HEC Montréal

École Affiliée à l'Université de Montréal

Trois essais sur les produits dérivés en finance

par

Pascal Létourneau

Département de Finance

Thèse présentée à HEC Montréal
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en Administration
option Finance

Mai, 2013

HEC Montréal

École Affiliée à l'Université de Montréal

Cette thèse intitulée:

Trois essais sur les produits dérivés en finance

présentée par

Pascal Létourneau

a été évaluée par un jury composé des personnes suivantes:

Tolga Cenesizoglu
(Président-rapporteur)

Lars Stentoft
(directeur de recherche)

Pascale Valéry
(co-directeur de recherche)

Christian Dorion
(membre du jury)

Gabriel Power
(membre externe)

Résumé

Cette thèse propose trois articles traitant de l'évaluation des produits dérivés en finance. Pour ce faire, nous avons trois articles différents et complémentaires. Le premier article s'intéresse à une méthode numérique permettant d'évaluer les options américaines; la méthode de Longstaff et Schwartz (2001). Plus particulièrement, nous ajoutons une technique simple afin d'imposer une structure dans notre estimation de la valeur de détention. Dans cet article, nous démontrons que l'ajout de contraintes au niveau des dérivés partielles permet de réduire de manière significative le biais d'estimation. De plus, les performances de notre technique se comparent avantageusement à celle de la technique originale.

Le deuxième chapitre s'intéresse aux facteurs de risque affectant les prix des options sur taux d'intérêt. La théorie nous indique qu'un modèle de taux d'intérêt sous la mesure risque neutre peut évaluer tous produits dérivés. Cependant, la présence de facteurs affectant uniquement un type de produit affectera l'efficacité de la couverture de ce produit. Dans cet article, nous démontrons que même après avoir contrôlé pour les effets non-linéaires de la structure par terme des taux d'intérêt, des facteurs externes sont nécessaires pour bien expliquer les variations des prix de caps. De plus, un de ces facteurs apparait indépendant.

Le troisième chapitre traite de l'utilisation des options réelles dans l'évaluation d'un projet d'investissement permettant plusieurs alternative et soumis à un risque politique. Une firme à l'option de construire une centrale électrique. La présence d'une politique "verte" sur l'émission de dioxyde de carbone (CO_2) et un risque que cette politique soit assouplie rend l'évaluation d'un tel projet complexe. Une version simplifiée du problème nous confirme qu'un risque d'assouplissement de la politique sur l'émission de CO_2 encourage les firmes à délaisser les alternatives "vertes". Un modèle plus sophistiqué est analysé à l'aide de la méthode de Least Squares Monte Carlo. Lorsque le prix des matières premières est sujet à une grande volatilité, les firmes investissement plus tôt dans le projet. Le risque d'assouplissement de la politique "verte" n'affecte pas le moment de l'investissement, mais affecte positivement les profits espérées.

Mots clés: Option Réelle, Option américaine, taux d'intérêt, risque, méthode numérique.

Summary

This thesis proposes three articles about the evaluation of financial derivatives. To do so, we have three different and complementary articles.

The first chapter aims at improving a numerical method used to price American type options; the Longstaff & Schwartz (2001) Least-Square Monte Carlo method. More specifically, we impose structure in the estimation of the holding value function. In this article, we show that adding constraints on partial derivatives in the estimation of the holding value function does reduce the pricing bias. Our technique is even more beneficial as the number of states variables is raised and the complexity of the pay-off function increases.

The second chapter address a very interesting question about unspanned risk factors in the interest rate derivative market. Interest rate derivatives have many flavors: bonds, swap, forwards, options, etc. All those derivatives share the same underlying. Theory suggests that once a model is calibrated on the dynamic of the interest rate under the risk neutral measure, all those derivatives can be priced. However, the presence of factors affecting only some derivatives, because of market particularities for instance, will affect the hedging process for those derivatives. In this article, we show that even after controlling for non-linear effects, external factors still help explain the variations in the implied volatility surface of interest rate options. Furthermore, one factor appears to be unspanned by the term structure of interest rate.

The third chapter focus on the use of real options in evaluating investment projects with multiple alternatives, under political risk. A firm has the option to build a power plant using coal or natural gas. The presence of carbon dioxide (CO_2) emission policy and a risk that it could be relaxed makes the evaluation of the value of the project complex. A simplified version of the problem confirms that the political risk reduces the effectiveness of the "green" policy. A more sophisticated model is evaluated using the Least Squares Monte Carlo method and shows that high volatility of commodity prices makes the firms invest sooner in the projects. The policy risk doesn't change the timing, but increases the expected profits.

Keywords: Real option, American option, interest rate, risk, numerical method.

Table des matières

Résumé **ii**

Summary **iv**

Liste des figures **ix**

Liste des tableaux **x**

Liste des acronymes **xi**

1 Refining the Least Squares Monte Carlo method by Imposing Structure **1**
Résumé . . . 1
1.1 Introduction . . . 2
1.2 American option price as an optimal stopping time problem . . . 6
1.2.1 The Longstaff & Schwartz algorithm . . . 7
1.2.2 Theoretical results . . . 9
1.2.3 Implementation . . . 10
1.3 Imposing structure . . . 13
1.3.1 Imposing constraints on partial derivatives . . . 13
1.3.2 Inequality constrained least squares . . . 16
1.3.3 Applying constraints in a multivariate setup . . . 19
1.4 Results . . . 22
1.4.1 Experimental setup . . . 23
1.4.2 Results in one dimension . . . 25
1.4.3 Results in three dimensions . . . 28
1.5 Conclusion . . . 32

2 Unspanned risk factors in the Cap volatility surface: a nonlinear approach. **35**
Résumé . . . 35
2.1 Introduction . . . 36
2.2 Data . . . 40

2.2.1 Cap market 41
2.2.2 The term structure of interest rates 47
2.2.3 Market factors 50
2.3 Modeling the surface 53
2.4 Results 57
2.5 Concluding remarks 63
Appendix - Extraction of the caplet volatilities 65

3 How "Animal Spirits" React to the Government Credibility Problem: A Real Option Analysis of Emission Permits Policy Risk 67
Résumé 67
3.1 Introduction 68
3.2 The real option 72
3.3 The data 75
3.4 A Net Present Value analysis 77
3.5 A simplified real option analysis using compound exchange options 79
3.6 LSM analyses 84
3.6.1 The credibility of government 84
3.6.2 Simulation cases 85
3.6.3 Numerical results and discussion 88
3.6.4 Sensitivity test 92
3.7 Conclusion 92
Appendix - Forward Curves Evolution 96

Bibliographie 99

Liste des figures

1.1 Holding value function approximations for $L = 3$ and $L = 8$ regressors. 12
1.2 Intrinsic value of an option on the arithmetic average of two assets. 20
1.3 American vanilla put option pricing using in-sample LSM method. . 26
1.4 American vanilla put option pricing using out-of-sample LSM method. 27
1.5 Put option on the arithmetic average of 3 assets using **in-sample and out-of-sample** LSM pricing method. 29
1.6 Put option on the arithmetic average of 3 assets using **in-sample and out-of-sample** LSM pricing method. 31
1.7 Put option on the arithmetic average of 3 assets using in-sample and out-of-sample LSM pricing method. 32

2.1 Typical Implied Volatility Surface 43
2.2 Surface of loadings on the 6 principal components of the IVS 45
2.3 Time series of LIBOR and SWAP rates. 48
2.4 Time series of 3 PC of the term structure of interest rates. 49
2.5 Time series of external factors. 52

3.1 The value of investment opportunities as a compound exchange option 80
3.2 Statistics of the CEO and SEO options as a function of volatility of generation cost 82
3.3 Electricity forward curve movement 98

Liste des tableaux

2.I Estimation of the lagged effect of the term structure factors on the IVS ($\hat{\Phi}_{CI}$) 59
2.II Estimation of the lagged effect of the external factors on the IVS ($\hat{\Phi}_{CE}$) 59
2.III Estimation of the lagged effect of the external factors on the term structure factors ($\hat{\Phi}_{IE}$) 61
2.IV Estimation of the lagged effect of the term structure factors on external factors ($\hat{\Phi}_{EI}$) 62
2.V Estimation of the lagged effect of the term structure factors on the IVS ($\hat{\Phi}_{CI}$) 62
2.VI Estimation of the lagged effect of the residuals of the external factors on the IVS ($\hat{\Phi}_{CE}$) 62
2.VII Estimation of the lagged effect of the residuals of the external factors on the term structure factors ($\hat{\Phi}_{IE}$) 63
2.VIII Estimation of the lagged effect of the term structure factors on residuals of the external factors ($\hat{\Phi}_{EI}$) 63

3.I Physical characteristics and capital costs 75
3.II Simple economics without the real option approach 78
3.III Inputs to LSM simulations 87
3.IV Outputs of LSM simulations 91
3.V Sensitivity test (high capital cost) 93
3.VI Sensitivity test (high emission cost) 94

Liste des acronymes

- **ARMA** auto-regressive moving average
- **ATM** at the money
- **BA** bid-ask spread
- **BSM** Black-Scholes-Merton
- **CBOE** Chicago board of option exchange
- **CCCT** water-cooled combined cycle combustion turbine
- **CEO** compound exchange option
- **CO_2** Carbon Dioxide
- **CRED** Credibility of the government
- **CV** constant volatility
- **DFT** default
- **DGS** Deuskar, Gupta & Subrahmanyam (2008)
- **GBM** geometric Borwnian motion
- **HJM** Heath, Jarrow & Morton (1992)
- **ICLS** inequality constrained least square
- **IGCC** integrated gasification combined turbine
- **ITM** in the money
- **IVS** implied volatility surface
- **LSM** least-square Monte Carlo method
- **LZ** Li & Zhao (2009)

- **ML** maximum likelihood
- **mmbtu** million British Termal Units
- **MWh** megawatt-hour
- **OLS** ordinary least square
- **OTM** out the money
- **PC** principal component
- **PCA** principal component analysis
- **PRES** Political pressure on the government
- **QTSM** quadratic term structure model
- **SEO** simple exchange option
- **VAR** vector auto-regressive
- **VARMA** vector auto-regressive moving average

Chapter 1

Refining the Least Squares Monte Carlo method by Imposing Structure

Pascal Létourneau[1] and Lars Stentoft[2]

ABSTRACT

The least squares Monte Carlo method of Longstaff and Schwartz (2001) has become a standard numerical method for option pricing with many potential risk factors. An important choice in the method is the number of regressors to use and using too few or too many regressors leads to biased results. This is so particularly when considering multiple risk factors or when simulation is computationally expensive and hence relatively few paths can be used. In this paper we show that by imposing structure in the regression problem we can improve the method by reducing the bias.

JEL classification: C15, G12, G13

Keywords: American options, bias reduction, constrained regression, simulation

1.1 Introduction

Financial derivatives markets have grown rapidly in size since the Chicago Board of Options Exchanges (CBOE) started its activities in 1973. The 1987 crash and particularly the 2007-2012 global financial crisis has changed the landscape of derivatives pricing, and market participants now realize that models should take into account more risk factors. However, when considering many risk factors, analytical solutions might not be available and if, on top of that, the option is of the American type, a numerical approach has to be considered. The numerical methods can be classified into three categories: trees, lattices, and simulation methods. The trees are very popular due to their simplicity and because they are able to incorporate the early exercise feature of American options. Lattices have also grown in popularity in the last decade and are still a subject of academic research. However, the main problem when using these two numerical methods is the exponential growth

[1]Department of Finance at HEC Montréal, 3000 Côte-Sainte-Catherine, H3T 2A7 Montréal (QC), Canada. Email: pascal.letourneau@hec.ca.

[2]Department of Finance at HEC Montréal, 3000 Côte-Sainte-Catherine, H3T 2A7 Montréal (QC), Canada, CIRANO, CIRPÉE and CREATES. Email: lars.stentoft@hec.ca.

We wish to thank Michel Denault and Pascale Valéry for constructive comments and participants of IFM2 Mathematical Finance Days 2010 Conference for useful discussions. We also thank the jury at the Sprott 7th annual PhD Symposium for relevant comments. Lars Stentoft gratefully appreciates financial support from CREATES (Center for Research in Econometric Analysis of Time Series, funded by the Danish National Research Foundation).All remaining errors are ours.

in computational complexity. In fact, because of this both trees and lattices are said to be plagued by the curse of dimensionality. Simulation methods do generally not suffer from the curse of dimensionality as the computational complexity only grows linearly in the number of stochastic factors. Thus, when the number of risk factors, i.e. the dimension of the problem, grows beyond a certain point, simulation based numerical methods are essentially the only available alternative.

At first, simulation based methods were used to price European type options and it was thought that the early exercise feature of American options could not be incorporated. Early attempts were made by Tilley (1993) and Barraquand and Martineau (1995) who used simulation to mimic the standard lattice method of determining the holding value function of the option. Carriere (1996) introduced the idea of using the information contained in all paths to approximate the holding value function by using regression techniques and a similar approach was used by Tsitsiklis and Van Roy (2001). However, it is probably fair to say that it was only with the method of Longstaff and Schwartz (2001) that the possibility of using simulation and regression methods for American option pricing became widely accepted.[1] Since then, their least squares Monte Carlo, or LSM, method has been analyzed in quite some detail and the method has proven to be very flexible when applied in various different settings. Moreover, Stentoft (2012b) showed that among the various proposed numerical methods based on simulation and regression the LSM method should be the one considered. In particular, the LSM method will have less accumulated errors and thus be less biased. Finally, as discussed in Stentoft (2012a) the regression based methods have clear advantages over sev-

[1] Other methods that use simulation are proposed in, e.g., Broadie and Glasserman (1997) and Broadie and Glasserman (2004). However, these methods either require additional subsampling or information about the transition densities to be implementable.

eral other simulation methods in terms of computational efficiency and asymptotic properties.

The LSM algorithm is of the optimal stopping time iteration kind. It uses the cross sectional information from simulated paths to approximate the holding value function by simple least squares regressions and from that deduce the optimal early exercise strategy. The quality of the estimation depends on the number of regressors, M, used in the cross sectional regression and the number of paths, N, used in the simulation (see, e.g., Moreno and Navas (2003) and Stentoft (2004a)). Only in the limit when both tend to infinity will the price estimate converge (see Stentoft (2004b)) whereas for finite choices the estimate is likely biased. The choice of the number of regressors is particularly important, and for a given number of simulated paths using too few regressors leads to low biased estimates whereas using too many regressors could lead to high biased results or even numerical problems.

In this paper we propose an innovative method for reducing the bias given the number of regressors and simulated paths. Our solution is to impose judicious constraints in the estimation of the holding value function. We show that by imposing such constraints the bias is reduced in the approximation of the holding value function. Therefore, the exercise policy will be less biased and that will reduce the low bias. Furthermore, imposing constraints also reduce the overfitting problem and thus the high bias. To illustrate our approach suppose that we want to price a put option. For this type of option, it is known that the holding value function is convex with respect to the underlying price and that the slope will be bounded by -1 and 0. Our approach takes this into account when performing the cross sectional regression in order to determine the holding value function. In order to accomplish this we need a way to impose constraints in the regression framework. Though non-parametric techniques exist, these are very costly in terms of computational burden. One of our contributions is to derive a simple linear inequality constrained least-squares estimator that is computationally much more efficient.

We compare the results from our inequality constrained least squares, or ICLS, method to the unconstrained ordinary least squares, or OLS, method originally proposed in Longstaff and Schwartz (2001) and show that the ICLS method in general has smaller bias than the OLS method. This holds across different maturities, for different categories of moneyness and for different types of option payoffs. The bias is also reduced when using the out-of-sample pricing approach of Longstaff and Schwartz (2001) which ensures low biased estimates. We also generalize our method to the multidimensional setting. Again the results show that our ICLS method generally leads to less biased estimates for a reasonable number of regressors. Moreover, whereas increasing the number of regressors in the simple OLS approach may lead to numerical problems and divergences in the price estimate, by imposing constraints in the ICLS method this is largely avoided.[2] Finally, we show that the ICLS method is often significantly more efficient than the regular OLS method.

It should be noted that several papers have proposed alternative refinements to this type of simulation methods. In particular, it has been suggested to use well known variance reduction techniques like antithetic simulation, control variates, importance sampling, as well as initial dispersion together with the LSM method (see among others Areal et al. (2008), Juneja and Kalra (2009), Lemieux and La (2005), Rasmussen (2005), and Wang and Caflisch (2010)). However, there are only

[2]As an added benefit the ICLS method generally produces estimates with smaller variance compared to the OLS method.

few alternative suggestions for how to reduce the bias of the simulated estimates of American option prices in general and of the LSM method in particular. One potentially interesting method, which could be combined with our method, is that of Kan et al. (2009), although this application is to the value function iteration method of, e.g., Carriere (1996) and Tsitsiklis and Van Roy (2001).

The remainder of the paper is structured as follows: Section 1.2 briefly describes the least squares Monte Carlo algorithm of Longstaff and Schwartz (2001) and motivates our approach. Section 1.3 provides the foundation for the constrained estimator and discusses the properties of the constrained LSM algorithm. Section 1.4 presents an extensive numerical analysis of the new constrained LSM algorithm in various contexts. Section 1.5 concludes.

1.2 American option price as an optimal stopping time problem

The problem of pricing American options involves the search for the optimal stopping time. In particular, the value of the option is given by

$$v\left(S_0, K\right) = \sup_{\tau} E\left[\rho_{0,\tau} \times v^e\left(S_\tau, K\right)\right], \tag{1.1}$$

where τ is a stopping time, S_0 and S_τ represent the asset price at time 0 and time τ respectively, $\rho_{0,\tau}$ represent the discount factor between 0 and τ, and $v^e\left(\cdot\right)$ is the exercise value of the option.[3]

[3]For an excellent introduction to this topic see Duffie (1996).

In this section, we first explain how the LSM method can be used to solve this problem. Next, we review some existing theoretical evidence on the method's convergence properties. Finally, we discuss some relevant issues related to the actual implementation of the method. These results are used to motivate our approach.

1.2.1 The Longstaff & Schwartz algorithm

In Longstaff and Schwartz (2001), the exercise strategy is determined by first approximating the holding value function and then comparing the holding value to the exercise value. Once the exercise strategy has been determined, it is then applied to the original set of simulated paths. Essentially, the LSM method proceeds according to the following three steps.

The first step in the algorithm is to simulate paths for the underlying asset using the appropriate stochastic model and to initialize the recursion. For this, one first of all needs to use as many time steps as exercise possibilities. In other words, the American option with continuous exercise must be approximated by a Bermudan option with sufficiently many exercise possibilities. Secondly, the payoff at maturity is determined. This is known with certainty and depends on the contractual terms of the option. As an example, the value of a vanilla put option on a stock with final price S_T and strike K is $v\left(S_T, K\right) = v^e\left(S_T, K\right) = \max\left[K - S_T; 0\right]$.

