , I and J along with the error measurement parameters γ^2 and η^2 for Model (4.14) and Model (4.16) respectively. The regression residuals $e_{t+1,t+2}^{(p)}$ can therefore be obtained as well as the R^2 associated with the regression.⁴

In Figure C.4, we present the R^2 results for the regression (C.14) in blue in the panels A and B, representing Models (4.14) and (4.16), respectively. The R^2 values observed in both panels indicate that a significant proportion of the moment is not explained by the variables on the right-hand side of the regression.

The GMM procedure uses the average of the residuals $e_{t+1,t+2}^{(p)}$ over the trajectory of each moments presented in Table 4.1 to estimate its parameters. We also compare the average of the residuals $e_{t+1,t+2}^{(p)}$ to the average of $\text{Var} \left[\xi_{t+1,t+2}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] - H \text{Var} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right]$ over the 1000 simulations in Figure C.5. Panel A corresponds to Model (4.14), and Panel B represents Model (4.16); both panels are run on Regression (C.14).

The key observation is that the sample average, denoted $E[e_{t+1,t+2}^{(p)}]$ (blue), exhibits

⁴The coefficients of the linear regression are not subject to estimation; rather, they are determined by the proofs provided in Appendix C.3. The computation of the coefficient of determination, denoted as R^2 , follows the formula $1 - \frac{SSE}{SST}$, where SSE represents the sum of squared residuals and SST stands for the total sum of squares. Given the potential disparity between sample averages and theoretical quantities, it's important to note that the classical definition of R^2 might not equate to the square of the correlation coefficient in our setup. We present here the square of the correlation coefficient.

significantly more variation compared to the error measurement (black). Furthermore, the black lines are near zero, which can pose challenges for the GMM to differentiate between the error measurement and residuals of the moment. These results are consistent for both Model (4.14) and Model (4.16).

The inclusion of moments VII) and VIII) from Table 4.1 significantly enhances the identification of the error measurement parameter in the model because $\text{Var} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right]$ is multiplied by $\tilde{H} = \alpha = (e^{-\kappa\Delta}) \approx (1 - \kappa\Delta)$, which is closer to one than zero, in opposition to $H - 1 = -2\kappa\Delta$. The regressions for moment VII) and VIII) are

$$RV_{t,t+1}RV_{t+1,t+2}RV_{t-1,t}^p = -\tilde{H} \text{Var} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right] RV_{t-1,t}^p + \tilde{H}RV_{t,t+1}^2RV_{t-1,t}^p \quad (\text{C.15}) \\ + \tilde{I}RV_{t,t+1}RV_{t-1,t}^p + \tilde{J}RV_{t-1,t}^p + \epsilon_{t+1,t+2}^{(p)}, \quad p \in \{0, 1, 2\}.$$

The black lines of both panels of Figure C.4 have substantially higher R^2 values, indicating that the residuals' variance is lower compared to the blue lines. This improvement in R^2 reflects an enhanced quality of parameter estimation for the model.

The results presented in Panel C of Figure C.5 offer insights of the contributions of $\tilde{e}_{t+1,t+2}^{(p)}$ and $-\tilde{H} \text{Var} \left[\xi_{t,t+1}^{(N)} \middle| \mathcal{F}_{t\Delta} \right]$ to the averaged moments condition. The latter (black lines) is distinct from the regression residuals (blue lines), enhancing the capacity to isolate the error measurement parameter in the GMM procedure, as opposed to panels A and B.

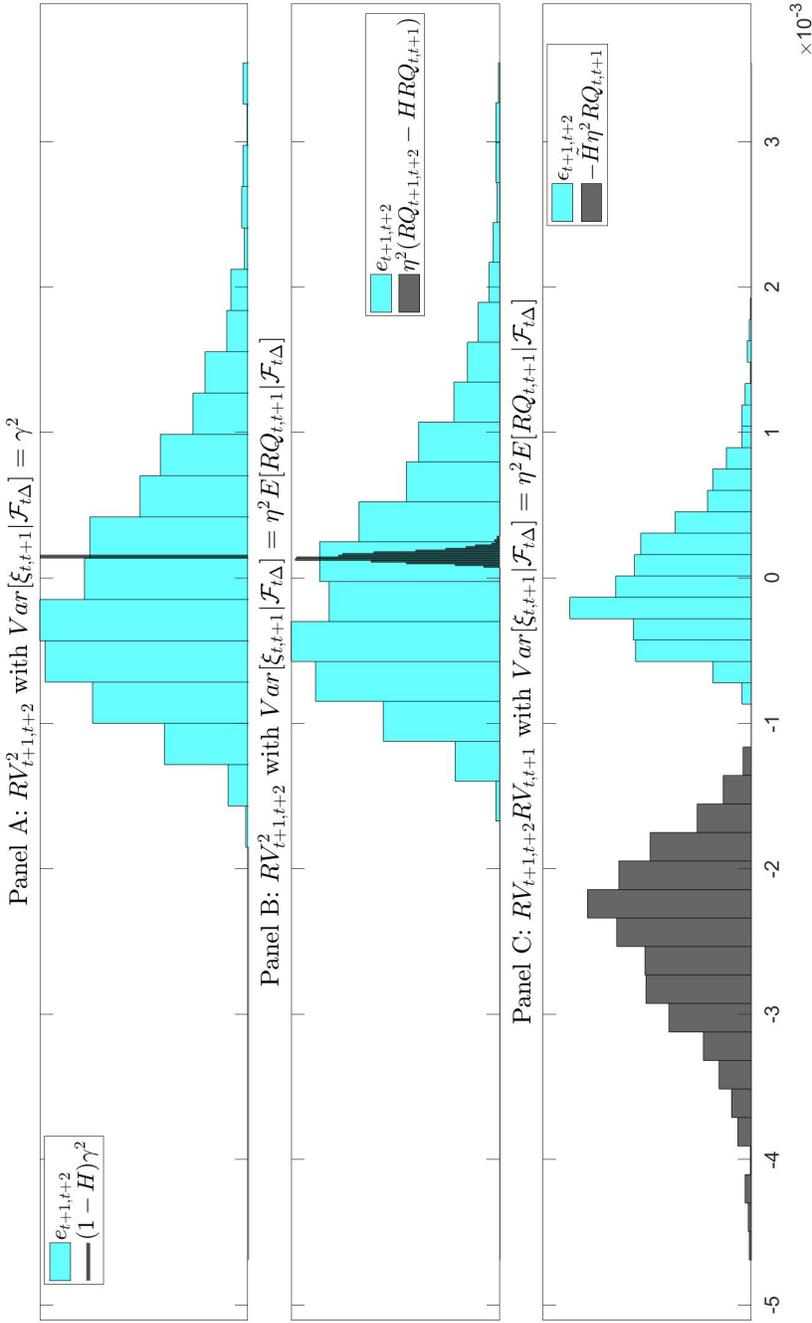
C.5.2 Two variances

We believe that the two moment conditions presented in Panel A of Table 4.6, as used in Bollerslev and Zhou, 2002, along with their instrumental variables, may not be sufficient to fully identify all the parameters of Model (4.21), even without any measurement error.

As derived in Appendix C.3.2, the first conditional moment in the case of two variance factors is, for $q \geq 1$ and $k \geq 1$,

$$(1 - \alpha_1 L)^q (1 - \alpha_2 L)^q E[\mathcal{V}_{t+k+2q,t+k+2q+1} | \mathcal{G}_{t\Delta}] = (1 - \alpha_1)^q (1 - \alpha_2)^q (\theta_1 + \theta_2) \Delta$$

By incorporating instrumental variables, we can identify, at most, the mean-reverting pa-



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Each panel shows the average over all the trajectory of the noise $e_{t+1,t+2}^{(p)}$ (A and B) or $\hat{e}_{t+1,t+2}^{(p)}$ (C) from Equations (C.14) and (C.15) in blue compared to the component related to the error measurement variance, which depends on the model. Panel A is based on Model (4.14), panel B and C on Model (4.16).

Figure C.5: Importance of the measurement error conditional variance in the GMM moments compared to the moment residuals

parameters κ_1 and κ_2 , along with the sum of the long-term variances $(\theta_1 + \theta_2)$ in the Model (4.21).

We propose two regressions related to the first moment of the integrated variance. The first is with $q = 1$ and $k = 3$, that is

$$\begin{aligned} \mathcal{V}_{t+5,t+6} &= (\alpha_1 + \alpha_2)\mathcal{V}_{t+4,t+5} - (\alpha_1\alpha_2)\mathcal{V}_{t+3,t+4} \\ &\quad - (1 - \alpha_1)(1 - \alpha_2)(\theta_1 + \theta_2)\Delta + \epsilon_{t+5,t+6}^{(1)}, \end{aligned} \quad (\text{C.16})$$

and the second, with $q = 2$ and $k = 1$, is

$$\begin{aligned} \mathcal{V}_{t+5,t+6} &= 2(\alpha_1 + \alpha_2)\mathcal{V}_{t+4,t+5} - (\alpha_1^2 + \alpha_2^2 + 4\alpha_1\alpha_2)\mathcal{V}_{t+3,t+4} \\ &\quad + 2(\alpha_1^2\alpha_2 + \alpha_2^2\alpha_1)\mathcal{V}_{t+2,t+3} - (\alpha_1^2\alpha_2^2)\mathcal{V}_{t+1,t+2} \\ &\quad - (1 - \alpha_1)^2(1 - \alpha_2)^2(\theta_1 + \theta_2)\Delta + \epsilon_{t+5,t+6}^{(2)}. \end{aligned} \quad (\text{C.17})$$

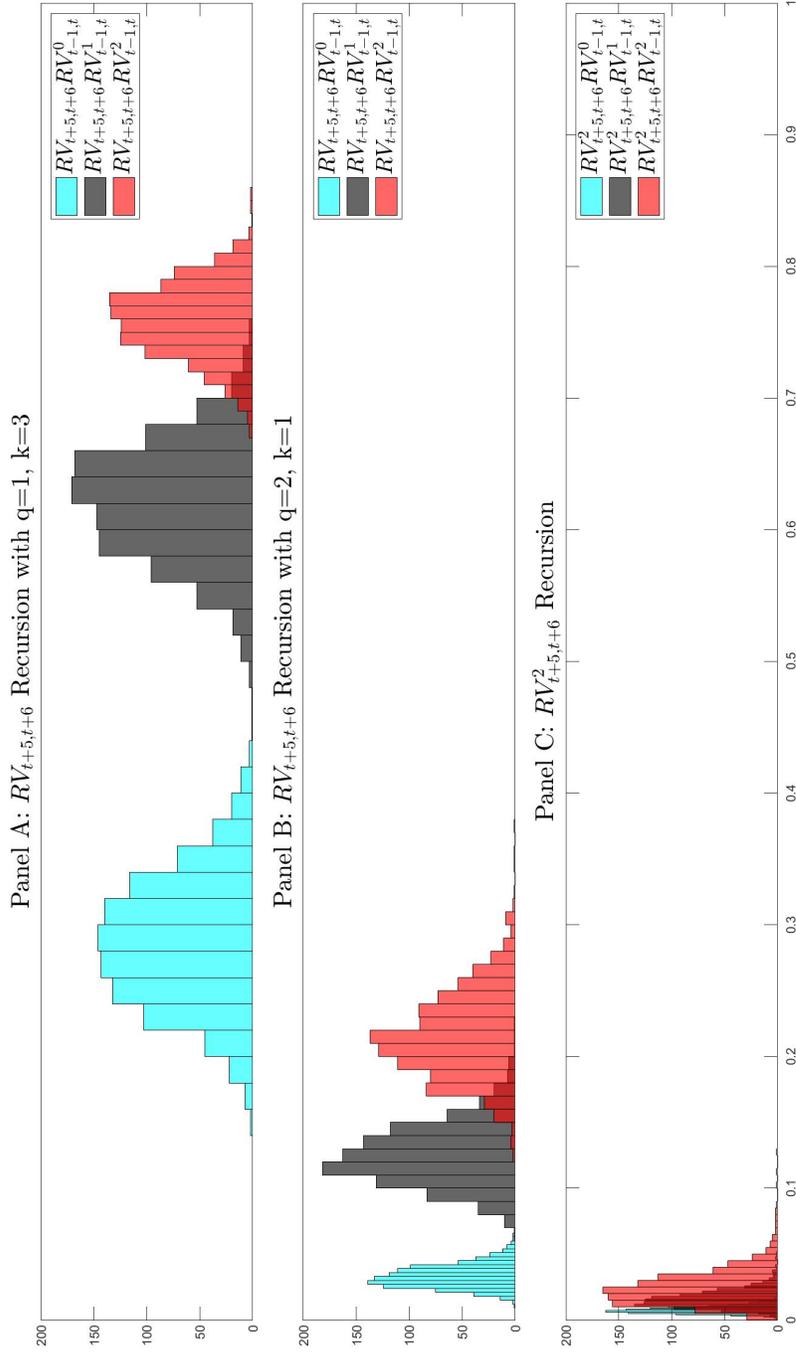
Figure C.6 panels A and B display the results for Regressions (C.16) and (C.17) respectively. We show our results with three types of instrumental variables. As we believe it is Panel B that is used in the BZ paper, we see that the R^2 are low.⁵

To further our point, we also consider the second conditional moment with the following result:

$$\mathbb{E} \left[\begin{array}{c} (1 - e^{-\kappa_1\Delta}L) (1 - e^{-\kappa_2\Delta}L) (1 - e^{-2\kappa_1\Delta}L) \\ (1 - e^{-2\kappa_2\Delta}L) (1 - e^{-\kappa_1\Delta}e^{-\kappa_2\Delta}L) \mathcal{V}_{t+5,t+6}^2 \end{array} \middle| \mathcal{F}_{t\Delta} \right] = f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2),$$

where L designate the lag operator and $f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2)$ is a function of all the variance parameters. With the use of instrumental variables, we can identify at most the mean reverting parameters and the constant $f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2)$. The constants of moments I) and II) of Table 4.6 Panel A are the only elements that include both the volatility of volatilities σ_k and long-term variances θ_k . Consequently, with only two sources of information, we cannot distinguish between the four parameters even with instrumental variables; thus

⁵Furthermore, as we calculate the R^2 with the sample correlation, we center the regression error; results calculated with $1 - \frac{SSE}{SST}$ have negative values indicating worse results.



1000 independent trajectories with $T = 7500$ daily steps of length $\Delta = 1/252$ are generated with parameter of specification (4.21) and two variance factors without drift and leverage effect: $\{\kappa_1 = 2, \kappa_2 = 1, \theta_1 = 0.05, \theta_2 = 0.02, \sigma_1 = 0.15, \sigma_2 = 0.12\}$. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Panel A are for the regression (C.16), panel B for regression (C.17) and panel C for regression (C.18).

Figure C.6: Regressions (C.16) to (C.18) R^2 comparison

leading to poor parameter identification in the case of two-variance factor model.

To show our intuition, we propose a regression based on

$$\begin{aligned} & (1 - \alpha_1 L)(1 - \alpha_2 L)(1 - H_1 L)(1 - H_2 L)(1 - \alpha_1 \alpha_2 L) \mathcal{V}_{t+5, t+6}^2 \\ & = f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2) + \epsilon_{t+5, t+6}^{(3)}. \end{aligned} \quad (\text{C.18})$$

Figure C.6 panel C displays the results for Regressions (C.18) where the R^2 are almost all below 10%, which is low, reinforcing our point. We also highlight that higher the R^2 in all three panels, better is the GMM parameter estimation is as the regression study indicates the amount of noise present in the moment conditions.

C.6 Optimization

This Appendix offers a comprehensive explanation of the two-step GMM procedure used in this paper.⁶ Additionally, it covers the scaling methodology and parameter transformations employed.

C.6.1 GMM procedure

Each moment condition is a function of parameters θ and observable variables at time t . If the model is correct, these entries should have an expectation of zero over the entire sample. Let $g_{t,i,\theta}$ be the entry of moment i at time t with parameter set θ , the average error of the moment condition is

$$\bar{g}_{i,\theta} = \frac{1}{T} \sum_{t=1}^T g_{t,i,\theta},$$

where all the moments conditions create a vector $\bar{\mathbf{g}}_\theta$. In a GMM, the objective is to find the parameters that align the theoretical moments of the model with the corresponding empirical moments observed in the data. Typically, these moments are weighted with a

⁶Supplementary materials on the background of GMM and the use of instrumental variables is available in the Online Appendix E.9

matrix \mathbf{W}_θ , where the weights indicate the relative importance of each moment condition. Therefore, the parameters that minimize the average weighted moment conditions are

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \bar{\mathbf{g}}_\theta^\top \mathbf{W}_\theta \bar{\mathbf{g}}_\theta.$$

As it is commonly done, we perform a two-step procedure where we first initialize \mathbf{W}_θ to the identity matrix.⁷ The scaling procedure discussed in the next subsection enables us to start with scaled observables of the same order of magnitude, significantly aiding the optimization process, with the identity matrix serving as a suitable initial guess.

We proceed to obtain the first estimate, denoted as $\hat{\boldsymbol{\theta}}^{(1)}$, from the first step of the GMM procedure⁸ with

$$\hat{\boldsymbol{\theta}}^{(1)} = \arg \min_{\boldsymbol{\theta} \in \Theta} \bar{\mathbf{g}}_\theta^\top \bar{\mathbf{g}}_\theta.$$

Subsequently, we compute the weighting matrix $\widehat{\mathbf{W}}_{\hat{\boldsymbol{\theta}}^{(1)}}$ using the estimate $\hat{\boldsymbol{\theta}}^{(1)}$ obtained from the first step and the Newey-West covariance matrix estimator (See Andrews, 1991; Newey and West, 1987) with a Bartlett-kernel and a lag⁹ L as it is commonly used in the literature to handle the heteroscedasticity.¹⁰ The weighting matrix is

$$\widehat{\mathbf{W}}_\theta = \left(\sum_{l=-L}^L \left(1 - \frac{l}{L+1} \right) \left(\sum_{t=1}^{T-l} \frac{\mathbf{g}_{t,\theta} \mathbf{g}_{t+l,\theta}^\top}{T} \right) \right)^{-1},$$

⁷Corradi and Distaso, 2006 derive a slightly different GMM procedure as they do one step where the weight matrix is set to the sample variance of the moments. Zhang and Wang, 2023 follows the same procedure but also describes an iterative approach that uses the identity matrix for the first step and a HAC matrix for the subsequent steps. This alternative approach is used by Bolko et al., 2023 and Garcia et al., 2011. Additionally, Andersen and Sørensen, 1996 use the sample moment weighting matrix in their first step, but in their second step, they use a kernel weighting matrix.

⁸We initialize the optimization of both steps at $\boldsymbol{\theta}^*$, the true values of the parameters as we are in a simulation framework. In a setup where the true parameters are not known, adjustments to the optimization procedure would be required.

⁹As in BZ, the lag is set to 5 and 60 for models with one and two variance factors respectively for CM approaches. For UM approaches, we follow Todorov, 2009 and set the lag to 80. Garcia et al., 2011 set a lag of 2.

¹⁰See Zhang and Wang, 2023 p.5 for a detailed steps description.

where $\mathbf{g}_{t,\theta}$ is a vector of all moment conditions $g_{t,i,\theta}$ at time t . Finally, we enter the second step of our GMM procedure using the updated weighting matrix $\widehat{\mathbf{W}}_{\hat{\theta}^{(1)}}$. This step refines our parameter estimates further,¹¹ resulting in the final estimate, denoted as $\hat{\theta}^{(2)}$,

$$\hat{\theta}^{(2)} = \arg \min_{\theta \in \Theta} \bar{\mathbf{g}}_{\theta}^T \widehat{\mathbf{W}}_{\hat{\theta}^{(1)}} \bar{\mathbf{g}}_{\theta}.$$

C.6.2 Scaling procedure

To prevent certain terms from dominating the objective function, observables for each trajectory s are scaled with the standard deviation of the realized variance, denoted as $\mathcal{S}^{(s)} = \text{std}(\mathbf{RV}^{(s)}/\Delta)$. Our observable transformations¹² are

$$RV_{t,t+1}^{(s)} \rightarrow \frac{RV_{t,t+1}^{(s)}}{\mathcal{S}^{(s)}}, \quad RQ_{t,t+1}^{(s)} \rightarrow \frac{RQ_{t,t+1}^{(s)}}{(\mathcal{S}^{(s)})^2}, \quad R_{t,t+1}^{(s)} \rightarrow \frac{R_{t,t+1}^{(s)}}{\sqrt{\mathcal{S}^{(s)}}}.$$

We must also scale the parameters before optimization¹³ with

$$\kappa \rightarrow \kappa, \quad \theta \rightarrow \frac{\theta}{\mathcal{S}^{(s)}}, \quad \sigma \rightarrow \frac{\sigma}{\sqrt{\mathcal{S}^{(s)}}}, \quad \rho \rightarrow \rho, \quad \left(\lambda - \frac{1}{2}\right) \rightarrow \left(\lambda - \frac{1}{2}\right) \sqrt{\mathcal{S}^{(s)}}.$$

Under specification (4.14), $\gamma^2 \rightarrow \frac{\gamma^2}{(\mathcal{S}^{(s)})^2}$, while under specification (4.16), η^2 remains unchanged. Additionally, we scale the risk-free rate by $\sqrt{\mathcal{S}^{(s)}}$. After the optimization is completed, we apply the reverse transformation to return to the unscaled parameters.

C.6.3 Rescaling the parameters based on the frequency Δ

In Bollerslev and Zhou, 2002, the parameters are presented on a daily basis ($\Delta = 1$). However, it is more common to present parameters annually ($\Delta = 1/252$) in continuous

¹¹In our optimization process, in the very rare case where the first step did not converge, we run an additional step $\hat{\theta}^{(3)} = \arg \min_{\theta \in \Theta} \bar{\mathbf{g}}_{\theta}^T \widehat{\mathbf{W}}_{\hat{\theta}^{(2)}} \bar{\mathbf{g}}_{\theta}$.

¹²When we use the unconditional moment based the realized quarticity, we additionally scale the moment condition by dividing with the standard deviation of the realized quarticity. This permits to have moment with the same order of magnitude.

¹³The scaling was determined by computing the SDE of $d(cV)$ and $d(\tilde{c}R)$, where c and \tilde{c} are constants.

time. Therefore, to align with this convention, we rescale the parameters from Bollerslev and Zhou, 2002 to annual values. This rescaling accounts for the impact of Δ , the time step, and transitioning from daily to annual instantaneous variance to ensure consistent integrated variance results. We rescale the daily parameters to

$$\kappa^{(d)} = \frac{\kappa^{(a)}}{252}, \quad \theta^{(d)} = \frac{\theta^{(a)}}{252}, \quad \sigma^{(d)} = \frac{\sigma^{(a)}}{252}.$$

However, the parameters on the return, ρ and λ are not affected by this rescaling. Regarding the error measurement parameter, its theoretical value must be recalculated with the new scaled parameters and time step Δ .

Furthermore, Bollerslev and Zhou, 2002 performs computations in percent, with returns multiplied by 100. Thus, we adjust the parameters accordingly to reflect this scaling.

C.6.4 Parameters optimization constraints

The constrained parameters directly used within the model require transformations from the unconstrained parameters given to the optimizer. Certain parameters necessitate positivity constraints, while others have bounds. Consequently, we implement the following parameter transformation:

$$\kappa_k = e^{\kappa_{k,unc}}, \quad \theta_k = e^{\theta_{k,unc}}, \quad \sigma_k = e^{\sigma_{k,unc}}, \quad \rho_k = \rho_{k,unc}, \quad \lambda_k = \lambda_{k,unc}.$$

We have purposely not constrained the leverage effect, as this approach allowed us to acknowledge the strengths and weaknesses of certain procedures.¹⁴ The unconstrained parameters are supplied to the optimizer. Regarding the error measurement parameter, when considering specification (4.16), $\eta^2 = \eta_{unc}^2$. However, in the case of specification (4.14), γ^2 is intricately linked to the other parameters as outlined in Equation (4.15). Therefore,

¹⁴Normally, the transformation would be: $\rho_k = \frac{2}{1+e^{-2\rho_{k,unc}}} - 1$,

we make the following transformation:

$$\gamma^2 = \gamma_{unc}^2 \left(\left(\sum_{k=1}^K \theta_k \right)^2 + \sum_{k=1}^K \frac{\sigma_k^2 \theta_k}{2\kappa_k} \right).$$

C.7 Supplementary Results

This Appendix presents the detailed results for parameter sets A , B and C for the case with one variance in Tables C.1 to C.3. The conclusions from the results remain consistent with those in the main paper. Panels A of Table C.1, C.2 and C.3 show estimation results for specification (4.14) under BZ.¹⁵ Specification (4.16) under CM moment selection is presented in panels B of the same Tables. An additional observation from these panels is the reduction of the bias¹⁶ on the mean-reverting parameter κ for the CM procedures. For the UM selection of Panel B of Table 4.1, the results are presented in the Tables C.1, C.2 and C.3 in panel C.

We conduct an in-depth analysis with Table C.4, focusing on the leverage effect and comparing our results with those from previous studies by Bregantini, 2013, and Garcia et al., 2011. By incorporating a leverage effect of -50% , in line with established literature, we evaluate the performance of our methodologies against these benchmarks.

Starting with Scenario A in Table C.4, we compare our findings with those in Garcia et al., 2011 Table 2 and Bregantini, 2013 Table 3. Bregantini, 2013 study, which uses GMM to analyze both integrated volatility and high-frequency returns, serves as a key reference point. Our analysis reveals that our methods significantly reduce bias in the leverage effect parameters and achieve a notably smaller RMSE. Specifically, when focusing on high-frequency returns and comparing with Panels A.2 and B.2 of our paper, where measurement error parameter is fixed, our methodology demonstrates a much lower bias and RMSE compared to Bregantini, 2013 results. However, they cover more complex case with jumps, extensions not reviewed in the paper.

¹⁵Replication of BZ tables are available in Online Appendix E.5.

¹⁶Bollerslev et al., 2016 discuss the bias on the persistence parameter when RV measure are used in the estimation and Hansen and Lunde, 2014 argue that adding instrumental variables help to reduce the bias.

Turning to Garcia et al., 2011 Table 2, which involves 960 observations and 5000 replications, we make comparisons with our scenario for $T = 1000$. Without the use of options, and focusing on Panels A.2 and B.2, our methods exhibit lower RMSE and bias. Although, their use of options improves their results, making them comparable to CM methods (with UM methods showing higher bias), our CM methods still outperform theirs in terms of RMSE. Even without resorting to additional calculations and data required for options, our methods achieve comparable or better results. It's important to note, however, that Garcia et al., 2011 ignore the measurement error, which could bias our analysis here.

In summary, we were able to provide, for the three parameter sets of Bollerslev and Zhou, 2002, an additional estimated parameter, the leverage effect, with accuracy. However, CM methods appear to perform better, as they exhibit lower RMSE and bias compared to UM methods.

Appendix C. GMM Parameter Estimation in Stochastic Volatility Models

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: BZ estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.6582	3.1856	3.5432	3.1644	1.3376	0.5974	0.9000	0.9330
$\theta = 0.2500$	0.2435	0.2481	0.2371	0.2468	0.0538	0.0275	0.9620	0.9470
$\sigma = 0.1000$	0.0982	0.0992	0.0981	0.0992	0.0079	0.0041	0.9490	0.9450
Panel B.1: CM estimations results with η fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.4938	3.0818	3.3454	3.0511	1.2696	0.5885	0.9300	0.9520
$\theta = 0.2500$	0.2421	0.2484	0.2365	0.2470	0.0540	0.0273	0.9610	0.9470
$\sigma = 0.1000$	0.0980	0.0994	0.0978	0.0993	0.0075	0.0040	0.9450	0.9450
Panel C.1: UM estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.9557	3.3453	3.8698	3.2903	1.5652	0.7432	0.8840	0.9160
$\theta = 0.2500$	0.2335	0.2436	0.2270	0.2421	0.0520	0.0273	0.9520	0.9440
$\sigma = 0.1000$	0.0986	0.0997	0.0984	0.0998	0.0073	0.0039	0.9500	0.9480
Panel A.2: BZ estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.7221	3.1934	3.6196	3.1605	1.3647	0.6114	0.9050	0.9380
$\theta = 0.2500$	0.2441	0.2483	0.2381	0.2470	0.0530	0.0274	0.9620	0.9480
$\sigma = 0.1000$	0.0917	0.0988	0.0916	0.0976	0.0242	0.0135	0.9340	0.9470
$10^2 \times \gamma^2 = 0.2541$	0.5006	0.1838	0.5115	0.3572	1.6873	1.1132	0.9490	0.9600
Panel B.2: CM estimations results with η estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.4428	3.0602	3.2784	3.0328	1.2712	0.5918	0.9190	0.9450
$\theta = 0.2500$	0.2436	0.2489	0.2381	0.2476	0.0538	0.0274	0.9630	0.9460
$\sigma = 0.1000$	0.0979	0.0992	0.0980	0.0993	0.0106	0.0055	0.9430	0.9500
$10^2 \times \eta^2 = 2.4390$	2.4472	2.4491	2.4397	2.4503	0.3185	0.1610	0.9500	0.9460
Panel C.2: UM estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$10^2 \times \kappa = 3.0000$	3.8804	3.2821	3.7641	3.2483	1.5360	0.7212	0.8900	0.9180
$\theta = 0.2500$	0.2406	0.2467	0.2351	0.2458	0.0522	0.0272	0.9600	0.9470
$\sigma = 0.1000$	0.0991	0.0999	0.0989	0.0998	0.0078	0.0040	0.9520	0.9510
$10^2 \times \gamma^2 = 0.2541$	0.2307	0.2454	0.2087	0.2398	0.1127	0.0591	0.9790	0.9610

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.1: Comparison results for parameter set A for specifications (4.14)-(4.16)

Appendix C. GMM Parameter Estimation in Stochastic Volatility Models

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: BZ estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1085	0.1030	0.1069	0.1028	0.0218	0.0105	0.9210	0.9440
$\theta = 0.2500$	0.2472	0.2491	0.2463	0.2492	0.0161	0.0081	0.9520	0.9480
$\sigma = 0.1000$	0.0987	0.0993	0.0986	0.0994	0.0070	0.0034	0.9500	0.9440
Panel B.1: CM estimations results with η fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1056	0.1010	0.1041	0.1006	0.0199	0.0096	0.9330	0.9480
$\theta = 0.2500$	0.2469	0.2492	0.2462	0.2493	0.0163	0.0081	0.9470	0.9500
$\sigma = 0.1000$	0.0988	0.0996	0.0987	0.0995	0.0061	0.0030	0.9460	0.9540
Panel C.1: UM estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1082	0.1022	0.1072	0.1015	0.0218	0.0107	0.9310	0.9470
$\theta = 0.2500$	0.2461	0.2487	0.2452	0.2486	0.0165	0.0082	0.9480	0.9460
$\sigma = 0.1000$	0.0994	0.0998	0.0992	0.0997	0.0062	0.0031	0.9510	0.9530
Panel A.2: BZ estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1085	0.1020	0.1068	0.1017	0.0219	0.0104	0.9330	0.9520
$\theta = 0.2500$	0.2478	0.2494	0.2469	0.2495	0.0160	0.0081	0.9550	0.9490
$\sigma = 0.1000$	0.0987	0.1044	0.0994	0.1045	0.0230	0.0122	0.9520	0.9330
$10^2 \times \gamma^2 = 0.1829$	0.1282	0.0484	0.1680	0.0688	0.4747	0.3020	0.9430	0.9230
Panel B.2: CM estimations results with η estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1053	0.1007	0.1037	0.1004	0.0215	0.0102	0.9310	0.9450
$\theta = 0.2500$	0.2475	0.2494	0.2468	0.2494	0.0162	0.0081	0.9500	0.9520
$\sigma = 0.1000$	0.0991	0.0996	0.0990	0.0995	0.0080	0.0039	0.9520	0.9430
$10^2 \times \eta^2 = 2.4390$	2.4308	2.4437	2.4176	2.4465	0.2582	0.1300	0.9480	0.9470
Panel C.2: UM estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1079	0.1019	0.1057	0.1013	0.0229	0.0109	0.9310	0.9480
$\theta = 0.2500$	0.2477	0.2493	0.2472	0.2495	0.0162	0.0082	0.9530	0.9510
$\sigma = 0.1000$	0.0999	0.0999	0.0995	0.0997	0.0071	0.0035	0.9600	0.9500
$10^2 \times \gamma^2 = 0.1829$	0.1778	0.1816	0.1765	0.1809	0.0299	0.0153	0.9450	0.9500

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.2: Comparison results for parameter set B for specifications (4.14)-(4.16)

Appendix C. GMM Parameter Estimation in Stochastic Volatility Models

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: BZ estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1141	0.1044	0.1124	0.1038	0.0264	0.0121	0.9020	0.9340
$\theta = 0.2500$	0.2398	0.2472	0.2375	0.2470	0.0334	0.0166	0.9490	0.9490
$\sigma = 0.2000$	0.1977	0.1989	0.1973	0.1987	0.0114	0.0056	0.9510	0.9490
Panel B.1: CM estimations results with η fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1102	0.1018	0.1091	0.1013	0.0248	0.0116	0.9240	0.9550
$\theta = 0.2500$	0.2386	0.2470	0.2358	0.2467	0.0339	0.0166	0.9420	0.9480
$\sigma = 0.2000$	0.1973	0.1992	0.1969	0.1990	0.0101	0.0051	0.9350	0.9480
Panel C.1: UM estimations results with γ fixed and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1151	0.1048	0.1132	0.1040	0.0297	0.0141	0.9120	0.9330
$\theta = 0.2500$	0.2392	0.2460	0.2371	0.2457	0.0332	0.0166	0.9530	0.9430
$\sigma = 0.2000$	0.1984	0.1993	0.1983	0.1991	0.0118	0.0061	0.9500	0.9470
Panel A.2: BZ estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1143	0.1040	0.1130	0.1031	0.0267	0.0123	0.9010	0.9390
$\theta = 0.2500$	0.2429	0.2479	0.2400	0.2477	0.0325	0.0164	0.9530	0.9500
$\sigma = 0.2000$	0.1920	0.2006	0.1885	0.1992	0.0306	0.0167	0.9500	0.9590
$10^2 \times \gamma^2 = 0.2744$	0.4095	0.1648	0.5581	0.2484	1.1906	0.7169	0.9560	0.9550
Panel B.2: CM estimations results with η estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1087	0.1012	0.1070	0.1009	0.0247	0.0117	0.9340	0.9540
$\theta = 0.2500$	0.2404	0.2477	0.2381	0.2473	0.0333	0.0165	0.9500	0.9500
$\sigma = 0.2000$	0.1980	0.1993	0.1974	0.1989	0.0131	0.0067	0.9480	0.9480
$10^2 \times \eta^2 = 2.4390$	2.4252	2.4461	2.4177	2.4452	0.4690	0.2415	0.9550	0.9480
Panel C.2: UM estimations results with γ estimated and $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \epsilon_{t,t+1}^{(N)}$								
$\kappa = 0.1000$	0.1139	0.1037	0.1119	0.1032	0.0322	0.0141	0.9320	0.9410
$\theta = 0.2500$	0.2435	0.2479	0.2418	0.2478	0.0325	0.0165	0.9580	0.9500
$\sigma = 0.2000$	0.1995	0.1997	0.1994	0.1995	0.0144	0.0073	0.9510	0.9570
$10^2 \times \gamma^2 = 0.2744$	0.2535	0.2683	0.2421	0.2658	0.0890	0.0460	0.9510	0.9520

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.3: Comparison results for parameter set C for specifications (4.14)-(4.16)

Appendix C. GMM Parameter Estimation in Stochastic Volatility Models

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: CM estimations results with no measurement error								
$10^2 \times \kappa = 3.0000$	3.5049	3.0711	3.4122	3.0604	1.1686	0.4646	0.9349	0.9420
$\theta = 0.2500$	0.2402	0.2470	0.2365	0.2458	0.0482	0.0242	0.9520	0.9520
$\sigma = 0.1000$	0.0993	0.0998	0.0994	0.0997	0.0027	0.0013	0.9359	0.9450
$\rho = -0.5000$	-0.4910	-0.4991	-0.4891	-0.4990	0.0646	0.0251	0.9469	0.9580
Panel B.1: UM estimations results with no measurement error								
$10^2 \times \kappa = 3.0000$	3.8067	3.2922	3.6841	3.2628	1.4878	0.7145	0.9030	0.9350
$\theta = 0.2500$	0.2332	0.2426	0.2291	0.2411	0.0502	0.0251	0.9470	0.9430
$\sigma = 0.1000$	0.0984	0.0993	0.0984	0.0995	0.0067	0.0035	0.9300	0.9430
$\rho = -0.5000$	-0.4880	-0.4908	-0.4692	-0.4860	0.1551	0.0818	0.9560	0.9590
Panel A.2: CM estimations results with η^2 fixed								
$10^2 \times \kappa = 3.0000$	3.4510	3.0627	3.3951	3.0551	1.1681	0.4909	0.9260	0.9540
$\theta = 0.2500$	0.2446	0.2488	0.2408	0.2477	0.0480	0.0240	0.9550	0.9550
$\sigma = 0.1000$	0.0986	0.0995	0.0984	0.0995	0.0073	0.0039	0.9380	0.9490
$\rho = -0.5000$	-0.4923	-0.4988	-0.4882	-0.4966	0.0862	0.0398	0.9460	0.9510
Panel B.2: UM estimations results with γ^2 fixed								
$10^2 \times \kappa = 3.0000$	3.8809	3.3762	3.7774	3.3419	1.5907	0.7124	0.9129	0.9000
$\theta = 0.2500$	0.2287	0.2391	0.2252	0.2368	0.0501	0.0260	0.9329	0.9260
$\sigma = 0.1000$	0.0981	0.0991	0.0981	0.0992	0.0078	0.0041	0.9429	0.9420
$\rho = -0.5000$	-0.4856	-0.4862	-0.4678	-0.4818	0.1592	0.0809	0.9550	0.9500
Panel A.3: CM estimations results with η^2 estimated								
$10^2 \times \kappa = 3.0000$	3.4265	3.0478	3.2831	3.0272	1.1798	0.5017	0.9290	0.9510
$\theta = 0.2500$	0.2459	0.2492	0.2430	0.2479	0.0486	0.0240	0.9560	0.9570
$\sigma = 0.1000$	0.0991	0.0994	0.0989	0.0993	0.0095	0.0056	0.9500	0.9470
$10^2 \times \eta^2 = 2.4390$	2.4301	2.4451	2.4164	2.4388	0.3073	0.1659	0.9510	0.9490
$\rho = -0.5000$	-0.4947	-0.5006	-0.4909	-0.4989	0.0907	0.0438	0.9420	0.9530
Panel B.3: UM estimations results with γ^2 estimated								
$10^2 \times \kappa = 3.0000$	3.8902	3.3397	3.8025	3.2646	1.5467	0.7012	0.8952	0.9020
$\theta = 0.2500$	0.2294	0.2411	0.2240	0.2404	0.0507	0.0258	0.9447	0.9342
$\sigma = 0.1000$	0.0982	0.0992	0.0983	0.0993	0.0082	0.0042	0.9447	0.9392
$10^2 \times \gamma^2 = 0.2541$	0.2086	0.2323	0.1910	0.2301	0.1008	0.0539	0.9620	0.9280
$\rho = -0.5000$	-0.4829	-0.4877	-0.4714	-0.4820	0.1583	0.0802	0.9574	0.9491

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.4: Comparison results for parameter set A for specification (4.20)

Appendix C. *GMM Parameter Estimation in Stochastic Volatility Models*

True value	Mean	Median	RMSE	CR
Panel A.1: <i>CM estimations results with no measurement error</i>				
$\kappa = 2.0000$	2.1064	2.0813	0.3617	0.9460
$10^2 \times \theta = 5.0000$	4.9820	4.9539	0.4357	0.9510
$\sigma = 0.1500$	0.1498	0.1499	0.0017	0.9380
$\rho = -0.8500$	-0.8482	-0.8483	0.0162	0.9570
$\lambda = 1.1500$	1.2541	1.2332	1.1591	0.9450
Panel B.1: <i>UM estimations results with no measurement error</i>				
$\kappa = 2.0000$	2.3844	2.3278	0.7363	0.8970
$10^2 \times \theta = 5.0000$	4.9591	4.9183	0.4283	0.9540
$\sigma = 0.1500$	0.1493	0.1494	0.0048	0.9420
$\rho = -0.8500$	-0.8111	-0.8010	0.1300	0.9590
$\lambda = 1.1500$	1.3061	1.2758	1.1836	0.9470
Panel A.2: <i>CM estimations results with η^2 fixed</i>				
$\kappa = 2.0000$	2.1617	2.1092	1.1175	0.9594
$10^2 \times \theta = 5.0000$	4.9483	4.9500	0.8312	0.9736
$\sigma = 0.1500$	0.1447	0.1496	0.0372	0.9421
$\rho = -0.8500$	-0.9550	-0.8425	0.4826	0.9594
$\lambda = 1.1500$	1.2559	1.2156	1.1570	0.9513
Panel B.2: <i>UM estimations results with γ^2 fixed</i>				
$\kappa = 2.0000$	2.4000	2.3088	0.8465	0.9120
$10^2 \times \theta = 5.0000$	4.9130	4.8689	0.4337	0.9490
$\sigma = 0.1500$	0.1479	0.1487	0.0139	0.9490
$\rho = -0.8500$	-0.8252	-0.8022	0.1488	0.9600
$\lambda = 1.1500$	1.3988	1.3286	1.2039	0.9440
Panel A.3: <i>CM estimations results with η^2 estimated</i>				
$\kappa = 2.0000$	2.3139	2.1585	1.4204	0.9472
$10^2 \times \theta = 5.0000$	5.0938	4.9501	1.3143	0.9658
$\sigma = 0.1500$	0.1493	0.1527	0.0443	0.9409
$10^4 \times \eta^2 = 0.9679$	0.9626	0.9628	0.0347	0.9534
$\rho = -0.8500$	-0.9650	-0.8194	0.5686	0.9544
$\lambda = 1.1500$	1.2521	1.2222	1.1612	0.9513
Panel B.3: <i>UM estimations results with γ^2 estimated</i>				
$\kappa = 2.0000$	2.3972	2.3187	0.8604	0.9160
$10^2 \times \theta = 5.0000$	4.9334	4.8900	0.4297	0.9500
$\sigma = 0.1500$	0.1485	0.1495	0.0151	0.9490
$10^8 \times \gamma^2 = 0.1068$	0.1033	0.1013	0.0187	0.9560
$\rho = -0.8500$	-0.8269	-0.8037	0.1527	0.9530
$\lambda = 1.1500$	1.3482	1.2980	1.1883	0.9500

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1/252$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.5: Comparison results for parameter set D for specification (4.20)

Appendix C. GMM Parameter Estimation in Stochastic Volatility Models

True value	Mean	Median	RMSE	CR
Panel A.1: <i>CM estimations results with no measurement error</i>				
$\kappa = 4.0000$	4.0783	4.0648	0.5174	0.9490
$10^2 \times \theta = 3.0000$	2.9787	2.9544	0.3348	0.9500
$\sigma = 0.3000$	0.2994	0.2993	0.0037	0.9500
$\rho = -0.9500$	-0.9480	-0.9476	0.0163	0.9550
$\lambda = 1.9500$	2.2074	2.1429	1.5200	0.9410
Panel B.1: <i>UM estimations results with no measurement error</i>				
$\kappa = 4.0000$	4.4556	4.3782	1.0470	0.9210
$10^2 \times \theta = 3.0000$	2.9463	2.9236	0.3265	0.9510
$\sigma = 0.3000$	0.2984	0.2985	0.0094	0.9410
$\rho = -0.9500$	-0.9301	-0.9194	0.0994	0.9520
$\lambda = 1.9500$	2.3497	2.2873	1.5537	0.9380
Panel A.2: <i>CM estimations results with η^2 fixed</i>				
$\kappa = 4.0000$	4.0722	4.0758	0.6960	0.9550
$10^2 \times \theta = 3.0000$	2.9869	2.9618	0.3363	0.9530
$\sigma = 0.3000$	0.2983	0.2984	0.0193	0.9480
$\rho = -0.9500$	-0.9490	-0.9478	0.0649	0.9520
$\lambda = 1.9500$	2.2123	2.1604	1.5132	0.9420
Panel B.2: <i>UM estimations results with γ^2 fixed</i>				
$\kappa = 4.0000$	4.6391	4.5963	1.1882	0.9110
$10^2 \times \theta = 3.0000$	2.8877	2.8674	0.3418	0.9350
$\sigma = 0.3000$	0.2980	0.2980	0.0138	0.9520
$\rho = -0.9500$	-0.9189	-0.9065	0.1075	0.9510
$\lambda = 1.9500$	2.6604	2.5987	1.7351	0.9150
Panel A.3: <i>CM estimations results with η^2 estimated</i>				
$\kappa = 4.0000$	4.0905	4.0860	0.7735	0.9460
$10^2 \times \theta = 3.0000$	2.9907	2.9630	0.3369	0.9520
$\sigma = 0.3000$	0.2997	0.2990	0.0275	0.9500
$10^4 \times \eta^2 = 0.9679$	0.9645	0.9639	0.0476	0.9540
$\rho = -0.9500$	-0.9495	-0.9445	0.0909	0.9470
$\lambda = 1.9500$	2.1893	2.1418	1.5049	0.9380
Panel B.3: <i>UM estimations results with γ^2 estimated</i>				
$\kappa = 4.0000$	4.5641	4.5215	1.1386	0.9170
$10^2 \times \theta = 3.0000$	2.9137	2.8953	0.3335	0.9440
$\sigma = 0.3000$	0.2980	0.2977	0.0141	0.9490
$10^{10} \times \gamma^2 = 4.7529$	4.4347	4.3307	1.0698	0.9530
$\rho = -0.9500$	-0.9241	-0.9104	0.1074	0.9610
$\lambda = 1.9500$	2.5157	2.4329	1.6467	0.9280

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1/252$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table C.6: Comparison results for parameter set E for specification (4.20)

References

- Andersen, T. G., & Sørensen, B. E. (1996). GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *Journal of Business and Economic Statistics*, 14(3), 328–352.
- Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59(3), 817–858.
- Bolko, A. E., Christensen, K., Pakkanen, M. S., & Veliyev, B. (2023). A GMM Approach to Estimate the Roughness of Stochastic Volatility. *Journal of Econometrics*, 235(2), 745–778.
- Bollerslev, T., Patton, A. J., & Quaedvlieg, R. (2016). Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting. *Journal of Econometrics*, 192(1), 1–18.
- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, 109(1), 33–65.
- Bregantini, D. (2013). Moment-based Estimation of Stochastic Volatility. *Journal of Banking and Finance*, 37(12), 4755–4764.
- Corradi, V., & Distaso, W. (2006). Semi-Parametric Comparison of Stochastic Volatility Models using Realized Measures. *Review of Economic Studies*, 73(3), 635–667.