Next, the algorithm moves back one time step and computes the expected cash flow for each path. This is simply done by discounting the cash flows. The holding value of the option could be approximated by this discounted cash flow, but it would be a very poor (and noisy) approximation based only on one simulated path. The idea suggested by Carriere (1996) and implemented in Longstaff and Schwartz (2001) is to use the information contained in the cross section of paths to improve the approximation. This is done by a simple ordinary least squares regression where the discounted cash flows are projected on, for example, a polynomial transformation of the state variables and the following problem is solved

$$\min_{\beta} \left(Y_t - P^L\left(X_t\right) \times \beta\right)' \times \left(Y_t - P^L\left(X_t\right) \times \beta\right), \tag{1.2}$$

where Y_t is the vector of the discounted cash flows, $P^L\left(\cdot\right)$ is the polynomial operator of order L, and X_t is a matrix with the state variables.[4] The holding value can then be approximated by

$$\hat{v}^h\left(S_t, K\right) = P^L\left(X_t\right) \times \hat{\beta}, \tag{1.3}$$

where $\hat{\beta}$ is the vector of estimated coefficients of the regression in (1.2). The exercise strategy can be deduced by comparing the holding value with the exercise value. If the expected value from keeping the option alive is lower than the exercise value, it is better to exercise. The exercise strategy is determined pathwise and that is the only information kept. Proceeding one time step further backward, the cash flows are discounted for each path taking the exercise strategy into account.

The algorithm then continues with the regression and proceeds backwards until $t = 1$, the first time step. Finally, at the initial time, $t = 0$, the estimated option value is calculated as the average of the discounted pathwise payoffs from following the optimal stopping time.

[4] In fact, Longstaff and Schwartz (2001) suggest to use only in the money paths for increased efficiency and we follow this procedure in our implementations.

1.2.2 Theoretical results

One of the major benefits of the LSM method is its simplicity and the ease with which it can be adapted to price various financial products. For example, the original paper contains several examples and applications can be found in the literature to price life insurance contracts (e.g., in Bacinello et al. (2010)), real-estate derivatives (e.g., in Longstaff (2005)), real options (e.g., in Gamba (2002)), which has several applications such as gas storage, mine expansion decisions, and timber harvest contracts, and executive stock options (e.g., in León and Vaello-Sebastià (2009)). The LSM method also has clear advantages over other methods that use simulation and regression for American option pricing like, e.g., those of Carriere (1996) and Tsitsiklis and Van Roy (2001) which use the value function directly for iteration. In particular, as shown in Stentoft (2012b) the LSM method, which iterates on the stopping time, is affected only by the errors in the exercise strategy and not the holding value functions. Because of this the LSM method can be expected to be less biased than these other methods.

More importantly though, the LSM method has nice convergence properties. For example, Stentoft (2004b) showed that the LSM algorithm converges to the true price when the number of paths, N, and the number of regressors, M, tend to infinity. The convergence has to first come from the approximation of the holding value function. The polynomial regression can by viewed as a series estimator and this estimator is shown to converge under certain assumptions. One of which, assumption 9 of Newey (1997), states that the smoothness of the function is important. This is reflected in Theorem 1 of Stentoft (2004b) which essentially states that if $M \to \infty$ and $M^3/N \to 0$, the rate of convergence is of order $O\left(M/N + M^{-2s/r}\right)$, where s is the number of continuous derivatives of the holding value function and r is the number of state variables. Convergence requires $M^3/N \to 0$, or $M/N \to 0$ as $N \to \infty$ (since $M \geq 1$). It also requires $M \to \infty$ to have $M^{-2s/r} \to 0$. In other words, increasing N alone will not guarantee convergence and instead it will translate into a biased estimate, though with very low variance.

As discussed in Stentoft (2004b) the term $M^{-2s/r}$ is essentially related to the bias of the approximation and as s increases, the function gets smoother and convergence is obtained more easily. However, for low values of s and as r increases, this term becomes more important and M needs to be increased faster to ensure convergence. In the worst case scenario with $s = 0$ convergence is not obtained. More generally, this means that when the volatility of the asset is low or when estimating the holding value function close to maturity, more regressors are needed because the function to approximate is less smooth.

1.2.3 Implementation

Although theory states that a function from a separable Hilbert space can be represented as a linear combination of a countable number of basis functions, in any actual application of the LSM method one uses a finite (and often relatively limited) number of regressors and a finite number of simulated paths and this leads to biased estimates. First of all, since the holding value function is approximated with a low order polynomial two cases may occur that bias the results: the option is exercised for trajectories where it should have been held, or it is held for trajectories where it should have been exercised. Both cases result in a sub-optimal exercise strategy and a low bias on the price estimate. Secondly, since the same trajectories are used to estimate the exercise strategy and to price the option, the exercise strategy will be optimal for that particular sample of trajectories. This in-sample overfitting introduces a high bias in the price estimate. To reduce the low bias one would need to increase the number of regressors used relative to the number of simulated paths and to reduce the high bias one would need to increase the number of simulated paths relative to the number of regressors used.

Based on the above discussion one approach is to increase the number of regressors to obtain convergence. In fact, Longstaff and Schwartz (2001) argue that one should increase the number of regressors until the price estimate starts to decrease. However, this may in fact provide a very poor estimate of the true price for two reasons. First of all, a large M can lead to numerical problems and the decrease in price could be caused by these problems. Secondly, though increasing M will increase the flexibility of the estimator, this will increase the in-sample overfitting and thus increase the high bias. The resulting effect of increasing M is hence in general unknown.[5] More generally, it might not always be possible to reduce both biases to a satisfactory level as this may lead to numerical problems as well as problems with memory management. This is particularly important when the dimension of the problem increases and if the smoothness of the holding value function is low so that many regressors are needed.

In Figure 1.1 we illustrate the potential issues with a very simple example: a put option with one early exercise and one year maturity in a Black-Scholes-Merton, or BSM, world. The figure shows first of all the intrinsic value, the solid black line, and the theoretical holding value, the solid green line, which in this setting

[5]The argument in Longstaff and Schwartz (2001) is based on their Proposition 1. However, as pointed out in Stentoft (2004b) this proposition does not hold when the coefficients are estimated and hence it is of little practical value.

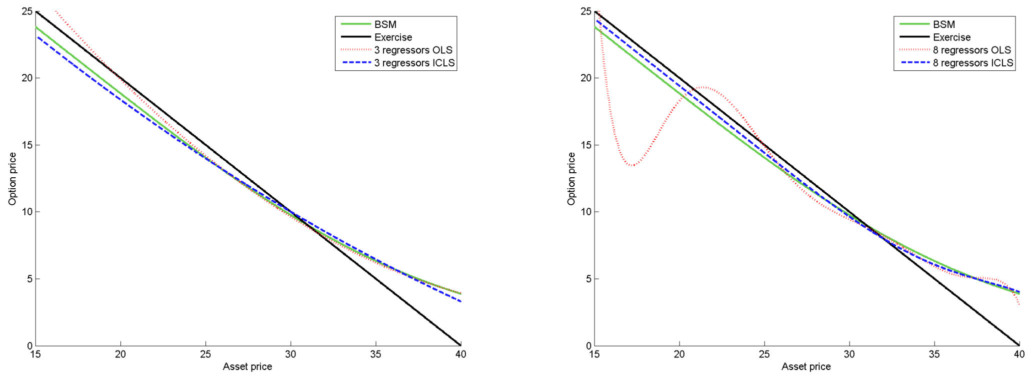

Figure 1.1: **Holding value function approximations for $L = 3$ and $L = 8$ regressors.**
Approximate holding value function with polynomial of order 2, left hand panel, and of order 7, right hand panel. The option matures in one year and has one early exercise in six months. The option characteristics are $S_0 = 40$, $K = 40$, $r = 6\%$, $\sigma = 40\%$ and S follows a GBM. Results are based on 1000 simulated paths.

equals the BSM European price of an option with half a year to maturity. The figure also shows the approximated holding value function obtained with different number of polynomials. In the left hand plot we use a polynomial of order 2 and in the right hand plot we use a polynomial of order 7. The results are shown with dotted lines labeled OLS. The left hand plot shows that with very few regressors the fit is poor in the tails because of the very constrained polynomial specification. In fact, because of this poor fit deep in the money paths do not lead to early exercise though this is clearly optimal. The right hand plot shows that with many regressors the fit is again poor, though this time due to overfitting.

The dashed lines labeled ICLS shows the approximate holding value function, again obtained with a polynomial of order 2 and 7, respectively, when imposing some simple constraints.[6] The figure shows that adding structure to the regression problem increases the fit in both cases. In particular, the left hand plot shows that once constraints are imposed the approximate holding value function no longer

[6] The actual constraints we impose are discussed in the next section.

leads to erroneous decisions for the deep in the money paths. The right hand plot shows that once constraints are imposed the poor fit is much improved as the function is now convex over the domain of interest.

1.3 Imposing structure

In Section 1.2 we explained how the LSM method can be used to price American options. We also discussed how the choice of regressors influences the performance of this method in real applications. In particular, when relatively few regressors are used this leads to underfitting of the approximate holding value function and when using relatively many regressors it leads to overfitting. Finally, we provided some evidence that these issues can be mitigated by imposing structure on the problem.

In this section we first discuss how constraints can be imposed based on partial derivatives. We then explain how this can be combined with series estimators in a simple linear regression framework and we discuss how to build the regression under constraints using a practical example. Finally, we discuss how the framework can be generalized to the multivariate case.

1.3.1 Imposing constraints on partial derivatives

Let us first introduce some notation and explain how to impose constraints on partial derivatives. Since we can impose convexity or monotonicity of a function by imposing constraints on partial derivatives, this method is a natural fit. To illustrate the method, suppose that an estimated function, $\hat{f}(X)$, exist and this is the function on which restrictions are to be imposed. The method used to imposed constraints on partial derivatives is related to Beresteanu (2007) who proposes a non-parametric estimator to which constraints can be imposed on partial derivatives over a grid of values. In this subsection, we use a multi-step procedure only to illustrate the method. In the next subsection we will present our simple one-step estimator.

The first step consists in estimating the function $\hat{f}(X)$. Then, we need to define a grid over which the regression function $\hat{f}(X)$ will be evaluated.

Definition 1. We define $\Gamma_g = (\gamma_1, \gamma_2, ..., \gamma_g)'$ to be a univariate grid on $\{\gamma_1 : \gamma_g\}$, containing g elements. The grid will be equidistant if $\gamma_i - \gamma_{i-1} = \gamma_j - \gamma_{j-1}$, $\forall i, j \in [2, g]$.

Definition 2. We define a grid on $\{\gamma_1 : \gamma_g\}$ in multiple dimensions by

$$\Gamma_{g\times r} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1g} \\ \gamma_{21} & \cdots & & \gamma_{21} \\ \vdots & \vdots & & \vdots \\ \gamma_{r1} & \cdots & & \gamma_{r1} \end{pmatrix}',$$

The grid is defined on $\{\gamma_{11} : \gamma_{1g}\}$ for the fixed values γ_{21} to γ_{r1} in the dimensions 2 to r. Equivalent grids can be built in other dimensions. The grid will be equidistant if $\gamma_{1i} - \gamma_{1(i-1)} = \gamma_{1j} - \gamma_{1(j-1)}$, $\forall i, j \in [2, g]$.

The second step consist in evaluating the regression function over a grid, $\hat{f}(\Gamma_g)$. A third step is required to impose constraints on discrete partial derivatives.

Definition 3. We define the first difference over a grid as

$$\left[\hat{f}(\gamma_{i+1}) - \hat{f}(\gamma_i)\right],$$

where $\hat{f}(\cdot)$ is the estimated function over the points of the grid.

The slope can be computed from the first difference by dividing it with the distance between the points in the grid.

Definition 4. We define the second difference over a grid as

$$\left[\left(\hat{f}(\gamma_{i+2}) - \hat{f}(\gamma_{i+1})\right) - \left(\hat{f}(\gamma_{i+1}) - \hat{f}(\gamma_i)\right)\right] = \left[\hat{f}(\gamma_{i+2}) - 2\hat{f}(\gamma_{i+1}) + \hat{f}(\gamma_i)\right],$$

where $\hat{f}(\cdot)$ is again the estimated function over the points of the grid.

One can verify if the function is discretely convex over three points of the grid by checking if the second difference is positive. The convexity can be computed from the second difference by dividing it with the square of the distance between the points in the grid.

The use of a grid enables us to use a differentiation matrix.

Definition 5. We define the differentiation matrix D_k as a $k \times (k+1)$ matrix given by

$$D_k = \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & & & \\ 0 & 0 & \ddots & \ddots & & \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

To get the first differences of a function evaluated over a grid, one simply applies the matrix multiplication: $D_{g-1} \times \hat{f}(\Gamma_g)$, where the differentiation matrix is adapted to the size of the grid. The second difference is obtained by successively applying two differentiation matrices: $D_{g-2} \times D_{g-1} \times \hat{f}(\Gamma_g)$.

Now, to apply the constraints to the estimated function in the univariate case, the problem is posed as

$$\min_{G} \left\| \hat{f}(\Gamma_g) - G \right\|$$
$$s.t. A \times G \geq 0,$$

where $\hat{f}(\Gamma_g)$ is a vector of the estimated function evaluated over the grid, G is a vector of the size g, A is a matrix with the finite difference constraints, and hence $A \times G$ is a vector with as many elements as constraints. Constraints are applied element by element. Without the constraints, the solution is trivial and $G^* = \hat{f}(\Gamma_g)$. Now, suppose we want to impose strict monotonicity over a grid of g elements, we let $A = D_{g-1}$. The solution of this minimization, provided it exists, is a new estimator G^* that respects the constraints.

1.3.2 Inequality constrained least squares

The estimator in the previous section has two particularities. First, it can be combined to a non-parametric estimator. Second, it proceeds in multiple steps. Using a non-parametric and non-linear regression could be implemented in a LSM algorithm, but it would be inoperable in terms of the required computer time since a regression has to be performed at every time step and it would be very slow. This is even more a problem in a multivariate context where non-parametric methods, which suffer from the curse of dimensionality, are problematic to implement. Instead, we propose an improved method based on series estimator using a linear regression which is much faster. Furthermore, here, we show how we can impose the constraints on the regression function in one simple step.

When using a series estimator with a polynomial transformation, the estimation problem can be represented as in (1.2). The estimated polynomial is easily evaluated over a grid Γ_g by $P^L(\Gamma_g) \times \hat{\beta}$ and the constraints can be verified using a constraint matrix A; $A \times P^L(\Gamma_g) \times \hat{\beta}$. We can estimate β and impose constraints simultaneously by defining the problem as

$$\begin{aligned} &\min_{\beta} \quad \left(Y_t - P^L(X_t) \times \beta\right)' \times \left(Y_t - P^L(X_t) \times \beta\right) \\ &s.t. \quad R \times \beta \geq b, \end{aligned} \tag{1.4}$$

where $R = A \times P^L(\Gamma_g)$ and b represent the boundaries on the constraints. The minimization problem in (1.4) is simply an inequality constrained least square, or ICLS, problem. If the constraints are not binding, the ICLS estimator reduces to an OLS estimator. If the constraints are binding, the ICLS estimator is the best linear unbiased estimator and has a truncated variance-covariance matrix.[7]

The constrained estimator in (1.4) can be estimated in a simple one step procedure and can be easily implemented using built-in functions available in modern software.[8] Alternatively, the problem can be expressed as a quadratic programming problem as

[7] See the discussion in Liew (1976)

[8] For example, in Matlab this can be done using *lsqlin*().

$$\begin{aligned} \min_{\beta} \quad & \frac{1}{2}\beta^{'} \times H \times \beta + f^{'} \times \beta \\ s.t. \quad & R \times \beta \geq b, \end{aligned} \tag{1.5}$$

where $H = \left(P^L\left(X_t\right)\right)'\left(P^L\left(X_t\right)\right)$ and $f = -\left(P^L\left(X_t\right)\right)' Y_t$. This quadratic programming problem can be solved using more efficient algorithms like the primal-dual interior-point.[9]

In (1.4) and (1.5) the regression is performed using all data available, but the constraints are imposed discretely over the grid. Note that this does not guarantee that the estimator will respect the constraints globally over the whole support. It is nevertheless a trade off we are willing to make to obtain an estimator that will converge in a reasonable time. Moreover, it is always possible to construct a finer grid such that it limits the possibility of breaking the constraints locally.[10]

Consider the pricing of a put option using the LSM method in a simple BSM model with only one state variable, that is the underlying asset price. We know that the holding value function we want to approximate is convex with a slope bounded by -1 and 0. A grid of g points, would generate $g-1$ slope constraints and $g-2$ convexity constraints. However, the convexity constraints insure the slope is monotonically increasing, thus all the slope constraints can be dropped, except at both ends. Consequently, it suffices to use $g-2$ convexity constraints and 2 slope constraints in the problem. The discrete convexity constraints are constructed over

[9]For example, one could use the collection of routines callable from Matlab provided at http://sigpromu.org/quadprog (see also Wills and Heath (2002)).

[10]More details on the choice of the grid is provided in Section 1.4.1

a grid Γ_g as

$$R_c = D_{g-2} \times D_{g-1} \times P^L \left(\Gamma_g \right),$$

and those constraints are compared to a vector B_c of size $g-2$ filled with zeros. Then, the constraints on the slopes can be constructed as

$$R_s = \begin{pmatrix} (\gamma_2 - \gamma_1) \times D_1 \times P^L \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \\ (\gamma_{g-1} - \gamma_g) \times D_1 \times P^L \begin{pmatrix} \gamma_{g-1} \\ \gamma_g \end{pmatrix} \end{pmatrix},$$

for which we use $B_s = [-1 \ 0]'$. Finally, the constraints matrix used in (1.4) is defined by the juxtaposition of R_c and R_s to be compared with the juxtaposition of B_c and B_s.

1.3.3 Applying constraints in a multivariate setup

As explained above, imposing constraints in the univariate case is straightforward when using the differentiation matrices. In the multidimensional case, applying convexity constraints is not as simple though. The problem comes from the fact that it is not possible to verify if a discretely multivariate function is globally convex (see, e.g., Yüceer (2002)).

As a practical solution, we propose to apply the convexity constraints with respect to one dimension at a time while keeping the other dimensions fixed.[11] That is, we do not verify the global convexity, but rather the convexity with respect to

[11] This not only makes the problem computationally feasible it also avoids the problem of an exponential growth in the number of constraints, the so-called curse of dimensionality.

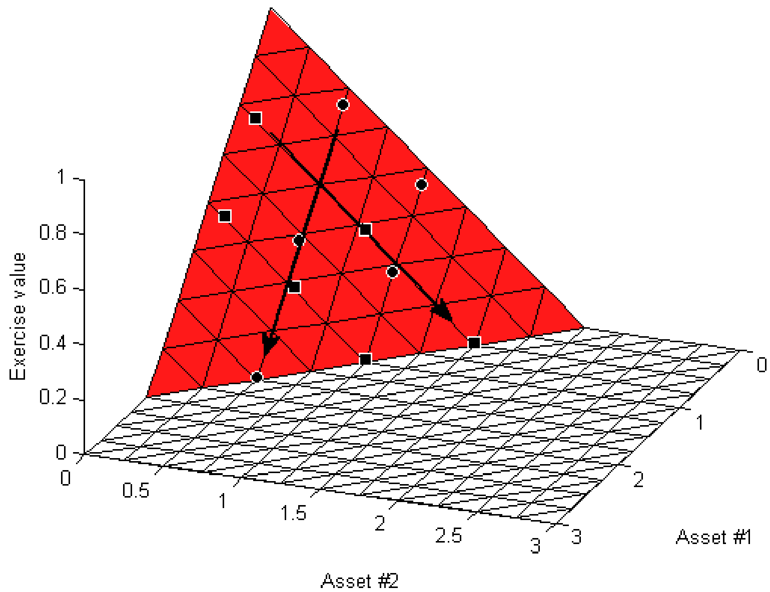

Figure 1.2: **Intrinsic value of an option on the arithmetic average of two assets.**
Intrinsic value for an option on the arithmetic average of two assets with a strike price $K = 1$. The dark area shows the regions over which constraints are imposed. The points of the grid are shown with the squares and circles, while the arrows show the directions in which the constraints are imposed.

each dimension only. This simplification enables us to impose different constraints on different dimensions. For example, it would be possible to impose convexity in one dimension and monotonicity in the other. That would be useful in the case of a vanilla put option with a stochastic volatility process.

As an example, we consider an option on the arithmetic average on two assets. This options has a payoff function that is convex with respect to both assets. Figure 1.2 shows the intrinsic value of this option. From this figure, we can observe that the in the money region is triangular. In particular, a choice has to be made in terms of the region where the constraints are to be imposed.

Figure 1.2 also shows two grids. The grid with circle (square) points consists of 2 sets of 3 points to impose convexity in the asset #1's (#2's) dimension. For each grid, we impose convexity and slope constraints in the direction of the black arrows. Consider imposing convexity constraints with respect to asset #1 for a single fixed value ($\gamma_{21} = 0.5$) of asset #2 using the following grid

$$\mathcal{G} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{21} & \gamma_{21} \end{pmatrix}',$$

convexity could be imposed using a simple constraints matrix as in

$$R_c = D_{g-2} \times D_{g-1} \times P^L(\mathcal{G}).$$

Now, by defining a bivariate grid over both sets of circle points

$$\Gamma^{\{c1\}} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} & \gamma_{15} & \gamma_{16} \\ \gamma_{21} & \gamma_{21} & \gamma_{21} & \gamma_{22} & \gamma_{22} & \gamma_{22} \end{pmatrix}',$$

the constraints are imposed with respect to asset #1 for each value of asset #2 by combining multiple constraints matrices using the Kronecker product ($\otimes$) as in

$$R_{c1} = I_2 \otimes (D_1 \times D_2) \times P^L\left(\Gamma^{\{c1\}}\right). \tag{1.6}$$

In a similar way the convexity constraints can be build in the second, or for that sake any, dimension by using an appropriately defined grid. Again this would lead to a formulation of constraints similar to that in (1.6).

The constraints on the slopes are build as in the univariate case by keeping the coordinates fixed in one dimension to impose the constraints in the other dimension.

For example, by defining the grid as

$$\Gamma^{\{s1\}} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{21} & \gamma_{22} & \gamma_{22} \end{pmatrix}',$$

one can impose the constraints of the lower bound on the slopes by using

$$R_{sl} = (I_2 \otimes D_1) \times P^L \left(\Gamma^{\{s1\}}\right). \tag{1.7}$$

The procedure is then repeated for the upper bound of the same dimension and for the upper and lower bound of the second dimension. The final constraints matrix will be composed of all constraints matrices with the corresponding limit values.

1.4 Results

In Section 1.3 we explained how to impose structure in the cross sectional regression in the LSM method. Imposing such structure should reduce the bias in the approximation, and a better approximation of the holding value function will improve the estimation of the stopping time process (or the exercise strategy). In the end, the algorithm should yield more precise price estimates.

In this section we present the results of imposing constraints in the LSM method and compare these to the unconstrained original method. First of all, we provide results in the univariate case when the number of regressors is increased. Next, we generalize these results to the multivariate setting and finally we report results in this case when the number of paths is increased.

1.4.1 Experimental setup

In order to analyze the performance of our proposed method we price a set of 9 different artificial put options in the univariate and 3 dimensional multivariate case with various moneyness and maturities. In particular, our sample includes in the money, at the money and out of the money (ITM, ATM, and OTM, respectively) options with 1, 3 and 6 months to maturity. The underlying asset follows a geometric Brownian motion with a risk free rate of 6% and volatility of 40%. The strike price is 40$ and the ITM, ATM and OTM options have an initial asset price of 36$, 40$, and 44$, respectively. In the multivariate case a correlation of zero is assumed. In all cases, we approximate the American option using a Bermudan option with one exercise possibility per trading day. We use the binomial method as the benchmark model.

The reason for considering different maturities is that we know that the holding value function is less smooth close to maturity, which in turn means that the first few regressions would need more regressors to reduce the bias. If we are approximating American options with Bermudan options with one exercise possibility per day in the estimation of a long maturity option, it is possible that the effect of the first few regressions is washed out over time. It is also possible that the errors made are amplified over time. By pricing options with maturities of 1, 3 and 6 months, we are able to test this. The reason for choosing different moneyness categories is that we know that this can have a numerical effect when pricing option using the LSM method. If the option is in the money, more simulated paths will be in the money and thus used in the cross sectional regression. The opposite is true if the option is out of the money at $t = 0$. Intuitively imposing structure is expected to be more important when fewer paths are available for regression. Thus, we should see a greater impact for the OTM options. By pricing options with different moneyness we are able to test this.

For the pricing we use 1,000 simulated paths in each simulation and we report the mean prices of 100 independent simulations. For the vanilla put option, we use polynomials of order 2 to 7.[12] For the option on the arithmetic average on three assets, we use polynomials of maximum order 2 to 5.[13] It is possible to experience numerical problems by adding more regressors, but we hope that by imposing structure, we will prevent that problem from happening. In each case, we compare the standard LSM method, denoted OLS, with the constrained LSM method, denoted ICLS for inequality constrained least square.[14]

To implement the constrained method, one needs to define the grid, select the constraints, and build the constraints matrix. In Section 1.3 we explained in details how convexity and slope constraints can be imposed for a given grid. Thus, all that is needed is to define the size of the grid and to choose the number of grid points. With respect to the size of the grid we opt for an adaptive approach which at each time step fixes the grid range from the lowest simulated path value, the path furthest in the money, to the limit of the ITM region. Though this range is adjusted at each time step before each regression the grid is kept constant when solving the

[12]Several authors have tried various basis in the LSM method (see, e.g., Moreno and Navas (2003) and Stentoft (2004a)). In general they come to the conclusion that the LSM method is robust to the choice of basis. Nevertheless, the choice of basis might have an effect on the approximation, but it is negligible compare to the choice of the number of regressors or the number of simulated paths. Note that our proposed regression method is independent of the choice of basis so it can be implemented with any preferred basis.

[13]The total number of regressors with a maximum order of at most m in r dimensions is given by $(m+r)!/(m!r!)$ (see also Feinerman and Newman (1973)).

[14]The experiment is implemented in Matlab. OLS regressions are executed using *mldivide* or Backslash operator "\", while the ICLS regressions are solved using the *qpip* function provided at http://sigpromu.org/quadprog.

optimization problem. Secondly, with respect to the number of grid points we choose to report results with a grid of 6 points.[15]

1.4.2 Results in one dimension

In Figure 1.3 we show the estimated prices for the ordinary least squares (OLS) method as well as our proposed inequality constrained least squares (ICLS) method for the 9 different options in the univariate American plain vanilla put option case. In this figure and the following ones, the red line with circles represent the LSM method using the OLS regressions, while the constrained method is represented by the blue line with squares. Each plot shows the prices obtained with polynomials of order 2 to 7. In the ICLS method we impose convexity and slope constraints on a grid of 6 points.

The figure shows that the OLS prices are always biased high compared to the true price, and this bias increases with the order of the polynomial used for approximation. This bias is a mix of the low bias from approximating the continuation value with a low dimensional polynomial and the high bias which stems from using the same paths to estimated the optimal stopping time and for pricing. When using only 1,000 simulated paths the latter of these dominates. The figure also shows that the ICLS prices are always lower than the OLS prices irrespectively of the moneyness and the maturity. In fact, when a low order polynomial is used the price obtained with the ICLS method may be low biased compared to the true price. The fact that the ICLS prices are always lower could be caused by two factors. First of all, the constrained estimator is less flexible and for a low number

[15] Similar results are, however, obtained with a grid of only 3 points.

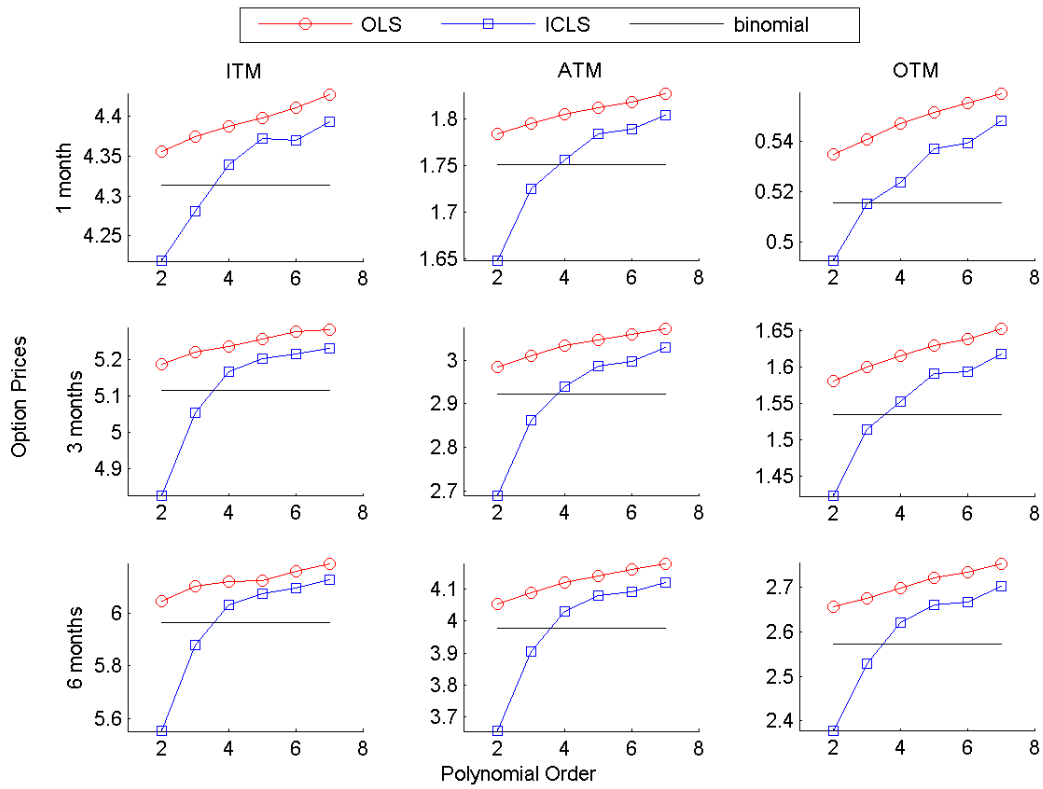

Figure 1.3: **American vanilla put option pricing using in-sample LSM method.**
ITM, ATM and OTM options are priced for maturities of 1, 3 and 6 months with daily exercise. The underlying asset follows a geometric Brownian motion with a risk free rate of 6% and volatility of 40%. The strike price is 40$ and the ITM, ATM and OTM options have an initial asset price of 36$, 40$ and 44$ respectively. All options are priced using polynomials of order 2 to 7, and the regressions are done using the paths that are ITM at the current time step. 1000 simulations are used and the mean prices of 100 repetitions are shown. The benchmark prices are obtained with the binomial model.

of regressors that can cause an increase in the low bias. Secondly, imposing constraints when the number of regressors is high will decrease the high bias due to overfitting.

The high bias stemming from overfitting can be eliminated by using a new simulation for pricing. For example, this method was suggested by Longstaff and Schwartz (2001) and referred to as "out-of-sample" pricing. In Figure 1.4 we show the estimated prices obtained with this approach. The figure shows that in this

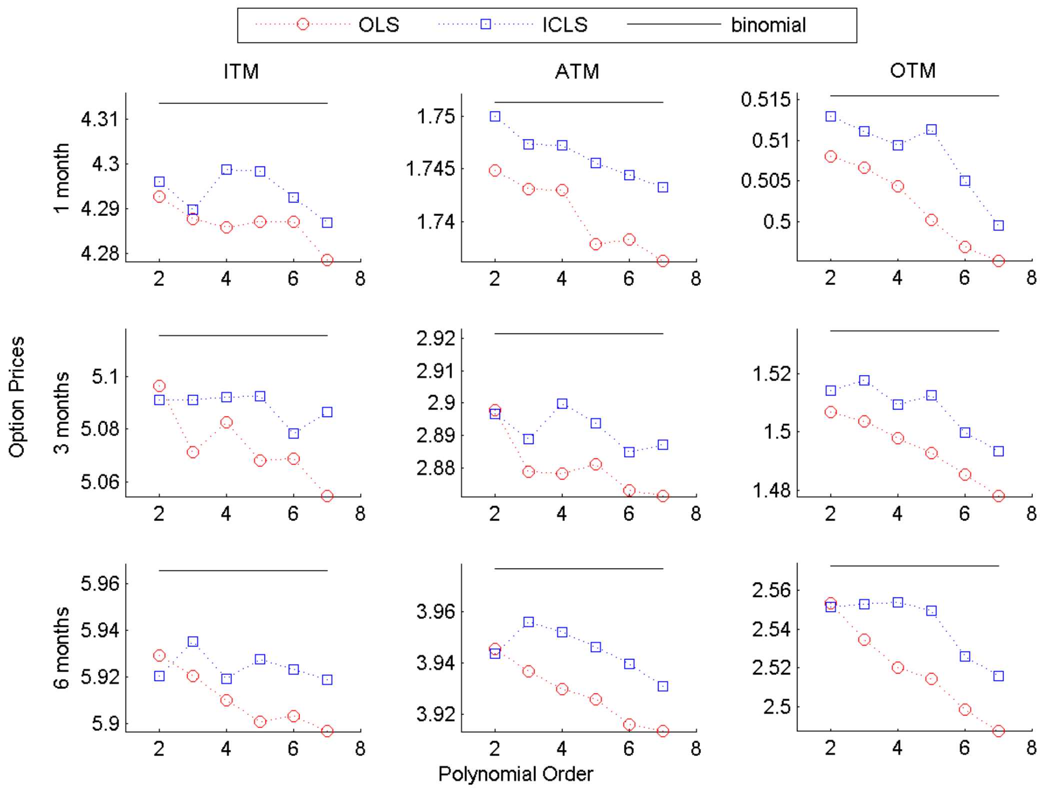

Figure 1.4: **American vanilla put option pricing using out-of-sample LSM method.**
ITM, ATM and OTM options are priced for maturities of 1, 3 and 6 months with daily exercise. The underlying asset follows a geometric Brownian motion with a risk free rate of 6% and volatility of 40%. The strike price is 40$ and the ITM, ATM and OTM options have an initial asset price of 36$, 40$ and 44$ respectively. All options are priced using polynomials of order 2 to 7, and the regressions are done using the paths that are ITM at the current time step. 1000 simulations are used and the mean prices of 100 repetitions are shown. The benchmark prices are obtained with the binomial model.

case the OLS prices are always biased low compared to the true price, and this bias increases with the order of the polynomial used for approximation. This is expected as it increases the overfitting to a particular sample. The ICLS method on the other hand generally produces prices which are higher than those from the OLS when using the out-of-sample approach.[16] Thus, imposing structure generally reduces the low bias resulting from approximating the conditional expectations.

[16]Only when the number of regressors is the lowest do we have a lower price for ICLS. This comes from the estimator being too rigid to fit the holding value, which increases the lower bias.

This is expected theoretically and illustrates the importance of adding structure to the regressions. In particular, the results show that the issue of overfitting is less important for the ICLS regression and that imposing constraints helps prevent this divergence. We conjecture that imposing further constraints would help even more.

1.4.3 Results in three dimensions

In Figure 1.5 we show the estimated prices for the ordinary least squares (OLS) method as well as our proposed inequality constrained least squares (ICLS) method for the 9 different options in the multivariate case with a put option on the arithmetic average of three underlying assets. The options have the same properties as in the univariate case and we assume a correlation of zero between the three assets. Each plot shows the prices obtained with polynomials of order 2 to 5 and we report both in-sample and out-of-sample results. In the ICLS method we impose convexity and slope constraints on a grid of 6 points.

Figure 1.5 first of all shows that in the multidimensional case the OLS method again has a high bias when the in-sample approach is used and a low bias when the out-of-sample method is used, and in both cases the bias increases with the order of the polynomial used for approximation. Secondly, the figure shows that in most cases imposing constraints leads to less biased estimates. For example, the in-sample estimates with ICLS are always lower than with the OLS method and they are generally closer to the true value. The few exceptions, for which the ICLS method produces estimates that are more biased than the OLS method are when the out-of-sample approach is used with very few polynomials, an indication that a

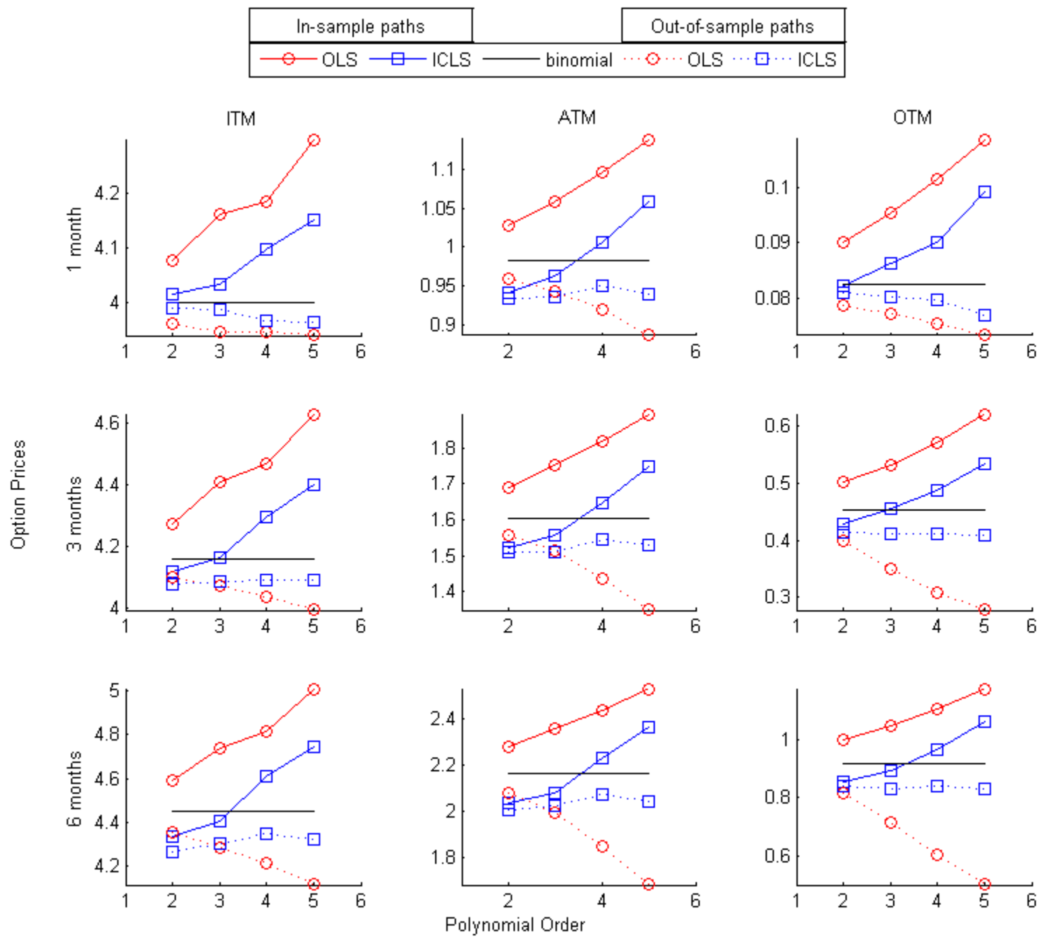

Figure 1.5: **Put option on the arithmetic average of 3 assets using in-sample and out-of-sample LSM pricing method.**
ITM, ATM and OTM options are priced for maturities of 1, 3 and 6 months with daily exercise. The underlying assets follow a geometric Brownian motion with a risk free rate of 6%, a volatility of 40% and no correlation. The strike price is 40$ and the ITM, ATM and OTM options have initial assets prices of 36$, 40$ and 44$ respectively. All options are priced using polynomials of order 2 to 5, and the regressions are done using the paths that are ITM at the current time step. 1000 simulations are used and the mean prices of 100 repetitions are shown. The benchmark prices are obtained with the binomial model. In-sample and out-of-sample pricing are represented using solid and dotted lines respectively.

polynomial of order 2 or 3 is not flexible enough to approximate the holding value properly.