- Garcia, R., Lewis, M.-A., Pastorello, S., & Renault, É. (2011). Estimation of Objective and Risk-Neutral Distributions based on Moments of Integrated Volatility. *Journal of Econometrics*, 160(1), 22–32.
- Hansen, P. R., & Lunde, A. (2014). Estimating the Persistence and the Autocorrelation Function of a Time Series that is Measured with Error. *Econometric Theory*, 30(1), 60–93.
- Meddahi, N. (2003). ARMA Representation of Integrated and Realized Variances. *Econometrics Journal*, 6(2), 335–356.
- Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708.
- Todorov, V. (2009). Estimation of Continuous-Time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics*, 148(2), 131–148.
- Zhang, L., & Wang, L. (2023). Generalized Method of Moments Estimation of Realized Stochastic Volatility Model. *Journal of Risk and Financial Management*, 16(8), 377.

Chapter D

Online Appendices of *Computationally-Efficient Variance Filtering in Multidimensional Affine Models*

D.1 The Forward Price

The time t value of the self-financing investment strategy that reinvests the dividends to buy more shares of the risky asset is

$$\exp\left(\int_0^t q_{i,u} du\right) S_{i,t}. \quad (\text{D.1})$$

From the Girsanov theorem, let $W_{k,t}^{\mathbb{Q}} = W_{k,t} + \int_0^t \gamma_{k,u} du$ and $B_{k,t}^{\mathbb{Q}} = B_{k,t} + \int_0^t \Gamma_{k,u} du$ where $\gamma_{k,u}$ and $\Gamma_{k,u}$ are predictable processes with respect to the continuous model filtration $(\mathcal{F}_t^C)_{t \geq 0}$. We assume that the Novikov condition

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^T \gamma_{k,u}^2 du \right) \right] < \infty \quad \text{and} \quad \mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^T \Gamma_{k,u}^2 du \right) \right] < \infty,$$

are satisfied. The associated Radon-Nikodym derivative is

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp \left(\sum_{k=1}^{N_V} \left(-\frac{1}{2} \int_0^T \gamma_{k,u}^2 du - \int_0^T \gamma_{k,u} dW_{k,u}^{\mathbb{P}} \right) + \sum_{k=1}^{N_V} \left(-\frac{1}{2} \int_0^T \Gamma_{k,u}^2 du - \int_0^T \Gamma_{k,u} dB_{k,u} \right) \right).$$

The absence of arbitrage opportunity requires that

$$\left\{ \exp \left(- \int_0^t r_u du \right) \exp \left(\int_0^t q_{i,u} du \right) S_{i,t} \right\}_{t \geq 0}$$

is a \mathbb{Q} -martingale. Therefore,

$$\begin{aligned} & d \exp \left(\int_0^t q_{i,u} - r_u du \right) S_{i,t} \tag{D.2} \\ &= S_{i,t} d \exp \left(\int_0^t q_{i,u} - r_u du \right) + \exp \left(\int_0^t q_{i,u} - r_u du \right) dS_{i,t} \\ &= \exp \left(\int_0^t q_{i,u} - r_u du \right) S_{i,t} \left(\begin{aligned} & (\alpha_{i,t} + q_{i,t} - r_t) dt \\ & + \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} d \left(\rho_k dW_{k,t} + \sqrt{1 - \rho_k^2} dB_{k,t} \right) \end{aligned} \right) \\ &= \exp \left(\int_0^t q_{i,u} - r_u du \right) S_{i,t} \underbrace{\left(\frac{\alpha_{i,t} + q_{i,t} - r_t - \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} \left(\rho_k \gamma_{k,t} + \sqrt{1 - \rho_k^2} \Gamma_{k,t} \right)}{=} \right)}_{=0} dt \\ &+ \exp \left(\int_0^t q_{i,u} - r_u du \right) S_{i,t} \left(\sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} d \left(\rho_k dW_{k,t}^{\mathbb{Q}} + \sqrt{1 - \rho_k^2} dB_{k,t}^{\mathbb{Q}} \right) \right), \end{aligned}$$

which implies that

$$\alpha_{i,t} = r_t - q_{i,t} + \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} \left(\rho_k \gamma_{k,t} + \sqrt{1 - \rho_k^2} \Gamma_{k,t} \right).$$

The model belongs to the exponential affine class if the prices of risk are defined as

$$\gamma_{k,t} = \tilde{\gamma}_k \sqrt{V_{k,t}} \text{ and } \Gamma_{k,t} = \tilde{\Gamma}_k \sqrt{V_{k,t}}. \tag{D.3}$$

Therefore,

$$\alpha_{i,t} = r_t - q_{i,t} + \sum_{k=1}^{N_V} c_{ik} \underbrace{\left(\rho_k \tilde{\gamma}_k + \sqrt{1 - \rho_k^2 \tilde{\Gamma}_k} \right)}_{\lambda_k} V_{k,t}.$$

The time t forward price $F_{i,t,T}$ of the contract of maturity T satisfies

$$0 = \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s \right) (S_{i,T} - F_{i,t,T}) \middle| \mathcal{F}_t^C \right].$$

Because

$$dS_{i,t} = (r_t - q_{i,t})S_{i,t}dt + S_{i,t} \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} d \left(\rho_k W_{k,t}^{\mathbb{Q}} + \sqrt{1 - \rho_k^2} B_{k,t}^{\mathbb{Q}} \right), \quad (\text{D.4})$$

the strong solution is

$$S_{i,T} = S_{i,t} \exp \left(\begin{aligned} & \int_t^T (r_u - q_{i,u} - \frac{1}{2} \sum_{k=1}^{N_V} c_{ik}^2 V_{k,u}) du \\ & + \sum_{k=1}^{N_V} c_{ik} \int_t^T \sqrt{V_{k,u}} d \left(\rho_k W_{k,u}^{\mathbb{Q}} + \sqrt{1 - \rho_k^2} B_{k,u}^{\mathbb{Q}} \right) \end{aligned} \right). \quad (\text{D.5})$$

The forward price satisfies

$$\begin{aligned} F_{i,t,T} &= \frac{\mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s ds \right) S_{i,T} \middle| \mathcal{F}_t^C \right]}{\mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t^C \right]} \\ &= \frac{S_{i,t} \mathbb{E}^{\mathbb{Q}} \left[\exp \left(\begin{aligned} & - \int_t^T q_{i,u} du - \frac{1}{2} \sum_{k=1}^{N_V} \int_t^T c_{ik}^2 V_{k,u} du \\ & + \sum_{k=1}^{N_V} c_{ik} \int_t^T \sqrt{V_{k,u}} d \left(\rho_k W_{k,u}^{\mathbb{Q}} + \sqrt{1 - \rho_k^2} B_{k,u}^{\mathbb{Q}} \right) \end{aligned} \right) \middle| \mathcal{F}_t^C \right]}{\mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t^C \right]}. \end{aligned}$$

If r_t and $q_{i,t}$ are non-stochastic, it simplifies to

$$F_{i,t,T} = S_{i,t} \exp \left(\int_t^T (r_s - q_{i,s}) ds \right).$$

Therefore, the SDE of the forward is

$$dF_{i,t,T} = F_{i,t,T} \left(\begin{array}{l} \left(\sum_{k=1}^{N_V} c_{ik} \lambda_k V_{k,t} \right) dt \\ + \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} d \left(\rho_k W_{k,t} + \sqrt{1 - \rho_k^2} B_{k,t} \right) \end{array} \right),$$

and the SDE of the log-forward is

$$d \log F_{i,t,T} = \left(\begin{array}{l} \left(\sum_{k=1}^{N_V} c_{ik} \lambda_k V_{k,t} - \frac{1}{2} \sum_{k=1}^{N_V} c_{ik}^2 V_{k,t} \right) dt \\ + \sum_{k=1}^{N_V} c_{ik} \sqrt{V_{k,t}} d \left(\rho_k W_{k,t} + \sqrt{1 - \rho_k^2} B_{k,t} \right) \end{array} \right),$$

where the subscript T is dropped in the main paper.

D.2 Realized Quarticity and the Measurement Error

This Online Appendix discusses the asymptotic variance of the realized covariance and compares it with the realized quarticity. We show why in a univariate case we can substitute the variance of the realized covariance by the realized coquarticity,¹ but not in the multivariate case.

From Equation (3.13),

$$\xi_{i,j,t,t+1} = RCV_{i,j,t,t+1} = \sum_{k=1}^{N_V} c_{i,k} c_{j,k} \mathcal{V}_{k,t,t+1} + \xi_{i,j,t,t+1} \quad (\text{D.6})$$

is the error measurement on the realized covariance. Equation (B.4) established that

$$\text{Cov} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] = \frac{\Delta}{N} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \Sigma_{\tilde{i},s} \Sigma_{\tilde{j},s} + \Sigma_{i\tilde{j},s} \Sigma_{j\tilde{i},s} ds \mid \mathcal{F}_{t\Delta} \right],$$

¹The relationship between realized coquarticity and realized quarticity is analogous to the one between realized covariance and realized variance.

where

$$\Sigma_{ij,s} = \sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k,s}. \quad (\text{D.7})$$

However, if $r_{i,t\Delta,n} = \left(\log \frac{S_{i,t\Delta+\frac{n\Delta}{N}}}{S_{i,t\Delta+\frac{(n-1)\Delta}{N}}} \right)$, the realized co-quarticity is

$$RQ_{i,j,\tilde{i},\tilde{j},t,t+1} = \frac{N}{3\Delta} \sum_{n=1}^N r_{i,t,n} r_{j,t,n} r_{\tilde{i},t,n} r_{\tilde{j},t,n}. \quad (\text{D.8})$$

Based on Remark 4-iii from p.897 of Barndorff-Nielsen and Shephard, 2004, Equation (D.8) asymptotically converges to

$$\frac{1}{3} \left(\int_{t\Delta}^{(t+1)\Delta} \Sigma_{ij,s} \Sigma_{\tilde{i}\tilde{j},s} + \Sigma_{\tilde{i}\tilde{i},s} \Sigma_{j\tilde{j},s} + \Sigma_{i\tilde{i},s} \Sigma_{j\tilde{i},s} ds \right).$$

Therefore, $RQ_{i,j,\tilde{i},\tilde{j},t,t+1}$ can not be used as a proxy for Equation (B.4) in a multivariate case as it does not have the same asymptotic convergence as the covariance between $\xi_{i,j,t,t+1}$ and $\xi_{\tilde{i},\tilde{j},t,t+1}$. However, in the univariate case where $N_V = 1$, the asymptotic theory simplifies to

$$\frac{\sqrt{\frac{N}{\Delta}} (\mathbf{RV}_{i,t,t+1} - \mathbf{QV}_{i,t,t+1})}{\sqrt{2 \int_{t\Delta}^{(t+1)\Delta} \Sigma_{i,i,s}^2 ds}} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1),$$

where we can replace the theoretical quantity in the denominator by the realized quarticity to obtain the well known result of Barndorff-Nielsen and Shephard, 2002:

$$\frac{\sqrt{\frac{N}{\Delta}} (\mathbf{RV}_{i,t,t+1} - \mathbf{QV}_{i,t,t+1})}{\sqrt{2RQ_{i,t,t+1}}} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

D.3 Measurement Error Additional Derivation

D.3.1 Derivation of $\mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}}$

We derive here the constant $\mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}}$ defined at Equation (B.6). Starting from Equation (B.4),

$$\begin{aligned} \text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}] &= \frac{\Delta}{N} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} \Sigma_{i\tilde{i},s} \Sigma_{j\tilde{j},s} + \Sigma_{i\tilde{j},s} \Sigma_{j\tilde{i},s} ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \eta \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right]. \end{aligned}$$

Because

$$\Sigma_{i\tilde{i},s} \Sigma_{j\tilde{j},s} = \left(\sum_{k=1}^{N_V} c_{i,k} c_{\tilde{i},k} V_{k,s} \right) \left(\sum_{\tilde{k}=1}^{N_V} c_{j,\tilde{k}} c_{\tilde{j},\tilde{k}} V_{\tilde{k},s} \right) = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{i,k} c_{\tilde{i},k} c_{j,\tilde{k}} c_{\tilde{j},\tilde{k}} V_{k,s} V_{\tilde{k},s},$$

and

$$\Sigma_{i\tilde{j},s} \Sigma_{j\tilde{i},s} = \left(\sum_{k=1}^{N_V} c_{i,k} c_{\tilde{j},k} V_{k,s} \right) \left(\sum_{\tilde{k}=1}^{N_V} c_{j,\tilde{k}} c_{\tilde{i},\tilde{k}} V_{\tilde{k},s} \right) = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{i,k} c_{\tilde{j},k} c_{j,\tilde{k}} c_{\tilde{i},\tilde{k}} V_{k,s} V_{\tilde{k},s},$$

it leads to

$$\mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} = c_{i,k} c_{\tilde{i},k} c_{j,\tilde{k}} c_{\tilde{j},\tilde{k}} + c_{i,k} c_{\tilde{j},k} c_{j,\tilde{k}} c_{\tilde{i},\tilde{k}}.$$

D.3.2 Derivation of $\mathbb{E} \left[\int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right]$

Lemma D.3.1. *The SDE of $V_k V_{\tilde{k}}$ is*

$$\begin{aligned} dV_{k,t} V_{\tilde{k},t} &= V_{\tilde{k},t} \kappa_k (\theta_k - V_{k,t}) dt + V_{\tilde{k},t} \sigma_k \sqrt{V_{k,t}} dW_{k,t} + V_{k,t} \kappa_{\tilde{k}} (\theta_{\tilde{k}} - V_{\tilde{k},t}) dt \\ &\quad + V_{k,t} \sigma_{\tilde{k}} \sqrt{V_{\tilde{k},t}} dW_{\tilde{k},t} + \mathbb{1}_{k=\tilde{k}} \sigma_k \sigma_{\tilde{k}} \sqrt{V_{k,t}} \sqrt{V_{\tilde{k},t}} dt. \end{aligned}$$

Proof. Direct application of Itô's lemma to $V_{k,t} V_{\tilde{k},t}$. □

Corollary D.3.2. For $0 < t < u$,

$$\begin{aligned}
 V_{k,u}V_{\bar{k},u} &= V_{k,t\Delta}V_{\bar{k},t\Delta}e^{-(\kappa_k+\kappa_{\bar{k}})(u-t\Delta)} \\
 &+ \kappa_k\theta_k \int_{t\Delta}^u V_{\bar{k},s}e^{-(\kappa_k+\kappa_{\bar{k}})(u-s)}ds + \sigma_k \int_{t\Delta}^u \sqrt{V_{k,s}}V_{\bar{k},s}e^{-(\kappa_k+\kappa_{\bar{k}})(u-s)}dW_{k,s} \\
 &+ \kappa_{\bar{k}}\theta_{\bar{k}} \int_{t\Delta}^u V_{k,s}e^{-(\kappa_k+\kappa_{\bar{k}})(u-s)}ds + \sigma_{\bar{k}} \int_{t\Delta}^u \sqrt{V_{\bar{k},s}}V_{k,s}e^{-(\kappa_k+\kappa_{\bar{k}})(u-s)}dW_{\bar{k},s} \\
 &+ \int_{t\Delta}^u \mathbb{1}_{k=\bar{k}}\sigma_k\sigma_{\bar{k}}\sqrt{V_{k,s}}\sqrt{V_{\bar{k},s}}e^{-(\kappa_k+\kappa_{\bar{k}})(u-s)}ds.
 \end{aligned}$$

Proof. Adding $V_{k,u}V_{\bar{k},u}(\kappa_k + \kappa_{\bar{k}})dt$ to Lemma D.3.1 on both sides and multiplying with $e^{t(\kappa_k+\kappa_{\bar{k}})}$ leads to

$$\begin{aligned}
 d(e^{t(\kappa_k+\kappa_{\bar{k}})}V_{k,t}V_{\bar{k},t}) &= e^{t(\kappa_k+\kappa_{\bar{k}})}V_{\bar{k},t}\kappa_k\theta_k dt + V_{\bar{k},t}e^{t(\kappa_k+\kappa_{\bar{k}})}\sigma_k\sqrt{V_{k,t}}dW_{k,t} \\
 &+ e^{t(\kappa_k+\kappa_{\bar{k}})}V_{k,t}\kappa_{\bar{k}}\theta_{\bar{k}} dt + V_{k,t}e^{t(\kappa_k+\kappa_{\bar{k}})}\sigma_{\bar{k}}\sqrt{V_{\bar{k},t}}dW_{\bar{k},t} \\
 &+ \mathbb{1}_{k=\bar{k}}\sigma_k\sigma_{\bar{k}}e^{t(\kappa_k+\kappa_{\bar{k}})}\sqrt{V_{k,t}}\sqrt{V_{\bar{k},t}}dt.
 \end{aligned}$$

Finally, we obtain the result through integration and multiplying with $e^{-(\kappa_k+\kappa_{\bar{k}})(u-t\Delta)}$. \square

Corollary D.3.3. For $0 < t < u$,

$$\int_{t\Delta}^u V_{k,s}V_{\bar{k},s}ds = \frac{1}{\kappa_k + \kappa_{\bar{k}}} \left(\begin{aligned} &V_{k,t\Delta}V_{\bar{k},t\Delta} - V_{k,u}V_{\bar{k},u} + \kappa_k\theta_k \int_{t\Delta}^u V_{\bar{k},s}ds + \kappa_{\bar{k}}\theta_{\bar{k}} \int_{t\Delta}^u V_{k,s}ds \\ &+ \sigma_k \int_{t\Delta}^u \sqrt{V_{k,s}}V_{\bar{k},s}dW_{k,s} + \sigma_{\bar{k}} \int_{t\Delta}^u \sqrt{V_{\bar{k},s}}V_{k,s}dW_{\bar{k},s} \\ &+ \int_{t\Delta}^u \mathbb{1}_{k=\bar{k}}\sigma_k\sigma_{\bar{k}}\sqrt{V_{k,s}}\sqrt{V_{\bar{k},s}}ds \end{aligned} \right).$$

Proof. Direct application of Lemma D.3.1 by integrating the result. \square

Lemma D.3.4. For $0 < t < u$,

$$\mathbb{E} \left[\int_{t\Delta}^u V_{k,s}V_{\bar{k},s}ds \middle| \mathcal{F}_{t\Delta} \right]$$

$$= \mathbb{E} \left[\frac{1}{\kappa_k + \kappa_{\bar{k}}} \left(\begin{array}{c} V_{k,t\Delta} V_{\bar{k},t\Delta} (1 - e^{-(\kappa_k + \kappa_{\bar{k}})(u-t\Delta)}) \\ + \kappa_k \theta_k \int_{t\Delta}^u (1 - e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)}) V_{\bar{k},s} ds \\ + \kappa_{\bar{k}} \theta_{\bar{k}} \int_{t\Delta}^u (1 - e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)}) V_{k,s} ds \\ + \int_{t\Delta}^u (1 - e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)}) \mathbb{1}_{k=\bar{k}} \sigma_k \sigma_{\bar{k}} \sqrt{V_{k,s}} \sqrt{V_{\bar{k},s}} ds \end{array} \right) \middle| \mathcal{F}_{t\Delta} \right].$$

Proof. Replacing $V_{k,u} V_{\bar{k},u}$ in Corollary D.3.3 by the results of Corollary D.3.2 and taking the expectation. \square

Lemma D.3.5. For $u > t\Delta$,

$$\begin{aligned} & \mathbb{E} \left[\int_{t\Delta}^u e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} V_{k,s} ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= (V_{k,t\Delta} - \theta_k) e^{-\kappa_k(u-t\Delta)} \frac{1 - e^{-\kappa_{\bar{k}}(u-t\Delta)}}{\kappa_{\bar{k}}} + \theta_k \frac{1 - e^{-(\kappa_k + \kappa_{\bar{k}})(u-t\Delta)}}{\kappa_k + \kappa_{\bar{k}}}, \end{aligned}$$

Proof. From Equation (B.8),

$$\begin{aligned} & \mathbb{E} \left[\int_{t\Delta}^u e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} V_{k,s} ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \mathbb{E} \left[\int_{t\Delta}^u e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} \left(\begin{array}{c} e^{-\kappa_k(s-t\Delta)} V_{k,t\Delta} + \theta_k (1 - e^{-\kappa_k(s-t\Delta)}) \\ + \sigma_k e^{-\kappa_k s} \int_{t\Delta}^s e^{\kappa_k r} \sqrt{V_{k,r}} dW_{k,r} \end{array} \right) ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \int_{t\Delta}^u e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} (e^{-\kappa_k(s-t\Delta)} V_{k,t\Delta} + \theta_k (1 - e^{-\kappa_k(s-t\Delta)})) ds \\ &= \int_{t\Delta}^u (V_{k,t\Delta} - \theta_k) e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} e^{-\kappa_k(s-t\Delta)} + \theta_k e^{-(\kappa_k + \kappa_{\bar{k}})(u-s)} ds, \end{aligned}$$

leading to the result. \square

Lemma D.3.6. For $0 < t < u$,

$$\mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\bar{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] = \mathcal{A}_{k,\bar{k}} + \mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta},$$

where

$$\begin{aligned}
 \mathcal{A}_{k,\tilde{k}} &= \frac{\kappa_{\tilde{k}}\theta_{\tilde{k}} + \mathbb{1}_{k=\tilde{k}}\sigma_k\sigma_{\tilde{k}}}{(\kappa_k + \kappa_{\tilde{k}})}\theta_k \left(\Delta - \frac{1 - e^{-(\kappa_k + \kappa_{\tilde{k}})\Delta}}{(\kappa_k + \kappa_{\tilde{k}})} \right) - \theta_k \mathcal{D}_{k,\tilde{k}} \\
 &\quad + \frac{\kappa_k\theta_k}{(\kappa_k + \kappa_{\tilde{k}})}\theta_{\tilde{k}} \left(\Delta - \frac{1 - e^{-(\kappa_k + \kappa_{\tilde{k}})\Delta}}{(\kappa_k + \kappa_{\tilde{k}})} \right) - \theta_{\tilde{k}} \mathcal{E}_{k,\tilde{k}}, \\
 \mathcal{B}_{k,\tilde{k}} &= \frac{(1 - e^{-(\kappa_k + \kappa_{\tilde{k}})\Delta})}{(\kappa_k + \kappa_{\tilde{k}})}, \\
 \mathcal{D}_{k,\tilde{k}} &= \frac{\kappa_{\tilde{k}}\theta_{\tilde{k}} + \mathbb{1}_{k=\tilde{k}}\sigma_k\sigma_{\tilde{k}}}{(\kappa_k + \kappa_{\tilde{k}})} \left(\frac{1 - e^{-\kappa_k\Delta}}{\kappa_k} - e^{-\kappa_k\Delta} \frac{1 - e^{-\kappa_{\tilde{k}}\Delta}}{\kappa_{\tilde{k}}} \right), \\
 \mathcal{E}_{k,\tilde{k}} &= \frac{\kappa_k\theta_k}{(\kappa_k + \kappa_{\tilde{k}})} \left(\frac{1 - e^{-\kappa_{\tilde{k}}\Delta}}{\kappa_{\tilde{k}}} - e^{-\kappa_{\tilde{k}}\Delta} \frac{1 - e^{-\kappa_k\Delta}}{\kappa_k} \right).
 \end{aligned}$$

Proof. Direct application from Lemmas D.3.4 and D.3.5 where

$$\begin{aligned}
 &\mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} (1 - e^{-(\kappa_k + \kappa_{\tilde{k}})((t+1)\Delta - s)}) V_{k,s} ds \middle| \mathcal{F}_{t\Delta} \right] \\
 &= (V_{k,t\Delta} - \theta_k) \left(\frac{1 - e^{-\kappa_k\Delta}}{\kappa_k} - e^{-\kappa_k\Delta} \frac{1 - e^{-\kappa_{\tilde{k}}\Delta}}{\kappa_{\tilde{k}}} \right) + \theta_k \left(\Delta - \frac{1 - e^{-(\kappa_k + \kappa_{\tilde{k}})\Delta}}{(\kappa_k + \kappa_{\tilde{k}})} \right). \quad \square
 \end{aligned}$$

D.3.3 Derivation of $\mathbb{E} \left[\int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} V_{\hat{k},s} ds \middle| \mathcal{F}_{t\Delta} \right]$

Lemma D.3.7. The SDE of $V_k V_{\tilde{k}} V_{\hat{k}}$ is

$$\begin{aligned}
 &dV_{k,t} V_{\tilde{k},t} V_{\hat{k},t} \\
 &= V_{\tilde{k},t} V_{\hat{k},t} \kappa_k (\theta_k - V_{k,t}) dt + V_{k,t} V_{\hat{k},t} \kappa_{\tilde{k}} (\theta_{\tilde{k}} - V_{\tilde{k},t}) dt + V_{k,t} V_{\tilde{k},t} \kappa_{\hat{k}} (\theta_{\hat{k}} - V_{\hat{k},t}) dt \\
 &\quad + V_{k,t} V_{\tilde{k},t} \sigma_{\hat{k}} \sqrt{V_{\hat{k},t}} dW_{\hat{k},t} + V_{\tilde{k},t} V_{\hat{k},t} \sigma_k \sqrt{V_{k,t}} dW_{k,t} + V_{k,t} V_{\hat{k},t} \sigma_{\tilde{k}} \sqrt{V_{\tilde{k},t}} dW_{\tilde{k},t} \\
 &\quad + V_{\tilde{k},t} \mathbb{1}_{k=\tilde{k}} \sigma_k \sigma_{\tilde{k}} \sqrt{V_{k,t}} \sqrt{V_{\tilde{k},t}} dt + V_{k,t} \mathbb{1}_{k=\hat{k}} \sigma_k \sigma_{\hat{k}} \sqrt{V_{k,t}} \sqrt{V_{\hat{k},t}} dt + V_{k,t} \mathbb{1}_{\hat{k}=\tilde{k}} \sigma_{\hat{k}} \sigma_{\tilde{k}} \sqrt{V_{\hat{k},t}} \sqrt{V_{\tilde{k},t}} dt.
 \end{aligned}$$

Proof. Direct application of Itô's lemma to $V_k V_{\tilde{k}} V_{\hat{k}}$. □

Corollary D.3.8. For $0 < t < u$,

$$\begin{aligned} & \mathbb{E} \left[\int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} V_{\hat{k},s} ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \frac{1}{\kappa_k + \kappa_{\tilde{k}} + \kappa_{\hat{k}}} \mathbb{E} \left[\begin{array}{c} V_{k,t\Delta} V_{\tilde{k},t\Delta} V_{\hat{k},t\Delta} - V_{k,u} V_{\tilde{k},u} V_{\hat{k},u} \\ + (\kappa_k \theta_k + \mathbb{1}_{k=\tilde{k}} \sigma_k \sigma_{\tilde{k}}) \int_{t\Delta}^u V_{\tilde{k},s} V_{\hat{k},s} ds \\ + (\kappa_{\tilde{k}} \theta_{\tilde{k}} + \mathbb{1}_{\tilde{k}=\hat{k}} \sigma_{\tilde{k}} \sigma_{\hat{k}}) \int_{t\Delta}^u V_{k,s} V_{\hat{k},s} ds \\ + (\kappa_{\hat{k}} \theta_{\hat{k}} + \mathbb{1}_{k=\hat{k}} \sigma_k \sigma_{\hat{k}}) \int_{t\Delta}^u V_{k,s} V_{\tilde{k},s} ds \end{array} \middle| \mathcal{F}_{t\Delta} \right]. \end{aligned}$$

Proof. Direct application of Lemma D.3.7 by integrating the result and taking the expectation. \square

D.4 Kalman Filter Comparison

This section delves into a comparative analysis between our filter and the standard Kalman Filter (KF) first introduced by Kalman, 1960.

The Kalman filter is designed for a linear model like

$$\mathbf{V}_{t+1} = \mathbf{A}_t + \mathbf{B}_t \mathbf{V}_t + \mathbf{C}_t \mathbf{E}_{t+1}, \quad (\text{D.9})$$

$$\mathbf{O}_{t+1} = \mathbf{a}_t + \mathbf{b}_t \mathbf{V}_t + \mathbf{c}_t \mathbf{e}_{t+1}. \quad (\text{D.10})$$

where $\mathbf{E}_{t+1} | \mathcal{G}_t$ and $\mathbf{e}_{t+1} | \mathcal{G}_t$ are independent multivariate random vectors. These noise vectors are also assumed to be independent from \mathbf{V}_t . The standard KF proposes an estimate $\mathbf{V}_{t+1|t+1}$ for the latent variable \mathbf{V}_{t+1} , accompanied by its conditional variance $\mathbf{P}_{t+1|t+1}$ with

$$\begin{aligned} \mathbf{V}_{t+1|t+1} &= \mathbb{E}[\mathbf{V}_{t+1} | \mathcal{G}_t] + \Sigma_t (\mathbf{O}_{t+1} - \mathbb{E}[\mathbf{O}_{t+1} | \mathcal{G}_t]), \\ \mathbf{P}_{t+1|t+1} &= \text{Var}[\mathbf{V}_{t+1} | \mathcal{G}_t] - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t]. \end{aligned}$$

As discuss in Uhlmann and Julier, 2022, the common assumption of a Gaussian distribution of the noise terms in the state and measurement equations is not a prerequisite for the Kalman Filter, although having a Gaussian distribution leads to an optimal estimator.

The update process, derived in Section 3.3, is computed by minimizing the mean squared error (MMSE) of Equation (3.25), which does not necessitate any distributional assumption. The update matrix, often referred to as the Kalman update, derived in Equation (3.26), is

$$\Sigma_t = \text{Cov}[\mathbf{V}_{t+1}, \mathbf{O}_{t+1} | \mathcal{G}_t] (\text{Var}[\mathbf{O}_{t+1} | \mathcal{G}_t])^{-1}. \quad (\text{D.11})$$

Proposition D.4.1. *Let $\mathbf{Y}_{t+1} = \mathbf{V}_{t+1} - \Sigma_t \mathbf{O}_{t+1}$ where Σ_t is defined in Equation (D.11).*

$$\begin{aligned} \text{Cov}[\mathbf{O}_{t+1}, \mathbf{Y}_{t+1} | \mathcal{G}_t] &= \mathbf{0}, \\ \text{E}[\mathbf{Y}_{t+1} | \mathcal{G}_t] &= \text{E}[\mathbf{V}_{t+1} | \mathcal{G}_t] - \Sigma_t \text{E}[\mathbf{O}_{t+1} | \mathcal{G}_t], \\ \text{Var}[\mathbf{Y}_{t+1} | \mathcal{G}_t] &= \text{Var}[\mathbf{V}_{t+1} | \mathcal{G}_t] - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t]. \end{aligned}$$

Proof: The covariance is

$$\begin{aligned} &\text{Cov}[\mathbf{O}_{t+1}, \mathbf{Y}_{t+1} | \mathcal{G}_t] \\ &= \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} - \Sigma_t \mathbf{O}_{t+1} | \mathcal{G}_t] \\ &= \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t] - \text{Var}[\mathbf{O}_{t+1} | \mathcal{G}_t] \Sigma_t^\top \\ &= \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t] - \text{Var}[\mathbf{O}_{t+1} | \mathcal{G}_t] (\text{Cov}[\mathbf{V}_{t+1}, \mathbf{O}_{t+1} | \mathcal{G}_t] (\text{Var}[\mathbf{O}_{t+1} | \mathcal{G}_t])^{-1})^\top \\ &= \mathbf{0}, \end{aligned}$$

and the variance

$$\begin{aligned} \text{Var}[\mathbf{Y}_{t+1} | \mathcal{G}_t] &= \text{Var}[\mathbf{V}_{t+1} | \mathcal{G}_t] + \Sigma_t \text{Var}[\mathbf{O}_{t+1} | \mathcal{G}_t] \Sigma_t^\top \\ &\quad - \text{Cov}[\mathbf{V}_{t+1}, \mathbf{O}_{t+1} | \mathcal{G}_t] \Sigma_t^\top - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t] \\ &= \text{Var}[\mathbf{V}_{t+1} | \mathcal{G}_t] - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1} | \mathcal{G}_t]. \quad \square \end{aligned}$$

Assumption D.4.2. *With Equations (D.9) and (D.10),*

$$\xi_{t+1} | \mathcal{G}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_k), \mathbf{e}_{t+1} | \mathcal{G}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_m),$$

and are independent from each other and from \mathbf{V}_t conditional to \mathcal{G}_t .

Corollary D.4.3. *Based on the linear framework of Equations (D.9) and (D.10), if \mathbf{V}_0 is normal and Assumption (D.4.2) holds, it implies that \mathbf{O}_{t+1} , \mathbf{V}_{t+1} and \mathbf{Y}_{t+1} conditional to \mathcal{G}_t follow a multivariate normal distribution.*

Proof: Direct application of the normal distribution properties. □

Corollary D.4.3 is based on the linear condition of the framework, which we do not have in the main paper.

Corollary D.4.4. *If $\mathbf{Y}_{t+1} = \mathbf{V}_{t+1} - \Sigma_t \mathbf{O}_{t+1}$, where Σ_t is defined in Equation (D.11), and Assumption D.4.2 is true, then \mathbf{Y}_{t+1} and \mathbf{O}_{t+1} are independent when conditional on $\mathcal{G}_{t\Delta}$.*

Proof: From Assumption D.4.2 and Corollary D.4.3, \mathbf{Y}_{t+1} and \mathbf{O}_{t+1} is multivariate normal with null conditional covariance from Proposition D.4.1. Therefore, it implies that \mathbf{Y}_{t+1} is independent from \mathbf{O}_{t+1} by the properties of the multivariate normal distribution. □

Corollary D.4.5. *If Assumption (D.4.2) is true, then*

$$\begin{aligned} \mathbb{E}[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}] &= \mathbb{E}[\mathbf{V}_{t+1}|\mathcal{G}_t] + \Sigma_t(\mathbf{O}_{t+1} - \mathbb{E}[\mathbf{O}_{t+1}|\mathcal{G}_t]) = \mathbf{V}_{t+1|t+1}, \\ \text{Var}[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}] &= \text{Var}[\mathbf{V}_{t+1}|\mathcal{G}_t] - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1}|\mathcal{G}_t] = \mathbf{P}_{t+1|t+1}. \end{aligned}$$

Proof: When conditioned on \mathcal{G}_t , \mathbf{Y}_{t+1} is independent of \mathbf{O}_{t+1} , which implies that

$$\begin{aligned} \mathbb{E}[\mathbf{Y}_{t+1}|\mathcal{G}_t] &= \mathbb{E}[\mathbf{Y}_{t+1}|\mathcal{G}_{t+1}], \\ \text{Var}[\mathbf{Y}_{t+1}|\mathcal{G}_t] &= \text{Var}[\mathbf{Y}_{t+1}|\mathcal{G}_{t+1}]. \end{aligned}$$

Therefore, from the first equality we conclude that

$$\begin{aligned} \mathbb{E}[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}] &= \mathbb{E}[\mathbf{Y}_{t+1} + \Sigma_t \mathbf{O}_{t+1}|\mathcal{G}_{t+1}] \\ &= \mathbb{E}[\mathbf{Y}_{t+1}|\mathcal{G}_t] + \Sigma_t \mathbf{O}_{t+1} \\ &= \mathbb{E}[\mathbf{V}_{t+1}|\mathcal{G}_t] + \Sigma_t (\mathbf{O}_{t+1} - \mathbb{E}[\mathbf{O}_{t+1}|\mathcal{G}_t]). \end{aligned}$$

Furthermore,

$$\begin{aligned}
 \text{Var}[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}] &= \text{Var}[\mathbf{Y}_{t+1} + \Sigma_t \mathbf{O}_{t+1}|\mathcal{G}_{t+1}] \\
 &= \text{Var}[\mathbf{Y}_{t+1}|\mathcal{G}_t] \\
 &= \text{Var}[\mathbf{V}_{t+1} - \Sigma_t \mathbf{O}_{t+1}|\mathcal{G}_t] \\
 &= \text{Var}[\mathbf{V}_{t+1}|\mathcal{G}_t] - \Sigma_t \text{Cov}[\mathbf{O}_{t+1}, \mathbf{V}_{t+1}|\mathcal{G}_t]. \quad \square
 \end{aligned}$$

If Assumption (D.4.2) is true, then the KF estimator $\mathbf{V}_{t+1|t+1}$ and its variance $\mathbf{P}_{t+1|t+1}$ equal to the exact conditional moments of the latent state $E[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}]$ and $\text{Var}[\mathbf{V}_{t+1}|\mathcal{G}_{t+1}]$ respectively. In Corollary D.4.5, the variance conditioned on \mathcal{G}_{t+1} is equal to the variance conditioned on \mathcal{G}_t , as opposed to Equation (3.27). This difference arises as the Assumption D.4.2 is not respected in our framework. The filter proposed in Section 3.3 minimized the mean squared error of our estimate. However, since it does not benefit from linearity and Gaussian distribution, it does not yield optimal estimates. In our context, the absence of the Assumption D.4.2 implies that we cannot conclude that $\mathbf{V}_{t+1} - \Sigma_t \mathbf{O}_{t+1}$ and \mathbf{O}_{t+1} are independent. They are simply uncorrelated.

To summarize, our filter derived in section 3.3 does not follow the standard KF in two ways 1) the noises are not Gaussian and 2) the state and measurement equations are not linear in the latent state. Nevertheless, we achieved to obtain analytical conditional moments of the latent state and observable variables, despite its non-linearity. Consequently, we can compute the update matrix based on the minimization of mean squared error, as derived in Equations (3.25) and (3.26), without necessitating any additional assumptions. Due to the absence of a Gaussian distribution, we derive an additional update for Equation (3.27), which is detailed in Online Appendix D.8.

D.5 Derivation of Third and Fourth Conditional Cumulant

This Online Appendix provides the complete derivations and proofs of Appendix B.4.1.3 and B.4.2.3, detailing how we compute the third and fourth cumulants including the mea-

surement error.

D.5.1 Measurement error third cumulant with respect to the observed filtration

We need to add the measurement error on the third central moments to transit from $\mathbf{O}_{(t+1)\Delta} - \mathbf{Q}\xi_{t,t+1}$ to $\mathbf{O}_{(t+1)\Delta}$. We then derive the third central moments involving to the measurement error with respect to the observed filtration with the following corollary:

Corollary D.5.1. *For $\mathcal{X}_{(t+1)\Delta} \in \{\mathbf{O}_{(t+1)\Delta} - \mathbf{Q}\xi_{t,t+1}, \mathbf{V}_{(t+1)\Delta}\}$, then*

$$\begin{aligned} & \text{Cum}^{(3)} [\xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{E} \left[\text{Cum}^{(3)} [\xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] \middle| \mathcal{G}_{t\Delta} \right], \\ & \text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] = \text{E} \left[\text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cov} [\text{Cov} [\xi_{t,t+1}, \xi_{t,t+1} | \mathcal{F}_{t\Delta}], \text{E} [\mathcal{X}_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}], \\ & \text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1} | \mathcal{G}_{t\Delta}] = \text{E} \left[\text{Cum}^{(3)} [\xi_{t,t+1}, \xi_{t,t+1}, \xi_{t,t+1} | \mathcal{F}_{t\Delta}] \middle| \mathcal{G}_{t\Delta} \right]. \end{aligned}$$

Proof: Direct application of the law of total cumulance (see Lemma B.4.2) and Corollary B.1.1. □

Corollary D.5.2. *For $\mathcal{X}_{l,(t+1)\Delta} \in \{\mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta}\}$, if*

$$\text{E} [\mathcal{X}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] = \beta_l \left(\sum_{\hat{k}=1}^{N_V} \alpha_{\hat{k}}^{(1)} V_{\hat{k},t\Delta} + \alpha_{\hat{k}}^{(0)} \right),$$

then

$$\begin{aligned} & \text{Cov} [\text{Cov} [\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} | \mathcal{F}_{t\Delta}], \text{E} [\mathcal{X}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \eta \text{Cov} \left[\beta_l \left(\sum_{\hat{k}=1}^{N_V} \alpha_{\hat{k}}^{(1)} V_{\hat{k},t\Delta} \right), \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} \left(\begin{array}{c} \mathcal{B}_{k,\tilde{k}} V_{k,t\Delta} V_{\tilde{k},t\Delta} \\ + \mathcal{D}_{k,\tilde{k}} V_{k,t\Delta} + \mathcal{E}_{k,\tilde{k}} V_{\tilde{k},t\Delta} \end{array} \right) \middle| \mathcal{G}_{t\Delta} \right], \end{aligned}$$

which is a linear sum of the moments of the latent variables.²

Proof: Direct application of Equation (B.5). □

D.5.2 Adding the measurement error third cumulant

The measurement error $\xi_{t,t+1}$ is included in the third cumulant with respect to the observed filtration as follows

$$\begin{aligned} & \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\mathbf{Q}_i \xi_{t,t+1}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}], \end{aligned}$$

$$\begin{aligned} & \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta}, \mathbf{O}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\mathbf{Q}_i \xi_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\mathbf{Q}_i \xi_{t,t+1}, \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}], \end{aligned}$$

$$\begin{aligned} & \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta}, \mathbf{O}_{j,(t+1)\Delta}, \mathbf{O}_{l,(t+1)\Delta} | \mathcal{G}_{t\Delta}] \\ &= \text{Cum}^{(3)} [\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \xi_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \xi_{t,t+1} | \mathcal{G}_{t\Delta}] \\ &+ \text{Cum}^{(3)} [\mathbf{Q}_i \xi_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \xi_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \xi_{t,t+1} | \mathcal{G}_{t\Delta}] \end{aligned}$$

²These equations can be rewritten as a function of the central moments:

$$\begin{aligned} \text{Var} [X, YZ] &= \text{E} [(X - \text{E}[X])(YZ - \text{E}[YZ])] \\ &= \text{E} [(X - \text{E}[X])(YZ)] \\ &= \text{E} [(X - \text{E}[X]) ((Y - \text{E}[Y])(Z - \text{E}[Z]) + Y\text{E}[Z] + Z\text{E}[Y] - \text{E}[Y]\text{E}[Z])] \\ &= \text{E} [(X - \text{E}[X]) ((Y - \text{E}[Y])(Z - \text{E}[Z]) + Y\text{E}[Z] + Z\text{E}[Y])] \\ &= \text{Cum}^{(3)} [X, Y, Z] + \text{E} [(X - \text{E}[X]) ((Y - \text{E}[Y])\text{E}[Z] + (Z - \text{E}[Z])\text{E}[Y])] \\ &= \text{Cum}^{(3)} [X, Y, Z] + \text{Var} [X, Y] \text{E}[Z] + \text{Var} [X, Z] \text{E}[Y]. \end{aligned}$$

$$\begin{aligned}
& + \text{Cum}^{(3)} [O_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, O_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{G}_{t\Delta}] \\
& + \text{Cum}^{(3)} [\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, O_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{G}_{t\Delta}] \\
& + \text{Cum}^{(3)} [\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, O_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{G}_{t\Delta}] \\
& + \text{Cum}^{(3)} [O_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{G}_{t\Delta}] \\
& + \text{Cum}^{(3)} [\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} | \mathcal{G}_{t\Delta}].
\end{aligned}$$

D.5.3 Measurement error fourth cumulant with respect to the observed filtration

We need to add the measurement error on the fourth cumulant to transit from $\mathbf{O}_{(t+1)\Delta} - \mathbf{Q}\boldsymbol{\xi}_{t,t+1}$ to $\mathbf{O}_{(t+1)\Delta}$. We then derive the fourth cumulant involving to the measurement error with respect to the observed filtration with the following corollary:

Corollary D.5.3. For $\boldsymbol{\mathcal{X}}_{(t+1)\Delta} \in \{\mathbf{O}_{(t+1)\Delta} - \mathbf{Q}\boldsymbol{\xi}_{t,t+1}, \mathbf{V}_{(t+1)\Delta}\}$, if Assumption 3.4.1 holds, then

$$\begin{aligned}
\text{Cum}^{(4)} [\boldsymbol{\xi}_{t,t+1}, \boldsymbol{\mathcal{X}}_{(t+1)\Delta}, \boldsymbol{\mathcal{X}}_{(t+1)\Delta}, \boldsymbol{\mathcal{X}}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] &= 0, \\
\text{Cum}^{(4)} [\boldsymbol{\xi}_{t,t+1}, \boldsymbol{\xi}_{t,t+1}, \boldsymbol{\xi}_{t,t+1}, \boldsymbol{\mathcal{X}}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}] &= 0.
\end{aligned}$$

Proof: Direct application of the law of total cumulance (see Lemma B.4.7). □

Corollary D.5.4. For $\boldsymbol{\mathcal{X}}_{l,(t+1)\Delta} \in \{O_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, V_{l,(t+1)\Delta}\}$, if

$$\begin{aligned}
\mathbb{E} [\boldsymbol{\mathcal{X}}_{l,(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \beta_l \left(\sum_{\hat{k}=1}^{N_V} \alpha_{\hat{k}}^{(1)} V_{\hat{k},t\Delta} + \alpha_{\hat{k}}^{(0)} \right), \\
\text{Var} [\boldsymbol{\mathcal{X}}_{l,(t+1)\Delta}, \boldsymbol{\mathcal{X}}_{m,(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \phi_{l,m} \left(\sum_{\bar{k}=1}^{N_V} \varphi_{\bar{k}}^{(1)} V_{\bar{k},t\Delta} + \varphi_{\bar{k}}^{(0)} \right),
\end{aligned}$$

and if Assumption 3.4.1 is respected, then

$$\text{Cum}^{(4)} [\xi_{i,j,t,t+1}, \xi_{\bar{i},\bar{j},t,t+1}, \boldsymbol{\mathcal{X}}_{l,(t+1)\Delta}, \boldsymbol{\mathcal{X}}_{m,(t+1)\Delta} | \mathcal{G}_{t\Delta}]$$

$$\begin{aligned}
 &= \eta \text{Cov} \left[\phi_{l,m} \left(\sum_{\bar{k}=1}^{N_V} \varphi_{\bar{k}}^{(1)} V_{\bar{k},t\Delta} \right), \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{\bar{k},\bar{k}}^{i,j,\tilde{i},\tilde{j}} \left(\begin{array}{c} \mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} \\ + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \end{array} \right) \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \eta \text{Cum}^{(3)} \left[\begin{array}{c} \beta_l \left(\sum_{\hat{k}=1}^{N_V} \alpha_{\hat{k}}^{(1)} V_{\hat{k},t\Delta} \right), \beta_m \left(\sum_{\bar{k}=1}^{N_V} \alpha_{\bar{k}}^{(1)} V_{\bar{k},t\Delta} \right), \\ \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{\bar{k},\bar{k}}^{i,j,\tilde{i},\tilde{j}} \left(\mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right],
 \end{aligned}$$

which is a linear sum of the moments of the latent variables.³

Proof: With the law of total cumulance (see Lemma B.4.7), the non-null terms are

$$\begin{aligned}
 &\text{Cum}^{(4)} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \mathcal{X}_{l,(t+1)\Delta}, \mathcal{X}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Var} \left[\text{Var} \left[\mathcal{X}_{l,(t+1)\Delta}, \mathcal{X}_{m,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right], \text{Var} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(3)} \left[\text{E} \left[\mathcal{X}_{l,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right], \text{E} \left[\mathcal{X}_{m,(t+1)\Delta} \middle| \mathcal{F}_{t\Delta} \right], \text{Var} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right],
 \end{aligned}$$

leading to the result. \square

Corollary D.5.5. *If Assumption 3.4.1 is respected, then*

$$\begin{aligned}
 &\text{Cum}^{(4)} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{l,m,t,t+1}, \xi_{\tilde{l},\tilde{m},t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \eta^2 \text{Cov} \left[\begin{array}{c} \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{\bar{k},\bar{k}}^{l,m,\tilde{l},\tilde{m}} \left(\mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \right), \\ \sum_{\hat{k}=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\tilde{k}}^{i,j,\tilde{i},\tilde{j}} \left(\mathcal{B}_{\hat{k},\tilde{k}} V_{\hat{k},t\Delta} V_{\tilde{k},t\Delta} + \mathcal{D}_{\hat{k},\tilde{k}} V_{\hat{k},t\Delta} + \mathcal{E}_{\hat{k},\tilde{k}} V_{\tilde{k},t\Delta} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \eta^2 \text{Cov} \left[\begin{array}{c} \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{\bar{k},\bar{k}}^{l,m,i,j} \left(\mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \right), \\ \sum_{\hat{k}=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\tilde{k}}^{l,\tilde{m},\tilde{i},\tilde{j}} \left(\mathcal{B}_{\hat{k},\tilde{k}} V_{\hat{k},t\Delta} V_{\tilde{k},t\Delta} + \mathcal{D}_{\hat{k},\tilde{k}} V_{\hat{k},t\Delta} + \mathcal{E}_{\hat{k},\tilde{k}} V_{\tilde{k},t\Delta} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right]
 \end{aligned}$$

³These equations can be rewritten as a function of the central moments:

$$\begin{aligned}
 &\text{Cum}^{(3)} [X, W, YZ] \\
 &= \text{E} [(X - \text{E}[X])(W - \text{E}[W])(YZ - \text{E}[YZ])] \\
 &= \text{E} \left[\begin{array}{c} (X - \text{E}[X])(W - \text{E}[W]) \\ ((Y - \text{E}[Y])(Z - \text{E}[Z]) - \text{E}[YZ] + Y\text{E}[Z] + Z\text{E}[Y] - \text{E}[Y]\text{E}[Z]) \end{array} \right] \\
 &= \text{E} [(X - \text{E}[X])(W - \text{E}[W])(Y - \text{E}[Y])(Z - \text{E}[Z])] \\
 &+ \text{Cum}^{(3)} [X, W, Y] \text{E}[Z] + \text{Cum}^{(3)} [X, W, Z] \text{E}[Y] - \text{Var}[X, W] \text{Var}[Y, Z].
 \end{aligned}$$

$$+ \eta^2 \text{Cov} \left[\begin{array}{c} \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{k,\bar{k}}^{l,m,\bar{i},\bar{j}} \left(\mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \right), \\ \sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{C}_{k,\bar{k}}^{i,j,\bar{l},\bar{m}} \left(\mathcal{B}_{k,\bar{k}} V_{k,t\Delta} V_{\bar{k},t\Delta} + \mathcal{D}_{k,\bar{k}} V_{k,t\Delta} + \mathcal{E}_{k,\bar{k}} V_{\bar{k},t\Delta} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right],$$

which is a linear sum of the moments of the latent variables.⁴

Proof: With the law of total cumulance (see Lemma B.4.7 for application), the non-null terms are

$$\begin{aligned} & \text{Cum}^{(4)} \left[\xi_{i,j,t,t+1}, \xi_{\bar{i},\bar{j},t,t+1}, \xi_{l,m,t,t+1}, \xi_{\bar{l},\bar{m},t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\ &= \text{Cov} \left[\text{Cov} \left[\xi_{l,m,t,t+1}, \xi_{\bar{l},\bar{m},t,t+1} \middle| \mathcal{F}_{t\Delta} \right], \text{Cov} \left[\xi_{i,j,t,t+1}, \xi_{\bar{i},\bar{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cov} \left[\text{Cov} \left[\xi_{i,j,t,t+1}, \xi_{\bar{l},\bar{m},t,t+1} \middle| \mathcal{F}_{t\Delta} \right], \text{Cov} \left[\xi_{l,m,t,t+1}, \xi_{\bar{i},\bar{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cov} \left[\text{Cov} \left[\xi_{\bar{i},\bar{j},t,t+1}, \xi_{\bar{l},\bar{m},t,t+1} \middle| \mathcal{F}_{t\Delta} \right], \text{Cov} \left[\xi_{i,j,t,t+1}, \xi_{l,m,t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \middle| \mathcal{G}_{t\Delta} \right], \end{aligned}$$

leading to the results. □

D.5.4 Adding the measurement error fourth cumulant

Therefore, the measurement error $\xi_{t,t+1}$ is included in the fourth cumulant with respect to the observed filtration as follows

$$\begin{aligned} & \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\ &= \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{V}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right], \end{aligned}$$

$$\begin{aligned} & \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta}, \mathbf{O}_{j,(t+1)\Delta}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\ &= \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \xi_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\ &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \xi_{t,t+1}, \mathbf{Q}_j \xi_{t,t+1}, \mathbf{V}_{l,(t+1)\Delta}, \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right], \end{aligned}$$

⁴These equations can be rewritten as a function of the central moments:

$$\text{Cov} [XW, YZ] = \text{Cum}^{(4)} [X, W, Y, Z] + \text{Cov} [X, Y] \text{Cov} [W, Z] + \text{Cov} [X, Z] \text{Cov} [W, Y]$$

$$\begin{aligned}
 & \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta}, \mathbf{O}_{j,(t+1)\Delta}, \mathbf{O}_{l,(t+1)\Delta} \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{V}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right],
 \end{aligned}$$

$$\begin{aligned}
 & \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta}, \mathbf{O}_{j,(t+1)\Delta}, \mathbf{O}_{l,(t+1)\Delta} \mathbf{O}_{m,(t+1)\Delta} \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Cum}^{(4)} \left[\begin{array}{l} \mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \\ \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{O}_{m,(t+1)\Delta} - \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \end{array} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{O}_{m,(t+1)\Delta} - \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{O}_{m,(t+1)\Delta} - \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1} \mathbf{O}_{m,(t+1)\Delta} - \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{l,(t+1)\Delta} - \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{O}_{i,(t+1)\Delta} - \mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{O}_{j,(t+1)\Delta} - \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right] \\
 &+ \text{Cum}^{(4)} \left[\mathbf{Q}_i \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_j \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_l \boldsymbol{\xi}_{t,t+1}, \mathbf{Q}_m \boldsymbol{\xi}_{t,t+1} \middle| \mathcal{G}_{t\Delta} \right].
 \end{aligned}$$

D.6 Higher Moments of Stochastic Integrals

D.6.1 General result

This Online Appendix derives the cumulant of key quantities. The subscript k is omitted for readability. We denote $\mathcal{W}_{t,t+\frac{\Delta}{u}} = \int_{t\Delta}^{t\Delta+u} \sqrt{V_s} dW_s$, $\mathcal{B}_{t,t+\frac{\Delta}{u}} = \int_{t\Delta}^{t\Delta+u} \sqrt{V_s} dB_s$ and $\mathcal{Z}_{t,t+\frac{\Delta}{u}} = \int_{t\Delta}^{t\Delta+u} e^{-\kappa((t\Delta+u)-s)} \sqrt{V_s} dW_s = e^{-\kappa(t\Delta+u)} \mathcal{Y}_{t,t+\frac{\Delta}{u}}$ where $\mathcal{Y}_{t,t+\frac{\Delta}{u}} = \int_{t\Delta}^{t\Delta+u} e^{\kappa s} \sqrt{V_s} dW_s$. To compute the higher moments related to observations and latent variables, we develop the following general result:

Lemma D.6.1. For $n, p, m, l \in \mathbb{N}$, and

$$\mathcal{X}_{t,t+\frac{u}{\Delta}} = \mathcal{Y}_{t,t+\frac{u}{\Delta}}^n \mathcal{W}_{t,t+\frac{u}{\Delta}}^p \mathcal{B}_{t,t+\frac{u}{\Delta}}^l e^{m\kappa u} V_{t\Delta+u}^m,$$

we have

$$\begin{aligned} & \mathbb{E}[\mathcal{X}_{t,t+1} | \mathcal{F}_{t\Delta}] - \mathcal{X}_{t,t} \\ &= \left(m\kappa\theta + \frac{1}{2}m(m-1)\sigma^2 \right) \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} V_{t\Delta+u}^{-1} | \mathcal{F}_{t\Delta}] du \\ &+ \frac{1}{2}n(n-1) \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-2} V_{t\Delta+u} | \mathcal{F}_{t\Delta}] e^{2\kappa su} du \\ &+ \frac{1}{2}p(p-1) \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-2} V_{t\Delta+u} | \mathcal{F}_{t\Delta}] du \\ &+ \frac{1}{2}l(l-1) \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{B}_{t,t+\frac{u}{\Delta}}^{-2} V_{t\Delta+u} | \mathcal{F}_{t\Delta}] du \\ &+ np \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1} V_{t\Delta+u} | \mathcal{F}_{t\Delta}] e^{\kappa su} du \\ &+ nm\sigma \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} | \mathcal{F}_{t\Delta}] e^{\kappa su} du \\ &+ pm\sigma \int_0^\Delta \mathbb{E}[\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1} | \mathcal{F}_{t\Delta}] du. \end{aligned}$$

Proof. The stochastic differential equation of $\mathcal{X}_{t,t+\frac{u}{\Delta}}$ is

$$\begin{aligned} & d\mathcal{X}_{t,t+\frac{u}{\Delta}} \\ &= n\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} d\mathcal{Y}_{t,t+\frac{u}{\Delta}} + p\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1} d\mathcal{W}_{t,t+\frac{u}{\Delta}} \\ &+ l\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{B}_{t,t+\frac{u}{\Delta}}^{-1} d\mathcal{B}_{t,t+\frac{u}{\Delta}} + m\mathcal{X}_{t,t+\frac{u}{\Delta}} V_{t\Delta+u}^{-1} dV_{t\Delta+u} + m\kappa\mathcal{X}_{t,t+\frac{u}{\Delta}} du \\ &+ \frac{1}{2}n(n-1)\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-2} d\langle \mathcal{Y}, \mathcal{Y} \rangle_u + \frac{1}{2}p(p-1)\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-2} d\langle \mathcal{W}, \mathcal{W} \rangle_u \\ &+ \frac{1}{2}l(l-1)\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{B}_{t,t+\frac{u}{\Delta}}^{-2} d\langle \mathcal{B}, \mathcal{B} \rangle_u + \frac{1}{2}m(m-1)\mathcal{X}_{t,t+\frac{u}{\Delta}} V_{t\Delta+u}^{-2} d\langle V, V \rangle_u \\ &+ np\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} \mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1} d\langle \mathcal{Y}, \mathcal{W} \rangle_u + nm\mathcal{X}_{t,t+\frac{u}{\Delta}} \mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} V_{t\Delta+u}^{-1} d\langle \mathcal{Y}, V \rangle_u \end{aligned}$$

$$\begin{aligned}
& + pm\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}V_{t\Delta+u}^{-1}d\langle\mathcal{W},V\rangle_u \\
& = n\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1}e^{\kappa u}\sqrt{V_{t\Delta+u}}dW_{t\Delta+u} + p\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}\sqrt{V_{t\Delta+u}}dW_{t\Delta+u} \\
& + l\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{B}_{t,t+\frac{u}{\Delta}}^{-1}\sqrt{V_{t\Delta+u}}dB_{t\Delta+u} + m\kappa\theta\mathcal{X}_{t,t+\frac{u}{\Delta}}V_{t\Delta+u}^{-1}du \\
& - m\kappa\mathcal{X}_{t,t+\frac{u}{\Delta}}du + m\mathcal{X}_{t,t+\frac{u}{\Delta}}V_{t\Delta+u}^{-1}\sigma\sqrt{V_{t\Delta+u}}dW_{t\Delta+u} + m\kappa\mathcal{X}_{t,t+\frac{u}{\Delta}}du \\
& + \frac{1}{2}n(n-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-2}e^{2\kappa u}V_{t\Delta+u}du + \frac{1}{2}p(p-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-2}V_{t\Delta+u}du \\
& + \frac{1}{2}l(l-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{B}_{t,t+\frac{u}{\Delta}}^{-2}V_{t\Delta+u}du + \frac{1}{2}m(m-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}V_{t\Delta+u}^{-2}\sigma^2V_{t\Delta+u}du \\
& + np\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}e^{\kappa u}V_{t\Delta+u}du \\
& + nm\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1}V_{t\Delta+u}^{-1}\sigma e^{\kappa u}V_{t\Delta+u}du + pm\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}V_{t\Delta+u}^{-1}\sigma V_{t\Delta+u}du.
\end{aligned}$$

The drift of $\mathcal{X}_{t,t+\frac{u}{\Delta}}$ is then composed of

$$\begin{aligned}
& \left(m\kappa\theta + \frac{1}{2}m(m-1)\sigma^2\right)\mathcal{X}_{t,t+\frac{u}{\Delta}}V_{t\Delta+u}^{-1} + \frac{1}{2}n(n-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}e^{2\kappa u}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-2}V_{t\Delta+u} \\
& + \frac{1}{2}p(p-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-2}V_{t\Delta+u} + \frac{1}{2}l(l-1)\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{B}_{t,t+\frac{u}{\Delta}}^{-2}V_{t\Delta+u} \\
& + np\mathcal{X}_{t,t+\frac{u}{\Delta}}e^{\kappa u}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}V_{t\Delta+u} + nm\sigma\mathcal{X}_{t,t+\frac{u}{\Delta}}e^{\kappa u}\mathcal{Y}_{t,t+\frac{u}{\Delta}}^{-1} \\
& + pm\sigma\mathcal{X}_{t,t+\frac{u}{\Delta}}\mathcal{W}_{t,t+\frac{u}{\Delta}}^{-1}.
\end{aligned}$$

Integrating the time variable u from 0 and Δ , taking the expectation from each side and using Fubini leads to the result. \square

D.6.2 Fourth moment calculations

This section details how to calculate the fourth cumulant of the stochastic integrals with respect to the model filtration. We first derive the relationship between the fourth cumulant and central moment.

Corollary D.6.2. *For the general stochastic integrals $\mathcal{X}_{i,t,t+1} \in \{\mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1}\mathcal{B}_{t,t+1}\}$,*

$\forall i \in \{1, 2, 3, 4\}$, the relationship between the fourth cumulant and centered moment is:

$$\begin{aligned}
 & \text{Cum}^{(4)} [\mathcal{X}_{1,t,t+1}, \mathcal{X}_{2,t,t+1}, \mathcal{X}_{3,t,t+1}, \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= \text{E} [\mathcal{X}_{1,t,t+1} \mathcal{X}_{2,t,t+1} \mathcal{X}_{3,t,t+1} \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\quad - \text{E} [\mathcal{X}_{1,t,t+1} \mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}] \text{E} [\mathcal{X}_{3,t,t+1} \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\quad - \text{E} [\mathcal{X}_{1,t,t+1} \mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}] \text{E} [\mathcal{X}_{2,t,t+1} \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &\quad - \text{E} [\mathcal{X}_{1,t,t+1} \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \text{E} [\mathcal{X}_{2,t,t+1} \mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}].
 \end{aligned}$$

Proof. By definition, the fourth cumulant of the random variables $\mathcal{X}_{1,t,t+1}$, $\mathcal{X}_{2,t,t+1}$, $\mathcal{X}_{3,t,t+1}$ and $\mathcal{X}_{4,t,t+1}$ is

$$\begin{aligned}
 & \text{Cum}^{(4)} [\mathcal{X}_{1,t,t+1}, \mathcal{X}_{2,t,t+1}, \mathcal{X}_{3,t,t+1}, \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= \text{E} \left[\begin{array}{l} (\mathcal{X}_{1,t,t+1} - \text{E} [\mathcal{X}_{1,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{2,t,t+1} - \text{E} [\mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}]) \\ (\mathcal{X}_{3,t,t+1} - \text{E} [\mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{4,t,t+1} - \text{E} [\mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}]) \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\
 &\quad - \left(\begin{array}{l} \text{E} [(\mathcal{X}_{1,t,t+1} - \text{E} [\mathcal{X}_{1,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{2,t,t+1} - \text{E} [\mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \\ \text{E} [(\mathcal{X}_{3,t,t+1} - \text{E} [\mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{4,t,t+1} - \text{E} [\mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \end{array} \right) \\
 &\quad - \left(\begin{array}{l} \text{E} [(\mathcal{X}_{1,t,t+1} - \text{E} [\mathcal{X}_{1,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{3,t,t+1} - \text{E} [\mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \\ \text{E} [(\mathcal{X}_{2,t,t+1} - \text{E} [\mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{4,t,t+1} - \text{E} [\mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \end{array} \right) \\
 &\quad - \left(\begin{array}{l} \text{E} [(\mathcal{X}_{1,t,t+1} - \text{E} [\mathcal{X}_{1,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{4,t,t+1} - \text{E} [\mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \\ \text{E} [(\mathcal{X}_{2,t,t+1} - \text{E} [\mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}]) (\mathcal{X}_{3,t,t+1} - \text{E} [\mathcal{X}_{3,t,t+1} | \mathcal{F}_{t\Delta}]) | \mathcal{F}_{t\Delta}] \end{array} \right).
 \end{aligned}$$

If these variables are centered at zero, the expression is further simplified, leading to the result. \square

Lemma D.6.3. For the stochastic integrals $\mathcal{X}_{i,t,t+1} \in \{\mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{B}_{t,t+1}\}$, $\forall i \in \{1, 2, 3, 4\}$, we have

$$\begin{aligned}
 & \text{E} [\mathcal{X}_{1,t,t+1} \mathcal{X}_{2,t,t+1} | \mathcal{F}_{t\Delta}] \text{E} [\mathcal{X}_{3,t,t+1} \mathcal{X}_{4,t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= \int_0^\Delta \text{E} [\mathcal{X}_{1,t,t+\frac{u}{\Delta}} \mathcal{X}_{2,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}] d\text{E} [\mathcal{X}_{3,t,t+\frac{u}{\Delta}} \mathcal{X}_{4,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}] \\
 &\quad + \int_0^\Delta \text{E} [\mathcal{X}_{3,t,t+\frac{u}{\Delta}} \mathcal{X}_{4,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}] d\text{E} [\mathcal{X}_{1,t,t+\frac{u}{\Delta}} \mathcal{X}_{2,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}].
 \end{aligned}$$

Proof. For any diffusion process,

$$dX_u Y_u = X_u dY_u + Y_u dX_u + d \langle Y, X \rangle_u,$$

where we set $Y_u = \mathbb{E} [\mathcal{X}_{1,t,t+\frac{u}{\Delta}} \mathcal{X}_{2,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}]$ and $X_u = \mathbb{E} [\mathcal{X}_{3,t,t+\frac{u}{\Delta}} \mathcal{X}_{4,t,t+\frac{u}{\Delta}} | \mathcal{F}_{t\Delta}]$. Based on Lemma D.6.1, the quadratic covariation term is null, leading to the result. \square

Corollary D.6.4.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] = 6 \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds.$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned} \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 6 \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds \\ &= 6 \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds + 6 \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds, \end{aligned}$$

and application of Lemma D.6.3 to

$$\mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] = 2 \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.5.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned} \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds \\ &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \underbrace{\mathbb{E} [\mathcal{B}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} ds + 3 \int_t^{t+1} \sigma \underbrace{\mathbb{E} [\mathcal{Z}_{t,s} \mathcal{W}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} ds. \end{aligned}$$

and where

$$\underbrace{\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{W}_{t,t+1}|\mathcal{F}_{t\Delta}]}_{=0}\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{B}_{t,t+1}|\mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.6.

$$\text{Cum}^{(4)}[\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}|\mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned} \mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{B}_{t,t+1}\mathcal{B}_{t,t+1}\mathcal{Z}_{t,t+1}|\mathcal{F}_{t\Delta}] &= 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E}[V_{s\Delta}\mathcal{Z}_{t,s}\mathcal{B}_{t,s}|\mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\ &= 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E}[V_{s\Delta}|\mathcal{F}_{t\Delta}] \underbrace{\mathbb{E}[\mathcal{B}_{t,s}\mathcal{Z}_{t,s}|\mathcal{F}_{t\Delta}]}_{=0} e^{\kappa s\Delta} ds \\ &+ 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \sigma \underbrace{\mathbb{E}[\mathcal{Z}_{t,s}\mathcal{Z}_{t,s}\mathcal{B}_{t,s}|\mathcal{F}_{t\Delta}]}_{=0} e^{\kappa s\Delta} ds. \end{aligned}$$

and where

$$\underbrace{\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{Z}_{t,t+1}|\mathcal{F}_{t\Delta}]}_{=0}\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{B}_{t,t+1}|\mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.7.

$$\text{Cum}^{(4)}[\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}|\mathcal{F}_{t\Delta}] = \int_t^{t+1} \sigma \mathbb{E}[\mathcal{Z}_{t,s}(\mathcal{W}_{t,s}\mathcal{W}_{t,s} + \mathcal{B}_{t,s}\mathcal{B}_{t,s})|\mathcal{F}_{t\Delta}] ds.$$

Proof. Direct application of Lemma D.6.1 to

$$\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{B}_{t,t+1}\mathcal{W}_{t,t+1}\mathcal{W}_{t,t+1}|\mathcal{F}_{t\Delta}]$$

$$\begin{aligned}
&= \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds + \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds \\
&= \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s} \mathcal{B}_{t,s} + \mathcal{W}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds \\
&+ \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{W}_{t,s} + \mathcal{B}_{t,s} \mathcal{B}_{t,s}) | \mathcal{F}_{t\Delta}] ds,
\end{aligned}$$

and application of Lemma D.6.3 to

$$\mathbb{E} [\mathcal{B}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s}^2 + \mathcal{W}_{t,s}^2 | \mathcal{F}_{t\Delta}] ds.$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.8.

$$\begin{aligned}
&\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
&= \sigma e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s} \mathcal{W}_{t,s} e^{\kappa s\Delta} + \mathcal{B}_{t,s}^2 e^{\kappa s\Delta}) | \mathcal{F}_{t\Delta}] ds.
\end{aligned}$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned}
&\mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
&= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
&+ e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
&= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s} \mathcal{B}_{t,s} + \mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
&+ e^{-\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s} \mathcal{W}_{t,s} + \mathcal{B}_{t,s} \mathcal{B}_{t,s}) | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds,
\end{aligned}$$

and application of Lemma D.6.3 to

$$\mathbb{E} [\mathcal{B}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}]$$

$$= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s}^2 + \mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.9.

$$\begin{aligned} & \text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s} \mathcal{Z}_{t,s} + \mathcal{B}_{t,s} \mathcal{B}_{t,s}) e^{2\kappa s\Delta} | \mathcal{F}_{t\Delta}] ds. \end{aligned}$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned} & \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &+ e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{B}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s} \mathcal{B}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &+ e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{Z}_{t,s} \mathcal{Z}_{t,s} + \mathcal{B}_{t,s} \mathcal{B}_{t,s}) | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds, \end{aligned}$$

and application of Lemma D.6.3 to

$$\begin{aligned} & \mathbb{E} [\mathcal{B}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{B}_{t,s}^2 + \mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds. \end{aligned}$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.10.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds \\
 &= 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \underbrace{\mathbb{E} [\mathcal{B}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} e^{3\kappa s\Delta} ds \\
 &+ e^{-3\kappa(t+1)\Delta} \int_t^{t+1} 3\sigma \underbrace{\mathbb{E} [\mathcal{Z}_{t,s} \mathcal{B}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} e^{3\kappa s\Delta} ds,
 \end{aligned}$$

and where

$$\mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.11.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] ds \\
 &= 3 \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \underbrace{\mathbb{E} [\mathcal{B}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} ds + 3 \int_t^{t+1} \sigma \underbrace{\mathbb{E} [\mathcal{Z}_{t,s} \mathcal{B}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} ds,
 \end{aligned}$$

and where

$$\mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.12.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned} & \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &+ 2e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\ &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \underbrace{\mathbb{E} [\mathcal{B}_{t,s} \mathcal{W}_{t,s} + 2\mathcal{B}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} e^{2\kappa s\Delta} ds \\ &+ e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \sigma \underbrace{\mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{B}_{t,s} \mathcal{W}_{t,s} + 2\mathcal{B}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}]}_{=0} e^{2\kappa s\Delta} ds, \end{aligned}$$

and where

$$\mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.13.

$$\text{Cum}^{(4)} [\mathcal{B}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Proof. Direct application of Lemma D.6.1 and Corollary B.4.3 to

$$\begin{aligned} & \mathbb{E} [\mathcal{B}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= 2e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\ &+ e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{B}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \end{aligned}$$

$$\begin{aligned}
 &= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E}[V_{s\Delta} | \mathcal{F}_{t\Delta}] \underbrace{\mathbb{E}[2\mathcal{B}_{t,s}\mathcal{W}_{t,s} + \mathcal{B}_{t,s}\mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}]}_{=0} e^{\kappa s\Delta} ds \\
 &+ e^{-\kappa(t+1)\Delta} \int_t^{t+1} \sigma \underbrace{\mathbb{E}[\mathcal{Z}_{t,s}(2\mathcal{B}_{t,s}\mathcal{W}_{t,s} + \mathcal{B}_{t,s}\mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}]}_{=0} e^{\kappa s\Delta} ds,
 \end{aligned}$$

and where

$$\mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad \mathbb{E}[\mathcal{B}_{t,t+1}\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 0.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.14.

$$\text{Cum}^{(4)}[\mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] = 6 \int_t^{t+1} \sigma \mathbb{E}[\mathcal{Z}_{t,s}\mathcal{W}_{t,s}\mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds.$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned}
 \mathbb{E}[\mathcal{W}_{t,t+1}\mathcal{W}_{t,t+1}\mathcal{W}_{t,t+1}\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 6 \int_t^{t+1} \mathbb{E}[V_{s\Delta}\mathcal{W}_{t,s}\mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds \\
 &= 6 \int_t^{t+1} \mathbb{E}[V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E}[\mathcal{W}_{t,s}\mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds + 6 \int_t^{t+1} \sigma \mathbb{E}[\mathcal{Z}_{t,s}\mathcal{W}_{t,s}\mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] ds,
 \end{aligned}$$

and application of Lemma D.6.3 to

$$\mathbb{E}[\mathcal{W}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E}[\mathcal{W}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = 2 \int_t^{t+1} \mathbb{E}[V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E}[\mathcal{W}_{t,s}^2 | \mathcal{F}_{t\Delta}] ds.$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.15.

$$\begin{aligned}
 &\text{Cum}^{(4)}[\mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E}[\mathcal{Z}_{t,s}(\mathcal{W}_{t,s}\mathcal{Z}_{t,s} + \mathcal{W}_{t,s}\mathcal{W}_{t,s}) | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds.
 \end{aligned}$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
 &+ 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
 &= 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{W}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds \\
 &+ 3e^{-\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{W}_{t,s} \mathcal{W}_{t,s}) | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds,
 \end{aligned}$$

and application of Lemma D.6.3 to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{W}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= e^{-\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] e^{\kappa s\Delta} ds.
 \end{aligned}$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.16.

$$\begin{aligned}
 & \text{Cum}^{(4)} [\mathcal{W}_{t,t+1}, \mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{W}_{t,s} + 4\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds.
 \end{aligned}$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
 &= e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{W}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\
 &+ 4e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds
 \end{aligned}$$

$$\begin{aligned}
& + e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\
& = e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [(\mathcal{W}_{t,s} \mathcal{W}_{t,s} + 4\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds \\
& + e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{W}_{t,s} + 4\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds,
\end{aligned}$$

and application of Lemma D.6.3 to

$$\mathbb{E} [\mathcal{Z}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = 2e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds,$$

and

$$\begin{aligned}
& \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
& = 2e^{-2\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{2\kappa s\Delta} ds.
\end{aligned}$$

Finally, Corollary D.6.2 is used to obtain the results. □

Corollary D.6.17.

$$\begin{aligned}
& \text{Cum}^{(4)} [\mathcal{W}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
& = 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds.
\end{aligned}$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned}
& \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \\
& = 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{W}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds \\
& + 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds \\
& = 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds
\end{aligned}$$

$$+ 3e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} (\mathcal{W}_{t,s} \mathcal{Z}_{t,s} + \mathcal{Z}_{t,s} \mathcal{Z}_{t,s}) | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds,$$

and application of Lemma D.6.3 to

$$\begin{aligned} & \mathbb{E} [\mathcal{W}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \\ &= e^{-3\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] e^{3\kappa s\Delta} ds. \end{aligned}$$

Finally, Corollary D.6.2 is used to obtain the results. \square

Corollary D.6.18.

$$\text{Cum}^{(4)} [\mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1}, \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 6e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s}^3 | \mathcal{F}_{t\Delta}] e^{4\kappa s\Delta} ds.$$

Proof. Direct application of Lemma D.6.1 to

$$\begin{aligned} & \mathbb{E} [\mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} \mathcal{Z}_{t,t+1} | \mathcal{F}_{t\Delta}] = 6e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{4\kappa s\Delta} ds \\ &= 6e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{4\kappa s\Delta} ds \\ &+ 6e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \sigma \mathbb{E} [\mathcal{Z}_{t,s} \mathcal{Z}_{t,s} \mathcal{Z}_{t,s} | \mathcal{F}_{t\Delta}] e^{4\kappa s\Delta} ds, \end{aligned}$$

and application of Lemma D.6.3 to

$$\begin{aligned} & \mathbb{E} [\mathcal{Z}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] \\ &= 2e^{-4\kappa(t+1)\Delta} \int_t^{t+1} \mathbb{E} [V_{s\Delta} | \mathcal{F}_{t\Delta}] \mathbb{E} [\mathcal{Z}_{t,s}^2 | \mathcal{F}_{t\Delta}] e^{4\kappa s\Delta} ds. \end{aligned}$$

Finally, Corollary D.6.2 is used to obtain the result. \square

D.7 Measurement Error under Assumption 3.4.2

D.7.1 General results

Definition D.7.1. Let X_i be a random variable where

$$X_i = \alpha_i + \sum_{n=1}^N X_{i,n},$$

where α_i is a constant and $X_{i,n}$ are random variable with $E[X_{i,n} | \mathcal{F}_{n-1}] = 0$ and

$$\mathcal{F}_N = \sigma\{\{X_{i,n}\}_{i \in \{1,2\}}\}_{n=1}^N.$$

The temporal structure is left general.

Lemma D.7.2.

$$E[X_i] = \alpha_i, \quad \text{Cov}[X_1, X_2] = \sum_{n=1}^N E[\text{Cov}[X_{1,n}, X_{2,n} | \mathcal{F}_{n-1}]].$$

Proof. With the help of the law of total covariance

$$\begin{aligned} \text{Cov}[X_1, X_2] &= \text{Cov}\left[\alpha_1 + \sum_{n=1}^N X_{1,n}, \alpha_2 + \sum_{n=1}^N X_{2,n}\right] \\ &= E\left[\text{Cov}\left[\alpha_1 + \sum_{n=1}^N X_{1,n}, \alpha_2 + \sum_{n=1}^N X_{2,n} \middle| \mathcal{F}_{N-1}\right]\right] \\ &\quad + \text{Cov}\left[E\left[\alpha_1 + \sum_{n=1}^N X_{1,n} \middle| \mathcal{F}_{N-1}\right], E\left[\alpha_2 + \sum_{n=1}^N X_{2,n} \middle| \mathcal{F}_{N-1}\right]\right] \\ &= E[\text{Cov}[X_{1,N}, X_{2,N} | \mathcal{F}_{N-1}]] + \text{Cov}\left[\sum_{n=1}^{N-1} X_{1,n}, \sum_{n=1}^{N-1} X_{2,n}\right]. \end{aligned}$$

Applying the law of total covariance $N - 1$ more times to

$$\text{Cov} \left[\sum_{n=1}^{N-1} X_{1,n}, \sum_{n=1}^{N-1} X_{2,n} \right],$$

leads to the result. □

Lemma D.7.3.

$$\begin{aligned} & \text{Cum} [X_1, X_2, X_3] \\ &= \sum_{n=1}^N \text{E} [\text{Cum} [X_{1,n}, X_{2,n}, X_{3,n} | \mathcal{F}_{n-1}]] \\ &+ \sum_{n=1}^N \text{Cov} \left[\sum_{m=1}^{n-1} X_{1,m}, \text{Cov} [X_{2,n}, X_{3,n} | \mathcal{F}_{n-1}] \right] \\ &+ \sum_{n=1}^N \text{Cov} \left[\sum_{m=1}^{n-1} X_{2,m}, \text{Cov} [X_{1,n}, X_{3,n} | \mathcal{F}_{n-1}] \right] \\ &+ \sum_{n=1}^N \text{Cov} \left[\sum_{m=1}^{n-1} X_{3,m}, \text{Cov} [X_{1,n}, X_{2,n} | \mathcal{F}_{n-1}] \right]. \end{aligned}$$

Proof. With the help of the law of total cumulance

$$\begin{aligned} \text{Cum} [X_1, X_2, X_3] &= \text{Cum} \left[\sum_{n=1}^N X_{1,n}, \sum_{n=1}^N X_{2,n}, \sum_{n=1}^N X_{3,n} \right] \\ &= \text{E} \left[\text{Cum} \left[\sum_{n=1}^N X_{1,n}, \sum_{n=1}^N X_{2,n}, \sum_{n=1}^N X_{3,n} \middle| \mathcal{F}_{N-1} \right] \right] \\ &+ \text{Cum} \left[\text{E} \left[\sum_{n=1}^N X_{1,n} \middle| \mathcal{F}_{N-1} \right], \text{E} \left[\sum_{n=1}^N X_{2,n} \middle| \mathcal{F}_{N-1} \right], \text{E} \left[\sum_{n=1}^N X_{3,n} \middle| \mathcal{F}_{N-1} \right] \right] \\ &+ \text{Cov} \left[\text{E} \left[\sum_{n=1}^N X_{1,n} \middle| \mathcal{F}_{N-1} \right], \text{Cov} \left[\sum_{n=1}^N X_{2,n}, \sum_{n=1}^N X_{3,n} \middle| \mathcal{F}_{N-1} \right] \right] \\ &+ \text{Cov} \left[\text{E} \left[\sum_{n=1}^N X_{2,n} \middle| \mathcal{F}_{N-1} \right], \text{Cov} \left[\sum_{n=1}^N X_{1,n}, \sum_{n=1}^N X_{3,n} \middle| \mathcal{F}_{N-1} \right] \right] \end{aligned}$$

$$\begin{aligned}
& + \text{Cov} \left[\mathbb{E} \left[\sum_{n=1}^N X_{3,n} \middle| \mathcal{F}_{N-1} \right], \text{Cov} \left[\sum_{n=1}^N X_{1,n}, \sum_{n=1}^N X_{2,n} \middle| \mathcal{F}_{N-1} \right] \right] \\
& = \mathbb{E} [\text{Cum} [X_{1,N}, X_{2,N}, X_{3,N} | \mathcal{F}_{N-1}]] \\
& + \text{Cum} \left[\sum_{n=1}^{N-1} X_{1,n}, \sum_{n=1}^{N-1} X_{2,n}, \sum_{n=1}^{N-1} X_{3,n} \right] + \text{Cov} \left[\sum_{n=1}^{N-1} X_{1,n}, \text{Cov} [X_{2,N}, X_{3,N} | \mathcal{F}_{N-1}] \right] \\
& + \text{Cov} \left[\sum_{n=1}^{N-1} X_{2,n}, \text{Cov} [X_{1,N}, X_{3,N} | \mathcal{F}_{N-1}] \right] + \text{Cov} \left[\sum_{n=1}^{N-1} X_{3,n}, \text{Cov} [X_{1,N}, X_{2,N} | \mathcal{F}_{N-1}] \right].
\end{aligned}$$

Applying the law of total cumulance $N - 1$ more times to

$$\text{Cum} \left[\sum_{n=1}^{N-1} X_{1,n}, \sum_{n=1}^{N-1} X_{2,n}, \sum_{n=1}^{N-1} X_{3,n} \right],$$

leads to the result. □

D.7.2 Approximations

Corollary D.7.4. *Based on Assumption 3.4.2, the centered variance $V_{k,(t+1)\Delta}$ is approximated with*

$$V_{k,(t+1)\Delta} - \mathbb{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \approx \sigma_k \sum_{n=1}^N \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}.$$

Proof. With Assumption 3.4.2 and Equation (3.9), it implies that

$$V_{k,(t+1)\Delta} \approx \underbrace{V_{k,t\Delta} e^{-\Delta\kappa_k} + \theta_k (1 - e^{-\Delta\kappa_k})}_{\mathbb{E}[V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]} + \sigma_k \sum_{n=1}^N e^{-\kappa_k \Delta (1 - \frac{n-1}{N})} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}.$$

We apply the Taylor expansion to $e^{-\kappa_k \Delta (1 - \frac{n-1}{N})} \cong 1 - \kappa_k \Delta (1 - \frac{n-1}{N}) \cong 1$, leading to the result. □

Corollary D.7.5. *Based on Assumption 3.4.2, the integrated variance $\mathcal{V}_{k,t,t+1}$ is approxi-*

mated with

$$\mathcal{V}_{k,t,t+1} - \mathbb{E}[\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] \approx \sigma_k \sum_{n=1}^N \Delta \frac{(N-n+1)}{N} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}.$$

The quadratic covariation is approximated with

$$\begin{aligned} & QCV_{i,j,t,t+1} - \mathbb{E}[QCV_{i,j,t,t+1} | \mathcal{F}_{t\Delta}] \\ & \approx \sum_{k=1}^{N_V} \sum_{n=1}^N \Delta \frac{(N-n+1)}{N} c_{i,k} c_{j,k} \sigma_k \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}. \end{aligned}$$

Proof. Using Assumption 3.4.2, Corollary D.7.4 and Taylor expansion, the integrated variance is

$$\begin{aligned} \mathcal{V}_{k,t,t+1} - \mathbb{E}[\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] & \approx \frac{\sigma_k}{\kappa_k} \sum_{n=1}^N \left(1 - e^{-\kappa_k \Delta (1 - \frac{n-1}{N})}\right) \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}, \\ & \approx \sigma_k \Delta \sum_{n=1}^N \frac{(N-n+1)}{N} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}}, \end{aligned}$$

since we apply the Taylor expansion to $\frac{1 - e^{-\kappa_k \Delta (1 - \frac{n-1}{N})}}{\kappa_k} \cong \Delta (1 - \frac{n-1}{N}) \cong \Delta \frac{(N-n+1)}{N}$. As defined in Equation (3.5), the quadratic covariation is a linear sum of the integrated variances. \square

Corollary D.7.6. Based on Assumption 3.4.2, the squared variance $V_{k,(t+1)\Delta} V_{\tilde{k},(t+1)\Delta}$ is approximated with

$$\begin{aligned} & V_{k,(t+1)\Delta} V_{\tilde{k},(t+1)\Delta} - \mathbb{E}\left[V_{k,(t+1)\Delta} V_{\tilde{k},(t+1)\Delta} \middle| \mathcal{F}_{t\Delta}\right] \\ & \approx \sigma_k \sum_{n=1}^N V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}} \\ & + \sigma_{\tilde{k}} \sum_{n=1}^N V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \sqrt{V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{\tilde{k},t\Delta + \frac{n\Delta}{N}}. \end{aligned}$$

Proof. Using Corollary (D.3.2) leads to the results. □

Corollary D.7.7. *Using Assumption 3.4.2, the log-returns is approximated with*

$$R_{i,t,t+1} - \mathbb{E}[R_{i,t,t+1} | \mathcal{F}_{t\Delta}] \approx \sum_{n=1}^N \left(\sum_{k=1}^{N_V} c_{i,k} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} z_{k,t\Delta + \frac{n\Delta}{N}} \right),$$

and, for the error measurement on the realized variance, we have

$$\xi_{i,j,t,t+1} - \mathbb{E}[\xi_{i,j,t,t+1} | \mathcal{F}_{t\Delta}] \approx \sum_{n=1}^N \left(\sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{k,t\Delta + \frac{n\Delta}{N}}^2 - 1) \right),$$

where $z_{k,t\Delta + \frac{n\Delta}{N}} = \rho_k w_{k,t\Delta + \frac{n\Delta}{N}} + \sqrt{1 - \rho_k^2} b_{k,t\Delta + \frac{n\Delta}{N}}$.

Proof. Using Assumption 3.4.2 and Equation (3.11) lead to the results for the log-returns as the drift is negligible at high-frequency compared to the stochastic integrals. Indeed,

$$R_{i,t,t+1} \approx \sum_{n=1}^N \left(\sum_{k=1}^{N_V} \left(c_{ik} \lambda_k - \frac{c_{ik}^2}{2} \right) V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N} + c_{i,k} \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} z_{k,t\Delta + \frac{n\Delta}{N}} \right).$$

The realized covariance uses the fact that

$$RCV_{i,j,t,t+1} = \sum_{n=1}^N R_{i,t + \frac{n-1}{N}, t + \frac{n}{N}} R_{j,t + \frac{n-1}{N}, t + \frac{n}{N}},$$

leading to the results with Equation (3.13). □

D.7.3 Application to second central moments

Corollary D.7.8. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cov} \left[\xi_{i,j,t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t + \frac{n-1}{N}, t + \frac{n}{N}}, \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} V_{k,t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}, \end{aligned}$$

with $\mathcal{C}_{k,\tilde{k}}^{i,i,\tilde{j},\tilde{j}}$ defined in Equation (B.6). Furthermore,

$$\text{Cov} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \approx \frac{\Delta}{N} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,i,\tilde{j},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \mid \mathcal{F}_{t\Delta} \right].$$

Proof. With Assumption 3.4.2, we have

$$\begin{aligned} & \text{Cov} \left[\xi_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \mathbb{E} \left[\left(\sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{k,t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \mid \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,i,\tilde{j},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}. \end{aligned}$$

Using Lemma D.7.2 with $X_1 = \xi_{i,j,t,t+1}$ and $X_2 = \xi_{\tilde{i},\tilde{j},t,t+1}$, leads to second the result. \square

Corollary D.7.9. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cov} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0, \\ & \text{Cov} \left[V_{k,t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0, \\ & \text{Cov} \left[QCV_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0, \end{aligned}$$

and

$$\begin{aligned} & \text{Cov} \left[R_{i,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \approx 0, \\ & \text{Cov} \left[V_{k,(t+1)\Delta}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \approx 0, \\ & \text{Cov} \left[QCV_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \approx 0. \end{aligned}$$

Proof. Based on Corollaries D.7.4, D.7.5 and D.7.7 along with Assumption 3.4.2, we have null covariance as any uneven moments of a standard multivariate normal distribution is null. Applying the first set of results in Lemma D.7.2 with $X_1 = R_{i,t,t+1}$, $X_1 = V_{k,(t+1)\Delta}$

or $X_1 = QCV_{i,j,t,t+1}$ and $X_2 = \xi_{\tilde{i},\tilde{j},t,t+1}$, leads to the second result. \square

D.7.4 Application to moments of $\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1}$

Corollary D.7.10. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cum} \left[\xi_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\tilde{j},\hat{i},\hat{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^3}{N^3}, \end{aligned}$$

with

$$\begin{aligned} \mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\tilde{j},\hat{i},\hat{j}} &= c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\hat{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{j},\hat{k}} + c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} \\ &+ c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} + c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} \\ &+ c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} + c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} \\ &+ c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}} + c_{i,k} c_{\tilde{i},k} c_{j,k} c_{\tilde{j},\tilde{k}} c_{\hat{j},\tilde{k}} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} c_{\hat{i},\hat{k}} c_{\hat{k}}. \end{aligned}$$

Proof. With Assumption 3.4.2 and Corollary D.7.7, we have

$$\begin{aligned} & \text{Cum} \left[\xi_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \text{E} \left[\begin{array}{c} \left(\sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{k,t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i},k} c_{\tilde{j},k} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \\ \left(\sum_{\hat{k}=1}^{N_V} c_{\hat{i},k} c_{\hat{j},k} V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\hat{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\tilde{j},\hat{i},\hat{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^3}{N^3}, \end{aligned}$$

where $\mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\tilde{j},\hat{i},\hat{j}}$ is found with Isserlis theorem. \square

Corollary D.7.11. *If Assumption 3.4.2 holds, based on Corollary D.7.6,*

$$\text{Cov} \left[\text{Cov} \left[\sum_{m=1}^{n-1} \xi_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}} \left| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right. \right] \left| \mathcal{F}_{t\Delta} \right. \right] \approx 0.$$

Proof. With Assumption 3.4.2 and Equation (B.6), we have

$$\begin{aligned} & \text{Cov} \left[\text{Cov} \left[\sum_{m=1}^{n-1} \xi_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}} \left| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right. \right] \left| \mathcal{F}_{t\Delta} \right. \right] \\ & \approx \text{Cov} \left[\left(\sum_{m=1}^{n-1} \left(\sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k,t\Delta+\frac{(m-1)\Delta}{N}} \frac{\Delta}{N} (z_{k,t\Delta+\frac{m\Delta}{N}}^2 - 1) \right) \right), \left| \mathcal{F}_{t\Delta} \right. \right] \\ & \quad \left(\sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\hat{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \right) \\ & = \frac{\Delta^3}{N^3} \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\hat{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} c_{i,k} c_{j,k} \text{Cov} \left[\left(V_{k,t\Delta+\frac{(m-1)\Delta}{N}} (z_{k,t\Delta+\frac{m\Delta}{N}}^2 - 1) \right), \left| \mathcal{F}_{t\Delta} \right. \right] \\ & \quad \left(V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \right) \\ & \approx 0, \end{aligned}$$

as we use Corollary D.7.6 and that all uneven moments of central normal variable are null. \square

Corollary D.7.12. *Based on Assumption 3.4.2,*

$$\begin{aligned} & \text{Cum} \left[\xi_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1} \left| \mathcal{F}_{t\Delta} \right. \right] \\ & \approx \frac{\Delta^2}{N^2} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{S}_{k,\tilde{k},\hat{k}}^{i,\tilde{i},\tilde{j},\hat{i},\hat{j}} \mathbf{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} V_{\hat{k},s} ds \left| \mathcal{F}_{t\Delta} \right. \right]. \end{aligned}$$

Proof. Applying the results of Corollaries D.7.10 and D.7.11 in Lemma D.7.3 with $X_1 = \xi_{i,j,t,t+1}$, $X_2 = \xi_{\tilde{i},\tilde{j},t,t+1}$ and $X_3 = \xi_{\hat{i},\hat{j},t,t+1}$ leads to the result.