Figure 1.5 also allow us to examine if the errors made in early regressions are washed away when pricing long maturity options and whether there are any differences in performance across moneyness. With respect to the maturity the figure shows that the bias for the 3 month options is about twice as large as the bias for the 1 month options, regardless of the initial conditions. This suggests that the errors made in early regressions might be amplified when pricing long term options and that it is even more important to impose structure in this case. With respect to moneyness the figure shows that there is indeed a tendency for decreasing performance of the OLS method when moving from the ITM to the OTM options. This is particularly clear for the out-of-sample results and for the longest maturity options. This issue, however, is much less prevalent for the ICLS method which does not deteriorate systematically when the option becomes out of the money when using the out-of-sample approach.

The results above show that there are clear benefits to the ICLS method when the order of the polynomial approximation and hence the number of regressors are increased for a fixed choice of N, the number of simulated paths. We now consider increasing the number of paths while holding the order of the polynomial fixed at 5. In Figure 1.6 we show the results when N is increased from $1,000$ to $10,000$, for the three dimensional options. The figure shows that the superior performance of the ICLS over the OLS which was observed with $N = 1,000$ in Figure 1.5 holds for other values of N also. In particular, irrespective of the value of N as this is increased the price estimates obtained with the ICLS method is closer to the true value, i.e. less biased, than that obtained with the OLS method. Moreover this holds when using both the in-sample and the out-of-sample approaches.

In terms of computing time, the ICLS method suffers from the additional overhead of building the constraints matrices and solving the quadratic programming problem. However, using efficient algorithms, like the primal-dual predictor-corrector, makes the ICLS method a strong competitor to the OLS method as the

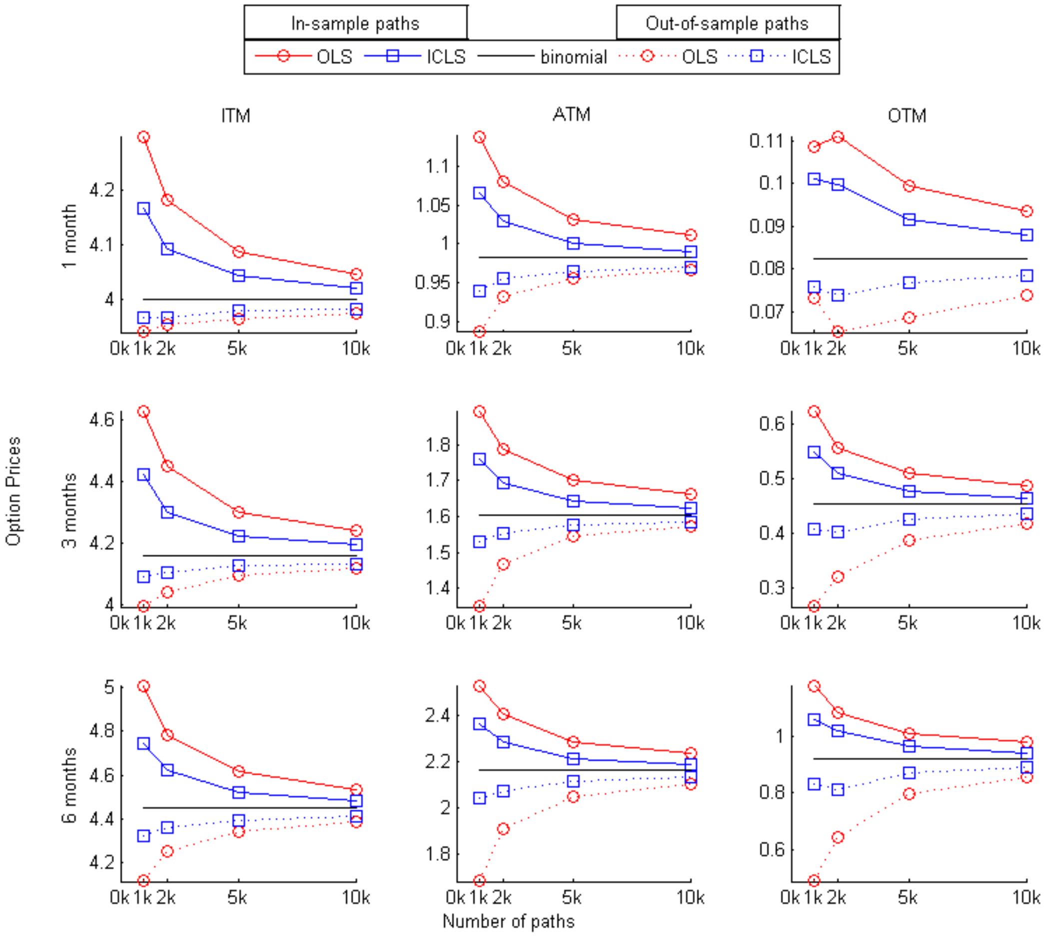

Figure 1.6: **Put option on the arithmetic average of 3 assets using in-sample and out-of-sample LSM pricing method.**
ITM, ATM and OTM options are priced for maturities of 1, 3 and 6 months with daily exercise. The underlying assets follow a geometric Brownian motion with a risk free rate of 6%, a volatility of 40% and no correlation. The strike price is 40$ and the ITM, ATM and OTM options have initial assets prices of 36$, 40$ and 44$ respectively. All options are priced using a polynomial of order 5 with increasing number of paths, N, and the regressions are done using the paths that are ITM at the current time step. The mean prices of 100 repetitions are shown. The benchmark prices are obtained with the binomial model. In-sample and out-of-sample pricing are represented using solid and dotted lines respectively.

additional computational time required for the ICLS method is compensated by the improved precision. Figure 1.7 shows the RMSE against the average time of both OLS and ICLS methods for various moneyness and maturities. The figure shows that the ICLS method, while being more time consuming, is in fact significantly more efficient than the OLS method in some cases.

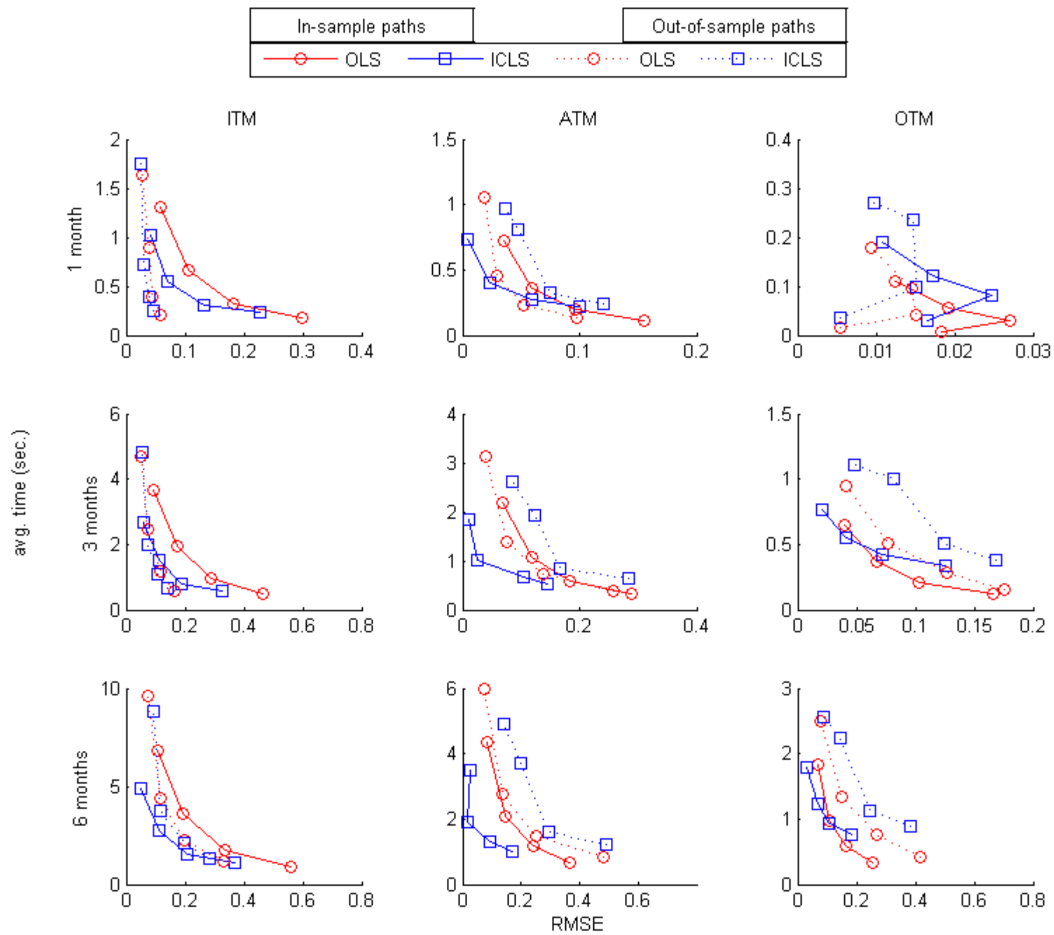

Figure 1.7: **Put option on the arithmetic average of 3 assets using in-sample and out-of-sample LSM pricing method.**
ITM, ATM and OTM options are priced for maturities of 1, 3 and 6 months with daily exercise. The underlying assets follow a geometric Brownian motion with a risk free rate of 6%, a volatility of 40% and no correlation. The strike price is 40$ and the ITM, ATM and OTM options have initial assets prices of 36$, 40$ and 44$ respectively. All options are priced using a polynomial of order 5 with increasing number of paths, N, and the regressions are done using the paths that are ITM at the current time step. The ICLS method is implemented using a grid of 3 points and constraints on slope. The RMSE error is computed using the average bias and standard deviation of 100 repetitions and the time is the average time over those 100 repetitions. In-sample and out-of-sample pricing are represented using solid and dotted lines respectively. The benchmark prices are obtained with the binomial model.

1.5 Conclusion

This paper refines the least squares Monte Carlo method of Longstaff and Schwartz (2001) which has become a standard numerical method for option pricing with many potential risk factors. An important choice in the method is the number of regressors to use and using too few or too many regressors leads to biased results. This is so particularly when considering multiple risk factors or when simulation is computationally expensive and hence relatively few paths can be used.

We propose to impose structure on the least squares problem and we show that this reduces the bias. Our approach uses an improved regression method that imposes constraints on partial derivatives in a simple linear setup. Specifically we propose an inequality constrained linear regression with a series estimator. Using our series estimator gives a flexible functional form, without the computational burden of a fully non-parametric method.

We compare the results from our inequality constrained least squares, or ICLS, method to the unconstrained ordinary least squares, or OLS, method originally proposed in Longstaff and Schwartz (2001) and show that the ICLS method in general has smaller bias than the OLS method. This holds across different maturities and for different categories of moneyness. The bias is also reduced when using the out-of-sample pricing approach of Longstaff and Schwartz (2001) which ensures low biased estimates.

We also generalize our method to the multidimensional setting. Again the results show that our ICLS method generally leads to less biased estimates for a reasonable number of regressors. Moreover, our results also show that, whereas increasing the number of regressors in the simple OLS approach may lead to numerical problems and divergences in the price estimate, by imposing constraints in the ICLS method this is largely avoided. These conclusions hold true as the number of simulated paths is increased. Finally, we show that the ICLS method is often significantly more efficient than the regular OLS method.

The method developed in this paper does not depend on the choice of basis, and examining the performance with other basis functions is an interesting area of future research. Moreover, the method can be used to price other types of options also, in particular in a multivariate context. Lastly, the ICLS method could be used with the value function iteration method of, e.g., Carriere (1996) and Tsitsiklis and Van Roy (2001) to reduce the bias of these methods.

Chapter 2

Unspanned risk factors in the Cap volatility surface: a nonlinear approach.

Pascal Létourneau[1] and Pascale Valéry[2]

ABSTRACT

Classical models for fixed income derivatives pricing are based on the principle that all derivatives of a same underlying share the same risk factors. Evidence in the literature suggests that factors unspanned by the term structure of interest rates might affect the prices of interest rate derivatives. In this paper, we contribute to the existing literature by first accounting for nonlinear effects essential to derivatives pricing. Second, we do not assume the additional factors are unspanned, but provide a testing methodology, and third, we consider the market as a whole using a VARMA model. We find evidence for the presence of one unspanned factor.

Thus, practitioners cannot identify the price of risk of one market by using a model of another.

JEL classification: G12, G13

Keywords: Derivatives, Financial Risks, Interest rates, Unspanned volatility, Volatility smile

2.1 Introduction

The fixed income option market is the most important in terms of nominal value,[1] and can be divided in two: the sell-side and the buy-side. Both, however, require different pricing models. Practitioners on the sell-side usually assume the market data as given. They compute their hedge ratios using flexible models that fit the observed option prices.[2] As a result, they must constantly recalibrate their models.

[1]Department of Finance at HEC Montréal, 3000 Côte-Sainte-Catherine, H3T 2A7 Montréal (QC), Canada. Email: pascal.letourneau@hec.ca.

[2]Department of Finance at HEC Montréal, CIRPÉE, 3000 Côte-Sainte-Catherine, H3T 2A7 Montréal (QC), Canada. Email: pascale.valery@hec.ca.

We thank Prof. Weidong Tian and XiaoZhou Zhou for commenting our paper at the EFA annual meeting and CIRPÉE's Student Day respectively. We thank Laura Coroneo, Jan Ericsson, Christian Dorion, Alexandre Jeanneret and Simon Van Norden for useful comments and suggestions. We are grateful to HEC Montréal and participants of Mathematical Finance Days of IFM[2] seminars for comments. We also thank seminar participants at Lingnan University in Hong Kong. We thank the Fond de Recherche sur la Société et la Culture (Québec) and Centre interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi (CIRPÉE) for financial support.

[1]source : Bank of International Settlement, nov.2011

[2]See, e.g., Hagan (2006), Rebonato (2007), Morini and Mercurio (2007) or Rebonato et al. (2010)

The bond portfolio managers and the primary lending banks on the buy-side estimate their models on historical quotes of interest rates or cross-section of bonds. They smooth the yield curve and can potentially exploit what their models identify as miss-pricing. Basic theory suggests that bond prices, interest rate option prices, and other interest rate products should all be integrated. Ideally, a model estimated on the historical yield curve should be able to price any interest rate product.[3] When using option prices to calibrate their model, the sell-side oftentimes violates basic theory because they arbitrarily fit time dependencies without any economic rationale.[4] However, the presence of factors not spanning different markets justifies such a procedure. The presence of an unspanned factor means that, e.g., when a primary lending bank uses a cap to hedge interest rate risk, it brings a new risk factor in its portfolio, affecting its risk measure and capital requirement. Tougher rules in the banking industry require financial institution to address such additional risk. A better understanding of the hedging market in relation with its underlying is thus very important.

Various studies focused on the relation between the yield curve and interest rate options. Litterman, Scheinkman, and Weiss (1991) show that the volatility of the interest rates is linked to curvature of the term structure, suggesting a butterfly spread be used to hedge volatility. Collin-Dufresne and Goldstein (2002) test this hedge using swaption straddle portfolios,[5] which are very sensitive to volatility, and swap rates. They find that swap rates explain only a small proportion of the variability of the portfolios. As a result, they conclude that swaps cannot be used to easily hedge volatility, therefore suggesting the presence of unspanned

[3] See, e.g., how Vasicek (1977) completes his market using a second derivative contract.

[4] See, e.g., Nawalkha (2009) or Nawalkha and Rebonato (2011)

[5] Buy an option on a buyer swap and an option on a seller swap with the same strike

factors. This is tested by Heidari and Wu (2003) using an approximate factor model of the implied volatility surface (IVS). They find that the three factors of Litterman and Scheinkman (1991) explain only about 60% of the variations of the IVS. According to their analysis, three additional factors are needed to explain the remaining variations. They thus conclude that the factors of the IVS are unspanned by the factors of the yield curve.

While the aforementioned studies suggest bonds cannot be used to hedge volatility, Fan et al. (2003) manage to hedge swaptions using bond portfolios. To do so, they model changes in the forward rates in a Heath, Jarrow, and Morton (1992) framework (HJM) using a statistical 4-factor model calibrated on bonds and derivatives. They specifically price options, and thus properly account for the strong nonlinear nature of the options. Besides an HJM specification that accounts for volatility explicitly, they use both markets, bonds and options, to calibrate their model. As a result, derivatives can be hedged using portfolios of bonds in this context.

Li and Zhao (2006), on the contrary, are not able to hedge cap/floor straddles using bonds under a 3-factor quadratic term structure model (QTSM) estimated on the historical yield curves (i.e. bond market). Their model can hedge bonds very well, explaining 95% of the variations, but poorly works when attempting to hedge cap/floor straddles. The difference between the methodologies of Fan et al. (2003) and Li and Zhao (2006) is that the former uses information from the derivatives market to estimate a HJM model, while the latter uses solely information from the bond market. Thus, factors specific to different markets are considered through the estimated parameters in Fan et al. (2003).