D.7.5 Application to moments of $\xi_{i,\hat{j},t,t+1}$

Corollary D.7.13. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cum} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, R_{j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{i,\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}, \end{aligned}$$

with $\mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}}$ defined in Equation (B.6).

Proof. With Assumption 3.4.2, we have

$$\begin{aligned} & \text{Cum} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, R_{j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{i,\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \text{Cum} \left[\begin{array}{c} \left(\sum_{k=1}^{N_V} c_{i,k} \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} z_{k,t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{k=1}^{N_V} c_{j,k} \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} z_{k,t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{i,\tilde{k}} c_{j,\tilde{k}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{k,t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{i,j,\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}. \quad \square \end{aligned}$$

Corollary D.7.14. *If Assumption 3.4.2 holds,*

$$\text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} R_{i,t+\frac{m-1}{N},t+\frac{m}{N}}, \\ \text{Cov} \left[R_{j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{i,\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \approx 0.$$

Proof. Direct application of Corollary D.7.9. □

Corollary D.7.15. *If Assumption 3.4.2 holds,*

$$\text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} \xi_{i,\tilde{j},t+\frac{m-1}{N},t+\frac{m}{N}}, \\ \text{Cov} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, R_{j,t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \approx 0.$$

Proof.

$$\begin{aligned}
 & \text{Cov} \left[\text{Cov} \left[\sum_{m=1}^{n-1} \xi_{\tilde{i}, \tilde{j}, t + \frac{m-1}{N}, t + \frac{m}{N}}, R_{i, t + \frac{n-1}{N}, t + \frac{n}{N}}, R_{j, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \middle| \mathcal{F}_{t\Delta} \right] \\
 & \approx \text{Cov} \left[\sum_{m=1}^{n-1} \xi_{\tilde{i}, \tilde{j}, t + \frac{m-1}{N}, t + \frac{m}{N}}, \sum_{k=1}^{N_V} c_{i,k} c_{j,k} V_{k, t\Delta + \frac{(n-1)\Delta}{N}} \middle| \mathcal{F}_{t\Delta} \right] \\
 & \approx \text{Cov} \left[\sum_{m=1}^{n-1} \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i}, \tilde{k}} c_{\tilde{j}, \tilde{k}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k}, t\Delta + \frac{m\Delta}{N}}^2 - 1) \right), \sum_{k=1}^{N_V} c_{i,k} c_{j,k} \left(\sum_{m=1}^{n-1} \left(\sigma_k \sqrt{V_{k, t\Delta + \frac{(m-1)\Delta}{N}} \frac{\Delta}{N}} w_{k, t\Delta + \frac{m\Delta}{N}} \right) \right) \frac{\Delta}{N} \middle| \mathcal{F}_{t\Delta} \right] \\
 & = 0,
 \end{aligned}$$

as any uneven moment of a standard Gaussian random variable is null. \square

Corollary D.7.16. *If Assumption 3.4.2 holds, then*

$$\begin{aligned}
 & \text{Cum} [R_{i, t, t+1}, R_{j, t, t+1}, \xi_{\tilde{i}, \tilde{j}, t, t+1} \middle| \mathcal{F}_{t\Delta}] \\
 & \approx \frac{\Delta}{N} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{k, \tilde{k}}^{i, \tilde{j}, \tilde{i}, \tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k, s} V_{\tilde{k}, s} ds \middle| \mathcal{F}_{t\Delta} \right].
 \end{aligned}$$

Proof. Applying the results of Corollaries D.7.13, D.7.14 and D.7.15 in Lemma D.7.3 with $X_1 = R_{i, t, t+1}$, $X_2 = R_{j, t, t+1}$ and $X_3 = \xi_{\tilde{i}, \tilde{j}, t, t+1}$, we have

$$\begin{aligned}
 & \text{Cum} [R_{i, t, t+1}, R_{j, t, t+1}, \xi_{\tilde{i}, \tilde{j}, t, t+1} \middle| \mathcal{F}_{t\Delta}] \\
 & \approx \sum_{n=1}^N \mathbb{E} \left[\sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} c_{k, \tilde{k}}^{i, \tilde{j}, \tilde{i}, \tilde{j}} V_{k, t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \middle| \mathcal{F}_{t\Delta} \right],
 \end{aligned}$$

leading to the result. \square

Corollary D.7.17.

$$\text{Cum} [V_{k, t+1}, R_{i, t, t+1}, \xi_{\tilde{i}, \tilde{j}, t, t+1} \middle| \mathcal{F}_{t\Delta}]$$

$$\approx \frac{\Delta}{N} \sum_{\tilde{k}=1}^{N_V} \mathcal{H}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right],$$

where

$$\mathcal{H}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} = \sigma_k c_{i,k} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}}^2 2\rho_k.$$

Proof. Based on Corollaries D.7.4 and D.7.7, we have

$$\begin{aligned} & \text{Cum} \left[V_{k,t\Delta+\frac{n\Delta}{N}}, R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \text{Cum} \left[\begin{array}{l} \left(\sigma_k \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\hat{k}=1}^{N_V} c_{i,\hat{k}} \sqrt{V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} z_{\hat{k},t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\ & = \sum_{\tilde{k}=1}^{N_V} \mathcal{H}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}, \end{aligned}$$

which $\mathcal{H}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}}$ is solved with Isserlis' Theorem:

$$\mathbb{E} \left[w_{k,t\Delta+\frac{n\Delta}{N}} z_{k,t\Delta+\frac{n\Delta}{N}} (z_{k,t\Delta+\frac{n\Delta}{N}}^2 - 1) \right] = 2\rho_k.$$

Following the the same steps detailed in Corollaries D.7.13 to D.7.16 leads to the results. \square

Corollary D.7.18.

$$\begin{aligned} & \text{Cum} \left[V_{k,t+1}, V_{\tilde{k},t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \frac{\Delta}{N} \sum_{\tilde{k}=1}^{N_V} \mathcal{I}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \end{aligned}$$

where

$$\mathcal{I}_{k,\tilde{k}}^{i,\tilde{i},\tilde{j}} = c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} \sigma_k^2 2\rho_k^2.$$

Proof. Based on Corollaries D.7.4 and D.7.7, we have

$$\begin{aligned}
 & \text{Cum} \left[V_{k,t\Delta+\frac{n\Delta}{N}}, V_{\hat{k},t\Delta+\frac{n\Delta}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\
 & \approx \text{Cum} \left[\begin{array}{c} \left(\sigma_k \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sigma_{\hat{k}} \sqrt{V_{\hat{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{\hat{k},t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\
 & = \sum_{\tilde{k}=1}^{N_V} \mathcal{I}_{k,\tilde{k}}^{\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2},
 \end{aligned}$$

which $\mathcal{I}_{k,\tilde{k}}^{\tilde{i},\tilde{j}}$ is solved with Isserlis' Theorem

$$\mathbb{E} \left[w_{k,t\Delta+\frac{n\Delta}{N}}^2 (z_{k,t\Delta+\frac{n\Delta}{N}}^2 - 1) \right] = 2\rho_k^2.$$

Following the the same steps detailed in Corollaries D.7.13 to D.7.16 leads to the results. \square

Corollary D.7.19.

$$\begin{aligned}
 & \text{Cum} \left[V_{k,t+1}, QCV_{\hat{i},\hat{j},t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\
 & \approx \frac{\Delta}{N} \frac{\Delta}{2} \sum_{\tilde{k}=1}^{N_V} \mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right],
 \end{aligned}$$

where

$$\mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}} = \sigma_k^2 c_{\hat{i},k} c_{\hat{j},k} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} 2\rho_k^2.$$

Proof. Based on Corollaries D.7.4, D.7.5 and D.7.7, we have

$$\text{Cum} \left[V_{k,t\Delta+\frac{n\Delta}{N}}, QCV_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right]$$

$$\begin{aligned}
 &\approx \text{Cum} \left[\begin{array}{c} \left(\sigma_k \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}}} w_{k,t\Delta + \frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} \Delta \frac{(N-n+1)}{N} c_{i,\tilde{k}} c_{j,\tilde{k}} \sigma_{\tilde{k}} \sqrt{V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}}} w_{\tilde{k},t\Delta + \frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{i,\tilde{k}} c_{j,\tilde{k}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} (z_{\tilde{k},t\Delta + \frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta + \frac{(n-1)\Delta}{N}} \right] \\
 &= \sum_{\tilde{k}=1}^{N_V} \frac{\Delta(N-n+1)}{N} \mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2},
 \end{aligned}$$

which $\mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}}$ is solved with Isserlis' Theorem

$$\mathbb{E} \left[w_{k,t\Delta + \frac{n\Delta}{N}}^2 (z_{\tilde{k},t\Delta + \frac{n\Delta}{N}}^2 - 1) \right] = 2\rho_k^2.$$

Following the the same steps detailed in Corollaries D.7.13 to D.7.16 and using Lemma D.7.3 with $X_1 = V_{k,(t+1)\Delta}$, $X_2 = QCV_{i,\hat{j},t,t+1}$ and $X_3 = \xi_{i,\tilde{j},t,t+1}$, we have

$$\begin{aligned}
 &\text{Cum} \left[V_{k,(t+1)\Delta}, QCV_{i,\hat{j},t,t+1}, \xi_{i,\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\
 &= \sum_{n=1}^N \mathbb{E} \left[\text{Cum} \left[V_{k,t\Delta + \frac{n\Delta}{N}}, QCV_{i,\hat{j},t + \frac{n-1}{N},t + \frac{n}{N}}, \xi_{i,\tilde{j},t + \frac{n-1}{N},t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \middle| \mathcal{F}_{t\Delta} \right] \\
 &\approx \sum_{n=1}^N \mathbb{E} \left[\sum_{\tilde{k}=1}^{N_V} \frac{\Delta(N-n+1)}{N} \mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \middle| \mathcal{F}_{t\Delta} \right] \\
 &\approx \frac{\Delta(N+1)}{2N} \sum_{n=1}^N \mathbb{E} \left[\sum_{\tilde{k}=1}^{N_V} \mathcal{L}_{k,\tilde{k}}^{\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \middle| \mathcal{F}_{t\Delta} \right],
 \end{aligned}$$

which leads to the result as we approximate $\frac{\Delta(N+1)}{2N}$ with $\frac{1}{2}$. □

Corollary D.7.20.

$$\begin{aligned}
 &\text{Cum} \left[R_{i,t,t+1}, QCV_{i,\hat{j},t,t+1}, \xi_{i,\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\
 &\approx \frac{\Delta}{N} \frac{\Delta}{2} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{M}_{k,\tilde{k}}^{i,\hat{i},\hat{j},\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right],
 \end{aligned}$$

where

$$\mathcal{M}_{k,\tilde{k}}^{i,\hat{i},\hat{j},\tilde{i},\tilde{j}} = \sigma_k c_{i,k} c_{\hat{i},k} c_{\hat{j},k} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} 2\rho_k.$$

Proof. Following the same steps as in Corollary D.7.18 leads to the results. \square

Corollary D.7.21.

$$\begin{aligned} & \text{Cum} \left[QCV_{i,j,t,t+1}, QCV_{\hat{i},\hat{j},t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \frac{\Delta}{N} \frac{\Delta^2}{3} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{J}_{k,\tilde{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \end{aligned}$$

where

$$\mathcal{J}_{k,\tilde{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}} = \sigma_k^2 c_{i,k} c_{j,k} c_{\hat{i},k} c_{\hat{j},k} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} 2\rho_k^2.$$

Proof. Based on Corollaries D.7.5 and D.7.7, we have

$$\begin{aligned} & \text{Cum} \left[QCV_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, QCV_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \\ & \approx \text{Cum} \left[\begin{array}{l} \left(\sum_{k=1}^{N_V} \Delta \frac{(N-n+1)}{N} c_{i,k} c_{j,k} \sigma_k \sqrt{V_{k,t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} \Delta \frac{(N-n+1)}{N} c_{\hat{i},\tilde{k}} c_{\hat{j},\tilde{k}} \sigma_{\tilde{k}} \sqrt{V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{\tilde{k},t\Delta+\frac{n\Delta}{N}} \right), \\ \left(\sum_{\tilde{k}=1}^{N_V} c_{\tilde{i},\tilde{k}} c_{\tilde{j},\tilde{k}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta}{N} (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right) \end{array} \middle| \mathcal{F}_{t\Delta+\frac{(n-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \frac{\Delta^2 (N-n+1)^2}{N^2} \mathcal{J}_{k,\tilde{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2}, \end{aligned}$$

which $\mathcal{J}_{k,\tilde{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}}$ is solved with Isserlis' Theorem

$$\mathbb{E} \left[w_{k,t\Delta+\frac{n\Delta}{N}}^2 (z_{\tilde{k},t\Delta+\frac{n\Delta}{N}}^2 - 1) \right] = 2\rho_k^2.$$

Following the the same steps detailed in Corollaries D.7.13 to D.7.16 and using Lemma

D.7.3 with $X_1 = QCV_{i,j,t,t+1}$, $X_2 = QCV_{\hat{i},\hat{j},t,t+1}$ and $X_3 = \xi_{\tilde{i},\tilde{j},t,t+1}$, we have

$$\begin{aligned}
 & \text{Cum} \left[QCV_{i,j,t,t+1}, QCV_{\hat{i},\hat{j},t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\
 &= \sum_{n=1}^N \mathbb{E} \left[\text{Cum} \left[QCV_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, QCV_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \mid \mathcal{F}_{t\Delta} \right] \\
 &\approx \sum_{n=1}^N \mathbb{E} \left[\sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \frac{\Delta^2(N-n+1)^2}{N^2} \mathcal{J}_{k,\bar{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\bar{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \mid \mathcal{F}_{t\Delta} \right] \\
 &\approx \frac{\Delta^2(N+1)(2N+1)}{6N^2} \sum_{n=1}^N \mathbb{E} \left[\sum_{k=1}^{N_V} \sum_{\bar{k}=1}^{N_V} \mathcal{J}_{k,\bar{k}}^{i,j,\hat{i},\hat{j},\tilde{i},\tilde{j}} V_{k,t\Delta+\frac{(n-1)\Delta}{N}} V_{\bar{k},t\Delta+\frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \mid \mathcal{F}_{t\Delta} \right],
 \end{aligned}$$

which leads to the result as we approximate $\frac{\Delta^2(N+1)(2N+1)}{6N^2}$ with $\frac{1}{3}$. \square

D.7.6 Application to moments of $\xi_{\tilde{i},\tilde{j},t,t+1}$ $\xi_{\hat{i},\hat{j},t,t+1}$

Corollary D.7.22. *If Assumption 3.4.2 holds,*

$$\begin{aligned}
 & \text{Cum} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0, \\
 & \text{Cum} \left[V_{k,t\Delta+\frac{n-1}{N}\Delta}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0, \\
 & \text{Cum} \left[QCV_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \approx 0.
 \end{aligned}$$

Proof. Based on Corollaries D.7.4, D.7.5, and D.7.7, the odd moments of a Gaussian random variable are null. \square

Corollary D.7.23. *If Assumption 3.4.2 holds,*

$$\begin{aligned}
 & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} \xi_{\tilde{i},\tilde{j},t+\frac{m-1}{N},t+\frac{m}{N}}, \\ \text{Cov} \left[V_{k,t\Delta+\frac{n-1}{N}\Delta}, \xi_{\hat{i},\hat{j},t+\frac{m-1}{N},t+\frac{m}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \end{array} \mid \mathcal{F}_{t\Delta} \right] \approx 0, \\
 & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} \xi_{\tilde{i},\tilde{j},t+\frac{m-1}{N},t+\frac{m}{N}}, \\ \text{Cov} \left[R_{i,t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{m-1}{N},t+\frac{m}{N}} \mid \mathcal{F}_{t\Delta+\frac{n-1}{N}\Delta} \right] \end{array} \mid \mathcal{F}_{t\Delta} \right] \approx 0,
 \end{aligned}$$

$$\text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} \xi_{\tilde{i}, \tilde{j}, t + \frac{m-1}{N}, t + \frac{m}{N}}, \\ \text{Cov} \left[QCV_{i, j, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{m-1}{N}, t + \frac{m}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \approx 0.$$

Proof. Direct application of Corollary D.7.9. \square

Corollary D.7.24. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} R_{i, t + \frac{m-1}{N}, t + \frac{m}{N}}, \\ \text{Cov} \left[\xi_{\tilde{i}, \tilde{j}, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{2} \mathbb{E} \left[(\sigma_k c_{i, k} \rho_k + \sigma_{\tilde{k}} \rho_{\tilde{k}} c_{i, \tilde{k}}) V_{k, t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right]. \end{aligned}$$

Proof.

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} R_{i, t + \frac{m-1}{N}, t + \frac{m}{N}}, \\ \text{Cov} \left[\xi_{\tilde{i}, \tilde{j}, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \mathbb{E} \left[\begin{array}{c} \left(\sum_{m=1}^{n-1} \left(\sum_{k=1}^{N_V} c_{i, k} \sqrt{V_{k, t\Delta + \frac{(n-1)\Delta}{N}}} \frac{\Delta}{N} z_{k, t\Delta + \frac{n\Delta}{N}} \right) \right) \\ \left(\sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{C}_{\hat{k}, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} V_{\tilde{k}, t\Delta + \frac{(n-1)\Delta}{N}} V_{\hat{k}, t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{C}_{\hat{k}, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} c_{i, k} \frac{\Delta^3}{N^3} \mathbb{E} \left[\begin{array}{c} \left(\sqrt{V_{k, t\Delta + \frac{(m-1)\Delta}{N}}} z_{k, t\Delta + \frac{m\Delta}{N}} \right) \\ \left(\begin{array}{c} \sigma_{\tilde{k}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}}} w_{\tilde{k}, t\Delta + \frac{m\Delta}{N}} \\ + \sigma_{\hat{k}} V_{\hat{k}, t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\hat{k}, t\Delta + \frac{(m-1)\Delta}{N}}} w_{\hat{k}, t\Delta + \frac{m\Delta}{N}} \end{array} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[(\sigma_k c_{i, k} \rho_k + \sigma_{\tilde{k}} \rho_{\tilde{k}} c_{i, \tilde{k}}) V_{k, t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[(N - n + 1) (\sigma_k c_{i, k} \rho_k + \sigma_{\tilde{k}} \rho_{\tilde{k}} c_{i, \tilde{k}}) V_{k, t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\tilde{i}, \tilde{j}, \hat{i}, \hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{2} \mathbb{E} \left[(\sigma_k c_{i, k} \rho_k + \sigma_{\tilde{k}} \rho_{\tilde{k}} c_{i, \tilde{k}}) V_{k, t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right]. \quad \square \end{aligned}$$

Corollary D.7.25. *If Assumption 3.4.2 holds,*

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} V_{k,t\Delta + \frac{m\Delta}{N}}, \\ \text{Cov} \left[\xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{2} \mathbb{E} \left[(\sigma_k \sigma_k + \sigma_{\tilde{k}} \sigma_{\tilde{k}}) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right]. \end{aligned}$$

Proof.

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} V_{k,t\Delta + \frac{m\Delta}{N}}, \\ \text{Cov} \left[\xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \mathbb{E} \left[\begin{array}{c} \left(\sum_{m=1}^{n-1} \left(\sigma_k \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}} \right) \right) \\ \left(\sum_{\hat{k}=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{\hat{k}, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} V_{\hat{k}, t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{\tilde{k}=1}^{N_V} \sum_{\hat{k}=1}^{N_V} \mathcal{C}_{\hat{k}, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} \sigma_k \frac{\Delta^3}{N^3} \mathbb{E} \left[\begin{array}{c} \left(\sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta}{N}} w_{k,t\Delta + \frac{n\Delta}{N}} \right) \\ \left(\begin{array}{c} \sigma_{\tilde{k}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\hat{k}, t\Delta + \frac{(m-1)\Delta}{N}} w_{\hat{k}, t\Delta + \frac{m\Delta}{N}} \\ + \sigma_{\hat{k}} V_{\hat{k}, t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} w_{\tilde{k}, t\Delta + \frac{m\Delta}{N}} \end{array} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[(\sigma_k \sigma_k + \sigma_{\tilde{k}} \sigma_{\tilde{k}}) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[(N - n + 1) (\sigma_k \sigma_k + \sigma_{\tilde{k}} \sigma_{\tilde{k}}) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k, \tilde{k}}^{\hat{i}, \hat{j}, \hat{i}, \hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{2} \mathbb{E} \left[(\sigma_k \sigma_k + \sigma_{\tilde{k}} \sigma_{\tilde{k}}) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k}, t\Delta + \frac{(m-1)\Delta}{N}} \right]. \quad \square \end{aligned}$$

Corollary D.7.26. *If Assumption 3.4.2 holds,*

$$\text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} QCV_{i,j,t + \frac{n-1}{N}, t + \frac{n}{N}}, \\ \text{Cov} \left[\xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}}, \xi_{\hat{i}, \hat{j}, t + \frac{n-1}{N}, t + \frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right]$$

$$\approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^3}{N^3} \frac{\Delta}{3} \mathbb{E} \left[\left(\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\tilde{k}}^2 c_{i,\tilde{k}} c_{j,\tilde{k}} \right) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \right].$$

Proof.

$$\begin{aligned} & \text{Cov} \left[\begin{array}{c} \sum_{m=1}^{n-1} Q C V_{i,j,t+\frac{n-1}{N},t+\frac{n}{N}}, \\ \text{Cov} \left[\xi_{\tilde{i},\tilde{j},t+\frac{n-1}{N},t+\frac{n}{N}}, \xi_{\hat{i},\hat{j},t+\frac{n-1}{N},t+\frac{n}{N}} \middle| \mathcal{F}_{t\Delta + \frac{n-1}{N}\Delta} \right] \end{array} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \mathbb{E} \left[\begin{array}{c} \left(\sum_{m=1}^{n-1} \left(\sum_{k=1}^{N_V} \Delta \frac{(N-n+1)}{N} c_{i,k} c_{j,k} \sigma_k \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}}} w_{k,t\Delta + \frac{n\Delta}{N}} \right) \right) \\ \left(\sum_{\tilde{k}=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(n-1)\Delta}{N}} \frac{\Delta^2}{N^2} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{\tilde{k},\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[\begin{array}{c} \left(\Delta \frac{(N-n+1)}{N} c_{i,k} c_{j,k} \sigma_k \sqrt{V_{k,t\Delta + \frac{(n-1)\Delta}{N}}} w_{k,t\Delta + \frac{n\Delta}{N}} \right) \\ \left(\begin{array}{c} \sigma_{\tilde{k}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}}} w_{\tilde{k},t\Delta + \frac{m\Delta}{N}} \\ + \sigma_{\tilde{k}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \sqrt{V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}}} w_{\tilde{k},t\Delta + \frac{m\Delta}{N}} \end{array} \right) \end{array} \right] \\ & = \sum_{m=1}^{n-1} \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \frac{\Delta(N-n+1)}{N} \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[\begin{array}{c} \left(\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\tilde{k}}^2 c_{i,\tilde{k}} c_{j,\tilde{k}} \right) \\ V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \end{array} \right] \\ & = \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^3}{N^3} \mathbb{E} \left[\frac{\Delta(N-n+1)^2}{N} \left(\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\tilde{k}}^2 c_{i,\tilde{k}} c_{j,\tilde{k}} \right) \right. \\ & \quad \left. V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^3}{N^3} \frac{\Delta}{3} \mathbb{E} \left[\left(\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\tilde{k}}^2 c_{i,\tilde{k}} c_{j,\tilde{k}} \right) V_{k,t\Delta + \frac{(m-1)\Delta}{N}} V_{\tilde{k},t\Delta + \frac{(m-1)\Delta}{N}} \right]. \quad \square \end{aligned}$$

Corollary D.7.27.

$$\begin{aligned} & \text{Cum} \left[R_{i,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \left(\sigma_k c_{i,k} \rho_k + \sigma_{\tilde{k}} \rho_{\tilde{k}} c_{i,\tilde{k}} \right) \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta}{N} \frac{\Delta}{2} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \\ & \text{Cum} \left[V_{k,(t+1)\Delta}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \left(\sigma_k^2 + \sigma_{\tilde{k}}^2 \right) \mathcal{C}_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta}{N} \frac{\Delta}{2} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \middle| \mathcal{F}_{t\Delta} \right], \end{aligned}$$

$$\begin{aligned} & \text{Cum} \left[QCV_{i,j,t,t+1}, \xi_{\tilde{i},\tilde{j},t,t+1}, \xi_{\hat{i},\hat{j},t,t+1} \mid \mathcal{F}_{t\Delta} \right] \\ & \approx \sum_{k=1}^{N_V} \sum_{\tilde{k}=1}^{N_V} \left(\sigma_k^2 c_{i,k} c_{j,k} + \sigma_{\tilde{k}}^2 c_{i,\tilde{k}} c_{j,\tilde{k}} \right) C_{k,\tilde{k}}^{\tilde{i},\tilde{j},\hat{i},\hat{j}} \frac{\Delta^2}{N^2} \frac{\Delta}{3} \mathbb{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_{k,s} V_{\tilde{k},s} ds \mid \mathcal{F}_{t\Delta} \right]. \end{aligned}$$

Proof. Application of Lemma D.7.3 and Corollaries D.7.22 to D.7.26.

D.8 Second Update

We outline our approach to manage the transition from the σ -algebra $\mathcal{G}_{t\Delta}$ to $\mathcal{G}_{(t+1)\Delta}$ for the conditional variance of the latent variable prevision error derived in Equation (3.27).

To update the variance matrix of the latent state estimator

$$\mathbb{E} \left[\left(\mathbf{V}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{V}_{(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right] \right) \left(\mathbf{V}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{V}_{(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right] \right)^\top \mid \mathcal{G}_{t\Delta} \right],$$

we define the variable

$$\begin{aligned} \mathbf{C}_{i,j,(t+1)\Delta} &= \left(\mathbf{V}_{i,(t+1)\Delta} - \mathbb{E} \left[\mathbf{V}_{i,(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right] \right) \left(\mathbf{V}_{j,(t+1)\Delta} - \mathbb{E} \left[\mathbf{V}_{j,(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right] \right) \\ &\simeq \left(\epsilon_{i,t+1|t}^{\mathbf{V}} - \Sigma_{i,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right) \left(\epsilon_{j,t+1|t}^{\mathbf{V}} - \Sigma_{j,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right), \end{aligned}$$

where $\Sigma_{i,t\Delta}$ is the i th line of the update matrix $\Sigma_{t\Delta}$ and

$$\mathbf{C}_{(t+1)\Delta} = [\mathbf{C}_{1,1,(t+1)\Delta}, \mathbf{C}_{1,2,(t+1)\Delta}, \dots, \mathbf{C}_{K,K,(t+1)\Delta}].$$

The a posteriori estimation⁵

$$\underbrace{\mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right]}_{\substack{\text{a posteriori estimation} \\ \text{of the latent variables} \\ \text{uncertainty}}} \simeq \underbrace{\mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} \mid \mathcal{G}_{t\Delta} \right]}_{\substack{\text{a priori prediction} \\ \text{of the latent variables} \\ \text{uncertainty}}} + \Gamma_{t\Delta} \underbrace{\left(\mathbf{O}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{O}_{(t+1)\Delta} \mid \mathcal{G}_{t\Delta} \right] \right)}_{\epsilon_{t+1|t}^{\mathbf{O}}}. \quad (\text{D.12})$$

The matrix $\Gamma_{t\Delta}$ is now constructed by minimizing the conditional variance of the latent

⁵In the very rare cases that the update does not yield to a semi definite positive covariance matrix, we set $\mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} \mid \mathcal{G}_{(t+1)\Delta} \right] \simeq \mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} \mid \mathcal{G}_{t\Delta} \right]$.

variable estimator given $\mathcal{G}_{t\Delta}$. Because

$$\underbrace{\mathbf{C}_{(t+1)\Delta} - \mathbb{E}[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta}]}_{\boldsymbol{\epsilon}_{t+1|t+1}^{\mathbf{C}}} \simeq \underbrace{\mathbf{C}_{(t+1)\Delta} - \mathbb{E}[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{t\Delta}]}_{\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}}} - \boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}},$$

the variance matrix of the estimator with respect to $\mathcal{G}_{t\Delta}$ is

$$\begin{aligned} & \mathbb{E} \left[\left(\mathbf{C}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta} \right] \right) \left(\mathbf{C}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta} \right] \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] \\ & \simeq \mathbb{E} \left[\left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} - \boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} - \boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] \\ & = \mathbb{E} \left[\left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \right) \left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] - \mathbb{E} \left[\left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \right) \left(\boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] \\ & \quad - \mathbb{E} \left[\left(\boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] + \mathbb{E} \left[\left(\boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\Gamma}_{t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] \\ & = \text{Var} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \middle| \mathcal{G}_{t\Delta} \right] + \boldsymbol{\Gamma}_{t\Delta} \text{Var} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \boldsymbol{\Gamma}_{t\Delta}^{\top} \\ & \quad - \text{Cov} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}}, \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \boldsymbol{\Gamma}_{t\Delta}^{\top} - \boldsymbol{\Gamma}_{t\Delta} \text{Cov} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}}, \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \middle| \mathcal{G}_{t\Delta} \right]. \end{aligned} \quad (\text{D.13})$$

Therefore, if the a posteriori covariance matrix is minimized (see proof in Appendix B.3) respective to $\boldsymbol{\Gamma}_{t\Delta}$, the second update matrix is

$$\boldsymbol{\Gamma}_{t\Delta} = \text{Cov} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}}, \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \left(\text{Var} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \right)^{-1}. \quad (\text{D.14})$$

Replacing (D.14) in (D.13) leads to

$$\begin{aligned} & \mathbb{E} \left[\left(\mathbf{C}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta} \right] \right) \left(\mathbf{C}_{(t+1)\Delta} - \mathbb{E} \left[\mathbf{C}_{(t+1)\Delta} | \mathcal{G}_{(t+1)\Delta} \right] \right)^{\top} \middle| \mathcal{G}_{t\Delta} \right] \\ & \simeq \text{Var} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \middle| \mathcal{G}_{t\Delta} \right] - \boldsymbol{\Gamma}_{t\Delta} \text{Cov} \left[\boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}}, \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{C}} \middle| \mathcal{G}_{t\Delta} \right]. \end{aligned}$$

The covariance term is

$$\begin{aligned} & \text{Cov} \left[\boldsymbol{\epsilon}_{i,j,t+1|t}^{\mathbf{C}}, \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \\ & \simeq \text{Cov} \left[\left(\boldsymbol{\epsilon}_{i,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{i,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{j,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{j,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right), \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \\ & = \mathbb{E} \left[\left(\boldsymbol{\epsilon}_{i,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{i,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \left(\boldsymbol{\epsilon}_{j,t+1|t}^{\mathbf{V}} - \boldsymbol{\Sigma}_{j,t\Delta} \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \right) \boldsymbol{\epsilon}_{t+1|t}^{\mathbf{O}} \middle| \mathcal{G}_{t\Delta} \right] \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} [\epsilon_{i,t+1|t}^{\mathbf{V}} \epsilon_{j,t+1|t}^{\mathbf{V}} \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] - \mathbb{E} [\epsilon_{i,t+1|t}^{\mathbf{V}} (\sum_{j,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}}) \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] \\
 &- \mathbb{E} [(\sum_{i,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}}) \epsilon_{j,t+1|t}^{\mathbf{V}} \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] + \mathbb{E} [(\sum_{i,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}}) (\sum_{j,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}}) \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}],
 \end{aligned}$$

which is a linear combination of third moments on the latent variable and observable variables (see Online Appendices D.5 and D.6 for calculations details). Since the state and measurement equations are not Gaussian, the covariance $\text{Cov} [\epsilon_{i,j,t+1|t}^{\mathbf{C}}, \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}]$ is not null.⁶

If the second update is used, based on Equation (D.12) and with Corollary B.4.8, we have instead

$$\begin{aligned}
 &\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \tag{D.15} \\
 &\simeq \mathbb{E} \left[\text{Cum}^{(4)} [\epsilon_{i,t+1|t+1}^{\mathbf{V}}, \epsilon_{j,t+1|t+1}^{\mathbf{V}}, \epsilon_{l,t+1|t+1}^{\mathbf{V}}, \epsilon_{m,t+1|t+1}^{\mathbf{V}} | \mathcal{G}_{(t+1)\Delta}] \middle| \mathcal{G}_{t\Delta} \right] \\
 &= \text{Cum}^{(4)} \left[\begin{array}{c} \left(\epsilon_{i,t+1|t}^{\mathbf{V}} - \sum_{i,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right), \left(\epsilon_{j,t+1|t}^{\mathbf{V}} - \sum_{j,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right), \\ \left(\epsilon_{l,t+1|t}^{\mathbf{V}} - \sum_{l,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right), \left(\epsilon_{m,t+1|t}^{\mathbf{V}} - \sum_{m,t\Delta} \epsilon_{t+1|t}^{\mathbf{O}} \right) \end{array} \middle| \mathcal{G}_{t\Delta} \right] \\
 &- \Gamma_{i,j,t\Delta} \text{Cov} [\epsilon_{t+1|t}^{\mathbf{O}}, \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] \Gamma_{l,m,t\Delta}^{\top} - \Gamma_{i,m,t\Delta} \text{Cov} [\epsilon_{t+1|t}^{\mathbf{O}}, \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] \Gamma_{j,l,t\Delta}^{\top} \\
 &- \Gamma_{i,l,t\Delta} \text{Cov} [\epsilon_{t+1|t}^{\mathbf{O}}, \epsilon_{t+1|t}^{\mathbf{O}} | \mathcal{G}_{t\Delta}] \Gamma_{j,m,t\Delta}^{\top},
 \end{aligned}$$

which is a linear combination of the second and fourth cumulant of $\mathbf{O}_{(t+1)\Delta}$ and fourth cumulant of $\mathbf{V}_{(t+1)\Delta}$ with respect to the information set $\mathcal{G}_{t\Delta}$ (see section B.4.2.3). However, contrary to the third cumulant, Equation (3.29) is affected directly by the second update as the matrix $\Gamma_{t\Delta}$ is present in the equation.

⁶Indeed, if the framework was Gaussian, the term $\Gamma_{t\Delta}$ would be null and the second update useless. A discussion on the Gaussian assumption is outlined in Online Appendix D.4.

D.9 Unconditional Moments

Corollary D.9.1. *If $V_{k,(t+1)\Delta}$ is stationary, then the first two unconditional moments of the variance are given by*

$$\mathbb{E}[V_{k,t\Delta}] = \theta_k \quad \text{and} \quad \text{Var}[V_{k,t\Delta}] = \frac{\theta_k \sigma_k^2}{2\kappa_k}.$$

Proof. The expected value is a direct application of Equation (B.8). The variance is based on calculations from Appendix B.2, where:

$$\begin{aligned} \text{Var}[V_{k,(t+1)\Delta}] &= \mathbb{E}[\text{Var}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}]] + \text{Var}[\mathbb{E}[V_{k,t\Delta}|\mathcal{F}_{t\Delta}]] \\ &= \mathbb{E}[\sigma_k^2 \text{Var}[\mathcal{Z}_{k,t,t+1}|\mathcal{F}_{t\Delta}]] + \text{Var}[e^{-\kappa_k \Delta}(V_{k,t\Delta} - \theta_k) + \theta_k] \\ &= \theta_k \sigma_k^2 \frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} + \text{Var}[V_{k,t\Delta}] e^{-2\kappa_k \Delta}, \end{aligned}$$

leading to the result. □

Corollary D.9.2. *If $\text{Var}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}] = C_k V_{k,t\Delta} + D_k$, $\text{Cum}^{(3)}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}] = P_k V_{k,t\Delta} + Q_k$ and $V_{k,(t+1)\Delta}$ is stationary, then the third unconditional cumulant of the variance is given by*

$$\text{Cum}^{(3)}[V_{k,t\Delta}] = \frac{P_k \theta_k + Q_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} e^{-\kappa_k \Delta} C_k}{1 - e^{-3\kappa_k \Delta}}.$$

Proof. Based on Appendix B.2 calculations and the law of total cumulance, the third unconditional cumulant of the variance is:

$$\begin{aligned} \text{Cum}^{(3)}[V_{k,(t+1)\Delta}] &= \mathbb{E}[\text{Cum}^{(3)}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}]] + \text{Cum}^{(3)}[\mathbb{E}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}]] \\ &\quad + 3\text{Var}[\mathbb{E}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}], \text{Var}[V_{k,(t+1)\Delta}|\mathcal{F}_{t\Delta}]] \\ &= P_k \theta_k + Q_k + e^{-3\kappa_k \Delta} \text{Cum}^{(3)}[V_{k,t\Delta}] + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k e^{-\kappa_k \Delta}, \end{aligned}$$

leading to the result. See Appendix B.2 for the second conditional cumulant of $V_{k,(t+1)\Delta}$ and Online Appendix D.5 for the third conditional cumulant of $V_{k,(t+1)\Delta}$. □

Corollary D.9.3. *If $\text{Var} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] = C_k V_{k,t\Delta} + D_k$, $\text{Cum}^{(3)} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] = P_k V_{k,t\Delta} + Q_k$, $\text{Cum}^{(4)} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] = M_k V_{k,t\Delta} + N_k$ and $V_{k,(t+1)\Delta}$ is stationary, then the fourth unconditional cumulant of the variance is given by*

$$\begin{aligned} & \text{Cum}^{(4)} [V_{k,t\Delta}] \\ &= \frac{M_k \theta + N_k + 4 \frac{\theta_k \sigma_k^2}{2\kappa_k} e^{-\kappa_k \Delta} P_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k^2 + 6C_k e^{-2\kappa_k \Delta} \text{Cum}^{(3)} [V_{k,t\Delta}]}{1 - e^{-4\kappa_k \Delta}}. \end{aligned}$$

Proof. Based on Appendix B.2 calculations and the law of total cumulance, the fourth unconditional cumulant of the variance is:

$$\begin{aligned} \text{Cum}^{(4)} [V_{k,(t+1)\Delta}] &= \text{E} \left[\text{Cum}^{(4)} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \right] + \text{Cum}^{(4)} \left[\text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \right] \\ &+ 4 \text{Var} \left[\text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Cum}^{(3)} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \right] \\ &+ 3 \text{Var} \left[\text{Var} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \right] \\ &+ 6 \text{Cum}^{(3)} \left[\text{Var} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] \right] \\ &= M_k \theta + N_k + e^{-4\kappa_k \Delta} \text{Cum}^{(4)} [V_{k,t\Delta}] \\ &+ 4 \frac{\theta_k \sigma_k^2}{2\kappa_k} e^{-\kappa_k \Delta} P_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k^2 + 6C_k e^{-2\kappa_k \Delta} \text{Cum}^{(3)} [V_{k,t\Delta}], \end{aligned}$$

which leads to the result by isolating $\text{Cum}^{(4)} [V_{k,t\Delta}]$. See Appendix B.2 for the second conditional cumulant of $V_{k,(t+1)\Delta}$ and Online Appendix D.5 for the third and fourth conditional cumulants of $V_{k,(t+1)\Delta}$. \square

D.10 Additional Results

D.10.1 Stochastic integral sample moments

Figures D.2 to D.3 show the sample skewness and kurtosis respectively of the stochastic integrals $\mathcal{Z}_{k,t,t+1}$, $\mathcal{W}_{k,t,t+1}$ and $\mathcal{B}_{k,t,t+1}$ for case two assets and three variances. $\mathcal{W}_{k,t,t+1}$ displays similar distributional properties as $\mathcal{Z}_{k,t,t+1}$. Despite this, their skewness and kurtosis are not 0 and 3, respectively, suggesting a distribution different from the Gaussian distribu-

Online Appendix D. *Variance filtering in multidimensional affine models*

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	
Assuming that $\xi_{i,j,t,t+1} \sim \text{Gaussian}$						$\xi_{i,j,t,t+1} = RCV_{i,j,t,t+1} - QCV_{i,j,t,t+1}$					
(1,1)	0.97	0.0233	-0.0005	-0.01	99.57	0.97	0.0233	-0.0005	-0.01	99.57	
(2,1)	0.93	-1.5488	-0.0760	-0.42	67.67	0.93	-1.5488	-0.0760	-0.42	67.67	
(3,1)	0.95	2.2272	-0.0332	-0.17	86.59	0.95	2.2272	-0.0332	-0.17	86.59	
(1,2)	0.95	0.0781	0.0019	0.26	79.55	0.95	0.0781	0.0019	0.26	79.55	
(2,2)	0.86	0.3496	-0.0001	0.00	99.60	0.86	0.3496	-0.0001	0.00	99.60	
(3,2)	0.91	0.0952	-0.0061	-0.14	88.77	0.91	0.0952	-0.0061	-0.14	88.77	
(1,3)	0.92	-0.0057	-0.0005	-0.46	64.86	0.92	0.0081	0.0005	0.47	64.08	
(2,3)	0.80	-0.1407	-0.0228	-2.49	1.28*	0.80	-0.0746	-0.0113	-1.23	21.71	
(3,3)	0.87	-0.2418	-0.0279	-1.12	26.15	0.87	-0.1068	-0.0098	-0.40	69.26	
(1,4)	0.89	0.0025	0.0002	0.99	32.33	0.89	0.0026	0.0002	1.03	30.14	
(2,4)	0.75	0.0383	0.0050	1.62	10.52	0.75	0.0390	0.0051	1.67	9.48	
(3,4)	0.83	-0.1223	-0.0178	-1.16	24.62	0.83	-0.1191	-0.0172	-1.12	26.10	
Joint	χ_{12}^2	22.90			2.85*	χ_{12}^2	13.52			33.26	

This Monte Carlo study simulates $L = 1000$ independent trajectories for the Model (3.1), using $M = 3900$ intraday steps over a period of $T = 4000$ days. The daily RCV is aggregated at the 5 minutes frequency with $N = 390$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell)} = RCV_{i,j,t,t+1}^{(\ell)} - QCV_{i,j,t,t+1}^{(\ell)}$. Filtering errors,

$$\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right],$$

are computed daily for each variance component k and each path ℓ . The filtering errors' higher moments are

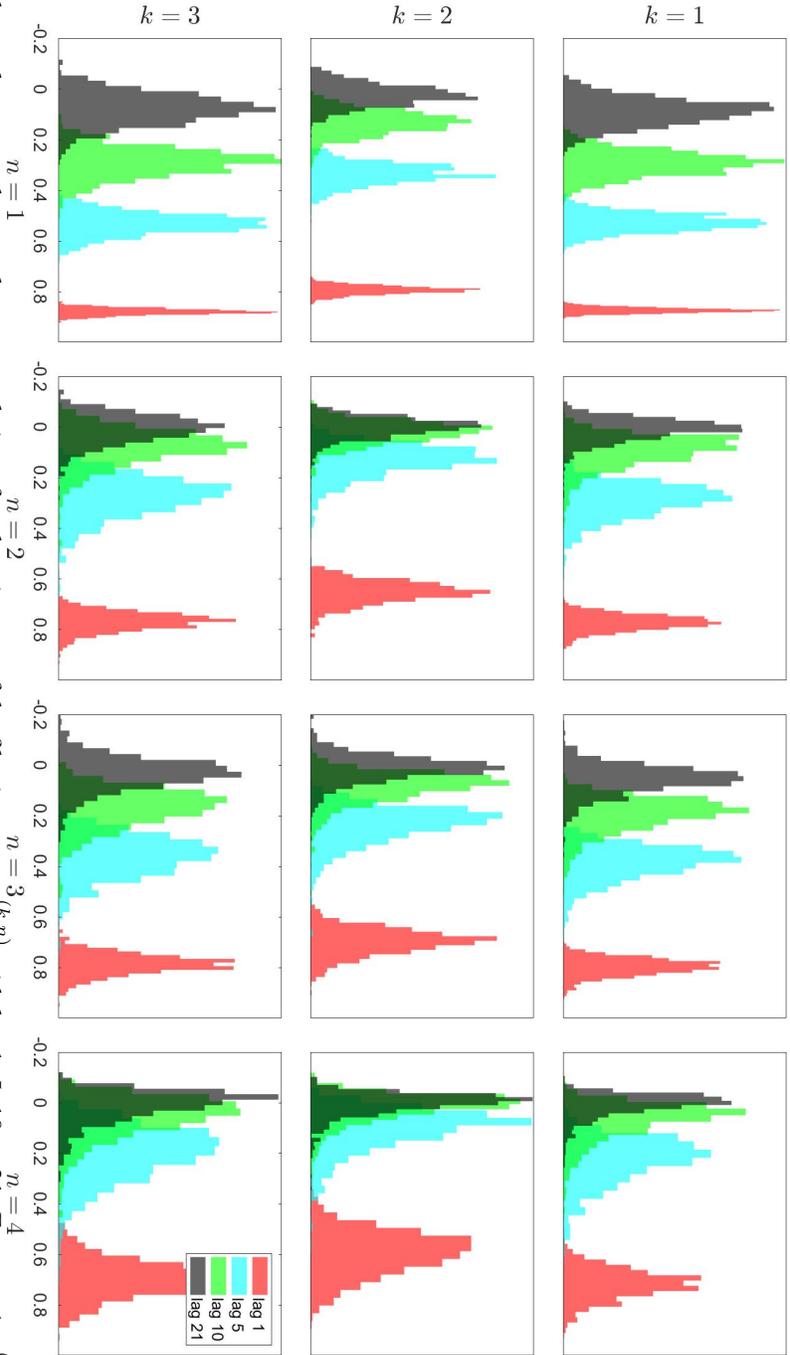
$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \mid \mathcal{G}_{t\Delta} \right], n \in \{2, 3, 4\}.$$

The average filtering error $\bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{k,t}^{(\ell,n)}$ is normalized by the long-term expected variance component θ_k raised to the appropriate power. This ratio, $\frac{\bar{\varepsilon}^{(k,n)}}{\theta_k^n}$, is expressed in basis points (bp). The panel regression

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}$$

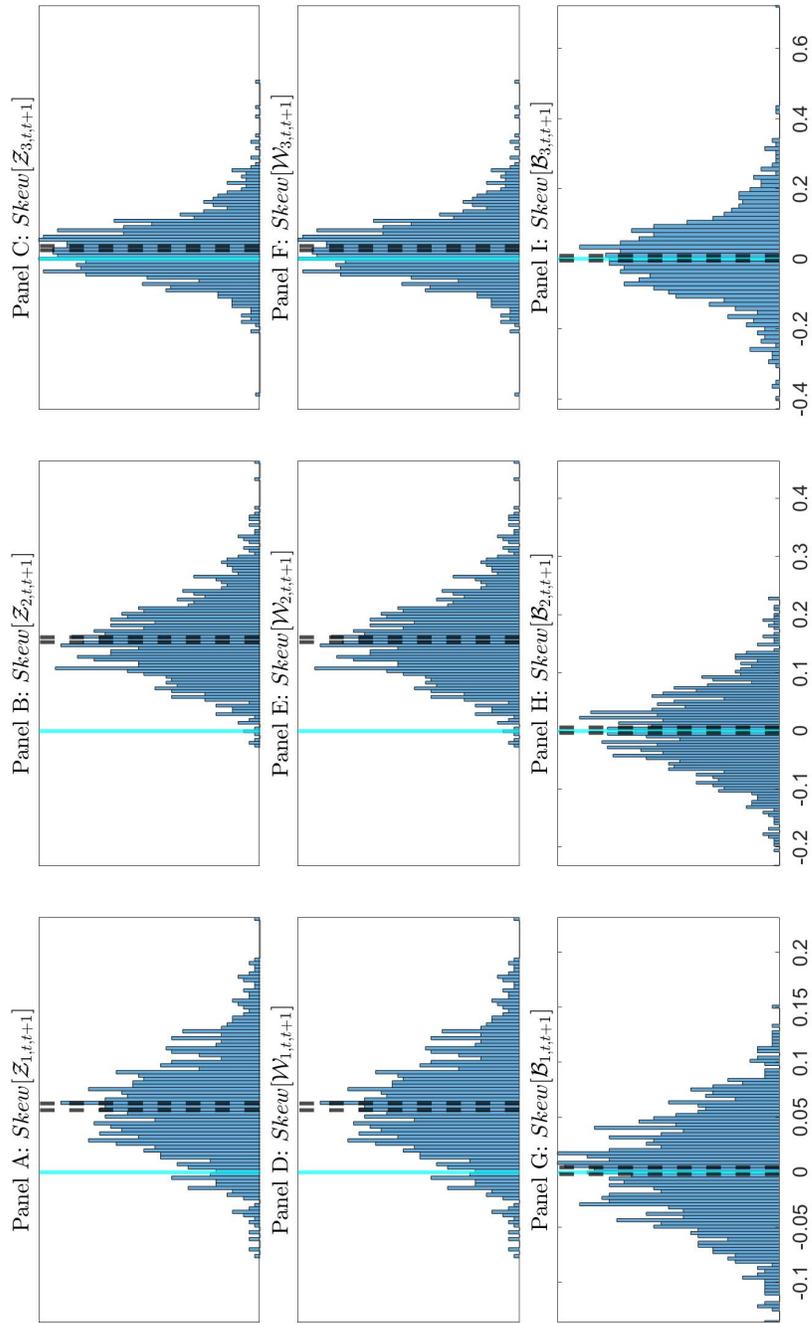
(as described in Section 3.4.2) accounts for autocorrelation and potential biases through the constants $\vartheta_k^{(n)}$. We test the absence of bias ($H_0 : \vartheta_k^{(n)} = 0$) and report the associated t-statistic (t-stat) and p-value. Standard errors are estimated considering spatial correlation, using the Newey-West estimator with a lag $Q = \lceil 4(\frac{T}{100})^{2/9} \rceil = 9$. An asterisk (*) denotes rejection of the null hypothesis at 5% and (**) at 1%. We provide the joint chi-square test statistic and p-value for the null hypothesis $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$. The average autocorrelation $\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{k,t}^{(\ell,n)}$ displays the cross-sectional and time-average of the autocorrelation.

Table D.1: Filtering error for case 2 assets and 3 variances ($N = 390$)



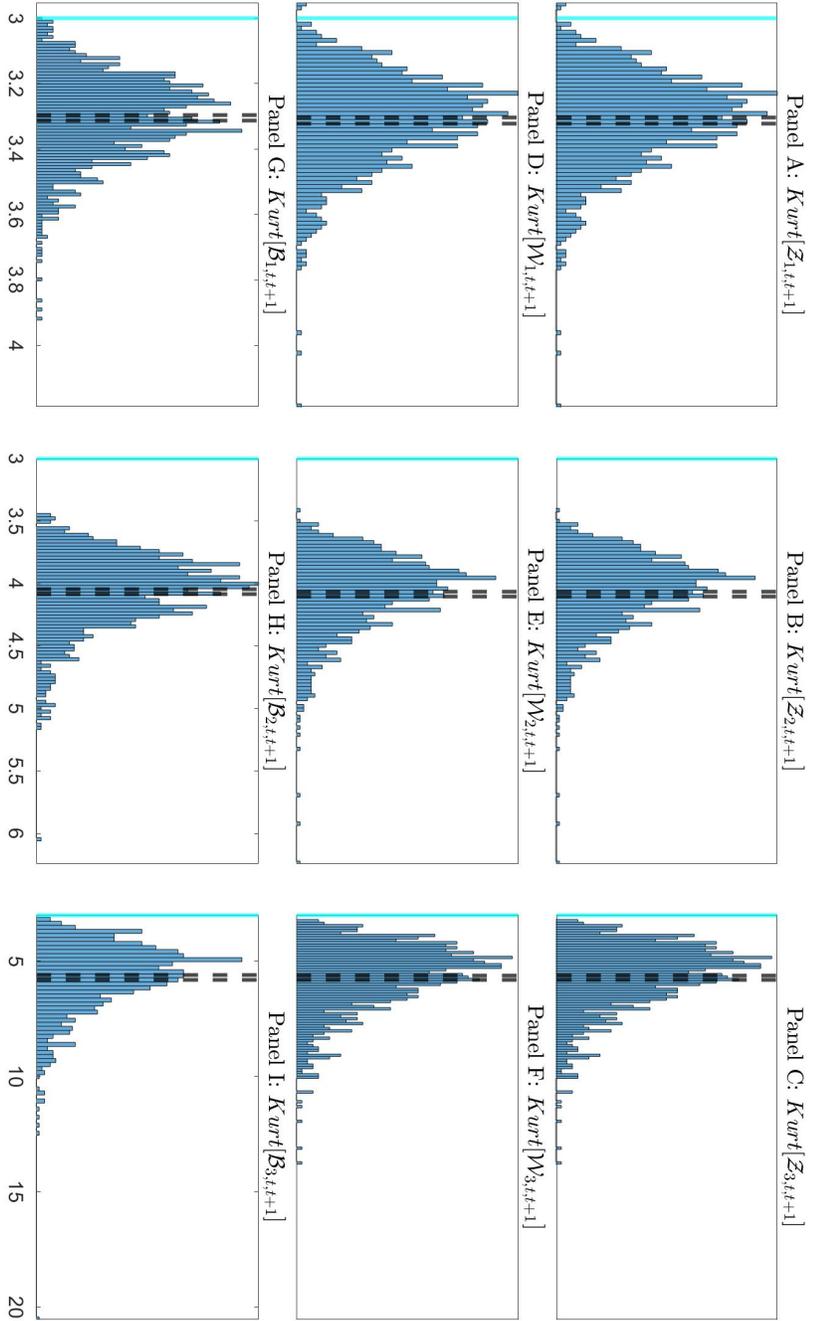
Each panel presents the sample autocorrelation of each trajectory of the filtering error $\varepsilon_{t,\ell}^{(k,n)}$ with lag 1, 5, 10 or 21. For $n = 1$, $\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right]$, where $V_{k,t\Delta}^{(\ell)}$ is simulated according to Section 3.4.1. $\mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \middle| \mathcal{G}_{t\Delta} \right]$ is the filtered expected value of the variance $V_{k,t\Delta}^{(\ell)}$ based on the filter derived in Section 3.3. For $n \in \{2, 3, 4\}$, $\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \middle| \mathcal{G}_{t\Delta} \right]$ where $\mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \middle| \mathcal{G}_{t\Delta} \right]$ is the conditional moment given by the filter. The realized variance is defined as $RCV_{i,j,t,t+1}^{(N)} = \sum_{n=1}^N R_{i,t+\frac{(n-1)}{N},t+\frac{n}{N}} R_{j,t+\frac{(n-1)}{N},t+\frac{n}{N}} = QCV_{i,j,t,t+1} + \xi_{i,j,t,t+1}^{(N)}$ and is filtered according to the correction as per Assumption 3.4.2; thus acknowledging that $\xi_{i,j,t,t+1}^{(N)}$ is not Gaussian.

Figure D.1: Autocorrelation of $\varepsilon_{t,\ell}^{(k,n)}$ for case 2 assets and 3 variances



This Monte Carlo study generates $L = 1000$ independent trajectories with $M = 3900$ intraday time steps for $T = 4000$ days for case 2 assets and 3 variances. The vertical blue lines correspond to the skewness of a standard normal random variable. The black dashed lines are the 95%-confidence interval around the average of the sample skewness.

Figure D.2: Sample skewness of stochastic integrals



This Monte Carlo study generates $L = 1000$ independent trajectories with $M = 3900$ intraday time steps for $T = 4000$ days for case 2 assets and 3 variances. The vertical blue lines correspond to the kurtosis of a standard normal random variable. The black dashed lines are the 95%-confidence interval around the average of the sample kurtosis.

Figure D.3: Sample kurtosis of stochastic integrals

tion. However, $\mathcal{B}_{k,t,t+1}$ has a null skewness since the Brownian motion $B_{k,t}$ is independent of the stochastic variances factors.

D.10.2 Sample correlation between the measurement error and stochastic integral

Based on Monte Carlo simulations, we empirically analyze Assumption B.1.1, which presumes that the stochastic integrals are uncorrelated with the error measurement. We present the sample correlations in Figures D.4,⁷ covering the most complex case, 2 assets and 3 variances, presented in this paper. The figures show that Assumption B.1.1 is empirically satisfied as the sample correlation between the measurement error and the stochastic integrals $\mathcal{W}_{t,t+1}$, $\mathcal{B}_{t,t+1}$ and $\mathcal{Z}_{t,t+1}$ are centered around zero.

D.10.3 Second update results under the asymptotic theory

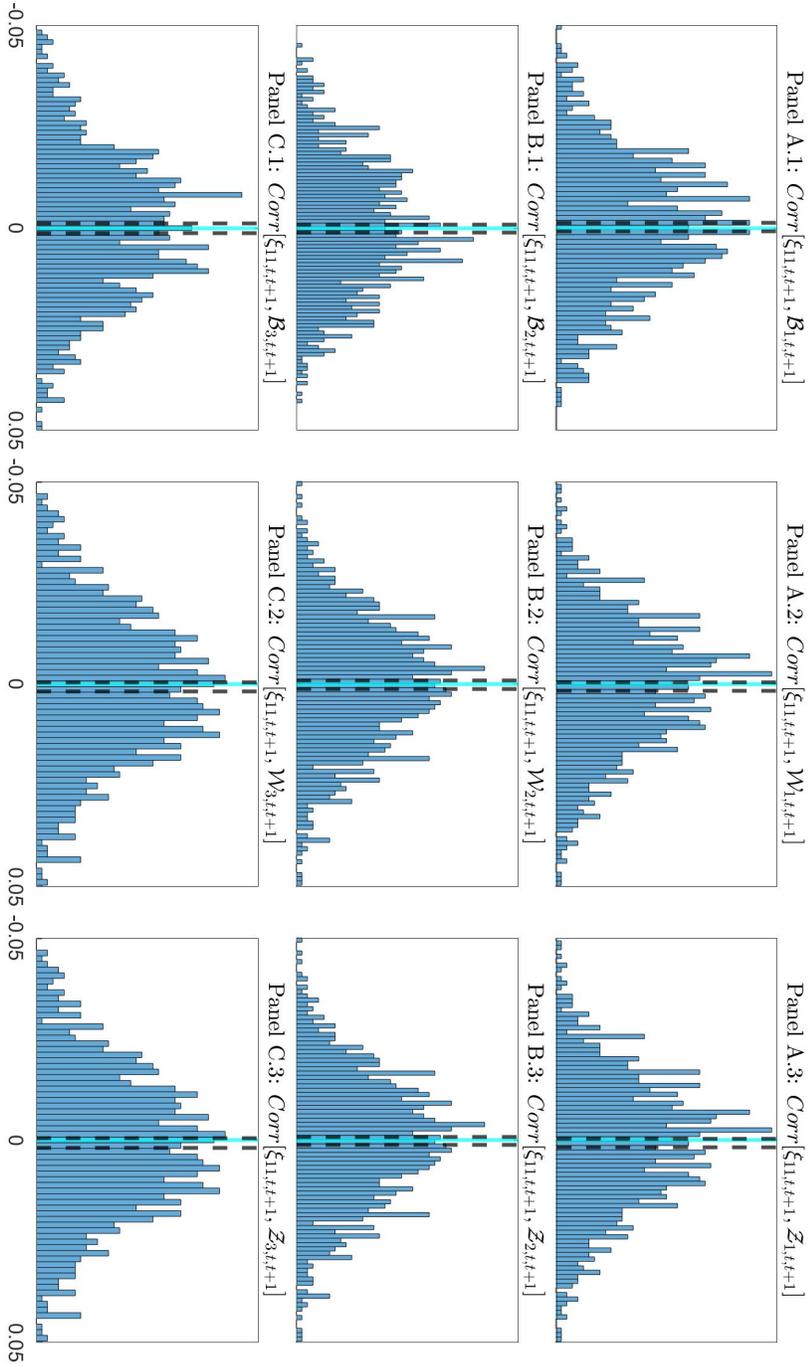
We present our results assuming the asymptotic theory is respected and also simulate an error according to Assumption (3.4.1). That is

$$RCV_{i,j,t,t+1}^{(\perp)} = QCV_{i,j,t,t+1} + \xi_{i,j,t,t+1}^{(\perp)}, \quad (\text{D.16})$$

where the measurement error $\xi_{i,j,t,t+1}^{(\perp)}$ is independently and normally distributed with zero mean and variance set to Assumption 3.4.1. The results are presented in tables D.2 to D.4.

For the right columns of tables D.2 to D.4, in addition to the measurement error simulated according to Assumption 3.4.1, we introduce the *Filter with second order update* (details in Online Appendix D.8). This additional step addresses the non-linear and non-Gaussian nature of our framework. If our model were completely linear and normally distributed, it would revert to a standard Kalman filter, which is optimal under Gaussianity assumptions. However, since we lack these properties, we added an extra step to reduce the gap between the conditional moments of our latent variables and the estimators obtained.

⁷Figures on measurement error $\xi_{1,2,t,t+1}$ and $\xi_{2,2,t,t+1}$ are available upon request.



This Monte Carlo study generates $L = 1000$ independent trajectories with $M = 3900$ intraday time steps for $T = 4000$ days. The daily RV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell,N)} = RCV_{i,j,t,t+1}^{(\ell,N)} - QCV_{i,j,t,t+1}^{(\theta)}$. The results are presented for case 2 assets and 3 variances. The vertical blue lines correspond to correlation of zero. The black dashed lines are the 95%-confidence interval around the sample mean.

Figure D.4: Sample correlation between $\xi_{1,t,t+1}$ and stochastic integrals

Online Appendix D. *Variance filtering in multidimensional affine models*

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)
<i>Filter with first order update</i>						<i>Filter with second order update</i>					
Panel A: Variance 1											
(1,1)	0.98	0.2705	0.0016	0.01	98.89	(1,1)	0.98	0.2862	0.0018	0.02	98.76
(1,2)	0.97	0.0359	0.0003	0.03	97.77	(1,2)	0.97	0.0618	0.0007	0.06	95.59
(1,3)	0.95	0.0236	0.0010	0.42	67.24	(1,3)	0.95	0.0247	0.0010	0.44	66.06
(1,4)	0.93	0.0025	0.0001	0.26	79.16	(1,4)	0.93	0.0037	0.0002	0.36	72.16
Joint	χ_4^2	0.34			98.73	Joint	χ_4^2	0.41			98.14
Panel B: Variance 2											
(1,1)	0.96	0.1310	0.0125	0.08	93.59	(1,1)	0.96	0.1794	0.0131	0.08	93.24
(1,2)	0.92	0.1647	0.0032	0.13	89.44	(1,2)	0.92	0.0767	0.0019	0.08	93.61
(1,3)	0.89	0.0622	0.0045	0.56	57.36	(1,3)	0.89	0.0602	0.0044	0.55	58.07
(1,4)	0.85	0.0097	0.0007	0.23	82.07	(1,4)	0.85	0.0051	0.0004	0.12	90.71
Joint	χ_4^2	0.41			98.15	Joint	χ_4^2	0.42			98.09
Panel C: Variance 3											
(1,1)	0.92	1.3299	0.0213	0.10	91.93	(1,1)	0.92	1.3502	0.0219	0.10	91.70
(1,2)	0.85	0.7841	0.0341	0.47	63.92	(1,2)	0.85	0.8782	0.0342	0.47	63.85
(1,3)	0.79	0.3334	0.0273	0.42	67.51	(1,3)	0.78	0.3465	0.0283	0.44	66.29
(1,4)	0.73	0.8320	0.0654	0.81	41.54	(1,4)	0.73	0.8774	0.0664	0.83	40.73
Joint	χ_4^2	0.67			95.44	Joint	χ_4^2	0.70			95.11

This Monte Carlo study simulates $L = 1000$ independent trajectories for the Model (3.1), using $M = 3900$ intraday steps over a period of $T = 4000$ days. The daily RCV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell)} = RCV_{i,j,t,t+1}^{(\ell)} - QCV_{i,j,t,t+1}^{(\ell)}$. Filtering errors,

$$\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right],$$

are computed daily for each variance component k and each path ℓ . The filtering errors' higher moments are

$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \mid \mathcal{G}_{t\Delta} \right], n \in \{2, 3, 4\}.$$

The average filtering error $\bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{k,t}^{(\ell,n)}$ is normalized by the long-term expected variance component θ_k raised to the appropriate power. This ratio, $\frac{\bar{\varepsilon}^{(k,n)}}{\theta_k^n}$, is expressed in basis points (bp). The panel regression

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}$$

(as described in Section 3.4.2) accounts for autocorrelation and potential biases through the constants $\vartheta_k^{(n)}$. We test the absence of bias ($H_0 : \vartheta_k^{(n)} = 0$) and report the associated t-statistic (t-stat) and p-value. Standard errors are estimated considering spatial correlation, using the Newey-West estimator with a lag $Q = \lfloor 4(\frac{T}{100})^{2/9} \rfloor = 9$. An asterisk (*) denotes rejection of the null hypothesis at 5% and (**) at 1%. We provide the joint chi-square test statistic and p-value for the null hypothesis $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$. The average autocorrelation $\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{k,t}^{(\ell,n)}$ displays the cross-sectional and time-average of the autocorrelation.

Table D.2: Filtering error for case 1 asset and 1 variance under Assumption 3.4.1

Online Appendix D. *Variance filtering in multidimensional affine models*

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)
<i>Filter with first order update</i>						<i>Filter with second order update</i>					
Panel A: Without covariation											
(1,1)	0.98	-0.6213	-0.0091	-0.08	93.57	(1,1)	0.98	-0.6243	-0.0092	-0.08	93.49
(2,1)	0.96	-1.1229	-0.0170	-0.11	91.29	(2,1)	0.96	-1.1200	-0.0169	-0.11	91.33
(1,2)	0.97	0.0887	0.0032	0.27	78.48	(1,2)	0.97	0.1035	0.0035	0.29	76.90
(2,2)	0.92	0.1146	0.0044	0.18	85.32	(2,2)	0.92	0.0549	0.0027	0.11	91.04
(1,3)	0.95	-0.0103	-0.0005	-0.23	81.42	(1,3)	0.95	-0.0098	-0.0005	-0.23	82.15
(2,3)	0.89	0.0173	0.0016	0.20	84.43	(2,3)	0.89	0.0150	0.0014	0.18	85.91
(1,4)	0.93	0.0027	0.0002	0.41	68.14	(1,4)	0.93	0.0033	0.0002	0.47	64.07
(2,4)	0.85	0.0142	0.0015	0.45	65.28	(2,4)	0.85	0.0108	0.0012	0.35	72.27
Joint	χ_8^2	0.66			99.96	Joint	χ_8^2	0.66			99.96
Panel B: With covariation											
(1,1)	0.98	-0.6245	-0.0092	-0.08	93.54	(1,1)	0.98	-0.6337	-0.0094	-0.08	93.37
(2,1)	0.96	-1.1218	-0.0184	-0.12	90.57	(2,1)	0.96	-1.1037	-0.0182	-0.12	90.64
(1,2)	0.97	0.0887	0.0032	0.27	78.48	(1,2)	0.97	0.1116	0.0035	0.29	76.88
(2,2)	0.92	0.1143	0.0046	0.20	84.49	(2,2)	0.92	0.0281	0.0032	0.14	89.23
(1,3)	0.95	-0.0104	-0.0005	-0.24	81.39	(1,3)	0.95	-0.0097	-0.0005	-0.22	82.20
(2,3)	0.89	0.0170	0.0016	0.20	84.12	(2,3)	0.89	0.0135	0.0014	0.17	86.11
(1,4)	0.93	0.0027	0.0002	0.41	68.13	(1,4)	0.93	0.0037	0.0002	0.49	62.16
(2,4)	0.85	0.0140	0.0015	0.46	64.24	(2,4)	0.85	0.0092	0.0011	0.35	72.93
Joint	χ_8^2	0.68			99.96	Joint	χ_8^2	0.69			99.95

See the notes of Table D.2. Panel A displays results where the two assets do not have a common variance factor, whereas Panel B shows results for two assets with a common variance factor.

Table D.3: Filtering error for case 2 assets and 2 variances under Assumption 3.4.1

Online Appendix D. *Variance filtering in multidimensional affine models*

(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)	(k, n)	$\bar{b}^{(\ell, n)}$	$\frac{\bar{\varepsilon}^{(k, n)}}{\theta_k^n}$ (bp)	$\frac{\vartheta_k^{(n)}}{\theta_k^n}$ (bp)	t-stat	p-value (%)
<i>Filter with first order update</i>						<i>Filter with second order update</i>					
(1,1)	0.98	-1.6895	-0.0339	-0.27	79.01	(1,1)	0.98	-2.1527	-0.0365	-0.29	77.44
(2,1)	0.96	-2.1337	-0.0593	-0.24	81.11	(2,1)	0.96	-1.6641	-0.0481	-0.19	84.60
(3,1)	0.98	3.0751	-0.0576	-0.23	82.09	(3,1)	0.98	2.3843	-0.0658	-0.26	79.57
(1,2)	0.97	0.2414	0.0074	0.48	63.09	(1,2)	0.97	0.4226	0.0096	0.63	53.12
(2,2)	0.91	0.8088	0.0161	0.29	76.97	(2,2)	0.91	0.1820	0.0120	0.22	82.62
(3,2)	0.96	0.2615	-0.0135	-0.15	88.24	(3,2)	0.96	-0.0748	-0.0052	-0.06	95.41
(1,3)	0.95	-0.0159	-0.0010	-0.31	75.92	(1,3)	0.95	-0.0273	-0.0014	-0.43	67.08
(2,3)	0.87	0.0383	0.0004	0.02	98.71	(2,3)	0.87	-0.0085	-0.0016	-0.06	95.21
(3,3)	0.93	-0.7818	-0.0451	-0.59	55.37	(3,3)	0.93	-0.8385	-0.0478	-0.63	52.85
(1,4)	0.93	0.0097	0.0006	0.71	47.86	(1,4)	0.93	0.0098	0.0004	0.49	62.11
(2,4)	0.84	0.1906	0.0189	1.38	16.84	(2,4)	0.84	0.0685	0.0086	0.64	52.26
(3,4)	0.91	0.2850	0.0110	0.13	89.42	(3,4)	0.91	0.2480	0.0117	0.14	88.66
Joint	χ_{12}^2	4.67			96.81	Joint	χ_{12}^2	1.94			99.95

This Monte Carlo study simulates $L = 1000$ independent trajectories for the Model (3.1), using $M = 3900$ intraday steps over a period of $T = 4000$ days. The daily RCV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell)} = RCV_{i,j,t,t+1}^{(\ell)} - QCV_{i,j,t,t+1}^{(\ell)}$. Filtering errors,

$$\varepsilon_{k,t}^{(\ell,1)} = V_{k,t\Delta}^{(\ell)} - \mathbb{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right],$$

are computed daily for each variance component k and each path ℓ . The filtering errors' higher moments are

$$\varepsilon_{k,t}^{(\ell,n)} = \left(\varepsilon_{k,t}^{(\ell,1)} \right)^n - \mathbb{E} \left[\left(\varepsilon_{k,t}^{(\ell,1)} \right)^n \mid \mathcal{G}_{t\Delta} \right], n \in \{2, 3, 4\}.$$

The average filtering error $\bar{\varepsilon}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T \varepsilon_{t,\ell}^{(k,n)}$ is normalized by the long-term expected variance component θ_k raised to the appropriate power. This ratio, $\frac{\bar{\varepsilon}^{(k,n)}}{\theta_k^n}$, is expressed in basis points (bp). The panel regression

$$\varepsilon_{k,t}^{(\ell,n)} = b_{k,t}^{(\ell,n)} \varepsilon_{k,t-1}^{(\ell,n)} + \vartheta_k^{(n)} + z_{k,t}^{(\ell,n)}$$

(as described in Section 3.4.2) accounts for autocorrelation and potential biases through the constants $\vartheta_k^{(n)}$. We test the absence of bias ($H_0 : \vartheta_k^{(n)} = 0$) and report the associated t-statistic (t-stat) and p-value. Standard errors are estimated considering spatial correlation, using the Newey-West estimator with a lag $Q = \lfloor 4(\frac{T}{100})^{2/9} \rfloor = 9$. An asterisk (*) denotes rejection of the null hypothesis at 5% and (**) at 1%. We provide the joint chi-square test statistic and p-value for the null hypothesis $H_0 : \vartheta_k^{(n)} = 0, n \in \{1, 2, 3, 4\}, k \in \{1, \dots, N_V\}$. The average autocorrelation $\bar{b}^{(k,n)} = \frac{1}{S} \frac{1}{T} \sum_{\ell=1}^S \sum_{t=1}^T b_{t,\ell}^{(k,n)}$ displays the cross-sectional and time-average of the autocorrelation.

Table D.4: Filtering error for case 2 assets and 3 variances under Assumption 3.4.1

Starting with Table D.2, and comparing it with Table 3.2 from the paper, we observe that the null hypothesis is never rejected. The joint test statistics are significantly lower, with values of 0.34, 0.41, and 0.67 for Panels A, B, and C, respectively. In contrast, Table 3.2 shows values of 3.21, 9.48, and 0.17. Despite Panel C being slightly higher, Panels A and B show a substantial reduction in the joint statistic. For Panel B, we do not reject the third moment, consistent with the need to correct for the third moment when the measurement error arises from the RV calculation. In the right column, we again find that the null hypothesis is never rejected. However, the second-order update does not provide significant additional benefit, as the results are very similar. This suggests that we have enough observables to accurately capture our latent variables and that the basic filter performs adequately without the extra step.

Moving to Table D.3, which examines the case of two assets and two variances, we compare it with Table 3.3. We find the same conclusion as the previous table: the null hypothesis is never rejected. The joint tests for Panels A and B are 0.66 and 0.68, much lower than the values in Table 3.3, which are 11.59 and 12.22. This underscores the significant impact of the measurement error Gaussianity assumption on the filter quality and reinforces the importance of our correction. The right panel, with the second-order update, again shows very similar results, indicating that this additional step does not add value in this case. The filter performs effectively with just one step when there are two latent variables and five observable variables.

Finally, we examine the most interesting table, Table D.4, which considers the case of two assets with three variances, and compare it with Table 3.4. In the left column, we see that no statistics fall within the rejection region. The global statistic decreases from 30.51 in Table D.4 to 4.67, highlighting the importance of Gaussian assumptions in measurement error. However, for the third variance's fourth centered moment, the p-value is 16.84%, double that of Table D.4, indicating potential difficulties.

Using the *Filter with second order update* on the right panel, we significantly correct the fourth moment of the second variance, reducing the global test from 4.67 to 1.94. This further improves the precision of our filter for higher-order moments. While this step may not be crucial in this context, it could be a valuable tool in other scenarios involving jumps

or more latent variables. This methodological development holds potential for future research.

D.11 Sample Moments Comparison

We present summary statistics on the sample moments (simulated and filtered) computed over each variance factor $k \in \{1, \dots, K\}$. For each trajectory $\ell \in \{1, \dots, 1000\}$ and for $n \in \{2, 3, 4\}$, we define the sample higher moments as

$$m_n(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) = \frac{\sum_{t=1}^T (\mathbf{X}_{t\Delta}^{(\ell)} - \mu(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}))^n}{T},$$

with the sample mean

$$\mu(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) = \frac{\sum_{t=1}^T \mathbf{X}_{t\Delta}^{(\ell)}}{T}. \quad (\text{D.17})$$

Based on Joanes and Gill, 1998, the unbiased sample cumulants are

$$\begin{aligned} \sigma^2(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) &= \frac{T}{T-1} m_2(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}), \\ \text{Cum}_3(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) &= \frac{T^2}{(T-1)(T-2)} m_3(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}), \\ \text{Cum}_4(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) &= \frac{T^2 \left((T+1)m_4(\mathbf{X}_{\Delta:T\Delta}^{(\ell)}) - 3(T-1)m_2(\mathbf{X}_{\Delta:T\Delta}^{(\ell)})^2 \right)}{(T-1)(T-2)(T-3)}. \end{aligned} \quad (\text{D.18})$$

where the third cumulant is the third centered moment and the fourth cumulant is comparable to the excess kurtosis, prior to undergoing standardization. These sample moments are applied to the simulated variance $V_{k,t\Delta}^{(\ell)}$ and are presented⁸ in Table D.5 panels A.1, B.1 and C.1 for case with 2 assets and three variances. We compare these sample moments to

⁸Our samples are not iid; however, we compare these sample moments to assess the distribution fit of the filtered variance compared to the simulated ones. We present our results on a subset of each trajectory (at each 20 steps) as the autocorrelation effect has practically vanished, as showed in Figure D.1. Furthermore, the filtering errors are expected to be autocorrelated as discussed in Online Appendix B.6.

the filtered variance first four conditional cumulants in Table D.5 panels A.2, B.2 and C.2.

We use the law of total cumulance as follows

$$\mu^*(\hat{\mathbf{V}}_{k,\Delta:T\Delta}^{(\ell)}) = \mu(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]), \quad (\text{D.19})$$

$$\sigma^{2*}(\hat{\mathbf{V}}_{k,\Delta:T\Delta}^{(\ell)}) = \mu(\text{Var}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) + \sigma^2(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]), \quad (\text{D.20})$$

to recover sample moment based on filtered moments. With the help of the law of total cumulance, the third cumulant of the filtered moment is

$$\begin{aligned} \text{Cum}_3^*(\hat{\mathbf{V}}_{k,\Delta:T\Delta}^{(\ell)}) &= \text{Cum}_3(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) + \mathbf{E}(\text{Cum}_3[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) \quad (\text{D.21}) \\ &\quad + 3\sigma^2(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}], \text{Var}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]), \end{aligned}$$

and the fourth cumulant of the filtered moment is

$$\begin{aligned} \text{Cum}_4^*(\hat{\mathbf{V}}_{k,\Delta:T\Delta}^{(\ell)}) &= \text{Cum}_4(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) + \mathbf{E}(\text{Cum}_4[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) \quad (\text{D.22}) \\ &\quad + 3\sigma^2(\text{Var}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}], \text{Var}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) \\ &\quad + 6\text{Cum}_3(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}], \mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}], \text{Var}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]) \\ &\quad + 4\text{Var}(\mathbf{E}[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}], \text{Cum}_3[\mathbf{V}_{k,\Delta:T\Delta}|\mathcal{G}_{\Delta:T\Delta}]). \end{aligned}$$

All the sample moments are compared to their theoretical quantities derived in Online Appendix D.9. Table D.5 shows that the sample moments based on the filter cumulants are similar to the sample cumulants of the simulated latent variable.

Online Appendix D. Variance filtering in multidimensional affine models

	Theo	Mean	Median	Std	\underline{Mean}	\overline{Mean}
Panel A.1: Simulated variance $V_{1,\Delta:T\Delta}$						
$\mu(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	1.0000	0.9990	0.9927	0.0840	0.9938	1.0042
$\sigma^2(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.1125	0.1037	0.0982	0.0315	0.1017	0.1056
$\text{Cum}_3(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.0253	0.0205	0.0153	0.0200	0.0193	0.0217
$\text{Cum}_4(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.0085	0.0038	0.0003	0.0149	0.0028	0.0047
Panel A.2: Filtered variance $V_{1,\Delta:T\Delta}$						
$\mu^*(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	1.0000	0.9991	0.9937	0.0822	0.9940	1.0042
$\sigma^{2*}(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.1125	0.1040	0.0982	0.0304	0.1021	0.1059
$\text{Cum}_3^*(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.0253	0.0206	0.0158	0.0189	0.0194	0.0218
$\text{Cum}_4^*(\mathbf{V}_{1\Delta:T\Delta}^{(\ell)}/\theta_1)$	0.0085	0.0041	0.0006	0.0139	0.0032	0.0049
Panel B.1: Simulated variance $V_{2,\Delta:T\Delta}$						
$\mu(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	1.0000	1.0093	1.0082	0.1111	1.0025	1.0162
$\sigma^2(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.3750	0.3692	0.3501	0.1100	0.3624	0.3761
$\text{Cum}_3(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.2812	0.2604	0.2102	0.1923	0.2485	0.2723
$\text{Cum}_4(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.3164	0.2534	0.1163	0.4360	0.2264	0.2804
Panel B.2: Filtered variance $V_{2,\Delta:T\Delta}$						
$\mu^*(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	1.0000	1.0095	1.0070	0.1089	1.0028	1.0163
$\sigma^{2*}(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.3750	0.3694	0.3520	0.1065	0.3628	0.3760
$\text{Cum}_3^*(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.2812	0.2598	0.2127	0.1861	0.2482	0.2713
$\text{Cum}_4^*(\mathbf{V}_{2\Delta:T\Delta}^{(\ell)}/\theta_2)$	0.3164	0.2565	0.1228	0.4277	0.2300	0.2831
Panel C.1: Simulated variance $V_{3,\Delta:T\Delta}$						
$\mu(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	1.0000	0.9729	0.7195	0.7823	0.9245	1.0214
$\sigma^2(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	1.8000	0.8341	0.4152	1.1567	0.7624	0.9058
$\text{Cum}_3(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	6.4851	0.9515	0.2301	2.4498	0.7997	1.1034
$\text{Cum}_4(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	35.0476	-0.3835	0.0147	6.6083	-0.7931	0.0261
Panel C.2: Filtered variance $V_{3,\Delta:T\Delta}$						
$\mu^*(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	1.0000	0.9726	0.7237	0.7776	0.9245	1.0208
$\sigma^{2*}(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	1.8000	0.8442	0.4282	1.1561	0.7725	0.9158
$\text{Cum}_3^*(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	6.4851	0.9865	0.2493	2.4532	0.8344	1.1386
$\text{Cum}_4^*(\mathbf{V}_{3\Delta:T\Delta}^{(\ell)}/\theta_3)$	35.0476	-0.2352	0.0249	6.3272	-0.6274	0.1569

This Monte Carlo study generates $L = 1000$ independent trajectories with $M = 3900$ intraday time steps for $T = 4000$ days. The daily RV is aggregated at the 5 minutes frequency with $N = 78$. The measurement error is calculated with Equation (3.32): $\xi_{i,j,t,t+1}^{(\ell,N)} = RCV_{i,j,t,t+1}^{(\ell,N)} - QCV_{i,j,t,t+1}^{(\ell)}$. For each trajectory ℓ , the observable variables $\mathbf{R}_{t,t+1}^{(\ell)}$ and $\mathbf{RCV}_{t,t+1}^{(\ell)}$ are used in the filter described in Section 3.3 to recover the conditional moments of the latent variable such as $\mathbb{E}^{\mathbb{P}} [V_{k,t\Delta}^{(\ell)} | \mathcal{G}_{t\Delta}]$ and $\text{Var}^{\mathbb{P}} [V_{k,t\Delta}^{(\ell)} | \mathcal{G}_{t\Delta}]$ for $k \in 1, 2, \dots, K$. Each column presents summary statistics computed over the 1000 trajectories of sample moments detailed in Equation (D.17) to (D.20). The columns with theoretical moments are derived in Online Appendix D.9 \underline{Mean} and \overline{Mean} are respectively the mean lower and upper bounds at a 5% confidence interval and are calculated across all the trajectories as $\underline{X} = \text{Avg}(X) - 1.96\text{std}(X)/\sqrt{1000}$ and $\overline{X} = \text{Avg}(X) + 1.96\text{std}(X)/\sqrt{1000}$ respectively. To ensure that the results were not skewed by autocorrelation, only information from every 20 daily time step was retained for all trajectories. The realized variance is filtered according to the correction as per Assumption 3.4.2; thus acknowledging that $\xi_{i,j,t,t+1}^{(N)}$ is not Gaussian.

Table D.5: Sample moments comparison for case 2 assets and 3 variances
cli

D.12 Filtering Error Derivation

This Online Appendix provides additional details on the filtering error $\varepsilon_{k,t\Delta}^{(\ell,n)}$ used in Section 3.4. We show that the expected value of $\varepsilon_{k,t\Delta}^{(\ell,n)}$ should be 0 if Assumption D.12.1 holds.

Assumption D.12.1. *Equations (3.23), (3.27), (3.28), (3.29) have an equality symbol instead of \simeq .*

We recall the definition of the filtering error for $n = 1$, that is

$$\varepsilon_{k,t\Delta}^{(\ell,1)} = \underbrace{V_{k,t\Delta}^{(\ell)}}_{\text{Simulated}} - \underbrace{\text{E} \left[V_{k,t\Delta}^{(\ell)} \mid \mathcal{G}_{t\Delta} \right]}_{\text{filter output}}.$$

If Assumption D.12.1 holds, $\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \mid \mathcal{G}_{t\Delta} \right] = 0$.

The first moment condition is

$$\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] = \text{E} \left[\underbrace{\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \mid \mathcal{G}_{t\Delta} \right]}_{=0} \right] = 0,$$

where we verify if the sample average of $\varepsilon_{k,t\Delta}^{(\ell,1)}$, which is an estimator for $\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right]$, is statistically different from zero.

With the law of total expectation and variance, the second moment of the filtering error is

$$\begin{aligned} \text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] &= \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] \right)^2 \right] = \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^2 \right] \\ &= \text{Var} \left[\underbrace{\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \mid \mathcal{G}_{t\Delta} \right]}_{=0} \right] + \text{E} \left[\underbrace{\text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \mid \mathcal{G}_{t\Delta} \right]}_{\text{filter output}} \right]. \end{aligned}$$

Therefore, the second panel regression is on

$$\varepsilon_{k,t\Delta}^{(\ell,2)} = \left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^2 - \underbrace{\text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right]}_{\text{filter output}},$$

since the sample average should be zero if Assumption D.12.1 holds.

For the third moment, using the law of total cumulance of Lemma B.4.2, we have:

$$\begin{aligned} \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] \right)^3 \right] &= \text{Cum}^3 \left[\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right] \right] + \text{E} \left[\text{Cum}^3 \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right] \right] \\ &+ 3\text{Cov} \left[\text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right], \text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right] \right] \\ &= \text{E} \left[\underbrace{\text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right] \right)^3 \middle| \mathcal{G}_{t\Delta} \right]}_{\text{filter output}} \right]. \end{aligned}$$

The third panel regression is therefore on

$$\varepsilon_{k,t\Delta}^{(\ell,3)} = \left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^3 - \underbrace{\text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \middle| \mathcal{G}_{t\Delta} \right] \right)^3 \middle| \mathcal{G}_{t\Delta} \right]}_{\text{filter output}}.$$

Finally, for the fourth moment, the application of Lemma B.4.7 leads to:

$$\begin{aligned} \text{Cum}_4(\varepsilon_{k,t\Delta}^{(\ell,1)}) &= \text{Cum}_4(\text{E}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) + \text{E}(\text{Cum}_4[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) \\ &+ 3\text{Var}(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}], \text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) \\ &+ 6\text{Cum}_3(\text{E}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}], \text{E}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}], \text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) \\ &+ 4\text{Var}(\text{E}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}], \text{Cum}_3[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) \\ &= \text{E}(\text{Cum}_4[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) + 3\text{Var}(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}], \text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]). \end{aligned}$$

Since

$$\text{Cum}_4(\varepsilon_{k,t\Delta}^{(\ell,1)}) = \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] \right)^4 \right] - 3\text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right]^2,$$

combining the two above leads to

$$\begin{aligned} & \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} - \text{E} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right] \right)^4 \right] \\ &= \text{E}(\text{Cum}_4[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) + 3 \left(\frac{\text{Var}(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}])}{\text{E} \left[(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}])^2 \right] - \text{E} \left[(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}]) \right]^2} + \text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right]^2 \right) \\ &= \text{E} \left[\text{Cum}_4[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right] + 3\text{E} \left[\left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right)^2 \right], \end{aligned}$$

since, based on Assumption D.12.1, we have

$$\begin{aligned} & \text{Var} \left[\varepsilon_{k,t\Delta}^{(\ell,1)} \right]^2 - \text{E} \left[\left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right)^2 \right] \\ &= \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^2 \right]^2 - \text{E} \left[\left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right)^2 \right] \\ &= \text{E} \left[\underbrace{\left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^2 - \left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right)}_{=0} \right] \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^2 + \left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right) \right]. \end{aligned}$$

Therefore, the fourth panel regression is

$$\begin{aligned} \varepsilon_{k,t\Delta}^{(\ell,4)} &= \left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^4 - \text{Cum}_4[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] - 3 \left(\text{Var}[\varepsilon_{k,t\Delta}^{(\ell,1)} | \mathcal{G}_{t\Delta}] \right)^2 \\ &= \left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^4 - \text{E} \left[\left(\varepsilon_{k,t\Delta}^{(\ell,1)} \right)^4 \middle| \mathcal{G}_{t\Delta} \right]. \end{aligned}$$

References

- Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating Quadratic Variation Using Realized Variance. *Journal of Applied Econometrics*, *17*(5), 457–477.
- Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric Analysis of Realized Covariation: High Frequency based Covariance, Regression, and Correlation in Financial Economics. *Econometrica*, *72*(3), 885–925.
- Joanes, D. N., & Gill, C. A. (1998). Comparing Measures of Sample Skewness and Kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(1), 183–189.
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems.
- Uhlmann, J., & Julier, S. J. (2022). Gaussianity and the Kalman Filter: A Simple Yet Complicated Relationship. *Journal de Ciencia e Ingeniería*, *14*(1), 21–26.

Chapter E

Online Appendices of *Enhancing Parameter Estimation in Stochastic Volatility Models Using Intraday Data: Generalized Method of Moments Comparative Analysis*

E.1 Risk-Neutralization Measure and the Drift

The time t value of the self-financing investment strategy that reinvests the dividends to buy more shares of the risky asset is

$$\exp\left(\int_0^t q_u du\right) S_t. \quad (\text{E.1})$$

From the Girsanov theorem, let $W_{k,t}^{\mathbb{Q}} = W_{k,t} + \int_0^t \Lambda_{k,u} du$ and $B_{k,t}^{\mathbb{Q}} = B_{k,t} + \int_0^t \Gamma_{k,u} du$ where $\Lambda_{k,u}$ and $\Gamma_{k,u}$ are predictable processes with respect to the continuous model filtration $(\mathcal{F}_t^C)_{t \geq 0}$ generated by $\mathcal{F}_t^C = \sigma(\{W_{k,s}, B_{k,s}\}_{k=1}^{N_V} | s \leq t)$. We assume that the Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \Lambda_{k,u}^2 du\right)\right] < \infty \quad \text{and} \quad \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \Gamma_{k,u}^2 du\right)\right] < \infty,$$

are satisfied. The associated Radon-Nikodym derivative is

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp \left(\sum_{k=1}^K \left(-\frac{1}{2} \int_0^T \Lambda_{k,u}^2 du - \int_0^T \Lambda_{k,u} dW_{k,u}^{\mathbb{P}} \right) + \sum_{k=1}^K \left(-\frac{1}{2} \int_0^T \Gamma_{k,u}^2 du - \int_0^T \Gamma_{k,u} dB_{k,u} \right) \right).$$

The absence of arbitrage opportunity requires that

$$\left\{ \exp \left(- \int_0^t r_u du \right) \exp \left(\int_0^t q_u du \right) S_t \right\}_{t \geq 0}$$

is a \mathbb{Q} -martingale. Therefore,

$$\begin{aligned} & d \exp \left(\int_0^t q_u - r_u du \right) S_t && \text{(E.2)} \\ &= S_t d \exp \left(\int_0^t q_u - r_u du \right) + \exp \left(\int_0^t q_u - r_u du \right) dS_t \\ &= \exp \left(\int_0^t q_u - r_u du \right) S_t \left(\begin{aligned} & (\alpha_t + q_t - r_u) dt \\ & + \sum_{k=1}^K \sqrt{V_{k,t}} d \left(\rho_k dW_{k,t} + \sqrt{1 - \rho_k^2} dB_{k,t} \right) \end{aligned} \right) \\ &= \exp \left(\int_0^t q_u - r_u du \right) S_t \underbrace{\left(\sum_{k=1}^K \sqrt{V_{k,t}} \left(\rho_k \Lambda_{k,t} + \sqrt{1 - \rho_k^2} \Gamma_{k,t} \right) \right)}_{=0} dt \\ &+ \exp \left(\int_0^t q_u - r_u du \right) S_t \left(\sum_{k=1}^K \sqrt{V_{k,t}} d \left(\rho_k dW_{k,t}^{\mathbb{Q}} + \sqrt{1 - \rho_k^2} dB_{k,t}^{\mathbb{Q}} \right) \right), \end{aligned}$$

which implies that

$$\alpha_t = r_t - q_t + \sum_{k=1}^K \sqrt{V_{k,t}} \left(\rho_k \Lambda_{k,t} + \sqrt{1 - \rho_k^2} \Gamma_{k,t} \right).$$

The model belongs to the exponential affine class if the prices of risk are defined as

$$\Lambda_{k,t} = \tilde{\Lambda}_k \sqrt{V_{k,t}} \text{ and } \Gamma_{k,t} = \tilde{\Gamma}_k \sqrt{V_{k,t}}. \quad \text{(E.3)}$$

Therefore,

$$\alpha_t = r_t - q_t + \sum_{k=1}^K \underbrace{\left(\rho_k \tilde{\Lambda}_k + \sqrt{1 - \rho_k^2} \tilde{\Gamma}_k \right)}_{\lambda_k} V_{k,t}.$$

E.2 Moment Selection

To determine the methodology employed in this paper for the case of one variance factor, we compare five distinct approaches. These various procedures are outlined in Table E.1, alongside comprehensive outcomes of statistical tests found in Tables E.2 and E.3. In addition to the Mean and Coverage Ratio tests, we also provide the percentage of trajectories that did not fail the Sargan–Hansen J-test. All tests are detailed in the Online Appendix E.10.