Collin-Dufresne, Goldstein, and Jones (2009) also address the presence of factors affecting volatility that are unspanned by the cross-section of bonds using a more flexible 4-factor quadratic model, which includes a nonlinear pricing formula for options. They show that some information necessary to hedge options can be extracted from the bond market as long as the time series variations are used and constraints are imposed to obtain an unspanned stochastic volatility model. Variations in the level of the IVS can be explained, though, they cannot explain the variations in skewness and kurtosis, which are crucial for derivatives pricing. Other papers also address the presence of unspanned factors and the possibility to hedge in the presence of such factors,[6] but very few articles focus on identifying those factors.

Among those are Deuskar, Gupta, and Subrahmanyam (2008) (DGS hereafter) who explore what those factors could be and find evidence that liquidity and default risks help explain the shape of the smile on the cap market. Li and Zhao (2009) (LZ hereafter) address the same question by considering the implied risk neutral distribution extracted from cap prices in conjunction with the mortgage market. They also find that external factors significantly affect cap prices after controlling for the yield curve.

Both of these articles limit the information contained in the yield curve to two or three factors and account only for linear effects. As shown before, such models fail when it comes to hedging. We revisit this question, and our contribution to the literature is threefold. First, we examine the nonlinear relation between the yield curve and the option prices by proposing a simple method to account for nonlinear effects based on the argument used in Longstaff and Schwartz (2001)[7]. Second, we

[6]See, e.g., Casassus et al. (2005), Driessen et al. (2009) or Gupta and Subrahmanyam (2005)

[7]A function that belongs to an Hilbert space can be approximated using a polynomial transform of the state variables

study the implied volatility surface as a whole as opposed to looking only at the smile for fixed maturities. Hence, unlike the articles aforementioned, we examine the shape of the surface in the maturity and strike dimensions. Considering the whole market at once is possible through the estimation of a vector model. The study of the whole surface in both dimensions gives more insight on factors effects and their dynamics. Third, our approach does not assume that the additional factors are unspanned by the term structure of interest rates. In fact, we propose an innovative methodology to test whether the factors are unspanned. Our findings first suggest that even after controlling for nonlinear effects of the yield curve, additional factors still help explain both the shape of the smile of implied volatilities in the strike dimension and the shape of the backbone in the maturity dimension. Second, we find that one additional factor is unspanned by the term structure of interest rates. This implies that caps are carrying more risk factors than the underlying interest rate, thus affecting the risk analysis of those using them as hedging tools.

The remaining of the paper is organized as follows. Section 2.2 describes the data and characteristics of the implied volatility surface, the yield curve and the economic factors. In Section 2.3, we introduce the model and the methodology used to estimate the effects of the different factors and test whether they are unspanned. Section 2.4 presents the results. Finally, Section 2.5 concludes.

2.2 Data

The data for this study comes from two sources, Bloomberg and Datastream, and is classified into three categories: the cap market, the term structure of interest rates, and the market factors. Once all series have been synchronized, the data spans 2004-02-25 to 2013-01-11 for a total of 2063 days of observations.

2.2.1 Cap market

The cap market is an over the counter market that quotes implied volatilities and not prices. Actual transactions are proprietary data. Bloomberg provides its users a consensus made from contributing financial institutions' quotes.[8] Quotes are available for at-the-money (ATM) caps with maturities of {0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10} years. They are also available for the following range of strikes {1, 1.5, 2, 3, 3.5, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14}%, and for the following range of maturities {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 30} years, for a total of 204 implied volatilities. There are missing quotes, especially at the beginning of the times series. For instance, up to 2005, the fixed strikes ranged from 1% to 8%. The range was gradually extended up to 14% only after 2008. Smoothing the surface is helpful to keep the time series coherent. During the smoothing phase, it is important to have information available on both sides of the ATM quotes. For some days of the time series, the ATM quote for the 1 year cap is the lowest available. Thus, there is a lack of information for the in the money (ITM) region for that maturity. For that reason, we drop the 1 year maturity caps from our analysis. Caps with a maturity of 15 or 30 years are also dropped in fear of a lack of liquidity. Measurement errors may exist in the IVS quotes, but the smoothing of the surface can help mitigate the unwanted effects.

[8] For the cap market, the historical data set contains only the mid quotes. So the information about the individual ask and bid implied volatilities is not available. In this study, we are interested in the general reaction of the market to economic factors and not the micro-market structure. Having only the mid-quotes is not a problem.

Let us now describe the implied volatility surface. Figure 2.1 shows the surface on a typical day. The curve A designates a smile exhibiting a convex and asymmetric shape. A flat line would indicate that the Black implied volatility is not a function of the strike; thus the distribution would be log-normal. The asymmetry of the smile relates to the skewness and the convexity to the kurtosis compared to the log-normal distribution. On the cap market, the smile will typically change into a smirk for longer maturities, as shown from curve B. Other patterns occur on the cap market; in fact, the surface can take many different shapes. Finally, the curve C identifies all the ATM quotes and is called the backbone. The backbone shows the shape taken by the surface in the maturity dimension. It represents the term structure of the expected volatility under the risk neutral measure for a continuum of maturities.

Smoothing of the cap implied volatility

On the cap market, the quotes are offered for a fixed range of strikes plus the ATM strikes. When we study the shape of the IVS, it is relevant to have information on both sides of the ATM quote. For example, knowing the effect of an increase in the term structure slope on ITM options is more informative than the effect on a 2% strike option. Information is available from the market, but the dispersion of quotes on each side of the ATM is irregular. The log-moneyness ratio is $\log(K/F)$, where K is the strike and F is the equivalent swap rate with the same maturity as the cap. This is standard for empirical studies of this nature.[9] A range of -0.75 to 0.75 log-moneyness ratios with 0.25 increments is used in our analysis.[10] This

[9]See for example : Pena et al. (1999), Cont and Da Fonseca (2002), Deuskar et al. (2008) or Li and Zhao (2009)

[10]$\ln(K/F) < 0 \implies ITM$, $\ln(K/F) = 0 \implies ATM$ and $\ln(K/F) > 0 \implies OTM$

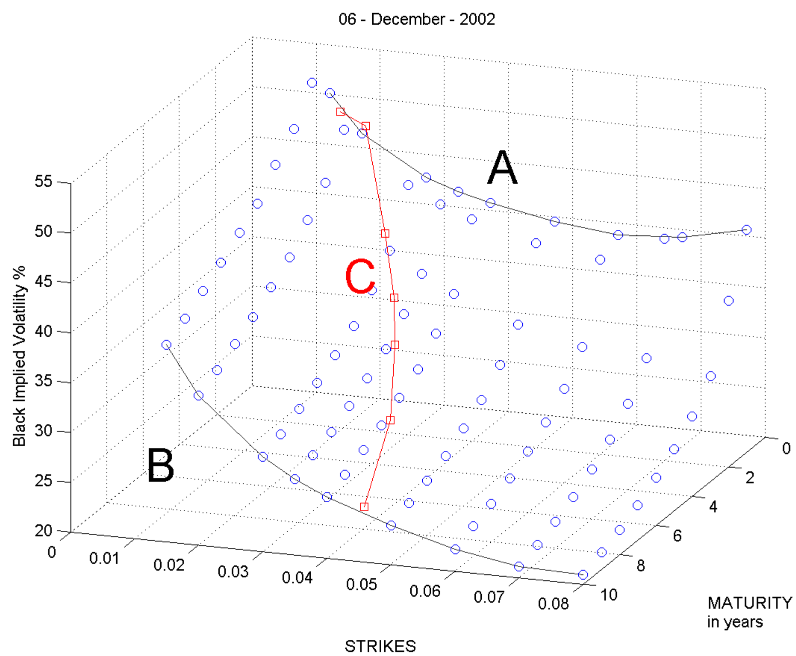

Figure 2.1: **Typical Implied Volatility Surface.**
This figure shows raw quotes of Black implied volatility on a typical day. Circles and squares represent quotes for fixed strikes and ATM respectively. A identifies the curve made by the 1 year caps of different strikes and exhibits a typical smile. The curve represented by B exhibits a smirk; usually observed at longer maturities. C identifies the curve made by ATM quotes and is called the backbone.

even spreading of the information on both sides of the ATM quotes simplifies the interpretation of the results.

Raw quotes sometime suggest the presence of arbitrage opportunities, but those might be the result of staled quotes or frictions. Arbitrage-free smoothing can help mitigate that problem. Ait-Sahalia and Duarte (2003) and Fengler (2009) propose an approach for smoothing the implied volatility smile in an arbitrage-free way. First, the options are priced using the implied volatilities, then constraints are imposed on slope and convexity of the pricing function. We adapt this approach in surface smoothing instead of curve smoothing of the smile. Effectively, we do one smoothing for the whole surface instead of doing each maturity independently. This has two advantages: first, it ensures the smoothed surface presents regularity in both dimensions. Second, smoothing in the maturity dimension permits the extraction of the caplet volatilities. Practitioners need to extract caplet implied volatilities from cap implied volatilities to calibrate their models. Basic methods are described in details in Hagan and Konikov (2004). The principle consists on smoothing in the maturity dimension for a selected strike, neglecting information from the ATM quotes and quotes of other strikes. The smoothing of the whole surface permits easy extraction of the caplet implied volatilities.[11]

The surface smoothing is executed using a bivariate kernel regression with polynomials of order 2, including cross products (see, e.g., Hardle (1990) or Bowman and Azzalini (1997)). In the kernel regression, the choice of the bandwidth is important. Low bandwidths generate less smoothing and the resulting surface is closer to raw quotes, but arbitrage opportunities are not eliminated. We select a random set of days in the times series and optimize the bandwidths in order to remove arbitrage opportunities in the strike dimension. We preclude negative caplet prices when they are extracted, while staying as close as possible to raw quotes. Once bandwidths are found for the random set, they are applied to the whole time series. This is an improvement on DGS who proceed directly to interpolation between available quotes for their log-moneyness range, because it provides better data to the analysis and mitigates the problem of errors in variables.

After the cap quotes have been smoothed over a log-moneyness range, we do not follow DGS who scale the quotes by the ATM quotes. Scaling the data removes the information related to the level of the surface. In our analysis, we keep the information on the level. As shown next, it is the principal dimension of the data. Knowing how the level of the surface is affected by variations in economic variables

[11] See Appendix for details on extracting caplet volatilities from the smoothed surface.

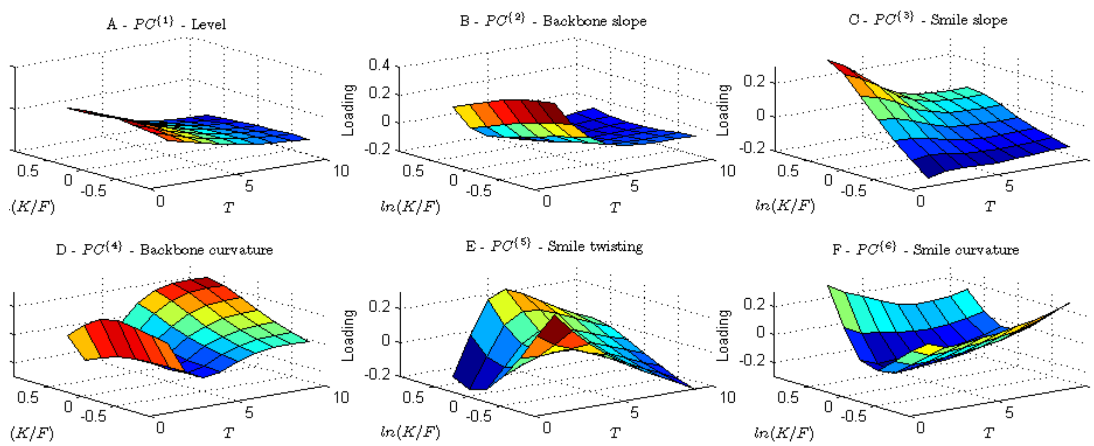

Figure 2.2: **Surface of loadings on the 6 principal components of the IVS.** A Principal component analysis is applied on the time series of the IVS using a grid of 7 strikes in log-moneyness ratios and 9 maturities. The six graphs show the loadings on the 6 principal components of the IVS. $\ln(K/F) < 0 \implies ITM$, $\ln(K/F) = 0 \implies ATM$ and $\ln(K/F) > 0 \implies OTM$. We interpret the components as follows: A-$PC^{\{1\}}$-Level, B-$PC^{\{2\}}$-Backbone slope, C-$PC^{\{3\}}$-Smile slope, D-$PC^{\{4\}}$-Backbone curvature, E-$PC^{\{5\}}$-Smile twisting and F-$PC^{\{6\}}$-Smile curvature

is important. Furthermore, according to Rogers and Tehranchi (2010) movements in the level have to be accompanied by movements in the shape also.

Describing the dynamics of the IVS

We want to understand the dynamics of the surface and visualize the principal axes on which the surface moves. In order to achieve this, a principal component analysis is performed over a whole range of points of the smoothed IVS. Then, for each component, a 3D graph of the factor loadings is provided. Looking at the first few components helps understand how the surface evolves and what are the major axes of deformation.

To proceed with the analysis, we select 63 points evenly spaced on the surface using the log-moneyness ratio scale (7 strikes and 9 maturities.)[12] Furthermore,

[12] As a robustness test, 90, 45 and 9 points were selected with similar results.

an ARMA(1,1) structure was applied to the implied volatilities to remove serial autocorrelation. A principal component analysis executed on the residuals, gives results similar to when the principal component analysis is executed directly on the points of the surface. The 6 principal components of the implied volatility surface explains 98% of the variations. Figure 2.2 show factor loadings on the first 6 components of the surface.

Intuitively, the factor loadings for different points of the surface appear to be smooth, creating a surface of factor loadings. In Figure 2.2-A we see that the first component has all positive and relatively flat factor loadings. That component is interpreted as the level of the surface. In Figure 2.2-B short maturities have positive loadings and long maturities have negative loadings. We thus relate the second component to the slope of the backbone, even though the surface of loadings is slightly convex. Figure 2.2-C shows positive loadings for out of the money caps and negative loadings for in the money caps. The third component clearly drives the slope or asymmetry of the smile for the whole surface. The fourth component is shown in Figure 2.2-D and clearly represents the backbone curvature. It also affects the skewness for the long maturities. In Figures 2.2-E, the fifth component appears a driver of the surface twisting. Finally, Figure 2.2-F the sixth component that drives the smile curvature.

We shall now select the number of components retained for the analysis. Different criteria are available, such as the Kaiser's stopping rule based on the eigenvalues, the Scree test or the cumulative explained variance. In this context, the Scree test suggests selecting only the first component while the Kaiser's stopping rule retains 11 components. However, following Heidari and Wu (2003) and Connor and Korajczyk (1993) we select the factors based on their economic significance.

In an approximate factor model, we observe the residuals after adding factors one by one. When six components are used, the residuals are generally smaller than the average bid-ask spread and no longer form a smooth surface. We argue that adding further components would only explain noise.

2.2.2 The term structure of interest rates

We need interest rates for two purposes: first, to price the caplets to check for arbitrage opportunities during the smoothing phase and second, to extract factors for the test model. Pricing caplets requires the forward rates curve. To build the term structure of interest rates and forward rates, we use daily LIBOR rates and swap rates from Bloomberg. We use all the available LIBOR rates ranging from the overnight rate to 1 year rate. As for swaps, we use quotes from 1 year to 15 years. Figure 2.3 shows the whole time series for LIBOR and swap rates. Finally, the extended Nelson-Siegel model of Svensson (1995) is used to get the term structure.This model is flexible and most importantly, does not produce negative forward rates.

Non-linear factors from the Term Structure

Due to the strong nonlinear nature of options, taking care of non-linearity is essential. First, we present a mathematical argument that justifies our approach. Then, we explain how we integrate non-linearity in the model. Finally, we explain how we select the number of factors to add in the model. Let the price of a caplet be expressed as a function of the level of the term structure, its distribution and the

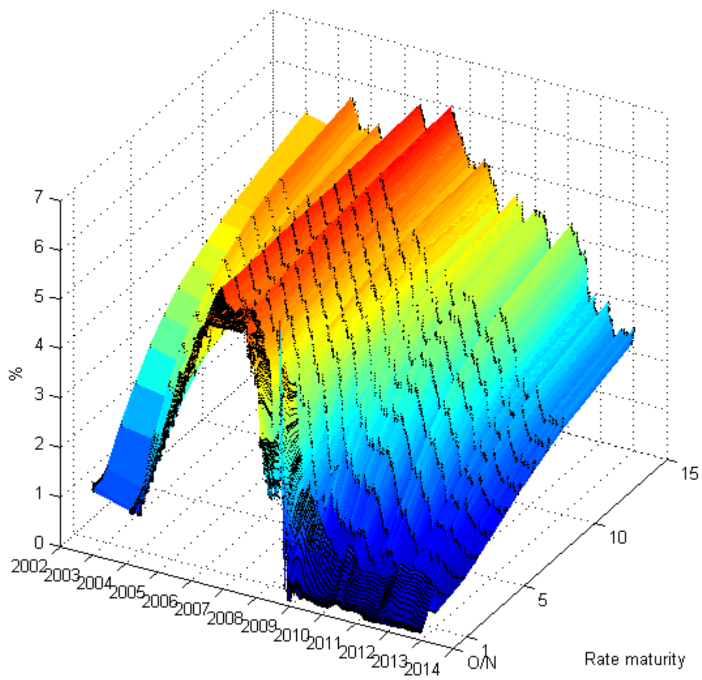

Figure 2.3: **Time series of LIBOR and SWAP rates**.
LIBOR rates have maturities ranging from overnight to 1 year and SWAP rates have maturities ranging from 1 year to 15 years. The rates shown are raw rates available from BLOOMBERG. The different quotes are joined by surfaces to allow for easy visualisation.

proper change of measure.[13] Then, consider the inverse of Black (1976)'s formula to retrieve the implied volatility. As shown below, the pay-off function is nonlinear, the distribution is nonlinear and the integration provides the price. Finally,

[13]For more details on option pricing and change of measure, see, e.g., Harrison and Kreps (1979) and Harrison and Pliska (1981). For a complete treatment on cap pricing see, e.g., Brigo and Mercurio (2006), Nawalkha et al. (2007) or Rebonato et al. (2010)