As discussed in this paper, we encountered two significant issues. First, the model misspecification, which we addressed by opting for specification (4.16) instead of specification (4.14). Second, for the problem concerning parameter identification, we included additional moments on the lagged value of the realized variance from Table E.1, namely VII) to IX), through different combinations.

Tables E.2 and E.3 show the three different tests for the five moment selections presented in Table E.1. The Mean columns refer to $(\bar{\theta} - \theta^*)^\top \hat{\Sigma}^{-1} (\bar{\theta} - \theta^*)$. The smaller the value, the smaller the averaged bias on the estimated parameters. The CR columns refer to the global coverage ratio across all parameters, while the last column indicates the percentage of trajectories that passed the Sargan–Hansen J-test, an over-identification test. A failed test suggests that the moment conditions are not satisfied by the data.

Table E.2 reveals that moment selections CM2 to CM5 exhibit shortcomings for specification (4.14). This is primarily attributed to model misspecification regarding moments VIII and IX), as discussed in Appendix C.2. Indeed, all three tests exhibit higher values than CM1, where the additional instrumental variables worsen the GMM estimator rather than improving it. This is due to the undesired dependency between the error measurement and the instrumental variables. The same problem would arise in moment conditions V) and VI). However, as discuss in Appendix C.5.1, the error measurement in moment

conditions IV) to VI) constitutes a non-significant portion of the moment condition. Consequently, it becomes apparent that adopting procedure CM1 for specification (4.14) is the most suitable option for facilitating a comparison with the BZ procedure. This very rationale underpins the using of CM1 for specification (4.14) in the main paper.

In a simulated framework, there's no need for estimating either γ^2 or η^2 since we have precise theoretical values for these parameters. However, when dealing with real-world data, we encounter additional complexities such as rounding errors and the influence of bid curves. Additionally, various intra-day statistical methods can be applied. This is why it's advantageous to have a methodology that facilitates the estimation of these parameters, allowing us to adapt to the specificity of real data.

Table E.3 highlights a noteworthy observation: all CM procedures exhibit reasonable moment selection, with the CR and Sargan-Hansen J tests achieving values higher than 90% for $T = 4000$. Furthermore, the Mean value is lower compared to those in Table E.2. Notably, CM3 minimizes the average test statistic in almost all panels, making it the preferred procedure for the main paper.

		BZ	CM1	CM2	CM3	CM4	CM5
I)	$RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta$	X	X	X	X	X	X
II)	$(RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t}$	X	X	X	X	X	X
III)	$(RV_{t+1,t+2} - \alpha RV_{t,t+1} - \beta \Delta) \times RV_{t-1,t}^2$	X	X	X		X	X
IV)	$(RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J$	X	X	X	X	X	X
V)	$((RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J) \times RV_{t-1,t}$	X	X	X	X	X	X
VI)	$((RV_{t+1,t+2}^2 - \eta^2 RQ_{t+1,t+2}) - H (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - IRV_{t,t+1} - J) \times RV_{t-1,t}^2$	X	X	X			X
VII)	$RV_{t,t+1} RV_{t+1,t+2} - \tilde{H} (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - \tilde{I} RV_{t,t+1} - \tilde{J}$	X	X	X	X	X	X
VIII)	$(RV_{t,t+1} RV_{t+1,t+2} - \tilde{H} (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - \tilde{I} RV_{t,t+1} - \tilde{J}) \times RV_{t-1,t}$	X	X	X	X	X	X
IX)	$(RV_{t,t+1} RV_{t+1,t+2} - \tilde{H} (RV_{t,t+1}^2 - \eta^2 RQ_{t,t+1}) - \tilde{I} RV_{t,t+1} - \tilde{J}) \times RV_{t-1,t}^2$	X					

cix

BZ stands for Bollerslev and Zhou, 2002 and CM corresponds to our moment selection. In specification (4.14), $\eta^2 RQ_{t,t+1}$ is replaced by γ^2 while specification (4.16) uses $\eta^2 RQ_{t,t+1}$. The moments are derived in Appendix C.3.1. Using a Taylor expansion, $\alpha \approx 1 - \kappa \Delta$, $\beta \approx \kappa \theta \Delta$, $H \approx 1 - 2\kappa \Delta$, $\tilde{H} \approx 1 - \kappa \Delta$, $I \approx (\sigma^2 + 2\kappa \theta) \Delta^2$, $\tilde{I} \approx \left(\frac{\sigma^2}{2} + 2\kappa \theta\right) \Delta^2$, $J \approx \left(\frac{\sigma^2 \theta \kappa}{3} + \kappa^2 \theta^2\right) \Delta^4$ and $\tilde{J} \approx -\left(\frac{\sigma^2 \kappa \theta}{6} + \frac{\kappa^2 \theta^2}{2}\right) \Delta^4$.

Table E.1: GMM moment selection

	Mean		Coverage Ratio		Sargan–Hansen J	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Panel A.1: Parameter set A with γ^2 estimated						
CM1	0.3914	0.2243	0.9440	0.9370	0.9820	0.9760
CM2	0.8987	0.9697	0.9430	0.9580	0.0020	0.0000
CM3	0.9501	0.9728	0.9490	0.9600	0.0160	0.0010
CM4	0.9093	0.9616	0.9470	0.9570	0.0180	0.0010
CM5	0.9053	0.9657	0.9460	0.9570	0.0290	0.0010
Panel B.1: Parameter set B with γ^2 estimated						
CM1	0.3100	0.2417	0.9530	0.9570	0.9820	0.9650
CM2	0.9075	0.9748	0.9620	0.9640	0.0010	0.0000
CM3	0.9202	0.9678	0.9630	0.9640	0.0120	0.0000
CM4	0.9201	0.9695	0.9550	0.9630	0.0170	0.0000
CM5	0.9153	0.9701	0.9560	0.9560	0.0220	0.0000
Panel C.1: Parameter set C with γ^2 estimated						
CM1	0.4056	0.2770	0.9540	0.9480	0.9550	0.9560
CM2	0.9594	0.9852	0.9580	0.9660	0.0340	0.0000
CM3	0.9502	0.9664	0.9610	0.9650	0.0920	0.0050
CM4	0.9500	0.9683	0.9530	0.9650	0.1290	0.0070
CM5	0.9503	0.9699	0.9610	0.9590	0.1640	0.0090
Panel A.2: Parameter set A with γ^2 fixed						
CM1	0.4121	0.2361	0.9520	0.9470	0.9880	0.9890
CM2	0.8799	0.9391	0.9550	0.9680	0.0010	0.0000
CM3	0.8445	0.8954	0.9640	0.9630	0.0140	0.0000
CM4	0.8024	0.8635	0.9620	0.9610	0.0240	0.0010
CM5	0.7609	0.8486	0.9590	0.9620	0.0300	0.0010
Panel B.2: Parameter set B with γ^2 fixed						
CM1	0.3262	0.2511	0.9440	0.9510	0.9900	0.9790
CM2	0.8623	0.9524	0.9590	0.9650	0.0010	0.0000
CM3	0.8387	0.9253	0.9640	0.9650	0.0070	0.0000
CM4	0.8421	0.9288	0.9610	0.9670	0.0120	0.0000
CM5	0.8225	0.9265	0.9550	0.9620	0.0080	0.0000
Panel C.2: Parameter set C with γ^2 fixed						
CM1	0.4250	0.2989	0.9540	0.9550	0.9650	0.9700
CM2	0.8855	0.9583	0.9610	0.9670	0.0070	0.0000
CM3	0.8393	0.9061	0.9590	0.9680	0.0550	0.0020
CM4	0.8440	0.9107	0.9650	0.9630	0.0690	0.0060
CM5	0.8353	0.9157	0.9710	0.9600	0.0640	0.0050

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean column is $(\hat{\theta} - \theta^*)^\top \hat{\Sigma}^{-1} (\hat{\theta} - \theta^*)$, and when multiplied by the number of simulations, it follows a chi-square distribution. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter vector. Sargan–Hansen J column is the percentage of trajectories that did not failed the overidentification test derived in Equation (E.12). See Online Appendix E.10 for details.

Table E.2: GMM comparison results for specification (4.14): $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$

	Mean		Coverage Ratio		Sargan–Hansen J	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Panel A.1: Parameter set A with η^2 estimated						
CM1	0.3592	0.1333	0.9490	0.9410	0.9730	0.9690
CM2	0.4099	0.1847	0.9420	0.9500	0.9500	0.9440
CM3	0.2538	0.0527	0.9460	0.9470	0.9250	0.9540
CM4	0.2531	0.0481	0.9480	0.9450	0.9330	0.9410
CM5	0.3614	0.1258	0.9460	0.9450	0.9420	0.9380
Panel B.1: Parameter set B with η^2 estimated						
CM1	0.2588	0.0861	0.9510	0.9510	0.9750	0.9570
CM2	0.3577	0.1529	0.9500	0.9560	0.9570	0.9410
CM3	0.1590	0.0392	0.9520	0.9600	0.9210	0.9470
CM4	0.1988	0.0506	0.9480	0.9500	0.9390	0.9430
CM5	0.2843	0.0980	0.9510	0.9540	0.9550	0.9380
Panel C.1: Parameter set C with η^2 estimated						
CM1	0.3617	0.1468	0.9490	0.9540	0.9360	0.9270
CM2	0.4595	0.2481	0.9480	0.9560	0.9060	0.9100
CM3	0.2073	0.0432	0.9440	0.9520	0.8570	0.9240
CM4	0.2766	0.0720	0.9440	0.9470	0.8820	0.8960
CM5	0.3971	0.1680	0.9550	0.9570	0.8870	0.9030
Panel A.2: Parameter set A with η^2 fixed						
CM1	0.3678	0.1355	0.9510	0.9370	0.9650	0.9640
CM2	0.4105	0.1737	0.9500	0.9430	0.9540	0.9510
CM3	0.2835	0.0597	0.9470	0.9480	0.9320	0.9470
CM4	0.2676	0.0548	0.9430	0.9450	0.9430	0.9500
CM5	0.3777	0.1308	0.9450	0.9410	0.9530	0.9520
Panel B.2: Parameter set B with η^2 fixed						
CM1	0.2687	0.0991	0.9520	0.9490	0.9580	0.9540
CM2	0.3576	0.1592	0.9510	0.9540	0.9400	0.9460
CM3	0.1866	0.0489	0.9530	0.9510	0.9400	0.9500
CM4	0.2129	0.0614	0.9470	0.9460	0.9370	0.9390
CM5	0.2933	0.1095	0.9470	0.9520	0.9440	0.9410
Panel C.2: Parameter set C with η^2 fixed						
CM1	0.3855	0.1750	0.9510	0.9570	0.9310	0.9240
CM2	0.4608	0.2537	0.9500	0.9560	0.9260	0.9150
CM3	0.2713	0.0759	0.9430	0.9510	0.8970	0.9140
CM4	0.3090	0.1004	0.9430	0.9460	0.9130	0.9030
CM5	0.4141	0.1907	0.9540	0.9580	0.9180	0.9060

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean column is $(\hat{\theta} - \theta^*)^\top \hat{\Sigma}^{-1} (\hat{\theta} - \theta^*)$, and when multiplied by the number of simulations, it follows a chi-square distribution. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter vector. Sargan–Hansen J column is the percentage of trajectories that did not failed the overidentification test derived in Equation (E.12). See Online Appendix E.10 for details.

Table E.3: GMM comparison results for specification (4.16): $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \varepsilon_{t,t+1}^{(N)}$

E.3 Constant Measurement Error Variance

In our study, we undertake a replication of Table 2 from BZ within the framework of specification (4.14). A detailed examination of moment selection for CM is provided in Online Appendix E.2, where we adopt a slightly different moment procedure, specifically CM1, as outlined in Table E.1. This decision is prompted by the bias in certain moments of Table 4.1 under specification (4.14), as discussed in Appendix C.2.

The comparative findings are presented in Table E.4, alongside our proposed procedure in Table E.5. In Table E.4, a significant increase in the root mean square error (RMSE) of the σ parameter¹ is noted when estimating the measurement error parameter. However, this issue is effectively addressed by employing the CM1 procedure, as demonstrated in Table E.5. Notably, the RMSE of σ remains consistent across the three scenarios, regardless of whether γ^2 is estimated or not.

The inclusion of additional moments notably enhances the quality of estimation. Consequently, all γ^2 values derived from the CM1 moment selection are positive, contrasting with the BZ moment selection's outcomes illustrated in Figure 4.1, panel A. Indeed, the parameter γ^2 cannot assume negative values, as it represents the variance of the measurement error.

Moreover, examining panel A of Figure² E.1, reveals a potential linear dependency in the estimation between γ^2 and σ , contrasting with the observations in Figures the panels B of the same three Figures.

We further investigate the estimated γ^2 by comparing it with simulated counterparts. Table E.6 provides summary statistics derived from simulated values, including the sample variance of the measurement error for each path denoted as $\text{Var} \left[\xi_{\Delta t, \Delta(t+1)}^{(N)} \right]$, the sample mean of the realized quarticity denoted as $E \left[RQ_{\Delta t, \Delta(t+1)}^{(N)} \right]$, and the sample mean of the integrated square variance represented by $E \left[\int_{t\Delta}^{(t+1)\Delta} V_s^2 ds \right]$. These metrics offer valuable insights into the theoretical value of the parameter γ^2 in specification (4.14) when the variance of the measurement error remains constant.

¹This is in concordance with footnote 13 of Brix and Lunde, 2015.

²Figures on other Scenario available upon request.

Across the three parameter sets (Panels A to C), the parameter γ_{GMM}^2 obtained through the CM1 procedure demonstrates similar summary statistics to those of the first three lines of each panel compared to the parameter γ_{GMM}^2 acquired through the BZ procedure. This observation underscores the robustness and consistency of the CM1 procedure in estimating γ^2 under various parameter configurations.

Online Appendix E. *GMM Parameter Estimation in Stochastic Volatility Models*

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: <i>Results with ($\Delta = 1, \kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 fixed</i>								
$10^2 \times \kappa = 3.0000$	3.6582	3.1856	3.5432	3.1644	1.3376	0.5974	0.9000	0.9330
$\theta = 0.2500$	0.2435	0.2481	0.2371	0.2468	0.0538	0.0275	0.9620	0.9470
$\sigma = 0.1000$	0.0982	0.0992	0.0981	0.0992	0.0079	0.0041	0.9490	0.9450
Panel A.2: <i>Results with ($\Delta = 1, \kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 estimated</i>								
$10^2 \times \kappa = 3.0000$	3.7221	3.1934	3.6196	3.1605	1.3647	0.6114	0.9050	0.9380
$\theta = 0.2500$	0.2441	0.2483	0.2381	0.2470	0.0530	0.0274	0.9620	0.9480
$\sigma = 0.1000$	0.0917	0.0988	0.0916	0.0976	0.0242	0.0135	0.9340	0.9470
$10^2 \times \gamma^2 = 0.2541$	0.5006	0.1838	0.5115	0.3572	1.6873	1.1132	0.9490	0.9600
Panel B.1: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 fixed</i>								
$\kappa = 0.1000$	0.1085	0.1030	0.1069	0.1028	0.0218	0.0105	0.9210	0.9440
$\theta = 0.2500$	0.2472	0.2491	0.2463	0.2492	0.0161	0.0081	0.9520	0.9480
$\sigma = 0.1000$	0.0987	0.0993	0.0986	0.0994	0.0070	0.0034	0.9500	0.9440
Panel B.2: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 estimated</i>								
$\kappa = 0.1000$	0.1085	0.1020	0.1068	0.1017	0.0219	0.0104	0.9330	0.9520
$\theta = 0.2500$	0.2478	0.2494	0.2469	0.2495	0.0160	0.0081	0.9550	0.9490
$\sigma = 0.1000$	0.0987	0.1044	0.0994	0.1045	0.0230	0.0122	0.9520	0.9330
$10^2 \times \gamma^2 = 0.1829$	0.1282	0.0484	0.1680	0.0688	0.4747	0.3020	0.9430	0.9230
Panel C.1: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$) and γ^2 fixed</i>								
$\kappa = 0.1000$	0.1141	0.1044	0.1124	0.1038	0.0264	0.0121	0.9020	0.9340
$\theta = 0.2500$	0.2398	0.2472	0.2375	0.2470	0.0334	0.0166	0.9490	0.9490
$\sigma = 0.2000$	0.1977	0.1989	0.1973	0.1987	0.0114	0.0056	0.9510	0.9490
Panel C.2: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$) and γ^2 estimated</i>								
$\kappa = 0.1000$	0.1143	0.1040	0.1130	0.1031	0.0267	0.0123	0.9010	0.9390
$\theta = 0.2500$	0.2429	0.2479	0.2400	0.2477	0.0325	0.0164	0.9530	0.9500
$\sigma = 0.2000$	0.1920	0.2006	0.1885	0.1992	0.0306	0.0167	0.9500	0.9590
$10^2 \times \gamma^2 = 0.2744$	0.4095	0.1648	0.5581	0.2484	1.1906	0.7169	0.9560	0.9550

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table E.4: GMM estimations results for specification (4.14) ($RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \gamma \epsilon_{t,t+1}^{(N)}$) with BZ moment selection

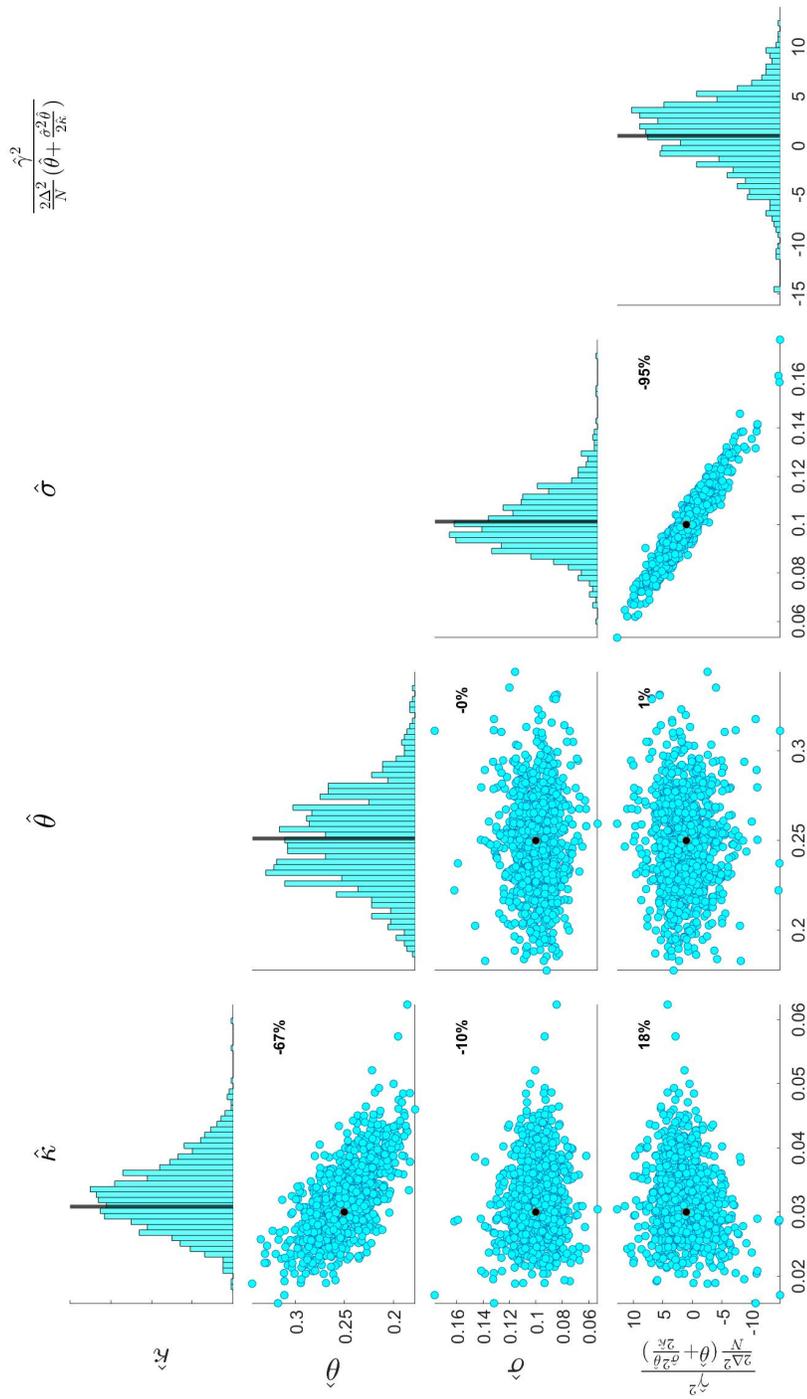
Online Appendix E. *GMM Parameter Estimation in Stochastic Volatility Models*

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A.1: <i>Results with ($\Delta = 1, \kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 fixed</i>								
$10^2 \times \kappa = 3.0000$	3.6962	3.2004	3.5393	3.1651	1.3749	0.6185	0.9040	0.9300
$\theta = 0.2500$	0.2400	0.2471	0.2335	0.2448	0.0531	0.0277	0.9600	0.9500
$\sigma = 0.1000$	0.0981	0.0992	0.0981	0.0992	0.0078	0.0040	0.9480	0.9480
Panel A.2: <i>Results with ($\Delta = 1, \kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 estimated</i>								
$10^2 \times \kappa = 3.0000$	3.6851	3.1907	3.5615	3.1711	1.3482	0.5995	0.8970	0.9320
$\theta = 0.2500$	0.2416	0.2478	0.2352	0.2461	0.0532	0.0274	0.9690	0.9470
$\sigma = 0.1000$	0.0982	0.0992	0.0980	0.0992	0.0082	0.0042	0.9460	0.9480
$10^2 \times \gamma^2 = 0.2541$	0.2334	0.2473	0.2098	0.2397	0.1115	0.0583	0.9700	0.9560
Panel B.1: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 fixed</i>								
$\kappa = 0.1000$	0.1079	0.1028	0.1069	0.1028	0.0217	0.0107	0.9260	0.9460
$\theta = 0.2500$	0.2466	0.2493	0.2454	0.2491	0.0174	0.0088	0.9490	0.9490
$\sigma = 0.1000$	0.0984	0.0993	0.0981	0.0992	0.0066	0.0032	0.9480	0.9470
Panel B.2: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$) and γ^2 estimated</i>								
$\kappa = 0.1000$	0.1085	0.1030	0.1069	0.1028	0.0217	0.0105	0.9200	0.9440
$\theta = 0.2500$	0.2471	0.2492	0.2464	0.2493	0.0161	0.0081	0.9510	0.9460
$\sigma = 0.1000$	0.0988	0.0993	0.0986	0.0993	0.0074	0.0036	0.9490	0.9430
$10^2 \times \gamma^2 = 0.1829$	0.1760	0.1812	0.1740	0.1805	0.0295	0.0151	0.9510	0.9560
Panel C.1: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$) and γ^2 fixed</i>								
$\kappa = 0.1000$	0.1140	0.1044	0.1128	0.1038	0.0269	0.0126	0.9020	0.9350
$\theta = 0.2500$	0.2392	0.2469	0.2361	0.2463	0.0345	0.0174	0.9450	0.9470
$\sigma = 0.2000$	0.1971	0.1987	0.1967	0.1986	0.0108	0.0054	0.9450	0.9510
Panel C.2: <i>Results with ($\Delta = 1, \kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$) and γ^2 estimated</i>								
$\kappa = 0.1000$	0.1144	0.1045	0.1130	0.1039	0.0265	0.0122	0.9010	0.9330
$\theta = 0.2500$	0.2391	0.2471	0.2370	0.2467	0.0334	0.0166	0.9430	0.9480
$\sigma = 0.2000$	0.1976	0.1989	0.1973	0.1988	0.0120	0.0059	0.9530	0.9450
$10^2 \times \gamma^2 = 0.2744$	0.2472	0.2653	0.2362	0.2619	0.0835	0.0438	0.9590	0.9560

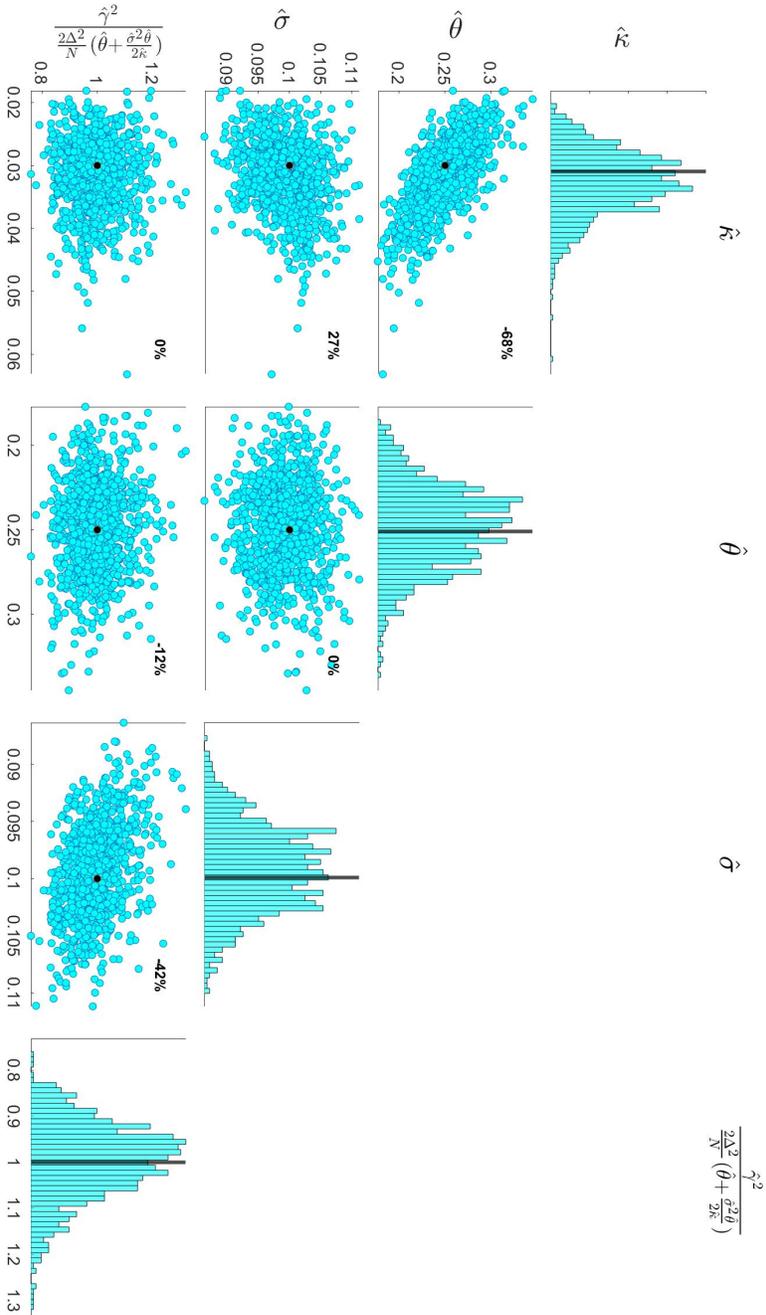
The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table E.5: GMM estimations results for specification (4.14) ($RV_{t,t+1}^{(N)} = \mathcal{V}_{1,t,t+1} + \gamma \epsilon_{t,t+1}^{(N)}$) with CM1 moment selection

Panel A: With moment selection BZ



Panel B: With moment selection CM1



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. Each diagonal panel presents the histogram of the point estimate parameters based on the GMM optimization described in Appendix C.6. The other panels show scatter plots between two parameters. The black line or dark dot represents the theoretical value of the parameters. The number in the right corner of each panel is the correlation between the parameter estimates.

Figure E.1: Results for specification (4.14) ($\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$) using parameters of scenario A where γ^2 estimated

	$\gamma^2 \times 100$		Mean		Median		RMSE		CR		
	theoretical	$T = 1000$	$T = 4000$								
Panel A. ($\kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$)											
$\text{Var} \left[\xi_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.2541	0.2490	0.2535	0.2225	0.2460	0.1133	0.0591	0.96	0.96	0.96	0.96
$2 \frac{\Delta}{N} \text{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_s^2 ds \right] \times 100$	0.2541	0.2485	0.2534	0.2247	0.2456	0.1120	0.0583	0.95	0.96	0.96	0.96
$2 \frac{\Delta}{N} \text{E} \left[RQ_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.2541	0.2487	0.2534	0.2249	0.2463	0.1124	0.0582	0.96	0.96	0.96	0.96
(1) $\gamma_{GMM}^2 \times 100$ BZ		0.5006	0.1838	0.5115	0.3572	1.6873	1.1132	0.95	0.96	0.96	0.96
(2) $\gamma_{GMM}^2 \times 100$ CM		0.2334	0.2473	0.2098	0.2397	0.1115	0.0583	0.97	0.96	0.96	0.96
Panel B. ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$)											
$\text{Var} \left[\xi_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.1829	0.1821	0.1830	0.1795	0.1827	0.0259	0.0132	0.95	0.95	0.95	0.95
$2 \frac{\Delta}{N} \text{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_s^2 ds \right] \times 100$	0.1829	0.1819	0.1829	0.1799	0.1826	0.0236	0.0122	0.96	0.96	0.96	0.96
$2 \frac{\Delta}{N} \text{E} \left[RQ_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.1829	0.1819	0.1828	0.1800	0.1828	0.0237	0.0123	0.96	0.96	0.96	0.96
(1) $\gamma_{GMM}^2 \times 100$ BZ		0.1282	0.0484	0.1680	0.0688	0.4747	0.3020	0.94	0.92	0.92	0.92
(2) $\gamma_{GMM}^2 \times 100$ CM		0.1760	0.1812	0.1740	0.1805	0.0295	0.0151	0.95	0.96	0.96	0.96
Panel C. ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$)											
$\text{Var} \left[\xi_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.2744	0.2710	0.2746	0.2573	0.2724	0.0769	0.0401	0.96	0.96	0.96	0.96
$2 \frac{\Delta}{N} \text{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_s^2 ds \right] \times 100$	0.2744	0.2710	0.2744	0.2581	0.2719	0.0733	0.0383	0.96	0.95	0.95	0.95
$2 \frac{\Delta}{N} \text{E} \left[RQ_{\Delta t, \Delta(t+1)}^{(N)} \right] \times 100$	0.2744	0.2709	0.2744	0.2586	0.2717	0.0735	0.0383	0.96	0.95	0.95	0.95
(1) $\gamma_{GMM}^2 \times 100$ BZ		0.4095	0.1648	0.5581	0.2484	1.1906	0.7169	0.96	0.95	0.95	0.95
(2) $\gamma_{GMM}^2 \times 100$ CM		0.2472	0.2653	0.2362	0.2619	0.0835	0.0438	0.96	0.96	0.96	0.96

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. $\text{Var} \left[\xi_{\Delta t, \Delta(t+1)}^{(N)} \right]$ is the sample variance of the error measurement for each path, $\text{E} \left[RQ_{\Delta t, \Delta(t+1)}^{(N)} \right]$ is the sample mean of the realized quarticity, and $\text{E} \left[\int_{t\Delta}^{(t+1)\Delta} V_s^2 ds \right]$ the sample mean of the integrated square variance. Finally, (1) and (2) represent the point estimates of the GMM optimization using the BZ and CM1 procedures, as documented in Tables 4.1 and E.1 respectively and in accordance with specification (4.14). Details of the choice of the procedure are provided in Appendix E.2. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. If estimated/fixe, the error measurement parameter is compared/set to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table E.6: Theoretical and empirical expected value for the measurement error γ^2 related to specification (4.14)

E.4 Moments under the Model Filtration $\{\mathcal{F}_t\}_{t \geq 0}$

This Online Appendix derives the conditional moments of the integrated variance and stochastic integrals. For readability, the subscript k is omitted and we denote the following:

$$\mathcal{V}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} V_s ds, \mathcal{W}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dW_s \text{ and } \mathcal{B}_{t,t+1} = \int_{t\Delta}^{(t+1)\Delta} \sqrt{V_s} dB_s.$$

E.4.1 Summary

$$\begin{aligned} \mathbb{E}[\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\ \text{Var}[\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \frac{\sigma^2}{\kappa^2} V_{t\Delta} \left(\frac{1 - e^{-2\kappa\Delta}}{\kappa} - 2e^{-\kappa\Delta} \Delta \right) \\ &\quad + \frac{\sigma^2}{\kappa^2} \theta \left(\Delta - \frac{1 - e^{-2\kappa\Delta}}{2\kappa} - 2 \frac{1 - e^{-\kappa\Delta}}{\kappa} + 2e^{-\kappa\Delta} \Delta \right), \\ \mathbb{E}[\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Var}[\mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\ \mathbb{E}[\mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Var}[\mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} \right) + \theta \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right), \\ \text{Cov}[\mathcal{W}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cov}[\mathcal{W}_{t,t+1}, \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] &= \frac{\sigma}{\kappa} V_{t\Delta} \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right) \\ &\quad + \frac{\sigma}{\kappa} \theta \left(\Delta - 2 \frac{1 - e^{-\kappa\Delta}}{\kappa} + e^{-\kappa\Delta} \Delta \right), \\ \text{Cov}[\mathcal{W}_{t,t+1}, V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \sigma V_{t\Delta} e^{-\kappa\Delta} \Delta + \sigma \theta \left(\frac{1 - e^{-\kappa\Delta}}{\kappa} - e^{-\kappa\Delta} \Delta \right), \\ \text{Cov}[\mathcal{V}_{t,t+1}, \mathcal{B}_{t,t+1} | \mathcal{F}_{t\Delta}] &= 0, \\ \text{Cov}[\mathcal{V}_{t,t+1}, V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \frac{\sigma^2}{\kappa} V_{t\Delta} e^{-\kappa\Delta} \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right) \\ &\quad + \frac{\sigma^2}{\kappa} \frac{\theta}{2\kappa} \left((1 - e^{-\kappa\Delta})^2 - 2e^{-\kappa\Delta} \left(\Delta - \frac{1 - e^{-\kappa\Delta}}{\kappa} \right) \right). \end{aligned}$$

E.4.2 Proofs

Lemma E.4.1. For $0 < s < t$,

$$V_t = V_s e^{-\kappa(t-s)} + \theta (1 - e^{-\kappa(t-s)}) + \sigma \int_s^t e^{-\kappa(t-u)} \sqrt{V_u} dW_u. \quad (\text{E.4})$$

Proof. Be $Y_t = \int_0^t e^{\kappa u} \sqrt{V_u} dW_u$ and note that $V_t = g(t, Y_t)$ where $g(t, y) = V_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}) + \sigma e^{-\kappa t} y$. Itô's lemma implies that

$$\begin{aligned} dV_t &= \frac{\partial g}{\partial t}(t, Y_t) dt + \frac{\partial g}{\partial y}(t, Y_t) dY_t + \frac{1}{2} \frac{\partial^2 g}{\partial y^2}(t, Y_t) d\langle Y \rangle_t \\ &= -\kappa (V_0 e^{-\kappa t} - \theta e^{-\kappa t} + \sigma e^{-\kappa t} Y_t) dt + \sigma e^{-\kappa t} e^{\kappa t} \sqrt{V_t} dW_t \\ &= \kappa (\theta - V_t) dt + \sigma \sqrt{V_t} dW_t, \end{aligned}$$

which means that Equation (E.4) holds when $s = 0$. Because $V_s = V_0 e^{-\kappa s} + \theta (1 - e^{-\kappa s}) + \sigma e^{-\kappa s} \int_0^s e^{\kappa u} \sqrt{V_u} dW_u$, then

$$V_0 = e^{\kappa s} V_s + \theta (1 - e^{\kappa s}) - \sigma \int_0^s e^{\kappa u} \sqrt{V_u} dW_u,$$

and

$$\begin{aligned} V_t &= V_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}) + \sigma e^{-\kappa t} \int_0^t e^{\kappa u} \sqrt{V_u} dW_u \\ &= \left(e^{\kappa s} V_s + \theta (1 - e^{\kappa s}) - \sigma \int_0^s e^{\kappa u} \sqrt{V_u} dW_u \right) e^{-\kappa t} \\ &\quad + \theta (1 - e^{-\kappa t}) + \sigma e^{-\kappa t} \int_0^t e^{\kappa u} \sqrt{V_u} dW_u \\ &= e^{-\kappa(t-s)} V_s + \theta (1 - e^{-\kappa(t-s)}) + \sigma e^{-\kappa t} \int_s^t e^{\kappa u} \sqrt{V_u} dW_u. \quad \square \end{aligned}$$

Lemma E.4.2.

$$\int_{t\Delta}^u V_s ds = \frac{1}{\kappa} \left(V_{t\Delta} (1 - e^{-\kappa(u-t\Delta)}) + \kappa \theta (u - t\Delta) - \theta (1 - e^{-\kappa(u-t\Delta)}) + \sigma \int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s \right). \quad (\text{E.5})$$

Proof. From the integral form of Equation (E.4),

$$V_u = V_{t\Delta} + \int_{t\Delta}^u \kappa (\theta - V_s) ds + \sigma \int_{t\Delta}^u \sqrt{V_s} dW_s,$$

which is equivalent to

$$\int_{t\Delta}^u V_s ds = \frac{-V_u + V_{t\Delta}}{\kappa} + \theta \Delta + \frac{\sigma}{\kappa} \int_{t\Delta}^u \sqrt{V_s} dW_s.$$

From Equation (E.4),

$$V_u = V_{t\Delta} e^{-\kappa(u-t\Delta)} + \theta (1 - e^{-\kappa(u-t\Delta)}) + \sigma \int_{t\Delta}^u e^{-\kappa(u-s)} \sqrt{V_s} dW_s.$$

By substitution,

$$\begin{aligned} \int_{t\Delta}^u V_s ds &= \frac{1}{\kappa} \left(\begin{aligned} &-V_{t\Delta} e^{-\kappa(u-t\Delta)} - \theta (1 - e^{-\kappa(u-t\Delta)}) - \sigma \int_{t\Delta}^u e^{-\kappa(u-s)} \sqrt{V_s} dW_s \\ &+ V_{t\Delta} + \kappa \theta \Delta + \sigma \int_{t\Delta}^u \sqrt{V_s} dW_s \end{aligned} \right) \\ &= \frac{1}{\kappa} \left(\begin{aligned} &V_{t\Delta} (1 - e^{-\kappa(u-t\Delta)}) + \kappa \theta (u - t\Delta) - \theta (1 - e^{-\kappa(u-t\Delta)}) \\ &+ \sigma \int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s \end{aligned} \right). \quad \square \end{aligned}$$

Corollary E.4.3.

$$\mathbb{E}[V_u | \mathcal{F}_{t\Delta}] = V_{t\Delta} e^{-\kappa(u-\Delta t)} + \theta (1 - e^{-\kappa(u-\Delta t)}).$$

Proof. Direct application of Lemma E.4.1. □

Corollary E.4.4.

$$\mathbb{E} \left[\int_{t\Delta}^u V_s ds \middle| \mathcal{F}_{t\Delta} \right] = V_{t\Delta} \left(\frac{1 - e^{-\kappa(u-\Delta t)}}{\kappa} \right) + \theta (u - \Delta t) - \theta \frac{1 - e^{-\kappa(u-\Delta t)}}{\kappa}.$$

Proof. Direct application of Lemma E.5. □

The next Lemma and Corollaries are useful to find the conditional covariance of the integrated variance and the stochastic integrals.

Lemma E.4.5. For $u > t\Delta$,

$$\mathbb{E} \left[\int_{t\Delta}^u e^{\bar{\kappa}s} V_s ds \middle| \mathcal{F}_{t\Delta} \right] = \begin{cases} (V_{t\Delta} - \theta) e^{\kappa t\Delta} \frac{e^{(\bar{\kappa}-\kappa)u} - e^{(\bar{\kappa}-\kappa)t\Delta}}{(\bar{\kappa}-\kappa)} + \theta \frac{e^{\bar{\kappa}u} - e^{\bar{\kappa}t\Delta}}{\bar{\kappa}} & \text{if } \bar{\kappa} > \kappa, \\ (V_{t\Delta} - \theta) e^{\kappa t\Delta} (u - t\Delta) + \theta \frac{e^{\bar{\kappa}u} - e^{\bar{\kappa}t\Delta}}{\bar{\kappa}} & \text{if } \bar{\kappa} = \kappa, \\ (V_{t\Delta} - \theta) \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} + \theta(u - t\Delta) & \text{if } \bar{\kappa} = 0. \end{cases}$$

Proof.

$$\begin{aligned} & \mathbb{E} \left[\int_{t\Delta}^u e^{\bar{\kappa}s} V_s ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \mathbb{E} \left[\int_{t\Delta}^u e^{\bar{\kappa}s} \left(e^{-\kappa(s-t\Delta)} V_{t\Delta} + \theta (1 - e^{-\kappa(s-t\Delta)}) + \sigma e^{-\kappa s} \int_{t\Delta}^s e^{\kappa r} \sqrt{V_r} dW_r \right) ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \int_{t\Delta}^u e^{\bar{\kappa}s} \left(e^{-\kappa(s-t\Delta)} V_{t\Delta} + \theta (1 - e^{-\kappa(s-t\Delta)}) \right) ds \\ &= \int_{t\Delta}^u (V_{t\Delta} - \theta) e^{\bar{\kappa}s} e^{-\kappa(s-t\Delta)} + \theta e^{\bar{\kappa}s} ds, \end{aligned}$$

leading to the result. □

Corollary E.4.6.

$$\begin{aligned} & \mathbb{E} \left[\int_{t\Delta}^u e^{-\bar{\kappa}(u-s)} V_s ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \begin{cases} (V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} \frac{1 - e^{-(\bar{\kappa}-\kappa)(u-t\Delta)}}{(\bar{\kappa}-\kappa)} + \theta \frac{1 - e^{-\bar{\kappa}(u-t\Delta)}}{\bar{\kappa}} & \text{if } \bar{\kappa} > \kappa, \\ (V_{t\Delta} - \theta) e^{-\bar{\kappa}(u-t\Delta)} (u - t\Delta) + \theta \frac{1 - e^{-\bar{\kappa}(u-t\Delta)}}{\bar{\kappa}} & \text{if } \bar{\kappa} = \kappa, \\ (V_{t\Delta} - \theta) \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} + \theta(u - t\Delta) & \text{if } \bar{\kappa} = 0. \end{cases} \end{aligned}$$

and

$$\mathbb{E} \left[\int_{t\Delta}^u (1 - e^{-\bar{\kappa}(u-s)}) V_s ds \middle| \mathcal{F}_{t\Delta} \right]$$

$$= \begin{cases} (V_{t\Delta} - \theta) \left(\frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} - e^{-\kappa(u-t\Delta)} \frac{1 - e^{-(\bar{\kappa}-\kappa)(u-t\Delta)}}{(\bar{\kappa}-\kappa)} \right) + \theta \left((u-t\Delta) - \frac{1 - e^{-\bar{\kappa}(u-t\Delta)}}{\bar{\kappa}} \right), & \text{if } \bar{\kappa} > \kappa, \\ (V_{t\Delta} - \theta) \left(\frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} - e^{-\bar{\kappa}(u-t\Delta)}(u-t\Delta) \right) + \theta \left((u-t\Delta) - \frac{1 - e^{-\bar{\kappa}(u-t\Delta)}}{\bar{\kappa}} \right) & \text{if } \bar{\kappa} = \kappa. \end{cases}$$

Proof. Direct application of Lemma E.4.5. □

Corollary E.4.7.

$$\begin{aligned} & e^{-\kappa u} \text{Cov} \left[\int_{t\Delta}^u e^{\kappa s} \sqrt{V_s} dW_s, \int_{t\Delta}^u \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \\ &= (V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} (u-t\Delta) + \theta \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa}, \\ & e^{-2\kappa u} \text{Var} \left[\int_{t\Delta}^u e^{\kappa s} \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \\ &= (V_{t\Delta} - \theta) \frac{e^{-\kappa(u-t\Delta)} - e^{-2\kappa(u-t\Delta)}}{\kappa} + \theta \frac{1 - e^{-2\kappa(u-t\Delta)}}{2\kappa}, \\ & \text{Var} \left[\int_{t\Delta}^u \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] = (V_{t\Delta} - \theta) \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} + \theta(u-t\Delta). \end{aligned}$$

Proof. Each conditional moment simplifies to

$$\begin{aligned} & e^{-\kappa u} \text{Cov} \left[\int_{t\Delta}^u e^{\kappa s} \sqrt{V_s} dW_s, \int_{t\Delta}^u \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] = e^{-\kappa u} \text{E} \left[\int_{t\Delta}^u e^{\kappa s} V_s ds \middle| \mathcal{F}_{t\Delta} \right], \\ & e^{-2\kappa u} \text{Var} \left[\int_{t\Delta}^u e^{\kappa s} \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] = e^{-2\kappa u} \left(\text{E} \left[\int_{t\Delta}^u e^{2\kappa s} V_s ds \middle| \mathcal{F}_{t\Delta} \right] \right), \\ & \text{Var} \left[\int_{t\Delta}^u \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] = \text{E} \left[\int_{t\Delta}^u V_s ds \middle| \mathcal{F}_{t\Delta} \right], \end{aligned}$$

leading to the results with the application of Corollary E.4.6. □

Corollary E.4.8.

$$\text{Cov} \left[\int_{t\Delta}^u \sqrt{V_s} dW_s, \int_{t\Delta}^u V_s ds \middle| \mathcal{F}_{t\Delta} \right] = \frac{\sigma}{\kappa} \theta \left((u-t\Delta) - \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right)$$

$$+ \frac{\sigma}{\kappa} (V_{t\Delta} - \theta) \left(\frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} - e^{-\kappa(u-t\Delta)}(u - t\Delta) \right).$$

Proof. From Equation (E.5)

$$\int_{t\Delta}^u V_s ds = \frac{1}{\kappa} \left(V_{t\Delta} (1 - e^{-\kappa(u-t\Delta)}) + \kappa\theta(u - t\Delta) - \theta(1 - e^{-\kappa(u-t\Delta)}) + \sigma \int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s \right).$$

Therefore, from Corollary E.4.7,

$$\begin{aligned} & \text{Cov} \left[\int_{t\Delta}^u \sqrt{V_s} dW_s, \int_{t\Delta}^u V_s ds \middle| \mathcal{F}_{t\Delta} \right] \\ &= \text{Cov} \left[\int_{t\Delta}^u \sqrt{V_s} dW_s, \frac{\sigma}{\kappa} \int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \\ &= \frac{\sigma}{\kappa} (V_{t\Delta} - \theta) \left(\frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} - e^{-\kappa(u-t\Delta)}(u - t\Delta) \right) \\ &+ \frac{\sigma}{\kappa} \theta \left((u - t\Delta) - \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right). \quad \square \end{aligned}$$

Corollary E.4.9.

$$\text{Var} [V_u | \mathcal{F}_{t\Delta}] = \sigma^2 \frac{e^{-\kappa(u-t\Delta)} - e^{-2\kappa(u-t\Delta)}}{\kappa} V_{t\Delta} + \frac{\theta\sigma^2}{2\kappa} (1 - e^{-\kappa(u-t\Delta)})^2.$$

Proof. From Lemma E.4.1 and Corollary E.4.7,

$$\begin{aligned} \text{Var} [V_u | \mathcal{F}_{t\Delta}] &= \sigma^2 e^{-2\kappa u} \text{Var} \left[\int_{t\Delta}^u e^{\kappa s} \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \\ &= \sigma^2 (V_{t\Delta} - \theta) \frac{e^{-\kappa(u-t\Delta)} - e^{-2\kappa(u-t\Delta)}}{\kappa} + \theta\sigma^2 \frac{1 - e^{-2\kappa(u-t\Delta)}}{2\kappa}. \quad \square \end{aligned}$$

Corollary E.4.10.

$$\text{Var} \left[\int_{t\Delta}^{(t+1)\Delta} V_s ds \middle| \mathcal{F}_{t\Delta} \right] = \frac{\sigma^2}{\kappa^2} (V_{t\Delta} - \theta) \left(\frac{1 - e^{-2\kappa(u-t\Delta)}}{\kappa} - 2e^{-\kappa(u-t\Delta)}(u - t\Delta) \right)$$

$$+ \frac{\sigma^2}{\kappa^2} \theta \left((u - t\Delta) + \frac{1 - e^{-2\kappa(u-t\Delta)}}{2\kappa} - 2 \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right).$$

Proof. From Lemma (E.5),

$$\int_{t\Delta}^{(t+1)\Delta} V_u du = \frac{1}{\kappa} \left(V_{t\Delta} (1 - e^{-\kappa\Delta}) + \kappa\theta\Delta - \theta (1 - e^{-\kappa\Delta}) + \sigma \int_{t\Delta}^{(t+1)\Delta} (1 - e^{-\kappa((t+1)\Delta-u)}) \sqrt{V_u} dW_u \right).$$

Therefore, from Corollary E.4.7,

$$\begin{aligned} \text{Var} \left[\int_{t\Delta}^u V_s ds \middle| \mathcal{F}_{t\Delta} \right] &= \frac{\sigma^2}{\kappa^2} \left(\text{Var} \left[\int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \right) \\ &= \frac{\sigma^2}{\kappa^2} (V_{t\Delta} - \theta) \left(\frac{\frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} + \frac{e^{-\kappa(u-t\Delta)} - e^{-2\kappa(u-t\Delta)}}{\kappa}}{-2e^{-\kappa(u-t\Delta)}(u - t\Delta)} \right) \\ &\quad + \frac{\sigma^2}{\kappa^2} \theta \left((u - t\Delta) + \frac{1 - e^{-2\kappa(u-t\Delta)}}{2\kappa} - 2 \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right). \quad \square \end{aligned}$$

Corollary E.4.11.

$$\begin{aligned} \text{Cov} \left[\int_{t\Delta}^u V_s ds, V_u \middle| \mathcal{F}_{t\Delta} \right] &= \frac{\sigma^2}{\kappa} (V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} \left((u - t\Delta) - \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right) \\ &\quad + \frac{\sigma^2}{\kappa} \frac{\theta}{2\kappa} (1 - e^{-\kappa(u-t\Delta)})^2. \end{aligned}$$

Proof. From Equations (E.4)-(E.5) and Corollary E.4.7

$$\begin{aligned} &\text{Cov} \left[\int_{t\Delta}^u V_s ds, V_u \middle| \mathcal{F}_{t\Delta} \right] \\ &= \text{Cov} \left[\frac{\sigma}{\kappa} \int_{t\Delta}^u (1 - e^{-\kappa(u-s)}) \sqrt{V_s} dW_s, \sigma \int_{t\Delta}^u e^{-\kappa(u-s)} \sqrt{V_s} dW_s \middle| \mathcal{F}_{t\Delta} \right] \\ &= \frac{\sigma^2}{\kappa} \left((V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} (u - t\Delta) + \theta \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right) \\ &\quad - \frac{\sigma^2}{\kappa} \left((V_{t\Delta} - \theta) \frac{e^{-\kappa(u-t\Delta)} - e^{-2\kappa(u-t\Delta)}}{\kappa} + \theta \frac{1 - e^{-2\kappa(u-t\Delta)}}{2\kappa} \right) \end{aligned}$$

$$\begin{aligned}
 &= \frac{\sigma^2}{\kappa} (V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} \left((u - t\Delta) - \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa} \right) \\
 &+ \frac{\sigma^2}{\kappa} \frac{\theta}{2\kappa} (1 - e^{-\kappa(u-t\Delta)})^2. \quad \square
 \end{aligned}$$

Corollary E.4.12.

$$\text{Cov} [\mathcal{W}_{t,u}, V_u | \mathcal{F}_{t\Delta}] = \sigma (V_{t\Delta} - \theta) e^{-\kappa(u-t\Delta)} (u - t\Delta) + \sigma \theta \frac{1 - e^{-\kappa(u-t\Delta)}}{\kappa}.$$

Proof. Direct application from Corollary E.4.7. □

E.5 BZ Replication

We conduct a comparative analysis between the original Table 1 from Bollerslev and Zhou, 2002 (Table E.7) and our replication (Table E.8). Similarly, we extend this comparison to Table 2 from Bollerslev and Zhou, 2002 (Table E.9) and our corresponding replication (Table E.10).