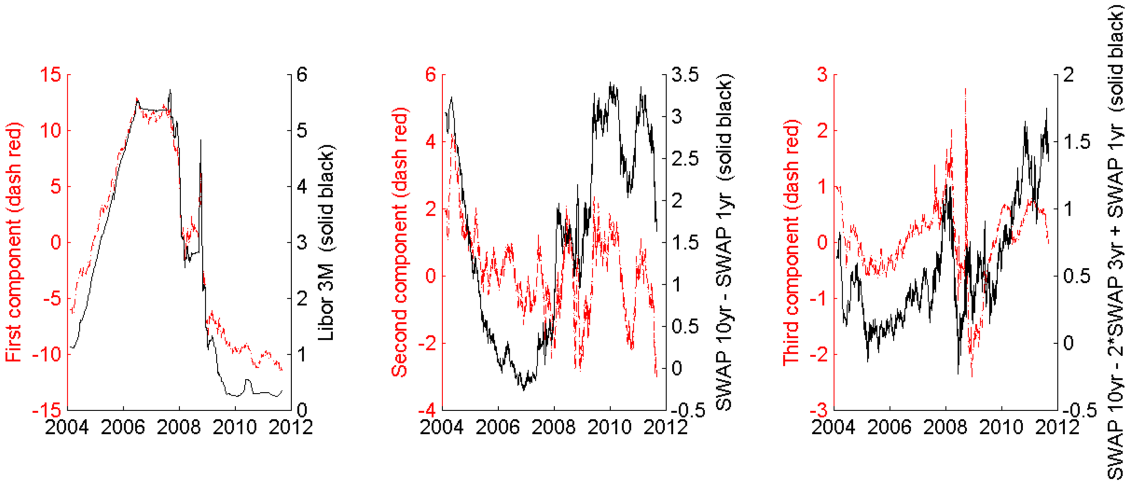

Figure 2.4: **Time series of 3 PC of the term structure of interest rates.** The 3 factors of Litterman and Scheinkman (1991) for our time series of interest rates compared to measures of level, slope and curvature. The principal components are shown in dash red lines and the replicating time series are shown using a solid black line.

we extract the implied volatility by inverting Black's formula.

$$
\begin{aligned}
Caplet\left(t, T_\beta, K, \sigma_K^{T_\beta}\right) &= E^{\mathbb{Q}_{T_\beta}}\left[\left(F\left(t, T_\alpha, T_\beta\right) - K\right)^+\right], \\
&= \int_{-\infty}^{+\infty} \left(F\left(t, T_\alpha, T_\beta\right) - K\right)^+ f_F\left(F, \sigma_K^{T_\beta}\right) \mathrm{d}F, \\
\sigma_K^{T_\beta} &= Black^{-1}\left(Caplet\left(t, T_\beta, K\right)\right)
\end{aligned}
$$

Thus, the relation between the forward rate and the implied volatility is clearly nonlinear. We want to approximate the relation using a flexible form. It can be shown that the pricing function belongs to a Hilbert space, so is the implied volatility. Thus, it can be approximated by a polynomial transformation of the state variables.[14] Consequently, we propose to use a polynomial transform of the factors determining the distribution of the interest rates. The first step consists in extracting principal components of the term structure of interest rates as in Litterman and Scheinkman (1991). As expected, three components explain the

[14] Longstaff and Schwartz (2001) uses the same argument.

majority of the variations in the interest rates. It is well known that the first component is related to the level of the interest rates, the second component to the slope and the third component to the curvature of the term structure. Figure 2.4 shows the times series of the three principal components and the replicating factors. We replicate the first component using the level of the 3 months US LIBOR rate, the second component using the difference between the 10 year and 1 year swap rates, and, finally, the third component with a discrete convexity combining the 10 year, 3 year and 1 year swap rates. Then, we build a polynomial of those three principal components.[15] We propose to keep the dimension of the problem relatively low by using a polynomial of order 3 including cross-products for a total of 19 factors.[16] The dimension needs to be further reduced. Usual criteria are irrelevant when selecting the number of factors to keep in the analysis because they are directly dependent on the choice of polynomial. There are thus no legitimate rule for selecting the number of factors in this context. Following Heidari and Wu (2003), we keep a total of 6 factors: the three principal components of the term structure and three principal components of the 16 remaining polynomial terms.[17] Those principal components are different from the 6 principal components of the IVS. To avoid confusion, we simply call them the factors from the term structure.

2.2.3 Market factors

Heidari and Wu (2003) suggest that factors outside the yield curve are necessary to fully explain the variations of the IVS. Nevertheless, few articles examine factors

[15] A high order polynomial will better approximate the function, but will introduce too many coefficients to estimate.

[16] Higher orders polynomials were tested, but the results were similar.

[17] The results presented in Section 2.4 are robust to other choices.

affecting cap prices. In this section, we focus on market factors as in DGS. In light of their findings, we argue that a friction and a risk factor are needed.

DGS explore potential economic determinants of the shape of the smile. The economic factors they use are measures of liquidity and default risk. Their argument comes from the trader's perspective – when a trader sells an option, he wants to hedge his exposure. Liquidity problems in the market will affect his hedging costs and therefore the price he will charge to sell the option. DGS test whether there is a different effect for ITM, ATM and out of the money (OTM) options. If the effect on prices depends on the degree of moneyness, then it affects the shape of the smile. The same is true for default risk. If default risk is more present in the market, hedging needs of financial institutions will increase. This will put more pressure on the option market and again, it increases hedging costs for traders. DGS find evidence that liquidity and default risks both affect the shape of the smile (i.e. the slope and curvature of the smile at different maturities).

It is well known that factors related to market frictions will affect option prices.[18] Friction problems will impact option prices by increasing the ask price and lowering the bid price. However, we test whether it will change the shape of the IVS. Another factor affecting option prices on the cap market is the risk of default. The caps are traded over the counter and the buyer bears the counter-party risk. Again, we test whether the effect varies with moneyness and maturity of implied volatilities. For our analysis, we keep one friction factor and one default factor. The bid-ask spread on the swaption market serves as our proxy for the friction factor (illiquidity risk). The measure will be computed from the mean relative bid-ask spread for the ATM, 1 to 5 years maturity swaptions. A general measure of risk of default in the

[18] See, e.g., Bollen and Whaley (2004), Chan et al. (2004) or Shiu et al. (2010) to name a few.

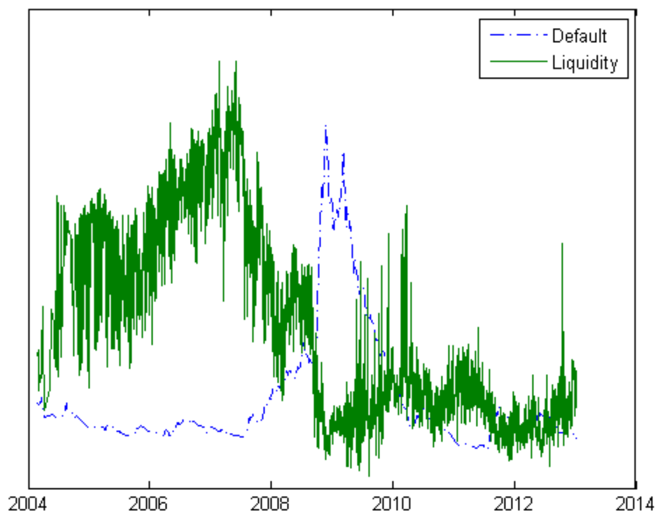

Figure 2.5: **Time series of external factors.** Both time series are scaled to be shown in the same graphic. Default is here modeled by the Fitch Index. The bid ask spread is the mean scaled bid-ask spread for different ATM swaptions.

economy is also needed. In this analysis, various alternatives are tested as default risk proxies, e.g., the Fitch probability of default index, the TED-spread and the average CDS spread of 10 large US Banks. A Fitch Default Index is an aggregation of the probability of default estimated for a large number of corporations across typically very liquid equity markets. Figure 2.5 shows the time series of the two proxies. Note that the relative bid-ask spread on the swaption market is rather volatile, but follows strong trends. As for the default index, it is stable, except for a sharp increase at the beginning of the financial crisis.

A priori, those proxies are external but not necessarily unspanned factors. This is why their dynamics is modeled along that of the IVS and the term structure. The results from the estimation of the model reveal whether the factors are unspanned.

2.3 Modeling the surface

In this section, we describe the vector model used for the test. We first consider the modeling of the IVS. Then, the factors explaining the variations of the IVS are introduced. Next, a dynamic model is proposed that integrates the interactions between all factors. Finally, we describe the specifications of the model and its interpretation.

We model the whole surface of implied volatilities to better understand the relation between the factors and the different deformation axes of the IVS. Previous literature divides the surface by maturities. DGS model the smile using a measure of the slope and a measure of convexity.[19] LZ extract the implied risk neutral densities and summarize the information using quantiles. In Section 2.2.1, we show how the principal components of the IVS are easily interpreted and can represent the surface. We explain most of the variations in the IVS and interpret the effects using the dynamics of the 6 principal components of the IVS.

While the principal components of the IVS are orthogonal by construction, they can share a joint dynamics. Since caps are options on interest rates, their dynamics are modeled along the dynamics of the term structure of interest rates. We explain in Subsection 2.2.2 how the cap implied volatilities are nonlinear functions of the state variables, and how they can be approximated using a polynomial transform of the 3 principal components of the term structure of interest rates. Thus, we will use the transformed principal components described in Subsection 2.2.2. Finally, external factors, are closely related to the interest rate markets and are also modeled

[19] We are able to replicate most of DGS results using their measures with our US data and mid quotes - except when the result for the Bid quotes is different for the Ask quotes.

jointly. To model the joint dynamics of all factors, we select a VARMA structure whose general form is:

$$Y_t = \sum_{i=1}^{p} \Phi_i Y_{t-i} + U_t - \sum_{j=1}^{q} \Theta_j U_{t-j}. \tag{2.1}$$

Y_t contains the principal components of the IVS, the term structure factors and the external factors, while U_t contains the innovations. The Φ_i's and Θ_j's capture the interrelations between the variables and the lagged innovations respectively. Equivalently, the model can be written in a more compact form

$$\Phi(L) Y_t = \Theta(L) U_t,$$

where

$$\begin{aligned} \Phi(L) &= I_k - \Phi_1 L - ... - \Phi_p L^p, \\ \Theta(L) &= I_k - \Theta_1 L - ... - \Theta_q L^q. \end{aligned}$$

Compared to DGS and LZ, this model exploits the interactions between the factors of the IVS. Those play an important role in the dynamics of the system.

Filtering individual implied volatilities with an ARMA(1,1) structure removes the serial auto-correlation. For this reason, we use a VARMA(1,1) structure for the vector model. Combined with a diagonal moving average matrix, Θ_1, this produces a parsimonious model. Furthermore, it permits identification of the parameters. Finally, the model is applied to the first difference of all factors to avoid post-test inference errors. We follow the estimation technique of Dufour and Pelletier (2008) using a diagonal MA representation; see also Lütkepohl (2005).

To simplify the interpretation, the model is reformulated using a specific vector for each constituent. Let the model be expressed as

$$\begin{pmatrix} Y_C \\ Y_I \\ Y_E \end{pmatrix}_t = \begin{pmatrix} \Phi_{CC} & \Phi_{CI} & \Phi_{CE} \\ \Phi_{IC} & \Phi_{II} & \Phi_{IE} \\ \Phi_{EC} & \Phi_{EI} & \Phi_{EE} \end{pmatrix} \begin{pmatrix} Y_C \\ Y_I \\ Y_E \end{pmatrix}_{t-1} + \begin{pmatrix} U_C \\ U_I \\ U_E \end{pmatrix}_t + \begin{pmatrix} \Theta_C & 0 & 0 \\ 0 & \Theta_I & 0 \\ 0 & 0 & \Theta_E \end{pmatrix} \begin{pmatrix} U_C \\ U_I \\ U_E \end{pmatrix}_{t-1}, \quad (2.2)$$

where Y_C denotes the vector of variations in the principal components of the IVS, Y_I the vector of variations in the transformed factors of the term structure of interest rates and Y_E the vector of variations in the external factors. Each Φ_{ij}, $i, j \in \{C, I, E\}$ is a matrix that captures the interactions between two constituents of the model. Each Θ_i, $i \in \{C, I, E\}$ is a diagonal matrix that captures the lagged effects of the innovations. This representation simplifies the interpretation of the model since each Φ_{ij} has a specific interpretation. For instance, Φ_{CC} captures the interactions between the principal components of the implied volatility surface. A non-diagonal matrix implies that deformation in one dimension of the surface also affects another one. Rogers and Tehranchi (2010) show that the implied volatility surface does not move by parallel shifts, thus Φ_{CC} should not be diagonal.[20] Additionally, we can test whether variations in external factors help predict variations in the principal components of the IVS by examining the individual ϕ_{ij} in Φ_{CE}.

Furthermore, this model allows testing for unspanned factors. So far, we claimed that factors unspanned by the term structure of interest rates are needed to better predict the variations in the IVS. However, proxy measures of those factors are used in the test model, which is free from constraints. The model thus permits interactions between external factors and factors from the term structure of interest

[20] Rogers and Tehranchi (2010) study the equity market, but the same statement should hold for the cap market.

rates. This interaction will show in Φ_{IE} and Φ_{EI}. Testing whether the coefficients in both of those matrices are nil will reveals whether the factors in Y_E are unspanned by the factors in Y_I.

Finally, the variations in principal components can be related to the variations in the shape of the IVS as in the following equation.

$$\sigma_t^{\{\kappa,T\}} = Y_{C,t}^{'} B^{\{\kappa,T\}} + \epsilon_t^{\{\kappa,T\}}, \tag{2.3}$$

where $\sigma_t^{\{\kappa,T\}}$ denotes the variations in the implied volatility for the cap of log-moneyness κ and maturity T, $Y_{C,t}$ the vector of variations in the principal components of the IVS, $B^{\{\kappa,T\}}$ the vector of loadings on the principal components for the cap of log-moneyness κ and maturity T, and $\epsilon_t^{\{\kappa,T\}}$ the error term. Now, to characterize the effects on the slope of the IVS, a measure of the slope as in $\left[\sigma_t^{\{-0.5,2\}} - \sigma_t^{\{0.5,2\}}\right]$, which is a proxy measure for the slope of the smile at a maturity of 2 years, can be used. From Figure 2.2 we argue that the main driver of the smile slope of the IVS is the third component. Neglecting the effects of the other components and using (2.3), we get

$$\left[\sigma_t^{\{-0.5,2\}} - \sigma_t^{\{0.5,2\}}\right] \simeq Y_{C_3,t}\left(\beta_3^{\{-0.5,2\}} - \beta_3^{\{0.5,2\}}\right) + \left(\epsilon_t^{\{-0.5,2\}} - \epsilon_t^{\{0.5,2\}}\right), \tag{2.4}$$

where $\beta_3^{\{\kappa,T\}}$ is the loading on the third component, and $Y_{C_3,t}$ the variations in the third component. From Figure 2.2, we observe that $\left(\beta_3^{\{-0.5,2\}} - \beta_3^{\{0.5,2\}}\right) < 0$. Therefore, an increase in the third component implies a decrease in the smile slope.

2.4 Results

This section is dedicated to analyzing the drivers of the IVS to understand how the surface reacts to variations of certain market factors. We provide estimation results of two VARMA(1,1) models, and discuss the effects from the external factors. Tables 2.I to 2.IV show the results of the estimation of the VARMA(1,1) model specified in (2.2) using the variations in the six factors from the IVS, the six factors from the term structure and 2 external factors. The results show that the external factors do help predict variations in the IVS, but also show that the external factors are not unspanned by the term structure. In an effort to disentangle the effect from the external factors from the interaction with the factors from the term structure, a VARMA(1,1) model is applied to the six factors from the term structure and the two external factors. The residuals from the two external factors are extracted. Then, the VARMA(1,1) model is estimated using the six factors of the IVS, the six factors from the term structure and the residuals of the two external factors from the previous step. Tables 2.V to 2.VIII show the results of that estimation with the residuals of the external factors.

Table 2.I shows the estimation of Φ_{CI}, which represents how variations of the term structure factors at $t-1$ affect the principal components of the IVS at t. We observe that variations in all dimensions of the implied volatility surface are driven in part by at least one factor from the term structure. Note, also, that all 6 factors from the term structure help explain variations in at least one principal components of the IVS.

Table 2.II shows the estimation of Φ_{CE}, which represents how the external factors at $t-1$ affect the variations of the principal components of the IVS at

t. First, let us examine the liquidity effect on the IVS (from BA_{t-1}). The bid-ask spread is positively and significantly related to the third component. The third component affects negatively the IVS smile slope. Hence, an increase in the bid-ask spread leads to a decrease in the smile slope.[21] Next, the bid-ask spread negatively and significantly affects the sixth component, which affects positively the IVS smile convexity. When the relative bid-ask spread increases, the convexity of the smile of the IVS decreases.

Second, we can describe the effects of the default factor (from DFT_{t-1}) on the IVS. An increase in default risks can have two different effects on the derivative market. First, it causes an increase in the overall level of risk, which should cause option prices or implied volatilities to increase. Second, on the OTC market, the buyer of the option bears the counter party risk. A discount is subtracted by the buyer when the market is more prone to default, thus an increase in default risks should produce lower implied volatilities. From Table 2.II, we observe that the default factor has a negative and significant relation with the first component, which positively affect the level of the IVS. Thus, an increase in default risks will produce a decrease in the level of the IVS. However, the default factor affects positively and significantly the fourth component, which impacts positively the curvature of the backbone of the IVS. From those two relations, we conclude that an increase in default risks will impact negatively the implied volatilities, except for the short term options. Finally, the default risk factor impacts negatively and significantly the fifth component which drives the surface twisting. This relation is more difficult to interpret, but the surface twisting might come from the fact that short term options react differently from the long term options to variations

[21] Remember that the analysis is applied to mid-quotes of cap implied volatilities, thus relatively little can be said on the effects on the bid and ask individually.

Table 2.I: **Estimation of the lagged effect of the term structure factors on the IVS ($\hat{\Phi}_{CI}$)**

IVS	$TS1_{t-1}$	$TS2_{t-1}$	$TS3_{t-1}$	$TS4_{t-1}$	$TS5_{t-1}$	$TS6_{t-1}$
$PC_t^{\{1\}}$	0.2884*** (0.000)	0.0205* (0.097)	0.0152** (0.014)	-0.0766*** (0.000)	0.0161** (0.027)	-0.1437*** (0.000)
$PC_t^{\{2\}}$	0.6622*** (0.000)	0.1409*** (0.000)	0.0052 (0.548)	-0.1688*** (0.003)	0.0353 (0.111)	-0.2458*** (0.007)
$PC_t^{\{3\}}$	-0.6088** (0.026)	-0.0736 (0.242)	-0.0853*** (0.000)	0.2137** (0.018)	-0.0550 (0.117)	0.3843*** (0.001)
$PC_t^{\{4\}}$	0.7685** (0.010)	0.0199 (0.737)	0.0201 (0.140)	0.0016 (0.984)	-0.0362 (0.277)	-0.1333 (0.298)
$PC_t^{\{5\}}$	1.7256*** (0.000)	0.4341*** (0.001)	0.1038*** (0.000)	-0.1672 (0.310)	-0.0146 (0.832)	-0.1701 (0.431)
$PC_t^{\{6\}}$	0.7937* (0.075)	0.3133*** (0.002)	0.0426 (0.361)	-0.1321 (0.432)	0.0575 (0.354)	-0.1784 (0.277)

This table shows $\hat{\Phi}_{CI}$ from the VARMA(1,1) estimation of the model expressed in (2.2). We can observe the effect of the variations in the factors from the term structure at $t-1$ ($TS1_{t-1}$, $TS2_{t-1}$, etc) on the variations of the 6 dimensions of the IVS at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis. The implied volatility surface is linked to the principal components by the following relations: $Level \propto PC^{\{1\}}$, $Backbone\,slope \propto PC^{\{2\}}$, $Smile\,slope \propto -PC^{\{3\}}$, $Backbone\,curvature \propto PC^{\{4\}}$, $Smile\,twisting \propto PC^{\{5\}}$, $Smile\,curvature \propto PC^{\{6\}}$.

Table 2.II: **Estimation of the lagged effect of the external factors on the IVS ($\hat{\Phi}_{CE}$)**

IVS	BA_{t-1}	DFT_{t-1}
$PC_t^{\{1\}}$	-0.0025 (0.253)	-0.1200* (0.075)
$PC_t^{\{2\}}$	-0.0029 (0.567)	-0.0998 (0.537)
$PC_t^{\{3\}}$	0.0230*** (0.002)	-0.0465 (0.821)
$PC_t^{\{4\}}$	-0.0143 (0.191)	0.3189** (0.042)
$PC_t^{\{5\}}$	0.0202 (0.282)	-0.9104* (0.059)
$PC_t^{\{6\}}$	-0.0519*** (0.004)	0.2096 (0.550)

This table shows $\hat{\Phi}_{CE}$ from the VARMA(1,1) estimation of the model expressed in (2.2). We can observe the effect of the variations in the external factors at $t-1$ (BA_{t-1}, the liquidity factor and DFT_{t-1}, the default risk factor) on the variations of the 6 dimensions of the IVS at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis. The implied volatility surface is linked to the principal components by the following relations: $Level \propto PC^{\{1\}}$, $Backbone\,slope \propto PC^{\{2\}}$, $Smile\,slope \propto -PC^{\{3\}}$, $Backbone\,curvature \propto PC^{\{4\}}$, $Smile\,twisting \propto PC^{\{5\}}$, $Smile\,curvature \propto PC^{\{6\}}$.

in default risks.

The VARMA(1,1) specification in (2.2) permits a straightforward Granger causality analysis because of the diagonal moving average matrix. We already confirmed that both our friction and default factors Granger cause some variations of the IVS components (see Table 2.II).[22]

Next, we want to check whether factors affecting cap prices are unspanned by factors of the interest rates term structure. Tables 2.III and 2.IV show how

[22] Based on the simplified conditions of Boudjellaba et al. (1994)

variations in the external factors can help predict variations in the factors from the term structure of interest rates and vice-versa. The friction factor, i.e. the bid-ask spread on the swaption market does not help predict the term structure factors (See Table 2.III), but is predicted by the fourth and fifth factors from the term structure (See Table 2.IV). Hence, the friction factor does not appear to be unspanned by the term structure factors. The default factor appears to help predict the term structure factors, but not the other way around. Thus the default factor is not unspanned either.

This raises an interesting question about the actual presence of unspanned factors. Therefore, we want to verify whether the effects of the two external factors come from the presence of unspanned factors or from spurious results. To do so, we first apply a VARMA(1,1) to the six factors from the term structure and the two external factors. From that estimation, the residuals of the two external factors are kept and used in the estimation of the full VARMA(1,1) model with the six factors of the IVS and the six factors from the term structure. Tables 2.V to 2.VIII show the results of that estimation.

Table 2.V shows that the interaction between the factors from the term structure at $t-1$ with the factors of the IVS at t are slightly different, but again, all six factors from the term structure are helpful to predict variations in the IVS. Table 2.VI shows that the residuals from the external factors still explain the variations in the IVS, though slightly differently. The two dimensions of the IVS that are explained by the external factors are the smile slope and the smile curvature. The third component of the IVS loads positively on the bid-ask spread, and the smile slope loads negatively on the third component. Thus when the bid-ask spread increases, the smile slope gets less asymmetric. The current data includes only the mid-quotes

Table 2.III: **Estimation of the lagged effect of the external factors on the term structure factors ($\hat{\Phi}_{IE}$)**

TS	...	BA_{t-1}		DFT_{t-1}	
TS1_t	...	-0.0001	(0.919)	0.0436***	(0.009)
TS2_t	...	0.0002	(0.970)	0.2019*	(0.064)
TS3_t	...	-0.0053	(0.561)	-0.2734*	(0.081)
TS4_t	...	0.0008	(0.724)	-0.0617*	(0.067)
TS5_t	...	0.0020	(0.759)	-0.2976**	(0.031)
TS6_t	...	-0.0017	(0.589)	-0.1339***	(0.000)

This table shows $\hat{\Phi}_{IE}$ from the VARMA(1,1) estimation of the model expressed in (2.2). We can observe the effect of the variations in the external factors at $t-1$ (BA_{t-1} and DFT_{t-1}) on the variations of the 6 principal components from polynomial transformation of the term structure factors at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis

of implied volatilities. The effect on the smile for the bid and the ask prices of the caps is thus unknown. The same coefficient for the default factor is negative. An increase in default risk thus results in an increase in the smile slope, or an increase in negative implied skewness. The sixth component of the IVS loads negatively on the bid-ask spread, and this suggests an increase in the bid-ask spread leads to a flatter smile. The opposite is true for the default factor. An increase in default risk thus accentuates the implied kurtosis on the cap market. Higher levels of implied kurtosis on the option market reflects the aversion of investors for extreme events.

Finally, Tables 2.VII and 2.VIII show the interactions between the residuals from the external factors and the factors from the term structure. The residuals of the bid-ask spread, that still explain variations in the IVS, appear to be unrelated to the factors of the term structure. This shows the presence of at least one factor unspanned by the term structure of interest rate. In contrast, the residuals of the default factor are still related to the factors of the term structure.

Table 2.IV: **Estimation of the lagged effect of the term structure factors on external factors ($\hat{\Phi}_{EI}$)**

IVS	$TS1_{t-1}$		$TS2_{t-1}$		$TS3_{t-1}$		$TS4_{t-1}$		$TS5_{t-1}$		$TS6_{t-1}$	
BA_t	0.0669	(0.835)	0.0442	(0.596)	-0.0095	(0.693)	0.3229**	(0.020)	-0.0997**	(0.020)	0.1560	(0.197)
DFT_t	-0.0108	(0.768)	0.0005	(0.928)	0.0023	(0.390)	0.0007	(0.940)	-0.0009	(0.737)	0.0083	(0.441)

This table shows $\hat{\Phi}_{EI}$ from the VARMA(1,1) estimation of the model expressed in (2.2). We can observe the effect of the variations of the 6 principal components from polynomial transformation of the term structure factors at $t-1$ (BA_{t-1} and DFT_{t-1}) on the variations of the external factors at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis

Table 2.V: **Estimation of the lagged effect of the term structure factors on the IVS ($\hat{\Phi}_{CI}$)**

IVS	$TS1_{t-1}$		$TS2_{t-1}$		$TS3_{t-1}$		$TS4_{t-1}$		$TS5_{t-1}$		$TS6_{t-1}$	
$PC_t^{\{1\}}$	0.1268***	(0.007)	0.0292**	(0.042)	0.0186***	(0.005)	-0.0019*	(0.068)	0.0029**	(0.015)	-0.0023	(0.121)
$PC_t^{\{2\}}$	0.3903***	(0.005)	0.1601***	(0.000)	0.0055	(0.608)	-0.0066***	(0.003)	0.0052**	(0.041)	-0.0019	(0.533)
$PC_t^{\{3\}}$	-0.3032	(0.188)	-0.0950	(0.144)	-0.1045***	(0.000)	0.0081**	(0.019)	-0.0070*	(0.087)	0.0147***	(0.000)
$PC_t^{\{4\}}$	0.5186**	(0.025)	0.0319	(0.614)	0.0134	(0.444)	0.0006	(0.860)	-0.0038	(0.325)	0.0033	(0.530)
$PC_t^{\{5\}}$	1.5326***	(0.001)	0.4612***	(0.000)	0.0872**	(0.010)	-0.0082	(0.190)	0.0026	(0.745)	0.0063	(0.435)
$PC_t^{\{6\}}$	0.8799*	(0.050)	0.3755***	(0.000)	0.0487	(0.238)	-0.0144**	(0.032)	0.0091	(0.243)	-0.0120*	(0.061)

This table shows $\hat{\Phi}_{CI}$ from the VARMA(1,1) estimation of the model expressed in (2.2). We can observe the effect of the variations in the factors from the term structure at $t-1$ ($TS1_{t-1}$, $TS2_{t-1}$, etc) on the variations of the 6 dimensions of the IVS at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis. The implied volatility surface is linked to the principal components by the following relations: $Level \propto PC^{\{1\}}$, $Backbone\,slope \propto PC^{\{2\}}$, $Smile\,slope \propto -PC^{\{3\}}$, $Backbone\,curvature \propto PC^{\{4\}}$, $Smile\,twisting \propto PC^{\{5\}}$, $Smile\,curvature \propto PC^{\{6\}}$.

Table 2.VI: **Estimation of the lagged effect of the residuals of the external factors on the IVS ($\hat{\Phi}_{CE}$)**

IVS	BA_{t-1}		DFT_{t-1}	
$PC_t^{\{1\}}$	-0.0009	(0.154)	0.0007	(0.419)
$PC_t^{\{2\}}$	-0.0014	(0.398)	-0.0027	(0.266)
$PC_t^{\{3\}}$	0.0072***	(0.001)	-0.0081***	(0.000)
$PC_t^{\{4\}}$	-0.0054	(0.104)	0.0015	(0.738)
$PC_t^{\{5\}}$	0.0032	(0.583)	-0.0092	(0.192)
$PC_t^{\{6\}}$	-0.0179***	(0.001)	0.0121***	(0.000)

This table shows $\hat{\Phi}_{CE}$ from the VARMA(1,1) estimation of the model expressed in (2.2) with the residuals from the external factors. We can observe the effect of the variations in the external factors at $t-1$ (BA_{t-1}, the liquidity factor and DFT_{t-1}, the default risk factor) on the variations of the 6 dimensions of the IVS at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis. The implied volatility surface is linked to the principal components by the following relations: $Level \propto PC^{\{1\}}$, $Backbone\,slope \propto PC^{\{2\}}$, $Smile\,slope \propto -PC^{\{3\}}$, $Backbone\,curvature \propto PC^{\{4\}}$, $Smile\,twisting \propto PC^{\{5\}}$, $Smile\,curvature \propto PC^{\{6\}}$.

Table 2.VII: **Estimation of the lagged effect of the residuals of the external factors on the term structure factors ($\hat{\Phi}_{IE}$)**

TS	...	BA_{t-1}		DFT_{t-1}	
$TS1_t$	...	-0.0001	(0.844)	0.0007*	(0.065)
$TS2_t$	...	0.0005	(0.752)	0.0014	(0.316)
$TS3_t$	...	-0.0021	(0.479)	0.0075***	(0.004)
$TS4_t$	...	0.0112	(0.376)	0.0013	(0.890)
$TS5_t$	...	-0.0041	(0.728)	-0.0135	(0.288)
$TS6_t$	...	-0.0060	(0.693)	0.0194**	(0.021)

This table shows $\hat{\Phi}_{IE}$ from the VARMA(1,1) estimation of the model expressed in (2.2) with the residuals from the external factors. We can observe the effect of the variations in the external factors at $t-1$ (BA_{t-1} and DFT_{t-1}) on the variations of the 6 principal components from polynomial transformation of the term structure factors at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis

Table 2.VIII: **Estimation of the lagged effect of the term structure factors on residuals of the external factors ($\hat{\Phi}_{EI}$)**

IVS	$TS1_{t-1}$		$TS2_{t-1}$		$TS3_{t-1}$		$TS4_{t-1}$		$TS5_{t-1}$		$TS6_{t-1}$	
BA_t	-0.1732	(0.870)	-0.3238	(0.189)	-0.0071	(0.940)	0.0263	(0.206)	-0.0310	(0.122)	-0.0079	(0.658)
DFT_t	0.1178	(0.901)	0.3464*	(0.096)	0.0323	(0.799)	-0.0156	(0.209)	0.0426**	(0.013)	-0.0142	(0.312)

This table shows $\hat{\Phi}_{EI}$ from the VARMA(1,1) estimation of the model expressed in (2.2) with the residuals from the external factors. We can observe the effect of the variations of the 6 principal components from polynomial transformation of the term structure factors at $t-1$ (BA_{t-1} and DFT_{t-1}) on the variations of the external factors at time t. (***, **, *) represent 1%, 5% and 10% significance respectively and the p-values are shown in parenthesis

2.5 Concluding remarks

In this research, we have examined the relation between the interest rates and the implied volatility surface of the cap market. We build upon previous literature by incorporating the nonlinear effects from the term structure in the modeling of the whole implied volatility surface. The dynamics of all factors is modeled jointly in a VARMA model and tests can be carried out on the external factors. We provide a simple way of constructing linear factors that can take into account the nonlinear information contained in the term structure of interest rates. We show that the nonlinear factors are in fact useful in explaining cap implied volatilities. Nonetheless, even after controlling for nonlinear effects and interactions between caps themselves, external factors are still needed to explain variations in cap implied volatilities. Finally, our model shows that there exists at least one factor that is unspanned by the term structure of interest rates. That unspanned factor mainly affects the smile slope and curvature, i.e., the implied volatilities of options ITM and OTM.

The presence of such a factor affects how market makers hedge their open positions on the market. It justifies the use of models specific for the option market. It also raises questions about the additional risk that is introduced in the portfolio of financial institutions using those options to hedge their interest rate risks. The unspanned factor will affect the price of the options and the value of the portfolio will vary differently from what is predicted by classical term structure models. If the options are not kept until maturity, this can have a financial impact on the value of the portfolio of those institutions.

A by-product of our research is a simple method to extract caplet implied volatilities from cap quotes. We extend arbitrage free smoothing of the implied volatility smile to the whole surface and show how the caplets implied volatilities can then be extracted.

Appendix

In this appendix we show how to extract the caplet volatilities by smoothing the whole implied volatility surface.

Extraction of the caplet volatilities

Let $i \in I = \{4, 8, 12, ...\}$ be the indices for cap quotes from the market;[23] $\sigma_K^{T_i}$ be the implied volatility quote for a cap of strike K and maturity T_i, then the price for a cap of strike K and maturity T_i can be expressed as the sum of the caplets evaluated with the cap implied volatility as in:

$$C_K^{T_i}\left(\sigma_K^{T_i}\right) = \sum_{j=2}^{T_i} Cl_K^{T_j}(\sigma_K^{T_i}), \tag{2.5}$$

where $C_K^{\tau}\left(\sigma_K^{\tau}\right)$ and $Cl_K^{\tau}(\sigma_K^{\tau})$ are the cap and caplet prices respectively with strike K and maturity τ.[24] Now suppose we have the individual caplet implied volatilities, noted γ. We can then price caps using a sum of individual caplets with their own implied volatilities as in:

$$C_K^{T_i}\left(\sigma_K^{T_i}\right) = \sum_{j=2}^{T_i} Cl_K^{T_j}(\gamma_K^{T_j}),$$

where $\gamma_K^{T_j}$ is the caplet volatility of strike K and maturity T_j.

[23] We use indices representing quarters because the US cap market uses the 3M US Libor as basis rate.

[24] Note that a 1 year cap contains only 3 caplets and the first one is noted $C_K^{T_2}$ and that is why we start at $j = 2$.

Now suppose we have the cap implied volatilities for all intermediate maturities, i.e. $i \in \mathcal{I} = \{1, 2, 3, ...\}$. We price a cap of maturity T_i using a cap of maturity T_{i-1} plus one caplet as in,

$$C_K^{T_i}\left(\sigma_K^{T_i}\right) = C_K^{T_{i-1}}\left(\sigma_K^{T_{i-1}}\right) + Cl_K^{T_i}(\gamma_K^{T_i}),$$
$$C_K^{T_i}\left(\sigma_K^{T_i}\right) = \sum_{j=2}^{T_{i-1}} Cl_K^{T_j}(\sigma_K^{T_{i-1}}) + Cl_K^{T_i}(\gamma_K^{T_i}). \tag{2.6}$$

Since we have quotes for all intermediate maturities, we are able to price each cap, so the only unknown in (2.6) is the caplet price. By smoothing the IVS and interpolating cap quotes for all intermediate maturities, we are able to extract caplet prices and invert Black's formula to get the caplet implied volatility.

Chapter 3

How "Animal Spirits" React to the Government Credibility Problem: A Real Option Analysis of Emission Permits Policy Risk

Sang Baum Kang[1] and Pascal Létourneau[2]

ABSTRACT

This paper demonstrates how to use the real option approach to make a physical capital investment decision under the presence of the government credibility problem. Specifically, we study a rational firm's optimal decision to build an electric power plant when the firm believes that there is a carbon dioxide (CO_2) policy risk. A simplified real option approach with analytic solutions finds that the time-inconsistency problem will decrease the degree that emission permits market encourages firms to invest in a green resource. A more sophisticated Least Squares Monte Carlo framework improves the analysis with the following predictions. First, considering the compound option nature of power plant investment decision, a rational firm invests in power plants early and the government credibility problem does not substantially increase such short investment timing. Second, the time-inconsistency problem does increase the expected profit of a rational firm. The approaches in this paper are applicable to other areas in finance.

JEL classification: D81, G13, Q50

Keywords: real option, government credibility, compound option, emission permit

3.1 Introduction

A carbon dioxide (hereafter, CO_2) emission permit is a policy instrument to reduce the amount of CO_2 pollution.[1] According to Helm et al. (2003), the ex-ante commitment of a government to keep the initial quantity of emission permits is important to achieve the target reduction. Otherwise, private firms, which can

[1]Stuart School of Business, Illinois Institute of Technology, 565 W. Adams St., #448, Chicago, IL, U.S.A. Email: sangbaum@gmail.com.

[2]HEC Montréal, 3000 Côte-Sainte-Catherine, H3T 2A7, Montréal, QC, Canada. Email: pascal.letourneau@hec.ca.

We are grateful to Robert Cairns, Stylianos Perrakis and participants of 2011 Administrative Science Association of Canada (ASAC) annual meeting and 2013 Midwest Finance Association (MFA) annual meeting for the valuable comments and advice. All errors are strictly ours.

[1]For the details, see, e.g., Ayres and Walter (1991), Nordhaus (2007) and Paolella and Taschini (2008).

make a choice between different carbon-emitting technologies, may expect credibility problems in the carbon policy, and they are less likely to invest in a resource emitting less amount of CO_2. It is a natural question to ask how "animal spirits", who can delay an investment, react to such policy uncertainty.

In the presence of the government credibility problem, three predictions can be made from the standard real option theory. First, a firm will make relatively fewer investments in a facility with relatively "greener" technology. Second, a firm will defer such a "green" investment in a facility with relatively "greener" technology because the less credible a government is, the more volatile an emission permit price will be. Third, a firm will make more (expected) profit as a result of the greater uncertainty. (For background, see, e.g., Majd and Pindyck (1987) and Pindyck (1991).) This article demonstrates how to extend the real option approach to analyze the optimal decisions of a rational firm under the time-inconsistency of government policy and the option to change the amount of electricity generation over time. Specifically, we propose two real option analyses.