The results in Table E.8 align with those in Table E.7. However, our attempt to replicate Table 2 from Bollerslev and Zhou, 2002, as shown in Table E.9, encountered challenges, particularly in estimating the measurement error parameter γ^2 . Furthermore, Bolko et al., 2023 suggested that Bollerslev and Zhou, 2002 estimated this parameter independently of the structural parameters, further supporting the validity of the obstacles we encountered during our replication process.

Online Appendix E. *GMM Parameter Estimation in Stochastic Volatility Models*

True value	Mean		Median		RMSE	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Panel A						
<i>GMM with quadratic variation from high-frequency return</i>						
$\kappa = 0.03$	0.0352	0.0313	0.0340	0.0310	0.0130	0.0054
$\theta = 0.25$	0.2430	0.2487	0.2355	0.2460	0.0523	0.0258
$\sigma = 0.10$	0.1016	0.1030	0.1018	0.1030	0.0080	0.0050
<i>GMM with integrated volatility</i>						
$\kappa = 0.03$	0.0382	0.0323	0.0374	0.0319	0.0139	0.0055
$\theta = 0.25$	0.2138	0.2456	0.2273	0.2437	0.0521	0.0257
$\sigma = 0.10$	0.0992	0.0999	0.0992	0.0098	0.0044	0.0020
Panel B						
<i>GMM with quadratic variation from high-frequency return</i>						
$\kappa = 0.10$	0.1057	0.1023	0.1048	0.1016	0.0214	0.0100
$\theta = 0.25$	0.2478	0.2491	0.2474	0.2489	0.0158	0.0078
$\sigma = 0.10$	0.1059	0.1073	0.1061	0.1072	0.0093	0.0082
<i>GMM with integrated volatility</i>						
$\kappa = 0.10$	0.1102	0.1032	0.1090	0.1027	0.0214	0.0091
$\theta = 0.25$	0.2460	0.2486	0.2459	0.2483	0.0163	0.0078
$\sigma = 0.10$	0.0994	0.1000	0.0995	0.0098	0.0042	0.0020
Panel C						
<i>GMM with quadratic variation from high-frequency return</i>						
$\kappa = 0.10$	0.1113	0.1035	0.1091	0.1035	0.0253	0.0111
$\theta = 0.25$	0.2389	0.2468	0.2364	0.2463	0.0326	0.0158
$\sigma = 0.20$	0.2031	0.2051	0.2030	0.2049	0.0122	0.0078
<i>GMM with integrated volatility</i>						
$\kappa = 0.10$	0.1153	0.1048	0.1131	0.1047	0.0270	0.0114
$\theta = 0.25$	0.2346	0.2455	0.2319	0.2449	0.0341	0.0160
$\sigma = 0.20$	0.1984	0.1997	0.1982	0.1995	0.0097	0.0046

This table is a direct copy of the Table 1 from Bollerslev and Zhou, 2002.

Table E.7: Table 1 from Bollerslev and Zhou, 2002

Online Appendix E. *GMM Parameter Estimation in Stochastic Volatility Models*

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A ($\kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$10^2 \times \kappa = 3.0000$	3.6247	3.1456	3.5060	3.1275	1.3173	0.5808	0.9040	0.9430
$\theta = 0.2500$	0.2421	0.2481	0.2357	0.2466	0.0537	0.0275	0.9680	0.9470
$\sigma = 0.1000$	0.1019	0.1028	0.1017	0.1028	0.0082	0.0049	0.9490	0.8980
<i>GMM with integrated volatility</i>								
$10^2 \times \kappa = 3.0000$	3.9285	3.2684	3.8246	3.2161	1.4586	0.6101	0.8810	0.9200
$\theta = 0.2500$	0.2355	0.2451	0.2293	0.2440	0.0597	0.0276	0.9780	0.9420
$\sigma = 0.1000$	0.0997	0.1000	0.0999	0.1000	0.0043	0.0021	0.9510	0.9490
Panel B ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$\kappa = 0.1000$	0.1071	0.1015	0.1057	0.1011	0.0212	0.0101	0.9300	0.9500
$\theta = 0.2500$	0.2468	0.2492	0.2462	0.2492	0.0162	0.0081	0.9500	0.9490
$\sigma = 0.1000$	0.1066	0.1070	0.1064	0.1070	0.0098	0.0079	0.8430	0.5000
<i>GMM with integrated volatility</i>								
$\kappa = 0.1000$	0.1101	0.1027	0.1088	0.1024	0.0205	0.0092	0.9060	0.9350
$\theta = 0.2500$	0.2458	0.2487	0.2453	0.2488	0.0165	0.0082	0.9460	0.9480
$\sigma = 0.1000$	0.0996	0.0999	0.0996	0.0999	0.0041	0.0020	0.9500	0.9470
Panel C ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$\kappa = 0.1000$	0.1133	0.1033	0.1120	0.1027	0.0260	0.0117	0.9090	0.9370
$\theta = 0.2500$	0.2383	0.2467	0.2361	0.2465	0.0339	0.0167	0.9430	0.9440
$\sigma = 0.2000$	0.2038	0.2051	0.2034	0.2049	0.0121	0.0076	0.9360	0.8580
<i>GMM with integrated volatility</i>								
$\kappa = 0.1000$	0.1165	0.1049	0.1151	0.1045	0.0278	0.0122	0.8840	0.9230
$\theta = 0.2500$	0.2357	0.2454	0.2333	0.2451	0.0346	0.0172	0.9390	0.9410
$\sigma = 0.2000$	0.1992	0.1999	0.1992	0.1999	0.0092	0.0045	0.9500	0.9440

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. It uses the moments of column BZ from Table 4.1 without any measurement error. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table E.8: Replication of Table 1 from Bollerslev and Zhou, 2002

True value	Mean		Median		RMSE	
	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$	$T = 1000$	$T = 4000$
Panel A: GMM with quadratic variation from high-frequency return						
$\kappa = 0.03$	0.0364	0.0317	0.0354	0.0315	0.0138	0.0056
$\theta = 0.25$	0.2456	0.2491	0.2384	0.2464	0.0520	0.0257
$\sigma = 0.10$	0.0909	0.0994	0.0905	0.0983	0.0230	0.0127
γ^2	0.0007	0.0004	0.0006	0.0004	0.0008	0.0005
Panel B: GMM with quadratic variation from high-frequency return						
$\kappa = 0.10$	0.1067	0.1027	0.1061	0.1023	0.0219	0.0104
$\theta = 0.25$	0.2489	0.2494	0.2484	0.2492	0.0157	0.0078
$\sigma = 0.10$	0.0990	0.1049	0.0986	0.1046	0.0214	0.0121
γ^2	0.0007	0.0004	0.0006	0.0003	0.0009	0.0005
Panel C: GMM with quadratic variation from high-frequency return						
$\kappa = 0.10$	0.1133	0.1042	0.1109	0.1043	0.0274	0.0119
$\theta = 0.25$	0.2435	0.2481	0.2400	0.2473	0.0314	0.0157
$\sigma = 0.20$	0.1893	0.1999	0.1884	0.1987	0.0303	0.0162
γ^2	0.0017	0.0010	0.0015	0.0009	0.0019	0.0013

This table is a direct copy of the Table 2 from Bollerslev and Zhou, 2002. The γ^2 in this paper is equivalent to the γ of BZ.

Table E.9: Table 2 from Bollerslev and Zhou, 2002

True value	Mean		Median		RMSE		CR	
	$T = 1000$	$T = 4000$						
Panel A ($\kappa = 0.03, \theta = 0.25$ & $\sigma = 0.10$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$10^2 \times \kappa = 3.0000$	3.7221	3.1934	3.6196	3.1605	1.3647	0.6114	0.9050	0.9380
$\theta = 0.2500$	0.2441	0.2483	0.2381	0.2470	0.0530	0.0274	0.9620	0.9480
$\sigma = 0.1000$	0.0917	0.0988	0.0916	0.0976	0.0242	0.0135	0.9340	0.9470
$10^2 \times \gamma^2 = 0.2541$	0.5006	0.1838	0.5115	0.3572	1.6873	1.1132	0.9490	0.9600
Panel B ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.10$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$\kappa = 0.1000$	0.1085	0.1020	0.1068	0.1017	0.0219	0.0104	0.9330	0.9520
$\theta = 0.2500$	0.2478	0.2494	0.2469	0.2495	0.0160	0.0081	0.9550	0.9490
$\sigma = 0.1000$	0.0987	0.1044	0.0994	0.1045	0.0230	0.0122	0.9520	0.9330
$10^2 \times \gamma^2 = 0.1829$	0.1282	0.0484	0.1680	0.0688	0.4747	0.3020	0.9430	0.9230
Panel C ($\kappa = 0.10, \theta = 0.25$ & $\sigma = 0.20$)								
<i>GMM with quadratic variation from high-frequency return</i>								
$\kappa = 0.1000$	0.1143	0.1040	0.1130	0.1031	0.0267	0.0123	0.9010	0.9390
$\theta = 0.2500$	0.2429	0.2479	0.2400	0.2477	0.0325	0.0164	0.9530	0.9500
$\sigma = 0.2000$	0.1920	0.2006	0.1885	0.1992	0.0306	0.0167	0.9500	0.9590
$10^2 \times \gamma^2 = 0.2744$	0.4095	0.1648	0.5581	0.2484	1.1906	0.7169	0.9560	0.9550

The table shows results obtained from GMM procedures based on Monte Carlo simulations with 1000 independent trajectories of T time steps of length $\Delta = 1$. The trajectories of the integrated variance are simulated with 820 intraday steps and the realized variance with 82 intraday steps, see Section 4.3 for details. It uses the moments of column BZ from Table 4.1 according to specification (4.14) with measurement error. The Mean (Median) reports the sample average (median) of the 1000 point estimates. The coverage ratio (CR) is the proportion of the 95% confidence interval that contains the true parameter. The error measurement parameter is here compared to its theoretical value. We employ a 2-step GMM procedure using the Newey and West, 1987 covariance matrix estimator with a Bartlett-kernel with a lag of 5, see Appendix C.6 for details.

Table E.10: Replication of Table 2 from Bollerslev and Zhou, 2002

E.6 Conditional Moment Derivation

E.6.1 One-variance factor

This Online Appendix details the results summarized in Appendix C.3.1.

Lemma E.6.1. *The recursive form of the integrated variance is*

$$\mathbb{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] = \alpha \mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \Delta\beta.$$

Proof. From Equations (C.3) and (C.4)

$$\begin{aligned} \mathbb{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] &= a \mathbb{E} [V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] + b \\ &= a (\alpha V_{t\Delta} + \beta) + b \\ &= a \left(\alpha \frac{\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] - b}{a} + \beta \right) + b \\ &= \alpha \mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + (1 - e^{-\kappa\Delta}) \theta \Delta \\ &= \alpha \mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \Delta\beta. \quad \square \end{aligned}$$

Corollary E.6.2.

$$\mathbb{E} [RV_{t+1,t+2} | \mathcal{G}_{t\Delta}] - \alpha \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] - \Delta\beta = 0.$$

Proof. Based on Equations (4.8)-(4.9) and Lemma E.6.1,

$$\begin{aligned} \mathbb{E} [RV_{t+1,t+2} | \mathcal{G}_{t\Delta}] &= \mathbb{E} [\mathbb{E} [RV_{t+1,t+2} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \mathbb{E} [\mathbb{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \alpha \mathbb{E} [\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] + \Delta\beta \\ &= \alpha \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + \Delta\beta. \quad \square \end{aligned}$$

Lemma E.6.3. *The recursive form of the squared integrated variance is*

$$\mathbf{E} [\mathcal{V}_{t+1,t+2}^2 | \mathcal{F}_{t\Delta}] = H \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + I \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + J,$$

where

$$\begin{aligned} H &= \alpha^2, \\ I &= \frac{1}{a}(a^2(C + 2\alpha\beta) + (\alpha - \alpha^2)(2ab + A)), \\ J &= -bI + (a^2(D + \beta^2) + \beta(2ab + A) + (1 - \alpha^2)(b^2 + B)). \end{aligned}$$

Proof. Using sequentially Equations (C.6), (C.4), (C.5) and (C.3)

$$\begin{aligned} \mathbf{E} [\mathcal{V}_{t+1,t+2}^2 | \mathcal{F}_{t\Delta}] &= \mathbf{Var} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] + \mathbf{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}]^2 & \text{(E.6)} \\ &= a^2 (\mathbf{E} [V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] + \mathbf{Var} [V_{(t+1)\Delta}^2 | \mathcal{F}_{t\Delta}]) + (A + 2ab) \mathbf{E} [V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}] + (B + b^2) \\ &= a^2(CV_{t\Delta} + D + \alpha^2V_{t\Delta}^2 + 2\alpha\beta V_{t\Delta} + \beta^2) + (A + 2ab) (\alpha V_{t\Delta} + \beta) + (B + b^2) \\ &= a^2\alpha^2V_{t\Delta}^2 + ((A + 2ab)\alpha + a^2C + a^22\alpha\beta) V_{t\Delta} \\ &\quad + (B + b^2) + (A + 2ab)\beta + a^2(D + \beta^2). \end{aligned}$$

Similarly, from Equations (C.4) and (C.6)

$$\begin{aligned} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] &= \mathbf{Var} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]^2 \\ &= AV_{t\Delta} + B + (aV_{t\Delta} + b)^2 \\ &= a^2V_{t\Delta}^2 + (A + 2ab) V_{t\Delta} + (B + b^2), \end{aligned}$$

which implies that

$$\alpha^2 a^2 V_{t\Delta}^2 = \alpha^2 \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] - \alpha^2 (A + 2ab) V_{t\Delta} - \alpha^2 (B + b^2).$$

Replacing back in Equation (E.6) leads to

$$\begin{aligned}
 & \mathbb{E} [\mathcal{V}_{t+1,t+2}^2 | \mathcal{F}_{t\Delta}] \\
 &= \alpha^2 \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + (B + b^2) + (A + 2ab) \beta + a^2(D + \beta^2) - \alpha^2 (B + b^2) \\
 &+ ((A + 2ab) \alpha + a^2 C + a^2 2\alpha\beta - \alpha^2 (A + 2ab)) V_{t\Delta} \\
 &= \alpha^2 \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + (1 - \alpha^2) (B + b^2) + (A + 2ab) \beta + a^2(D + \beta^2) \\
 &+ ((\alpha - \alpha^2) (A + 2ab) + a^2(C + 2\alpha\beta)) \left(\frac{\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] - b}{a} \right) \quad (\text{from Eq. (C.4)}) \\
 &= H \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + I \mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + J. \quad \square
 \end{aligned}$$

Corollary E.6.4. *If $\mathbb{E} [\mathcal{V}_{t,t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0, \quad \forall t \in \{1, \dots, T\}$, then*

$$\begin{aligned}
 & \mathbb{E} [RV_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] \\
 & - H (\mathbb{E} [RV_{t,t+1}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{G}_{t\Delta}]) - I \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] - J = 0.
 \end{aligned}$$

Proof. From Equation (4.8),

$$\mathbb{E} [RV_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = \mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \underbrace{2\mathbb{E} [\mathcal{V}_{t,t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}]}_{=0},$$

which implies that $\mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] = \mathbb{E} [RV_{t,t+1}^2 | \mathcal{F}_{t\Delta}] - \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{F}_{t\Delta}]$. This relation holds for all t . Therefore, based on Lemma E.6.3,

$$\begin{aligned}
 & \mathbb{E} [RV_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t+1,t+2}^2 | \mathcal{G}_{t\Delta}] \\
 &= \mathbb{E} [\mathbb{E} [\mathbb{E} [RV_{t+1,t+2}^2 - \xi_{t+1,t+2}^2 | \mathcal{F}_{(t+1)\Delta}] | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\
 &= \mathbb{E} [\mathbb{E} [\mathcal{V}_{t+1,t+2}^2 | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\
 &= H \mathbb{E} [\mathbb{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] + I \mathbb{E} [\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] + J \\
 &= H (\mathbb{E} [RV_{t,t+1}^2 | \mathcal{G}_{t\Delta}] - \mathbb{E} [\xi_{t,t+1}^2 | \mathcal{G}_{t\Delta}]) + I \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + J. \quad \square
 \end{aligned}$$

Lemma E.6.5. *The recursive form of the squared lagged integrated variance is*

$$\mathbf{E} [\mathcal{V}_{t+1,t+2} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] = \tilde{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \tilde{I} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \tilde{J},$$

where

$$\begin{aligned} \tilde{H} &= \alpha, \\ \tilde{I} &= \frac{1}{a}(a(F + b + \alpha b + a\beta) - \alpha(2ab + A)), \\ \tilde{J} &= -b\tilde{I} + (aG + b^2 + ab\beta - \alpha(B + b^2)). \end{aligned}$$

Proof. Using sequentially Equations (C.4) and (C.7)

$$\begin{aligned} &\mathbf{E} [\mathcal{V}_{t+1,t+2} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \mathbf{E} [\mathbf{E} [\mathcal{V}_{(t+1)\Delta,(t+2)\Delta} | \mathcal{F}_{(t+1)\Delta}] \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \mathbf{E} [(aV_{(t+1)\Delta} + b) \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= a \mathbf{E} [V_{(t+1)\Delta} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + b \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= a(FV_{t\Delta} + G + (\alpha V_{t\Delta} + \beta)(aV_{t\Delta} + b)) + b(aV_{t\Delta} + b) \\ &= \alpha a^2 V_{t\Delta}^2 + (aF + ab + a\alpha b + a^2\beta) V_{t\Delta} + (aG + b^2 + ab\beta). \end{aligned} \quad (\text{E.7})$$

From Equations Equations (C.4) and (C.6)

$$\begin{aligned} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] &= \text{Var} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]^2 \\ &= AV_{t\Delta} + B + (aV_{t\Delta} + b)^2 \\ &= a^2 V_{t\Delta}^2 + (A + 2ab) V_{t\Delta} + (B + b^2), \end{aligned}$$

which implies that

$$\alpha a^2 V_{t\Delta}^2 = \alpha \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] - \alpha(A + 2ab) V_{t\Delta} - \alpha(B + b^2).$$

Replacing back in Equation (E.7) leads to

$$\begin{aligned}
 & \mathbf{E} [\mathcal{V}_{t+1,t+2}\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}] \\
 &= \alpha \mathbf{E} [\mathcal{V}_{t,t+1}^2|\mathcal{F}_{t\Delta}] + (aG + b^2 + ab\beta - \alpha B - \alpha b^2) \\
 &+ (aF + ab + a\alpha b + a^2\beta - \alpha A - \alpha 2ab) V_{t\Delta} \\
 &= \alpha \mathbf{E} [\mathcal{V}_{t,t+1}^2|\mathcal{F}_{t\Delta}] + (aG + b^2 + ab\beta - \alpha B - \alpha b^2) \\
 &+ (aF + ab + a\alpha b + a^2\beta - \alpha A - \alpha 2ab) \left(\frac{\mathbf{E} [\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}] - b}{a} \right) \\
 &= \tilde{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2|\mathcal{F}_{t\Delta}] + \tilde{I} \mathbf{E} [\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}] + \tilde{J}. \quad \square
 \end{aligned}$$

Corollary E.6.6. *If $\mathbf{E} [\mathcal{V}_{t,t+1}\xi_{t,t+1}|\mathcal{F}_{t\Delta}] = 0$ and $\mathbf{E} [V_{t+1}\xi_{t,t+1}|\mathcal{F}_{t\Delta}] = 0 \quad \forall t \in \{1, \dots, T\}$, then*

$$\begin{aligned}
 & \mathbf{E} [RV_{t+1,t+2}RV_{t,t+1}|\mathcal{G}_{t\Delta}] \\
 & - \tilde{H} (\mathbf{E} [RV_{t,t+1}^2|\mathcal{G}_{t\Delta}] - \mathbf{E} [\xi_{t,t+1}^2|\mathcal{G}_{t\Delta}]) - \tilde{I} \mathbf{E} [RV_{t,t+1}|\mathcal{G}_{t\Delta}] - \tilde{J} = 0.
 \end{aligned}$$

Proof. Based on Equations (4.8) and (4.9), Equation (C.4) and Lemma E.6.5

$$\begin{aligned}
 & \mathbf{E} [RV_{t+1,t+2}RV_{t,t+1}|\mathcal{G}_{t\Delta}] = \mathbf{E} [\mathbf{E} [RV_{t+1,t+2}RV_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &= \mathbf{E} [\mathbf{E} [(\mathcal{V}_{t+1,t+2} + \xi_{t+1,t+2})(\mathcal{V}_{t,t+1} + \xi_{t,t+1})|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &= \mathbf{E} [\mathbf{E} [\mathcal{V}_{t+1,t+2}\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] + \mathbf{E} [\mathbf{E} [\mathbf{E} [\xi_{t+1,t+2}|\mathcal{F}_{(t+1)\Delta}] \mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &+ \mathbf{E} [\mathbf{E} [\mathbf{E} [\mathcal{V}_{t+1,t+2}\xi_{t,t+1}|\mathcal{F}_{(t+1)\Delta}]|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &+ \mathbf{E} [\mathbf{E} [\mathbf{E} [\xi_{t+1,t+2}|\mathcal{F}_{(t+1)\Delta}] \xi_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &= \mathbf{E} [\mathbf{E} [\mathcal{V}_{t+1,t+2}\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] + \mathbf{E} [\mathbf{E} [(aV_{t+1} + b) \xi_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &= \mathbf{E} [\mathbf{E} [\mathcal{V}_{t+1,t+2}\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] \\
 &= \tilde{H} \mathbf{E} [\mathbf{E} [\mathcal{V}_{t,t+1}^2|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] + \tilde{I} \mathbf{E} [\mathbf{E} [\mathcal{V}_{t,t+1}|\mathcal{F}_{t\Delta}]|\mathcal{G}_{t\Delta}] + \tilde{J} \\
 &= \tilde{H} (\mathbf{E} [RV_{t,t+1}^2|\mathcal{G}_{t\Delta}] - \mathbf{E} [\xi_{t,t+1}^2|\mathcal{G}_{t\Delta}]) + \tilde{I} \mathbf{E} [RV_{t,t+1}|\mathcal{G}_{t\Delta}] + \tilde{J}. \quad \square
 \end{aligned}$$

Lemma E.6.7. *The returns expectation conditional to the model filtration is*

$$\mathbf{E} [R_{t,t+1} | \mathcal{F}_{t\Delta}] = \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + r\Delta.$$

Proof. Direct application of Equation (4.4). □

Corollary E.6.8. *The returns expectation conditional to the observed filtration is*

$$\mathbf{E} [R_{t,t+1} | \mathcal{G}_{t\Delta}] = \left(\lambda - \frac{1}{2} \right) \mathbf{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + r\Delta.$$

Proof. Direct application of Lemma E.6.7. □

Lemma E.6.9. *The cross-moment between the returns and the integrated variance is*

$$\mathbf{E} [R_{t,t+1} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] = \widehat{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \widehat{I} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \widehat{J},$$

where

$$\begin{aligned} \widehat{H} &= \left(\lambda - \frac{1}{2} \right), \\ \widehat{I} &= \left(\frac{\rho \widetilde{F}}{a} + r\Delta \right) \approx \left(\rho \frac{\sigma}{2} + r \right) \Delta, \\ \widehat{J} &= \left(\rho \widetilde{G} - \rho \widetilde{F} \frac{b}{a} \right) \approx -\frac{1}{12} \rho \Delta^3 \kappa \theta \sigma. \end{aligned}$$

Proof. From Equations (4.4), (C.4) and (C.8),

$$\begin{aligned} &\mathbf{E} [(R_{t,t+1} - r\Delta) \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \rho \mathbf{E} [\mathcal{V}_{t,t+1} \mathcal{W}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \rho \left(\widetilde{F} V_{t\Delta} + \widetilde{G} \right) \\ &= \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \rho \left(\widetilde{F} \left(\frac{\mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] - b}{a} \right) + \widetilde{G} \right) \end{aligned}$$

$$= \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \frac{\rho \tilde{F}}{a} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \left(\rho \tilde{G} - \rho \tilde{F} \frac{b}{a} \right).$$

Therefore,

$$\begin{aligned} & \mathbf{E} [R_{t,t+1} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &= \left(\lambda - \frac{1}{2} \right) \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \left(\frac{\rho \tilde{F}}{a} + r\Delta \right) \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \left(\rho \tilde{G} - \rho \tilde{F} \frac{b}{a} \right) \\ &= \hat{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \hat{I} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \hat{J}. \quad \square \end{aligned}$$

Corollary E.6.10. *If $\mathbf{E} [R_{t,t+1} \xi_{t,t+1} | \mathcal{F}_{t\Delta}] = 0 \quad \forall t \in \{1, \dots, T\}$, the expectation conditional to the observed filtration of the return multiplied by the realized variance is*

$$\begin{aligned} & \mathbf{E} [R_{t,t+1} RV_{t,t+1} | \mathcal{G}_{t\Delta}] \\ &= \hat{H} \left(\mathbf{E} [RV_{t,t+1}^2 | \mathcal{G}_{t\Delta}] - \mathbf{E} [\xi_{t,t+1}^2 | \mathcal{G}_{t\Delta}] \right) + \hat{I} \mathbf{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + \hat{J}. \end{aligned}$$

Proof. Based on Equations (4.8) and (4.9) and the Lemma E.6.9

$$\begin{aligned} & \mathbf{E} [R_{t,t+1} RV_{t,t+1} | \mathcal{G}_{t\Delta}] \\ &= \mathbf{E} [\mathbf{E} [R_{t,t+1} RV_{t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \mathbf{E} [\mathbf{E} [R_{t,t+1} (\mathcal{V}_{t,t+1} + \xi_{t,t+1}) | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \mathbf{E} [\mathbf{E} [R_{t,t+1} \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\ &= \mathbf{E} \left[\hat{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{F}_{t\Delta}] + \hat{I} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \hat{J} \middle| \mathcal{G}_{t\Delta} \right] \\ &= \hat{H} \mathbf{E} [\mathcal{V}_{t,t+1}^2 | \mathcal{G}_{t\Delta}] + \hat{I} \mathbf{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + \hat{J}. \quad \square \end{aligned}$$

Lemma E.6.11. *The cross-moment between the returns and the lagged integrated variance is*

$$\begin{aligned} & \mathbf{E} \left[\left(R_{t,t+1} - r\Delta - \left(\lambda - \frac{1}{2} \right) \mathcal{V}_{t,t+1} \right) \mathcal{V}_{t+1,t+2} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \bar{I} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] + \bar{J}, \end{aligned}$$

where

$$\bar{I} = \rho\bar{F}, \bar{J} = a\rho\bar{G} - b\rho\bar{F}.$$

Proof. From Equations (4.4), (C.4) and (C.9),

$$\begin{aligned} & \mathbf{E} \left[\left(R_{t,t+1} - r\Delta - \left(\lambda - \frac{1}{2} \right) \mathcal{V}_{t,t+1} \right), \mathcal{V}_{t+1,t+2} \middle| \mathcal{F}_{t\Delta} \right] \\ &= \rho \mathbf{Cov} \left[aV_{(t+1)\Delta}, \mathcal{W}_{t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \\ &= a\rho (\bar{F}V_{t\Delta} + \bar{G}) \\ &= a\rho\bar{G} + a\rho\bar{F} \frac{\mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] - b}{a} \\ &= a\rho\bar{G} - b\rho\bar{F} + \rho\bar{F} \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]. \end{aligned}$$

leading to the result. □

Corollary E.6.12. *The expectation conditional to the observed filtration of the return multiplied by the lagged realized variance is*

$$\begin{aligned} & \mathbf{E} \left[\left(R_{t,t+1} - r\Delta - \left(\lambda - \frac{1}{2} \right) RV_{t,t+1} \right) RV_{t+1,t+2} \middle| \mathcal{G}_{t\Delta} \right] \\ &= \bar{I} \mathbf{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] + \bar{J}. \end{aligned}$$

Proof. Direct application of Lemma E.6.11.

E.6.2 Two-variance factor

This Online Appendix details the moment condition for the two-variance factors as in Bollerslev and Zhou, 2002.

Corollary E.6.13. *In the case of two variances, the integrated variance is defined as*

$$\mathcal{V}_{t,t+1} = \mathcal{V}_{1,t,t+1} + \mathcal{V}_{2,t,t+1},$$

and for $k \in \{1, 2\}$,

$$\begin{aligned} \mathbf{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}] &= \alpha_k V_{k,t\Delta} + \beta_k, \\ \mathbf{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] &= a_k V_{k,t\Delta} + b_k, \\ \mathbf{E} [\mathcal{V}_{k,t+1,t+2} | \mathcal{F}_{t\Delta}] &= \alpha_k \mathbf{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] + \beta_k \Delta, \\ \mathbf{E} [\mathcal{V}_{k,t+1,t+2}^2 | \mathcal{F}_{t\Delta}] &= H_k \mathbf{E} [\mathcal{V}_{k,t,t+1}^2 | \mathcal{F}_{t\Delta}] + I_k \mathbf{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] + J_k. \end{aligned}$$

where $\alpha_k, \beta_k, a_k, b_k, H_k, I_k, J_k$ are the analogous of Equations (C.3)-(C.12) with parameters θ_k, κ_k and σ_k .

Proof. Direct generalization the results of section C.3.1. □

Lemma E.6.14. *The recursive form of the integrated variance with two-variance factors is*

$$\begin{aligned} \mathbf{E} [\mathcal{V}_{t+2,t+3} | \mathcal{F}_{t\Delta}] &= (\alpha_1 + \alpha_2) \mathbf{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] - (\alpha_1 \alpha_2) \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] \\ &\quad + (1 - \alpha_1) \beta_2 \Delta + (1 - \alpha_2) \beta_1 \Delta. \end{aligned}$$

Proof. From Lemma E.6.13

$$\begin{aligned} &\mathbf{E} [\mathcal{V}_{t+2,t+3} | \mathcal{F}_{t\Delta}] - \beta_1 \Delta - \beta_2 \Delta \\ &= \alpha_1 \mathbf{E} [\mathcal{V}_{1,t+1,t+2} | \mathcal{F}_{t\Delta}] + \alpha_2 \mathbf{E} [\mathcal{V}_{2,t+1,t+2} | \mathcal{F}_{t\Delta}] \\ &= (\alpha_1 + \alpha_2) \mathbf{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] - \alpha_2 \mathbf{E} [\mathcal{V}_{1,t+1,t+2} | \mathcal{F}_{t\Delta}] - \alpha_1 \mathbf{E} [\mathcal{V}_{2,t+1,t+2} | \mathcal{F}_{t\Delta}] \\ &= (\alpha_1 + \alpha_2) \mathbf{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] - \alpha_1 \alpha_2 \mathbf{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] - \alpha_1 \beta_2 \Delta - \alpha_2 \beta_1 \Delta. \quad \square \end{aligned}$$

Corollary E.6.15.

$$\begin{aligned} \mathbf{E} [RV_{t+2,t+3} | \mathcal{G}_{t\Delta}] &= (\alpha_1 + \alpha_2) \mathbf{E} [RV_{t+1,t+2} | \mathcal{G}_{t\Delta}] - (\alpha_1 \alpha_2) \mathbf{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] \\ &\quad + (1 - \alpha_1) \beta_2 \Delta + (1 - \alpha_2) \beta_1 \Delta. \end{aligned}$$

Proof. Based on Equations (4.8)-(4.9) and Lemma E.6.14,

$$\begin{aligned}
 \mathbb{E} [RV_{t+2,t+3} | \mathcal{G}_{t\Delta}] &= \mathbb{E} [\mathbb{E} [\mathcal{V}_{t+2,t+3} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\
 &= (\alpha_1 + \alpha_2) \mathbb{E} [\mathbb{E} [\mathcal{V}_{t+1,t+2} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] - (\alpha_1\alpha_2) \mathbb{E} [\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}] | \mathcal{G}_{t\Delta}] \\
 &\quad + (1 - \alpha_1)\beta_2\Delta + (1 - \alpha_2)\beta_1\Delta \\
 &= (\alpha_1 + \alpha_2) \mathbb{E} [RV_{t+1,t+2} | \mathcal{G}_{t\Delta}] - (\alpha_1\alpha_2) \mathbb{E} [RV_{t,t+1} | \mathcal{G}_{t\Delta}] \\
 &\quad + (1 - \alpha_1)\beta_2\Delta + (1 - \alpha_2)\beta_1\Delta. \quad \square
 \end{aligned}$$

Lemma E.6.16. *The recursive form of the square integrated variance with two variance factors is*

$$\mathbb{E} \left[\begin{array}{c} (1 - \alpha_1 L)(1 - \alpha_2 L)(1 - H_1 L)(1 - H_2 L) \\ (1 - \alpha_1\alpha_2 L) \mathcal{V}_{t+5,t+6}^2 \end{array} \middle| \mathcal{F}_{t\Delta} \right] = f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2), \quad (\text{E.8})$$

where L designate the lag operator and $f(\kappa_1, \kappa_2, \theta_1, \theta_2, \sigma_1, \sigma_2)$ a function depending on the variance parameters.

Proof. See Bollerslev and Zhou, 2002 Appendix B.1.2 for details and Equation (B.8) for the function details. \square

E.7 Unconditional Moment Derivation

This Online Appendix details the summarized results in Appendix C.4.

Corollary E.7.1. *The first two unconditional moments of the variance are*

$$\mathbb{E} [V_{k,t\Delta}] = \theta_k \quad \text{and} \quad \text{Var} [V_{k,t\Delta}] = \frac{\theta_k \sigma_k^2}{2\kappa_k}.$$

Proof. The expected value is a direct application of Corollary E.4.3, The variance is based

on Corollaries E.4.3 and E.4.9, that is

$$\begin{aligned}\text{Var} [V_{k,t\Delta}] &= \text{E} [\text{Var} [V_{k,t\Delta} | \mathcal{F}_{t\Delta}] + \text{Var} [\text{E} [V_{k,t\Delta} | \mathcal{F}_{t\Delta}]] \\ &= \theta_k \sigma_k^2 \frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} + \text{Var} [V_{k,t\Delta}] e^{-2\kappa_k \Delta} \\ &\Rightarrow \text{Var} [V_{k,t\Delta}] = \frac{\theta_k \sigma_k^2}{2\kappa_k}. \quad \square\end{aligned}$$

Lemma E.7.2. *The unconditional expected value of the realized variance with K variance factors is*

$$\text{E} [RV_{t,t+1}] = \sum_{k=1}^K \theta_k \Delta.$$

Proof. With Corollary E.7.1 and Lemma E.5

$$\begin{aligned}\text{E} [RV_{t,t+1}] &= \text{E} [\mathcal{V}_{t,t+1}] = \sum_{k=1}^K \text{E} [\text{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] \\ &= \sum_{k=1}^K \text{E} [a_k V_{k,t\Delta} + b_k] = \sum_{k=1}^K \theta_k \Delta. \quad \square\end{aligned}$$

Lemma E.7.3. *The unconditional expected value of the realized quarticity with K variance factors is*

$$\text{E} [RQ_{t,t+1}] = \left(\sum_{k=1}^K \theta_k \right)^2 \Delta + \sum_{k=1}^K \frac{\theta_k \sigma_k^2}{2\kappa_k} \Delta.$$

Proof. See Appendix C.1. □

Corollary E.7.4. *The expected value of the conditional variance of the integrated variance is*

$$\text{E} [\text{Var} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] = \frac{\sigma_k^2}{\kappa_k^2} \theta_k \left(\Delta + \frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} - 2 \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right) = \frac{\theta_k \sigma_k^2}{2\kappa_k} f(\Delta, \kappa_k),$$

where $f(\Delta, \kappa_k) = \frac{2}{\kappa_k} \left(\Delta + \frac{1 - e^{-2\kappa_k \Delta}}{2\kappa_k} - 2 \frac{1 - e^{-\kappa_k \Delta}}{\kappa_k} \right)$.

Proof. Direct application from Corollary E.4.10 □

Corollary E.7.5. *The expected value of the conditional covariance between the integrated variance and the instantaneous variance is*

$$\mathbb{E} [\text{Cov} [\mathcal{V}_{k,t,t+1}, V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] = \frac{\theta_k \sigma_k^2}{2\kappa_k^2} (1 - e^{-\kappa_k \Delta})^2 = \frac{\theta_k \sigma_k^2}{2\kappa_k} g(\Delta, \kappa_k),$$

where $g(\Delta, \kappa_k) = \frac{1}{\kappa_k} (1 - e^{-\kappa_k \Delta})^2$.

Proof. Direct application from Corollary E.4.11. □

Corollary E.7.6. *The expected values of conditional covariances between $\mathcal{W}_{t,t+1}$ and both the integrated variance and the instantaneous variance are:*

$$\mathbb{E} [\text{Cov} [\mathcal{V}_{k,t,t+1}, \mathcal{W}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] = \frac{\sigma_k}{\kappa_k} \theta_k (\Delta - a_k),$$

and

$$\mathbb{E} [\text{Cov} [\mathcal{W}_{k,t,(t+1)\Delta}, V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] = \sigma_k \theta_k a_k.$$

Proof. Direct application from Corollaries E.4.8 and E.4.12. □

Lemma E.7.7. *The unconditional variance of the integrated variance with K variance factors is*

$$\text{Var} [\mathcal{V}_{t,t+1}] = \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} \left(\frac{2}{\kappa_k} (\Delta - a_k) \right).$$

Proof. With Corollaries E.4.4, E.7.1, and E.7.4,

$$\begin{aligned} \text{Var} [\mathcal{V}_{t,t+1}] &= \mathbb{E} [\text{Var} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]] + \text{Var} [\mathbb{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]] \\ &= \sum_k \mathbb{E} [\text{Var} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] + \text{Var} [\mathbb{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] \end{aligned}$$

$$\begin{aligned}
 &= \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} f(\Delta, \kappa_k) + \text{Var} [a_k V_{k,t\Delta} + b_k] \\
 &= \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} (f(\Delta, \kappa_k) + a_k^2),
 \end{aligned}$$

where, since $a_k^2 = \frac{2(1-e^{-\kappa_k\Delta})-(1-e^{-2\kappa_k\Delta})}{\kappa_k^2}$, we have

$$\begin{aligned}
 (f(\Delta, \kappa_k) + a_k^2) &= \frac{2}{\kappa_k} \Delta + \frac{1 - e^{-2\kappa_k\Delta}}{\kappa_k^2} - 4 \frac{1 - e^{-\kappa_k\Delta}}{\kappa_k^2} + \frac{2(1 - e^{-\kappa_k\Delta}) - (1 - e^{-2\kappa_k\Delta})}{\kappa_k^2} \\
 &= \Delta \frac{2}{\kappa_k} - 2 \frac{1 - e^{-\kappa_k\Delta}}{\kappa_k^2} \\
 &= \frac{2}{\kappa_k} (\Delta - a_k) \\
 &\approx \Delta \frac{2}{\kappa_k} - \frac{2}{\kappa_k^2} (\kappa_k \Delta - \frac{1}{2} \kappa_k^2 \Delta^2) = \Delta^2. \quad \square
 \end{aligned}$$

Corollary E.7.8. *If $E[\mathcal{V}_{k,t,t+1}\xi_{t,t+1}] = 0$, then the unconditional variance of the realized variance with K variance factors is*

$$\text{Var} [RV_{t,t+1}] = E[\xi_{t,t+1}^2] + \sum_k \frac{\theta_k \sigma_k^2}{2\kappa_k} \left(\frac{2}{\kappa_k} (\Delta - a_k) \right).$$

Proof. Direct application of Lemma E.7.7 and Equation (4.9). □

Lemma E.7.9. *For $h \in \{1, 2, \dots\}$, the unconditional covariance between lagged integrated variance is*

$$\text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] = \sum_k \alpha_k^{h-1} a_k^2 \frac{\theta_k \sigma_k^2}{2\kappa_k}.$$

Proof. With Corollaries E.4.11, E.7.1 and E.7.5 along with Lemmas E.4.1 and E.5,

$$\begin{aligned}
 &\text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] \\
 &= E[\text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1} | \mathcal{F}_{t\Delta}]] + \text{Cov} [E[\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}], E[\mathcal{V}_{t+h,t+h+1} | \mathcal{F}_{t\Delta}]]
 \end{aligned}$$

$$\begin{aligned}
&= \sum_k \mathbf{E} [\text{Cov} [\mathcal{V}_{k,t,t+1}, a_k \alpha_k^{h-1} V_{(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\
&+ \sum_k \text{Cov} [\mathbf{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}], \mathbf{E} [a_k \alpha_k^{h-1} V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\
&= \sum_k a_k \alpha_k^{h-1} \left(g(\Delta, \kappa_k) \frac{\theta_k \sigma_k^2}{2\kappa_k} + (a_k \alpha_k) \frac{\theta_k \sigma_k^2}{2\kappa_k} \right) \\
&= \sum_k \alpha_k^{h-1} a_k^2 \frac{\theta_k \sigma_k^2}{2\kappa_k},
\end{aligned}$$

since

$$\begin{aligned}
(g(\Delta, \kappa_k) + (a_k \alpha_k)) &= \frac{1}{\kappa_k} (1 - e^{-\kappa_k \Delta})^2 + \left(\frac{(1 - e^{-\kappa_k \Delta})}{\kappa_k} e^{-\kappa_k \Delta} \right) \\
&= \frac{(1 - e^{-\kappa_k \Delta})}{\kappa_k} (1 - e^{-\kappa_k \Delta} + e^{-\kappa_k \Delta}) \\
&= \frac{(1 - e^{-\kappa_k \Delta})}{\kappa_k} \\
&= a_k. \quad \square
\end{aligned}$$

Corollary E.7.10. *If $\mathbf{E} [\xi_{t,t+1} V_{k,(t+1)\Delta}] = 0$, then for $h \in \{1, 2, \dots\}$, the unconditional covariance between lagged realized variance is*

$$\text{Cov} [RV_{t,t+1}, RV_{t+h,t+h+1}] = \sum_k \alpha_k^{h-1} a_k^2 \frac{\theta_k \sigma_k^2}{2\kappa_k}.$$

Proof. From Equation (4.9) and Lemma E.7.9, we have

$$\begin{aligned}
&\text{Cov} [RV_{t,t+1}, RV_{t+h,t+h+1}] \\
&= \text{Cov} [\mathcal{V}_{t,t+1} + \xi_{t,t+1}, \mathcal{V}_{t+h,t+h+1} + \xi_{t+h,t+h+1}] \\
&= \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] + \mathbf{E} [\xi_{t,t+1} \mathcal{V}_{t+h,t+h+1}] \\
&+ \mathbf{E} [\mathcal{V}_{t,t+1} \xi_{t+h,t+h+1}] + \mathbf{E} [\xi_{t,t+1} \xi_{t+h,t+h+1}]
\end{aligned}$$

$$\begin{aligned}
 &= \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] + \text{E} \left[\xi_{t,t+1} \left(\sum_k \alpha_k^{h-1} a_k V_{k,(t+1)\Delta} \right) \right] \\
 &+ \text{E} [\mathcal{V}_{t,t+1} \text{E} [\xi_{t+h,t+h+1} | \mathcal{F}_{(t+h)\Delta}]] + \text{E} [\xi_{t,t+1} \text{E} [\xi_{t+h,t+h+1} | \mathcal{F}_{(t+h)\Delta}]] \\
 &= \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] \\
 &= \sum_k \alpha_k^{h-1} a_k^2 \frac{\theta_k \sigma_k^2}{2\kappa_k}. \quad \square
 \end{aligned}$$

Lemma E.7.11. *For $n \in \{1, 2, \dots\}$, the unconditional covariance recursion between lagged integrated variance for a model with two variances is*

$$\begin{aligned}
 \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1+2n,t+2+2n}] &= (\alpha_1^n + \alpha_2^n) \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1+n,t+2+n}] \\
 &- (\alpha_1^n \alpha_2^n) \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1,t+2}].
 \end{aligned}$$

Proof. With Lemma E.7.9,

$$\text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] = \alpha_1^{h-1} a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} + \alpha_2^{h-1} a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2}.$$

Therefore,

$$\begin{aligned}
 \text{Cov} [\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1+2n,t+2+2n}] &= \alpha_1^{2n} a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} + \alpha_2^{2n} a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2} \\
 &= (\alpha_1^n + \alpha_2^n) \underbrace{\left(\alpha_1^n a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} + \alpha_2^n a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2} \right)}_{\text{Cov}[\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1+n,t+2+n}]} - (\alpha_1^{2n} \alpha_2^{2n}) \underbrace{\left(a_1^2 \frac{\theta_1 \sigma_1^2}{2\kappa_1} + a_2^2 \frac{\theta_2 \sigma_2^2}{2\kappa_2} \right)}_{\text{Cov}[\mathcal{V}_{t,t+1}, \mathcal{V}_{t+1,t+2}]} . \quad \square
 \end{aligned}$$

Corollary E.7.12. *If $\text{E} [\xi_{t,t+1} V_{k,(t+1)\Delta}] = 0$, then for $n \in \{1, 2, \dots\}$, the unconditional covariance recursion between lagged realized variance for a model with two variances is*

$$\begin{aligned}
 &\text{Cov} [RV_{t,t+1}, RV_{t+1+2n,t+2+2n}] \\
 &= (\alpha_1^n + \alpha_2^n) \text{Cov} [RV_{t,t+1}, RV_{t+1+n,t+2+n}] - (\alpha_1^n \alpha_2^n) \text{Cov} [RV_{t,t+1}, RV_{t+1,t+2}].
 \end{aligned}$$

Proof. Direct application of Lemma E.7.11, Corollary E.7.10 and Equation (4.9). □

Lemma E.7.13. *The unconditional covariance between the integrated variance and return is*

$$\text{Cov} [\mathcal{V}_{t,t+1}, R_{t,t+1}] = \sum_k \left(\lambda_k - \frac{1}{2} \right) (a_k^2 + f(\Delta, \kappa_k)) \frac{\sigma_k^2 \theta_k}{2\kappa_k} + \rho_k \frac{\sigma_k}{\kappa_k} \theta_k (\Delta - a_k).$$

Proof. From Corollary E.7.6

$$\begin{aligned} & \text{Cov} [R_{t,t+1}, \mathcal{V}_{t,t+1}] \\ &= \text{Cov} [\text{E} [R_{t,t+1} | \mathcal{F}_{t\Delta}], \text{E} [\mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]] + \text{E} [\text{Cov} [R_{t,t+1}, \mathcal{V}_{t,t+1} | \mathcal{F}_{t\Delta}]] \\ &= \sum_k \text{Cov} \left[\left(\lambda_k - \frac{1}{2} \right) \text{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}], \text{E} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] \right] \\ &+ \sum_k \text{E} \left[\text{Cov} \left[\left(\lambda_k - \frac{1}{2} \right) \mathcal{V}_{k,t,t+1} + \rho_k \mathcal{W}_{k,t,t+1}, \mathcal{V}_{k,t,t+1} \middle| \mathcal{F}_{t\Delta} \right] \right] \\ &= \sum_k \left(\lambda_k - \frac{1}{2} \right) a_k^2 \text{Var} [V_{k,t\Delta}] \\ &+ \sum_k \left(\lambda_k - \frac{1}{2} \right) \text{E} [\text{Var} [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] + \rho_k \text{E} [\text{Cov} [\mathcal{V}_{k,t,t+1}, \mathcal{W}_{k,t,t+1} | \mathcal{F}_{t\Delta}]] \\ &= \sum_k \left(\lambda_k - \frac{1}{2} \right) (a_k^2 + f(\Delta, \kappa_k)) \frac{\sigma_k^2 \theta_k}{2\kappa_k} + \rho_k \frac{\sigma_k}{\kappa_k} \theta_k (\Delta - a_k), \end{aligned}$$

leading to the result. □

Corollary E.7.14. *If $\text{E} [R_{t,t+1} \xi_{t,t+1}] = 0$, then the unconditional covariance between the realized variance and return is*

$$\text{Cov} [RV_{t,t+1}, R_{t,t+1}] = \sum_k \left(\lambda_k - \frac{1}{2} \right) (a_k^2 + f(\Delta, \kappa_k)) \frac{\sigma_k^2 \theta_k}{2\kappa_k} + \rho_k \frac{\sigma_k}{\kappa_k} \theta_k (\Delta - a_k).$$

Proof. We have

$$\begin{aligned} \text{Cov} [RV_{t,t+1}, R_{t,t+1}] &= \text{Cov} [R_{t,t+1}, \mathcal{V}_{t,t+1} + \xi_{t,t+1}] \\ &= \text{Cov} [R_{t,t+1}, \mathcal{V}_{t,t+1}] + \underbrace{\text{E} [R_{t,t+1} \xi_{t,t+1}]}_{=0}, \end{aligned}$$

which leads to the result with Lemma E.7.13. \square

Lemma E.7.15. *The unconditional covariance between the lagged integrated variance and return for $h \in \{1, 2, \dots\}$ is*

$$\text{Cov} [\mathcal{V}_{t+h,t+h+1}, R_{t,t+1}] = \sum_k \alpha_k^{h-1} a_k^2 \left(\left(\lambda_k - \frac{1}{2} \right) \frac{\theta_k \sigma_k^2}{2\kappa_k} + \rho_k \sigma_k \theta_k \right).$$

Proof. From Corollary E.7.6

$$\begin{aligned} & \text{Cov} [R_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] \\ &= \sum_k \text{Cov} \left[\left(\lambda_k - \frac{1}{2} \right) \mathcal{V}_{k,t,t+1} + \rho_k \mathcal{W}_{k,t,t+1}, \mathcal{V}_{k,t+h,t+h+1} \right] \\ &= \sum_k \left(\lambda_k - \frac{1}{2} \right) \text{Cov} [\mathcal{V}_{k,t,t+1}, \mathcal{V}_{k,t+h,t+h+1}] + \rho_k \text{Cov} [\mathcal{W}_{k,t,t+1}, \mathcal{V}_{k,t+h,t+h+1}] \\ &= \sum_k \left(\lambda_k - \frac{1}{2} \right) \text{Cov} [\mathcal{V}_{k,t,t+1}, \alpha_k^{h-1} a_k V_{(t+1)\Delta}] + \rho_k \text{Cov} [\mathcal{W}_{k,t,t+1}, \alpha_k^{h-1} a_k V_{(t+1)\Delta}] \\ &= \sum_k \left(\lambda_k - \frac{1}{2} \right) \alpha_k^{h-1} a_k \left(\begin{array}{l} \text{Cov} [\text{E} [\mathcal{V}_{k,t,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\ + \text{E} [\text{Cov} [\mathcal{V}_{k,t,(t+1)\Delta}, V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \end{array} \right) \\ &+ \rho_k a_k \alpha_k^{h-1} \text{E} [\text{Cov} [\mathcal{W}_{k,t,(t+1)\Delta}, V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\ &= \sum_k \alpha_k^{h-1} a_k \left(\left(\lambda_k - \frac{1}{2} \right) \frac{\theta_k \sigma_k^2}{2\kappa_k} (g(\Delta, \kappa_k) + \alpha_k a_k) + a_k \rho_k \sigma_k \theta_k \right) \\ &= \sum_k \alpha_k^{h-1} a_k^2 \left(\left(\lambda_k - \frac{1}{2} \right) \frac{\theta_k \sigma_k^2}{2\kappa_k} + \rho_k \sigma_k \theta_k \right), \end{aligned}$$

leading to the result. \square

Corollary E.7.16. *The unconditional covariance between the lagged realized variance and return for $h \in \{1, 2, \dots\}$ is*

$$\text{Cov} [RV_{t+h,t+h+1}, R_{t,t+1}] = \sum_k \alpha_k^{h-1} a_k^2 \left(\left(\lambda_k - \frac{1}{2} \right) \frac{\theta_k \sigma_k^2}{2\kappa_k} + \rho_k \sigma_k \theta_k \right).$$

Proof. We have

$$\begin{aligned}
 & \text{Cov} [RV_{t+h,t+h+1}, R_{t,t+1}] \\
 &= \text{Cov} [R_{t,t+1}, \mathcal{V}_{t+h,t+h+1} + \xi_{t+h,t+h+1}] \\
 &= \text{Cov} [R_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] + \text{E} [R_{t,t+1}\xi_{t+h,t+h+1}] \\
 &= \text{Cov} [R_{t,t+1}, \mathcal{V}_{t+h,t+h+1}] + \text{E} \left[R_{t,t+1} \underbrace{\text{E} [\xi_{t+h,t+h+1} | \mathcal{F}_{(t+h)\Delta}]}_{=0} \right],
 \end{aligned}$$

which leads to the result with Lemma E.7.15. \square

Lemma E.7.17. *The third conditional central moment of the instantaneous variance is*

$$\text{E} \left[(V_{k,(t+1)\Delta} - \text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}])^3 \middle| \mathcal{F}_{t\Delta} \right] = T_k (V_{k,t\Delta} - \theta_k) + Q_k \theta_k,$$

where

$$\begin{aligned}
 T_k &= 3\sigma_k^4 \int_{t\Delta}^{(t+1)\Delta} e^{-\kappa_k(s-t\Delta)} \frac{1 - e^{-\kappa_k(s-t\Delta)}}{\kappa_k} e^{-3\kappa_k((t+1)\Delta-s)} ds, \\
 Q_k &= 3\sigma_k^4 \int_{t\Delta}^{(t+1)\Delta} \frac{1 - e^{-2\kappa_k(s-t\Delta)}}{2\kappa_k} e^{-3\kappa_k((t+1)\Delta-s)} ds.
 \end{aligned}$$

Proof. We have

$$\begin{aligned}
 & \text{E} \left[(V_{k,(t+1)\Delta} - \text{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}])^3 \middle| \mathcal{F}_{t\Delta} \right] \\
 &= \sigma_k^3 \text{E} \left[\left(\underbrace{\int_{t\Delta}^{(t+1)\Delta} e^{-\kappa_k((t+1)\Delta-s)} \sqrt{V_{k,s}} dW_{k,s}}_{\mathcal{Z}_{k,t,t+1}} \right)^3 \middle| \mathcal{F}_{t\Delta} \right] \\
 &= 3\sigma_k^3 e^{-3\kappa_k(t+1)\Delta} \int_{t\Delta}^{(t+1)\Delta} \text{E} [V_{k,s} \mathcal{Z}_{k,t,s/\Delta} | \mathcal{F}_{t\Delta}] e^{3\kappa_k s} ds. \tag{E.9}
 \end{aligned}$$

The last step is obtained with Fubini and Itô's Lemma, as with the stochastic process

$\mathcal{Y}_{k,t,u} = \int_{t\Delta}^u e^{\kappa_k s} \sqrt{V_{k,s}} dW_{k,s}$, we have

$$d\mathcal{Y}_{k,t,u}^n = n\mathcal{Y}_{k,t,u}^{n-1} d\mathcal{Y}_{k,t,u} + \frac{n(n-1)}{2} \mathcal{Y}_{k,t,u}^{n-2} e^{2\kappa_k u} V_{k,u} du.$$

Furthermore,

$$\begin{aligned} & \mathbb{E} [V_{k,s} \mathcal{Z}_{k,t,s/\Delta} | \mathcal{F}_{t\Delta}] \\ &= \sigma_k \mathbb{E} [\mathcal{Z}_{k,t,s/\Delta}^2 | \mathcal{F}_{t\Delta}] \\ &= \sigma_k \int_{t\Delta}^s \mathbb{E} [V_{k,r} | \mathcal{F}_{t\Delta}] e^{-2\kappa_k(s-r)} dr \quad \text{Fubini} \\ &= \sigma_k \int_{t\Delta}^s (e^{-\kappa_k(r-t\Delta)} (V_{k,t\Delta} - \theta_k) + \theta_k) e^{-2\kappa_k(s-r)} dr \\ &= \sigma_k \theta_k \int_{t\Delta}^s e^{-2\kappa_k(s-r)} dr + \sigma_k (V_{k,t\Delta} - \theta_k) e^{-\kappa_k(s-t\Delta)} \int_{t\Delta}^s e^{-\kappa_k(s-r)} dr \\ &= \sigma_k \theta_k \frac{1 - e^{-2\kappa_k(s-t\Delta)}}{2\kappa_k} + \sigma_k (V_{k,t\Delta} - \theta_k) e^{-\kappa_k(s-t\Delta)} \frac{1 - e^{-\kappa_k(s-t\Delta)}}{\kappa_k}, \end{aligned}$$

and replacing in Equation (E.9) leads to the results. \square

Corollary E.7.18. *The third unconditional cumulant of the instantaneous variance is*

$$\text{Cum}_3 [V_{k,(t+1)\Delta}] = \frac{Q_k \theta_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k \alpha_k}{1 - \alpha_k^3},$$

where Q_k is defined in Lemma E.7.17.

Proof. Based on Lemma E.7.17 along with Equations (C.3) and (C.5), we have

$$\begin{aligned} \text{Cum}_3 [V_{k,(t+1)\Delta}] &= \mathbb{E} [\text{Cum}_3 [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] + \text{Cum}_3 [\mathbb{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\ &\quad + 3 \text{Var} [\mathbb{E} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}], \text{Var} [V_{k,(t+1)\Delta} | \mathcal{F}_{t\Delta}]] \\ &= Q_k \theta_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k \alpha_k + \alpha_k^3 \text{Cum}_3 [V_{k,(t+1)\Delta}] \\ \Rightarrow \text{Cum}_3 [V_{k,(t+1)\Delta}] &= \frac{Q_k \theta_k + 3 \frac{\theta_k \sigma_k^2}{2\kappa_k} C_k \alpha_k}{1 - \alpha_k^3}. \quad \square \end{aligned}$$

Lemma E.7.19. *The third conditional cumulant of the integrated variance is*

$$\text{Cum}_3 [\mathcal{V}_{k,t,t+1} | \mathcal{F}_{t\Delta}] = M_k V_{k,t\Delta} + \theta_k N_k = M_k (V_{k,t\Delta} - \theta_k) + \theta_k P_k,$$

where

$$M_k = \frac{3\sigma_k^4 (-\alpha_k^3 + 2 - 2\alpha_k^2(1 + 2\Delta\kappa_k) + \alpha_k(1 - 2\Delta\kappa_k(1 + \Delta\kappa_k)))}{2\kappa_k^5},$$

$$N_k = \frac{\sigma_k^4 (\alpha_k^3 + 6\alpha_k^2(1 + \Delta\kappa_k) + 2(-11 + 3\Delta\kappa_k) + 3\alpha_k(5 + 2\Delta\kappa_k(3 + \Delta\kappa_k)))}{2\kappa_k^5},$$

$$P_k = \frac{\sigma_k^4 (-2\alpha_k^3 + 18\alpha_k - 16 + 6\Delta\kappa_k + 12\alpha_k\Delta\kappa_k - 6\alpha_k^2\Delta\kappa_k)}{2\kappa_k^5}.$$

Proof. See Garcia et al., 2011. □

E.8 Robustness Checks

We conducted robustness tests on the moments analyzed in Tables 4.1 and E.1, focusing on two key areas: the mean over 1,000 trajectories with 4,000 time steps (Panels A) and the sample autocorrelation of the conditional moments (Panels B) derived in Section 4.4.

The results for the notable CM moments are presented in Figures E.2 and E.3, assuming constant conditional variance for the measurement error (Specification 4.14). Figure E.2 shows no bias, except for Panel A.2, where the blue bar almost falls outside the 95% confidence interval. This observation is consistent with Appendix C.5, where, if the measurement error is misspecified, its impact on the moments is negligible. Consequently, the lack of independence between instrumental variables and measurement error does not significantly affect these moments.

However, Figure E.3 reveals clear biases in the moments, particularly in Panels 2 and 3. While Panel 1 shows no bias due to the absence of instrumental variables, Panels 2 and 3 indicate that the lack of independence between the measurement error and instrumental variables leads to evident biases. When correcting the model specification, as shown in Figure E.4, the moments in Panels A are centered around zero, confirming that the condi-

tional moments are appropriately evaluated for GMM estimation.

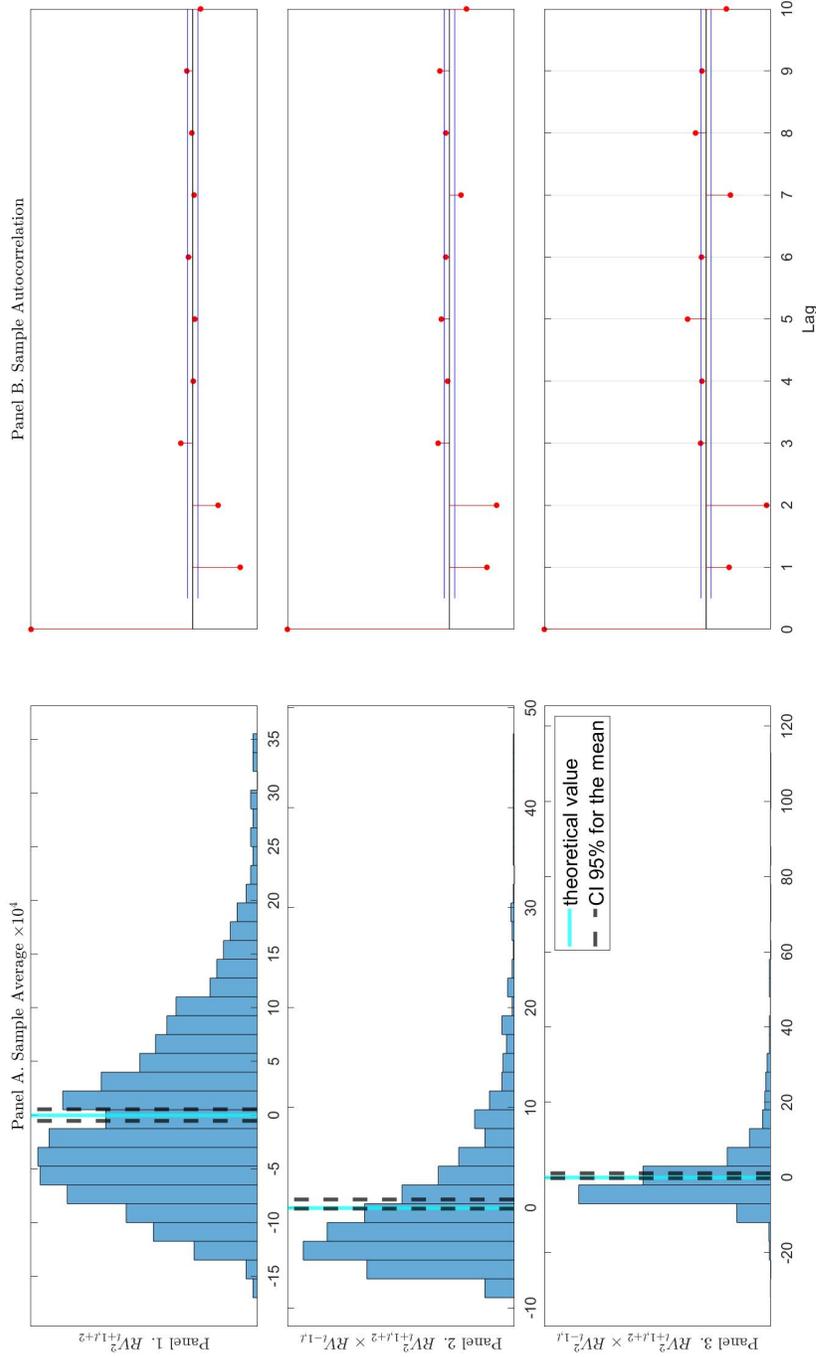
For the UM procedures, represented in Figure E.5, we also include the moment related to Realized Quarticity for completeness. All moments are well-centered around zero, indicating no issues with the UM moments.³

Additionally, the second important point in these Figures is showed in Panels B. In a GMM framework, we use two steps and the second step involves a weighting matrix based on Newey-West HAC matrix. Various papers have proposed different lags to use when dealing with realized variances. Bollerslev and Zhou, 2002 suggested a lag of 5 for one variance and 60 for two variances (CM); Todorov, 2009 suggests a lag of 80 (UM).

These figures are insightful as they reveal the autocorrelation structure in our GMM moments, validating the chosen assumptions. Figures E.2 to E.4 show some autocorrelation, but this diminishes to near zero at lag 5, primarily present at lags 1 and 2. This is expected because subtracting the conditional expectation at time t removes significant information, reducing autocorrelation, which is not observed in Figure E.5. Here, we subtract the mean over the entire trajectory, not eliminating the temporal dependence structure at each time step, leading to substantial autocorrelation in our GMM moments.

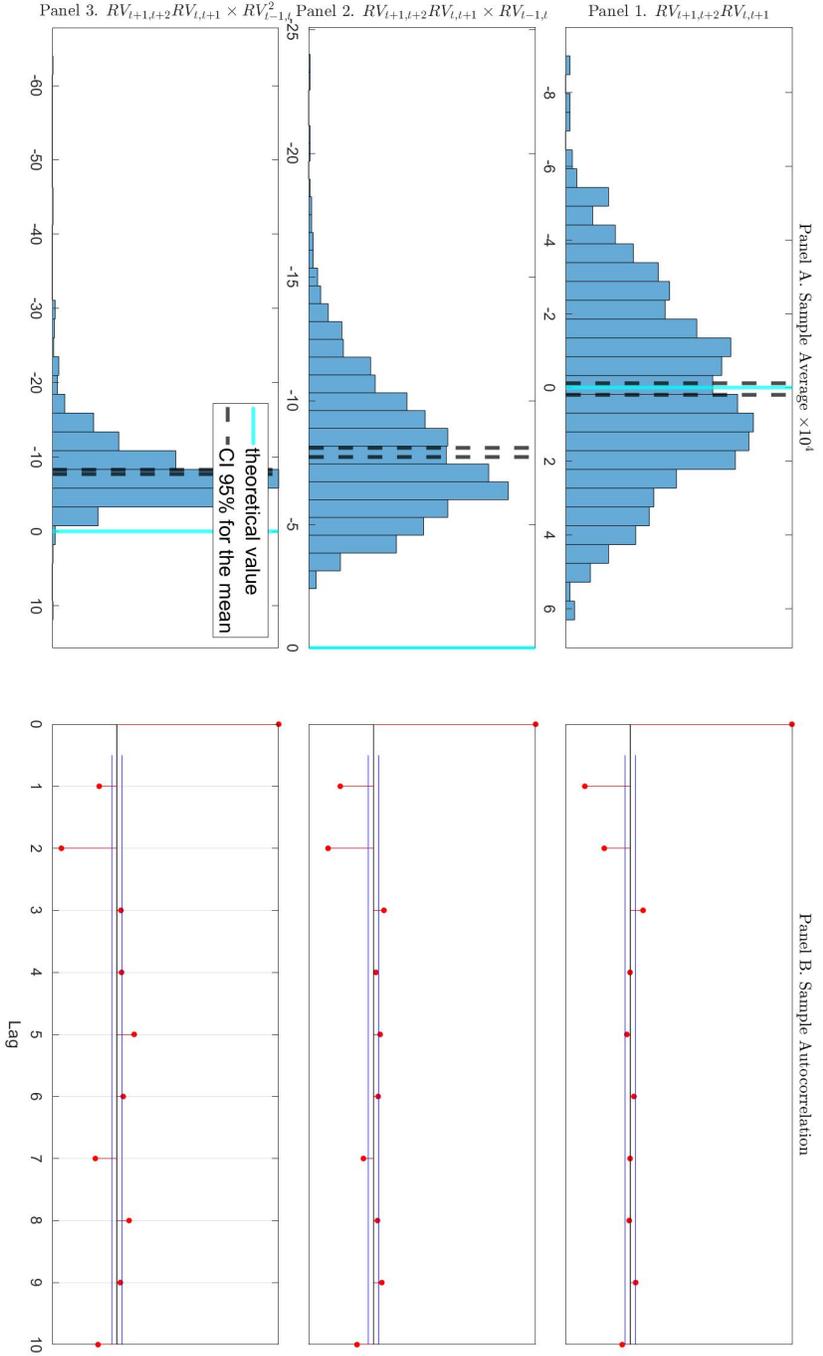
This highlights the strength of conditional methods over unconditional methods, as they produce moments with minimal autocorrelation, resulting in more robust outcomes. However, for other reasons discussed in the text, UM methods have important advantages depending on the variance nature.

³Figures related to moments condition not presented are available upon request.



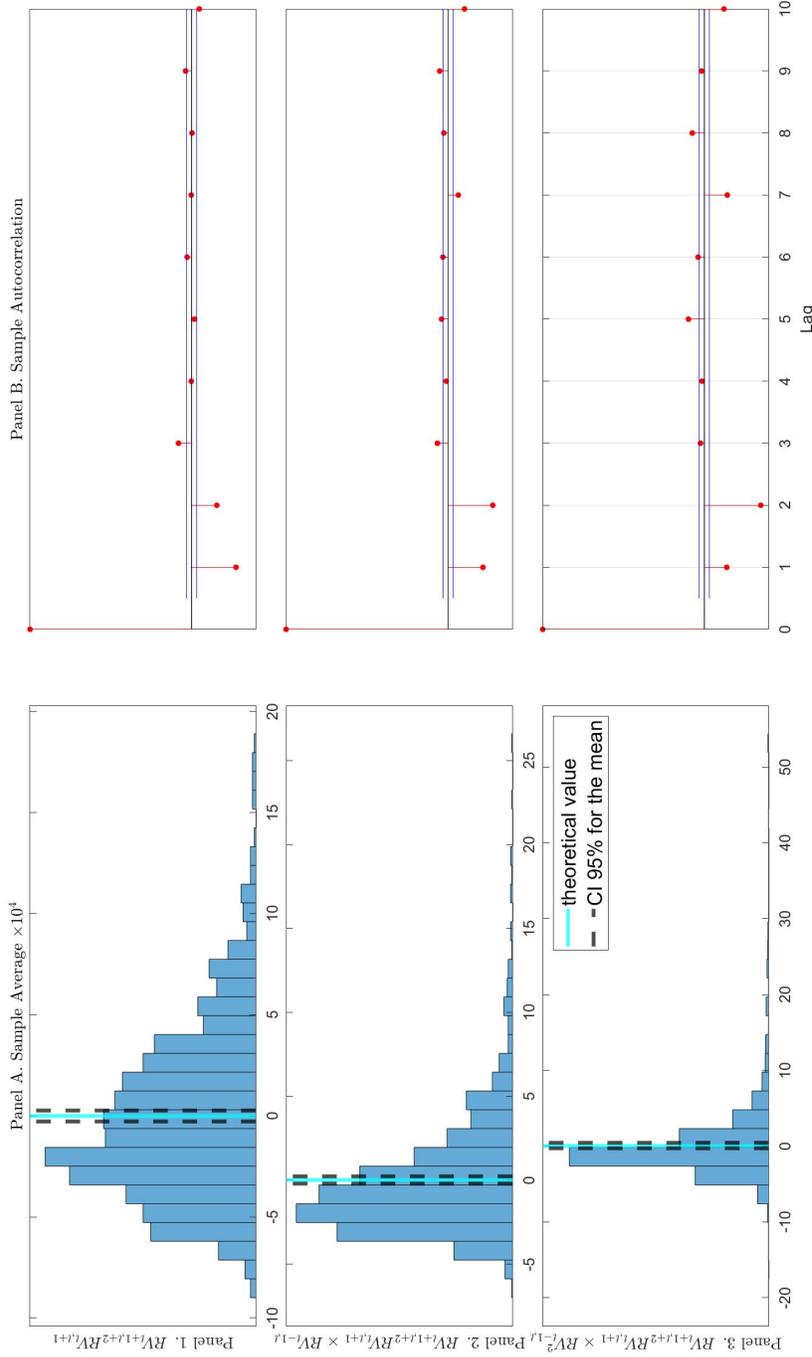
1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. The moments are detailed in either tables 4.1 or E.1.

Figure E.2: CM moments IV)-VI) with $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$



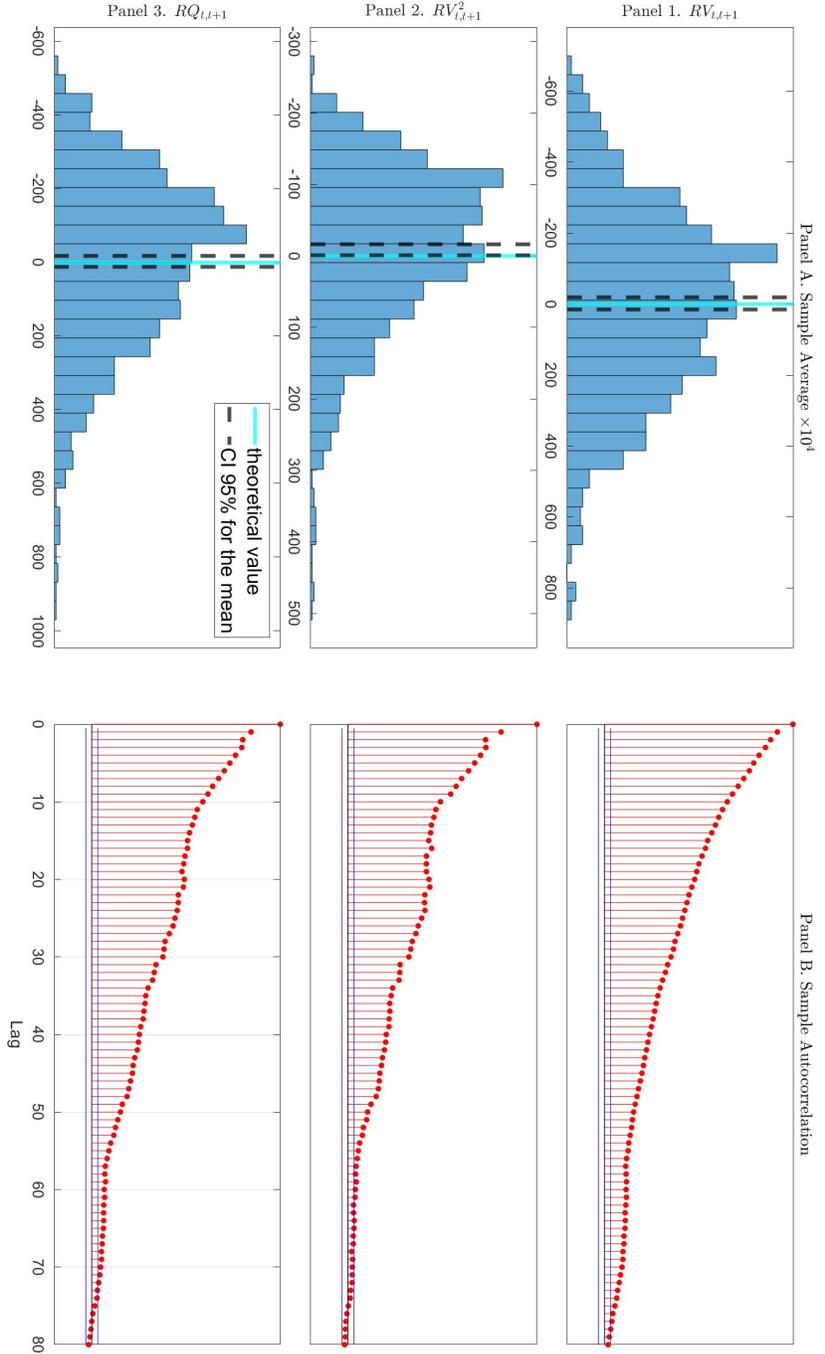
1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. The moments are detailed in either tables 4.1 or E.1.

Figure E.3: CM moments VIII-IX with $\xi_{t,t+1}^{(N)} = \gamma \epsilon_{t,t+1}^{(N)}$



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. The moments are detailed in either tables 4.1 or E.1.

Figure E.4: CM moments VII-IX) with $\xi_{t,t+1}^{(N)} = \eta \sqrt{RQ_{t,t+1}^{(N)}} \varepsilon_{t,t+1}^{(N)}$



1000 independent trajectories with $T = 4000$ daily steps of length $\Delta = 1$ are generated with parameter of scenario A $\{\kappa = 0.03, \theta = 0.25, \sigma = 0.10\}$ and one variance factor without drift and leverage effect. For each path, we simulate the realized variance with 82 intraday steps and the integrated variance with 820 intraday steps. All simulation details are provided in Section 4.3. The moments are detailed in either tables 4.1 or E.1.

Figure E.5: UM moments with $\xi_{t,t+1}^{(N)} = \eta\sqrt{RQ_{t,t+1}^{(N)}}\epsilon_{t,t+1}^{(N)}$

E.9 Background on the Generalized Methods of Moments

This Online Appendix provides an introductory exposition on the Generalized Method of Moments (GMM) framework, particularly emphasizing its utilization of instrumental variables for estimating coefficients and their associated standard errors. While initially discussed in the context of linear models, its applicability extends to nonlinear models, showcasing its versatility and robustness.

The single linear problem is formulated as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}, \quad (\text{E.10})$$

where \mathbf{y} represents the observed data and $\boldsymbol{\eta}$ denotes the error term, both with dimensions $N \times 1$. The matrix \mathbf{X} , comprising the regressors, is of size $N \times K$, while $\boldsymbol{\beta}$, our parameter vector, is $K \times 1$, where K signifies the number of coefficients awaiting estimation. Additionally, $E[\boldsymbol{\eta}|\mathbf{X}] = \mathbf{0}$ and $\text{Cov}[\boldsymbol{\eta}|\mathbf{X}] = \boldsymbol{\Omega}$. The specification is intentionally left general to encompass a complex error structure in the covariance matrix $\boldsymbol{\Omega}$.

The moment conditions for GMM are expressed as a set of equations leveraging on the orthogonality between a vector of sample moments and a vector of instrumental variables. In a general sense, for a sample of size N , the moment conditions can be written as

$$h_n(\boldsymbol{\beta}) = \mathbf{Z}_n^\top (y_n - \mathbf{X}_n\boldsymbol{\beta}) = \mathbf{Z}_n^\top \eta_n, \quad (\text{E.11})$$

where $h_n(\boldsymbol{\beta})$ characterizes the moment conditions and \mathbf{Z}_n represents the instrumental variables of dimension $N \times L$ with $L \geq K$ and $\text{Cov}[\mathbf{Z}, \boldsymbol{\eta}] = \mathbf{0}$. The moment conditions should satisfy $E[h_n(\boldsymbol{\beta})] = \mathbf{0}$ with

$$h(\hat{\boldsymbol{\beta}}) = \frac{1}{N} \sum_{n=1}^N h_n(\hat{\boldsymbol{\beta}}) = \frac{1}{N} \mathbf{Z}^\top \hat{\boldsymbol{\eta}}.$$

In a linear regression model, \mathbf{Z} is replaced with \mathbf{X} .⁴ The estimated parameters from the

⁴See Chapter 13.6 of Greene, 2017 for complete details.

moment conditions derived in Equation (E.11) are

$$\hat{\beta} = \arg_{\beta} \min h(\beta)^\top \mathbf{W} h(\beta) = [(\mathbf{X}^\top \mathbf{Z}) \mathbf{W} (\mathbf{Z}^\top \mathbf{X})]^{-1} (\mathbf{X}^\top \mathbf{Z}) \mathbf{W} (\mathbf{Z}^\top \mathbf{y}),$$

where \mathbf{W} is a weighting matrix of dimension $L \times L$. The parameters have following asymptotic property:

$$\sqrt{N} (\hat{\beta} - \beta) \rightarrow \mathcal{N} \left(0, [(\mathbf{X}^\top \mathbf{Z}) \mathbf{W} (\mathbf{Z}^\top \mathbf{X})]^{-1} (\mathbf{X}^\top \mathbf{Z}) \mathbf{W} \tilde{\Omega} \mathbf{W} (\mathbf{Z}^\top \mathbf{X}) [(\mathbf{X}^\top \mathbf{Z}) \mathbf{W} (\mathbf{Z}^\top \mathbf{X})]^{-1} \right),$$

where $\tilde{\Omega} = \mathbf{Z}^\top \Omega \mathbf{Z}$. If \mathbf{W} is chosen to be proportional to $\tilde{\Omega}^{-1}$, the asymptotic expression simplifies to

$$\sqrt{N} (\hat{\beta} - \beta) \rightarrow \mathcal{N} \left(0, [(\mathbf{X}^\top \mathbf{Z}) \tilde{\Omega}^{-1} (\mathbf{Z}^\top \mathbf{X})]^{-1} \right).$$

In practice, since Ω is typically unknown, an iterative procedure is employed. Initially, we set $\hat{\mathbf{W}}^{(0)} = (\mathbf{Z}^\top \mathbf{Z})^{-1}$ and obtain

$$\hat{\beta}^{(0)} = [(\mathbf{X}^\top \mathbf{Z}) (\mathbf{Z}^\top \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{X})]^{-1} (\mathbf{X}^\top \mathbf{Z}) (\mathbf{Z}^\top \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{y}).$$

With $\hat{\mathbf{W}}^{(1)} = (\mathbf{Z}^\top \hat{\Omega}(\hat{\eta}^{(1)}) \mathbf{Z})^{-1}$, where $\hat{\Omega}(\hat{\eta})$ is a function of the residuals, the subsequent step estimator would be

$$\hat{\beta}^{(1)} = [(\mathbf{X}^\top \mathbf{Z}) (\mathbf{Z}^\top \hat{\Omega}(\hat{\eta}^{(1)}) \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{X})]^{-1} (\mathbf{X}^\top \mathbf{Z}) (\mathbf{Z}^\top \hat{\Omega}(\hat{\eta}^{(1)}) \mathbf{Z})^{-1} (\mathbf{Z}^\top \mathbf{y}).$$

If $\mathbf{Z} = \mathbf{X}$, each step (i) estimates simplifies to

$$\hat{\beta}^{(i)} = (\mathbf{X}^\top \mathbf{X})^{-1} (\mathbf{X}^\top \mathbf{y}),$$

since the choice of the weighting matrix is inconsequential.⁵ Consequently, only a single

⁵When the GMM problem is exactly identified, there is a unique solution; therefore, the weighting matrix is irrelevant.

step is necessary, resulting in the Ordinary Least Squares (OLS) estimation.

E.10 Confidence Interval and Ellipse

This Online Appendix explains how we compute our p-value for the parameter average and the coverage ratio, along with the Sargan–Hansen J-test. After conducting simulations for $S = 1000$ independent trajectories using the procedure described in Section 4.3 with the parameter values θ^* ($1 \times p$), we obtain the observables $\mathbf{O}_{t+1} = \{RV_{t,t+1}, R_{t,t+1}, RQ_{t,t+1}\}$ at each time step. These observables, along with the GMM procedure outlined in Appendix C.6, allow us to estimate our parameters depending on the specific model used. Consequently, we obtain S sets of parameter estimates $\hat{\theta}$ ($S \times p$) corresponding to the true parameters θ^* used in the simulations.

E.10.1 Parameter average

If $\hat{\theta}_s$ ($1 \times p$) is the point estimate for the s^{th} simulated path, we define the estimator for θ^* as the average of the optimized parameters $\hat{\theta}_s$, that is

$$\bar{\theta} = \frac{1}{S} \sum_{s=1}^S \hat{\theta}_s.$$

We estimate the covariance matrix, denoted as $\hat{\Sigma}$ ($p \times p$), with

$$\hat{\Sigma} = \frac{1}{S-1} \sum_{s=1}^S (\hat{\theta}_s - \bar{\theta})^\top (\hat{\theta}_s - \bar{\theta}).$$

For S large enough, the estimator $\bar{\theta}$ tends to a multivariate normal distribution, that is

$$\sqrt{S} \hat{\Sigma}^{-1/2} (\bar{\theta} - \theta^*) \rightarrow \mathcal{N}(0, 1).$$

Thus, the quantity

$$S (\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\top \hat{\boldsymbol{\Sigma}}^{-1} (\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*),$$

follows a chi-squared distribution with p degrees of freedom, which can be used for hypothesis testing and constructing confidence intervals. The p-value is calculated with

$$\text{p-value}_{avg} = 1 - F_{\mathcal{X},p} \left(S (\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)^\top \hat{\boldsymbol{\Sigma}}^{-1} (\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) \right),$$

where $F_{\mathcal{X},p}$ is the cumulative function of the chi-squared distribution with p degrees of freedom.

E.10.2 Coverage ratio

To compute the coverage ratio on all the parameters, we individually test each simulation s to see if it rejects the chi-squared distribution at a confidence level of $\alpha = 0.05$, which is

$$\widehat{CR}_s = \left(\hat{\boldsymbol{\theta}}_s - \boldsymbol{\theta}^* \right)^\top \hat{\boldsymbol{\Sigma}}^{-1} \left(\hat{\boldsymbol{\theta}}_s - \boldsymbol{\theta}^* \right) < F_{\mathcal{X},p}^{-1}(1 - \alpha),$$

where $F_{\mathcal{X},p}^{-1}(1 - \alpha)$ represents the inverse of a chi-square distribution with probability $1 - \alpha$ and degrees of freedom p . The estimator for the overall coverage ratio $(1 - \alpha)$, denoted as \widehat{CR} , is calculated as follows

$$\widehat{CR} = \frac{1}{S} \sum_{s=1}^S \widehat{CR}_s.$$

Since we have S independent coverage ratios, we can compute a p-value for \widehat{CR}

$$\text{p-value}_{CR} = 2 \left(1 - \Phi \left(\frac{|\widehat{CR} - (1 - \alpha)|}{\sqrt{(1 - \alpha) \alpha / S}} \right) \right).$$

E.10.3 Sargan–Hansen J-test

For each trajectory and under some standard regularity conditions, we do as in Bollerslev and Zhou, 2002 and define the test⁶ defined as

$$T (\bar{\mathbf{g}}_{\hat{\theta}_s})^\top \widehat{\mathbf{W}}_{\hat{\theta}_s}^{-1} (\bar{\mathbf{g}}_{\hat{\theta}_s}) \rightarrow \chi_{m-p}^2, \quad (\text{E.12})$$

with $\bar{\mathbf{g}}_{\hat{\theta}_s} = \frac{1}{T} \sum_{t=1}^T \mathbf{g}_{t,\hat{\theta}_s}$, p as the number of parameters and m the number of moments in the GMM. $\widehat{\mathbf{W}}_{\hat{\theta}_s}^{-1}$ is the weighting matrix of the second step of the GMM for trajectory s . $\mathbf{g}_{t,\hat{\theta}_s}$ represents the vector of moment conditions at each time step t , see Section 4.4 for examples.

⁶The test is also derived in Jiang and Oomen, 2007 and Bolko et al., 2023. See Lemma 4.2 of Hansen, 1982 for additional reference.

References

- Bolko, A. E., Christensen, K., Pakkanen, M. S., & Veliyev, B. (2023). A GMM Approach to Estimate the Roughness of Stochastic Volatility. *Journal of Econometrics*, 235(2), 745–778.
- Bollerslev, T., & Zhou, H. (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics*, 109(1), 33–65.
- Brix, A. F., & Lunde, A. (2015). Prediction-based Estimating Functions for Stochastic Volatility Models with Noisy Data: Comparison with a GMM Alternative. *AStA Advances in Statistical Analysis*, 99, 433–465.
- Garcia, R., Lewis, M.-A., Pastorello, S., & Renault, É. (2011). Estimation of Objective and Risk-Neutral Distributions based on Moments of Integrated Volatility. *Journal of Econometrics*, 160(1), 22–32.
- Greene, W. (2017). *Econometric Analysis (8th ed)*. Pearson.
- Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.
- Jiang, G. J., & Oomen, R. C. (2007). Estimating Latent Variables and Jump Diffusion Models using High-Frequency Data. *Journal of Financial Econometrics*, 5(1), 1–30.
- Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3), 703–708.

Todorov, V. (2009). Estimation of Continuous-Time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics*, 148(2), 131–148.