First, motivated by Paxson (2007) who studies real options of property rights, we study the value of an investment opportunity for a power plant using a compound exchange option (Carr (1988)) (hereafter, CEO). A firm generating electric power makes a sequence of two separate but related decisions. The first one is whether the firm invests in a physical capital and the second one is whether the firm dispatches the electric power plant. Because an exchange option (Margrabe (1978)) can model the second decision, a compound option on an exchange option (CEO), can model the first decision. Calculating CEO premia[2], we find that the

[2] One can quantify not only the intrinsic value of such an investment opportunity but also the extrinsic value; the probability of an increase in electricity prices relative to the generation cost creates this additional value.

government credibility problem will decrease the degree that emission permits market encourages firms to invest in a green resource. However, this simple framework does not account for the time series nature of the decision to dispatch and cannot formally model the government policy uncertainty.

Second, because of the aforementioned two limitations of the simple CEO framework, we implement a Least Squares Monte Carlo simulation of Longstaff and Schwartz (2001) (hereafter, LSM) and investigate the investment decisions of an electricity company that faces government credibility risk and provide intuition with regards to the investment timing, the investment choice and the profitability.[3]

To analyze the effect of government policy uncertainty on the real option, this paper proposes adding two ingredients to the standard real option framework. First, political pressure on a government is modeled as a Brownian motion and the submission of government to such pressure as a first passage time.[4] Pawlina and Kort (2005) use the first passage time of the value of the entire investment project to study investment under uncertainty and policy change. We contribute to the literature by proposing to use a separate latent stochastic process for which the threshold of the first passage time can be easily calibrated.

Second, we model a firm's option to build an electric power plant as an American compound option, because a plant is a time-series of options representing a choice between producing a positive amount of output and not producing any output. Yang et al. (2008) study the plant investment options under the uncertainty of climate policy change, but they ignore one of the most important features within our methodology, which is that a power plant itself is a time-series of spread options.

[3]Our analysis is partly motivated by Schwartz and Trolle (2007) who study the real option under expropriation risk.

[4]For the details of a Brownian motion first passage time, see, e.g., Shreve (2004).

Yang et al. (2008) conclude that a firm should wait if it expects the climate policy uncertainty. However, considering the compound option nature of an investment opportunity, we find that such a conclusion is not always true. This constitutes another contribution of our research.

The complexity of this compound option raises an interesting question. What is the implication of the current high volatilities of energy commodities on a rational producer's choice of investment timing? On the one hand, the higher the energy commodities volatilities are, the more valuable a physical plant in place will be (larger intrinsic value). A rational firm thus has incentives to invest in a power plant earlier rather than later. On the other hand, the higher the energy commodities' volatilities, the higher the volatility of a plant value will be. Hence, the extrinsic value of a real option increases and a rational firm has incentives to postpone investments. What is the trade-off between these two opposite effects? To our knowledge, the LSM is the only way to assess such a trade-off.

The numerical results of LSM show that the effect of higher plant value (larger intrinsic value) dominates the effect of higher plant value volatility (larger extrinsic value), leading to a relatively small extrinsic value. Such dominance is too strong for the incremental uncertainty, as a result of the government time-inconsistency problem, to delay the investment timing substantially. In addition, the time-inconsistency problem does increase the investment into less-green plants and the expected profit of a rational firm.

The remainder of this paper is structured as follows. Section 3.2 documents the real option of an investment opportunity for a power plant. Section 3.3 and 3.4 discusses data and a preliminary numerical analysis, respectively. Section 3.5

reports a simplified real option analysis using CEOs and section 3.6 documents our real option analyses using LSM simulations. Section 3.7 concludes.

3.2 The real option

The methodology in this paper is relevant for unregulated wholesale electricity markets, such as the United States and the United Kingdom, where forward markets of electricity and natural gas are available for trading. It does not necessarily require electricity prices to be unregulated at the retail market level. An electricity company, which is a price taker of a wholesale market electricity and fuel prices, has an option to dispatch a plant or to purchase electricity from a wholesale market.

Consider the following simple numerical example. Assume that the wholesale electricity price is \$50 per megawatt-hour (hereafter, MWh), the wholesale fuel price is \$4.00 per million British Thermal Units (hereafter, mmbtu), and the heat rate of a generator that the company owns is 10 mmbtu/MWh. Dispatching the generator costs \$40/MWh (=\$4.00/mmbtu * 10 mmbtu/MWh). On the other hand, purchasing electricity from the wholesale market costs \$50/MWh. As dispatching the generator (\$40/MWh) costs less than purchasing electricity from the wholesale market, a rational firm chooses to dispatch its generator. For another example, assume that the wholesale electricity price is \$40/MWh, the wholesale fuel price is \$5.00/mmbtu, and the heat rate of the generator is the same as before. Dispatching the generator then costs \$50/MWh (=\$5.00/mmbtu * 10 mmbtu/MWh). On the other hand, purchasing electricity from the wholesale market costs only \$40/MWh. Therefore, a rational firm chooses to stop running the plant and purchases electricity from the wholesale market.

Let us describe a market where there is a price associated with the emission of CO_2. Let E denote the wholesale electricity price in \$/MWh, F the wholesale fuel price in \$/mmbtu, H the heat rate in mmbtu/MWh, M the emission amount of CO_2 (llb/mmbtu), and P the emission cost (\$/llb) of a generator that a firm owns. The electricity cost for the firm is then:

$$Electricity\,Cost_t = \min\left(E_t, H \times (F_t + M \times P_t)\right). \tag{3.1}$$

Suppose that the firm does not have any electric generator. The electricity cost is then E, which is greater than or equal to (3.1). Hence, the payoff of owning an electric generation plant is as follows:

$$Payoff_t = E - \min\left(E_t, H \times F_t\right) = \max\left(0, E_t - H \times (F_t + M \times P_t)\right). \tag{3.2}$$

As H is constant[5], and E_t, F_t and P_t are random variables in the wholesale market, (3.2) is the same as the payoff function of a spread option (a.k.a. exchange option). The parameters of the spread option model are wholesale electricity and fuel prices, their forward volatilities and correlation and an interest rate. The value of the electricity generation plant is given as the expectation of a time series of exchange options:

$$Plant\,Value_t = \tilde{E}_t\left[\sum_{i=t+1}^{t+T} exp\left(-r_i\,(i-t)\right) \times Payoff_i\right]. \tag{3.3}$$

where T is the operating life of the plant; r_t is the forward interest rate at time t; $\tilde{E}$ is a risk neutral expectation.[6]

[5]In reality, H is not a constant because the physical heat rate depends on the age of a generator, temperature, air pressure and many other variables.

[6]Section 3.5 proposes a simplified version where the decisions to dispatch are grouped at the half-life of the plant.

As discussed in the introduction, the company has an option to choose the timing of an irreversible investment to the best of its economic payoff. The time-0 value of an investment opportunity for a power plant is given as an American compound option on a time series of exchange options.

$$
\begin{aligned}
&(The\, value\, of\; a\, single\, plant\, investment\, opportunity)_{t=0}\\
&= \sup_{\tau} \tilde{E}_{t=0}\left[exp\left(-r_{\tau}\tau\right) \times \max\left[0, Plant\, Value_{\tau} - K_{\tau}\right]\right]
\end{aligned}
\tag{3.4}
$$

where τ is the optimal investment time and K_{τ} is capital investment at time τ. Finally, the value of an investment opportunity of a firm which can make a choice between a (less green) coal plant and a (more green) natural gas plant is given as an exotic American compound exchange option:

$$
\begin{aligned}
&(The\, value\, of\; a\, two-plants\, investment\, opportunity)_{t=0}\\
&= \sup_{\tau} \tilde{E}_{t=0}\left[\max\left[\begin{array}{c} 0 \\ \sum_{i=\tau+1}^{\tau+T}\left[\exp\left(-r_i i\right)\tilde{E}\left[\max\left[0, E_i - H\left(F_i^C + M^C P_i\right)\right]\right]\right] - K_{\tau}^C \\ \sum_{i=\tau+1}^{\tau+T}\left[\exp\left(-r_i i\right)\tilde{E}\left[\max\left[0, E_i - H\left(F_i^G + M^G P_i\right)\right]\right]\right] - K_{\tau}^G \end{array}\right]\right]
\end{aligned}
\tag{3.5}
$$

where τ is the optimal investment time; F_t^C and F_t^G are forward prices (\$/mmbtu) with delivery time t of coal and natural gas, respectively; M^C and M^G are plant emission amounts (llb/mmbtu) for coal and natural gas; P_t is emission cost (\$/llb); K_t^C and K_t^G are the capital investments at time t for a coal plant and a natural gas plant, respectively.

Table 3.I: Physical characteristics and capital costs

Plant	Coal plant	Natural gas plant
Detail	An integrated gasification combined turbine with minimum carbon preparation and level II control	A water-cooled combined cycle combustion turbine
Heat rate	8.732 mmbtu/MWh	7.223 mmbtu/MWh
CO2 emission	205.35 pound/mmbtu	118.00 pound/mmbtu
Capital cost	$2,479 per kilowatt	$895 per kilowatt

3.3 The data

To study how the government time-inconsistency affects investment choice, investment timing, and profitability of a firm, we consider a relatively "greener" natural gas electric generation plant versus a relatively "less-green" coal plant. The physical and economic data of each plant are from the Integrated Resource Plan (hereafter, PacifiCorp (2007)) that PacifiCorp, a multi-billion dollar electric power utility serving six western states in the United States, filed to the Oregon Public Utility Commission.

An integrated gasification combined turbine (hereafter, IGCC) with minimum carbon preparation and level II control has been chosen as the example coal plant because, given the current public awareness of air pollution, the choice of other "traditional" coal plants may lead to deterioration in public relations. A water-cooled combined cycle combustion turbine (hereafter, CCCT) is selected as the natural gas plant because of its flexibility and popularity. Table 3.I summarizes the physical characteristics and cost information of the coal plant and the natural gas plant. The physical heat rate of the selected coal plant (8.732 mmbtu/MWh) is less favorable than that of the selected natural gas plant (7.223 mmbtu/MWh). The CO_2 emission of the coal (205.35 pound/mmbtu) is also less favorable than that of the natural gas plant (118.00 pound/mmbtu). Finally, the capital cost of the coal plant ($2,479 per kilowatt) is less favorable than that of the natural gas plant ($895 per kilowatt). A typical range of coal prices is from $0.20/mmbtu to $1.20/mmbtu, and natural gas prices typically range from $2.00/mmbtu to $20.00/mmbtu. In the present simulation, the initial coal price is assumed to be $0.50/mmbtu and the initial natural gas price is assumed to be $8.50/mmbtu.

A typical electricity price ranges from $10 to $150. In the simulations in this study, the electricity price is assumed to be $70. To compute the forward prices, we assume an interest rate of 4%. Electricity and natural gas forward volatilities are modeled as a weakly decreasing function in tenor starting from 0.6. (For the detail of decreasing term-structure of volatility, see, e.g., Kang and Klein (2005).) As the commoditization of coal is *very* limited[7] relative to that of natural gas, coal volatility is assumed to be zero. The forward correlations between electricity prices and natural gas prices are modeled as a weakly increasing function of tenor starting from 0.6. Finally, from Ayres and Walter (1991) and PacifiCorp (2007), we use $35 per ton or $0.0175 per pound as an estimate for an emission permit price and index it with the interest rate. Price discovery from emissions futures markets[8] is neglected because it is dubious that the additional volatility from the emissions market will change the main conclusion of this paper.

[7]Coal in the Powder River Basin and the Appalachian area is somewhat commoditized, but the levels of standardization and liquidity are much lower than those of the natural gas commodity. Furthermore, long-term (e.g., 30 years) fixed-price coal purchase contracts are very common.

[8]See, e.g., Carmona and Hinz (2011).

3.4 A Net Present Value analysis

Consider the example of an electric power utility that evaluates three alternatives: purchasing power from a wholesale market, building a natural gas plant, and building a coal plant. The natural gas plant is "greener" than the coal plant because the natural gas plant emits 118.00 pounds of CO_2 per mmbtu , whereas the coal plant emits 205.35 pounds of CO_2 per mmbtu; about twice as much as the natural gas plant. Furthermore, the natural gas plant is about 17% more fuel efficient than the coal plant. To generate 1 MWh of electric power, a natural gas plant burns 7.223 mmbtu of fossil fuel, whereas a coal plant burns 8.732 mmbtu.[9]

Finally, capital costs are higher when building a coal plant than a natural gas plant. PacifiCorp (2007) calculates that building a coal plant costs $2,479 per kilowatt, whereas building a natural gas plant costs only $895 per kilowatt; a coal plant is about three times as expensive as a natural gas plant with regards to the amount of capital spending.

Table 3.II summarizes the simple economics of each plant, "not assuming" and "assuming" a CO_2 permit cost. Observe in Panel A, which does not assume a CO_2 permit cost, that the natural gas plant is twice as expensive as the coal plant, notwithstanding the superior "greenness" and fuel-efficiency. In Panel B, which does assume a CO_2 permit cost, observe that the relative economics favorable to the coal plant deteriorates. According to the preliminary analysis in the previous section, building and dispatching the coal plant ($76.41/MWh) or the natural gas plant ($95.49/MWh) is more expensive than purchasing electricity from the wholesale market ($70/MWh). A Net Present Value analysis would reject both projects.

[9] The data was taken from PacifiCorp (2007).

Table 3.II: Simple economics without the real option approach

Panel A: Assuming no emission permits cost

Plant	**Coal plant**	**Natural gas plant**
Fuel price	$0.50/mmbtu	$8.50/mmbtu
Fuel cost	$4.37/MWh (=8.732mmbtu/MWh * $0.50/mmbtu)	$61.40/MWh (=7.223 mmbtu/MWh * 8.50/mmbtu)
Capital cost*	$36.06/MWh	$19.18/MWh
Cost per MWh	$40.43/MWh (=$4.37/MWh +36.06/MWh)	$80.58/MWh (=$61.40/MWh * 8.50/MWh)

* Contains capital costs and fixed operation and management costs per calculation in PacifiCorp (2007).

Panel B: Assuming $35/ton emission permits cost

Plant	**Coal plant**	**Natural gas plant**
Emission amount	2,056 pound/MWh (=8.732 mmbtu/MWh *205.35 pound/mmbtu)	852 pound/MWh (=7.223 mmbtu/MWh *118.00 pound/mmbtu)
Emission price	$0.0175/pound (=$35/ton divided by 2000 pound/ton)	$0.0175/pound (=$35/ton divided by 2000 pound/ton)
Emission cost	$35.98/MWh (=2,056 pound/MWh * $0.0175/pound)	$14.92/MWh (=852 pound/MWh * $0.0175/pound)
Cost per MWh	$76.41/MWh (=$40.43/MWh+ $35.98/MWh)	$95.49/MWh (=$80.58/MWh+ $14.92/MWh)

This, however, would neglect the value of delaying the investment (see McDonald and Siegel (1986)).

3.5 A simplified real option analysis using compound exchange options

To improve on the net present value analysis, we use real options to evaluate the value of the investment opportunity. Starting with a simplified model, we use a CEO to model the investment alternatives. As we will discuss more in this section, the CEO premia of these two investment opportunities reflects the expected positive future payoff and suggest that both real options are economically valuable.

Consider a single dispatch decision made at time τ_S and a simple exchange option (hereafter, SEO) with a payoff of $S(\tau_S) \equiv max[0, E_{\tau_S} - H \cdot F_{\tau_S}]$. Furthermore, an investment decision is made at time $\tau_C \leq \tau_S$ modeled by a CEO with a payoff of $max[0, S(\tau_S) - K]$ expiring at time τ_C where K, a per-MWh physical capital cost for building a plant, is the strike price of the CEO. Using a bivariate geometric Brownian motion, Carr (1988) proposes a closed-form solution for a CEO price. Adapting his solution, we calculate a CEO price as

$$CEO_0 = E_0 N_2\left(d_1\left(\frac{P}{P^*}, \tau_C\right), d1\left(P, \tau_S\right)\right) - HF_0 N\left(d_2\left(\frac{P}{P^*}, \tau_C\right) d_2\left(P, \tau_S\right)\right) - KN_1\left(d_2\left(\frac{P}{P^*}, \tau_C\right)\right) \quad (3.6)$$

where $P \equiv E_0/(HF_0)$ is the ratio between the electricity price and the generation cost; σ_E is the electricity price volatility; σ_F is the generation cost volatility; ρ is the

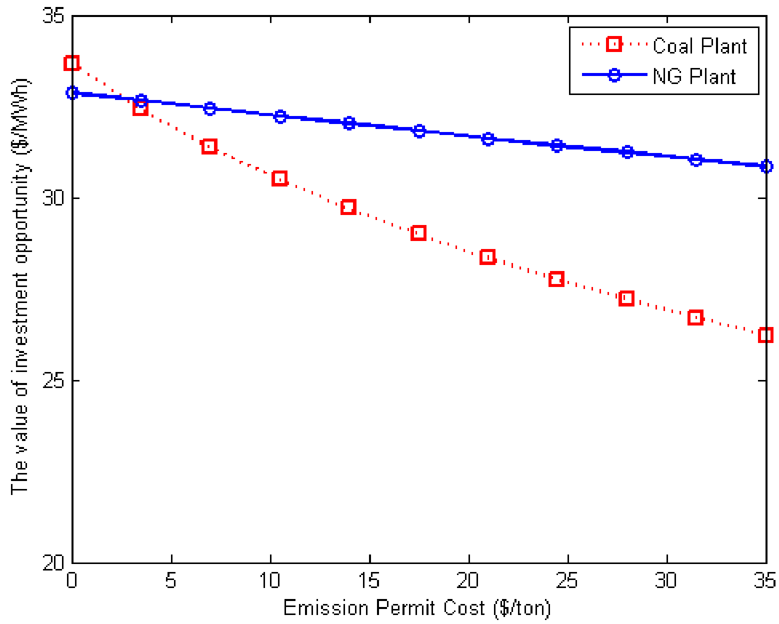

Figure 3.1: **The value of investment opportunities as a compound exchange option.**
This figure reports the CEO premia of the coal plant and the natural gas plant. From table 1, the heat rate for coal (natural gas) plant is 8.732 (7.233) mmbtu/MWh. From table 3, the electricity price=$70/MWh, the coal price=$0.50/mmbtu, the natural gas price=$8.50/mmbtu, the electricity volatility=0.60, the natural gas volatility=0.60, the coal volatility=0 and the correlation between the natural gas and electricity=0.6. Finally, from table 2, the capital cost for coal (natural gas) plant is $36.06/MWh ($19.18/MWh). The X-axis (Y-axis) represents the fuel plus emission cost (the value of investment opportunities). The red (blue) curve is for the coal (natural gas) plant. The right (left) end of each curve assumes the full $35/ton (zero) emissions cost.

correlation between these two; the volatility of $P \triangleq \bar{\sigma} \equiv \sqrt{\sigma_E^2 + H^2\sigma_F^2 - 2H\rho\sigma_E\sigma_F}$; $d_1(y,\tau) \equiv (ln(y) + 0.5\sigma^2\tau)/(\sigma\sqrt{\tau})$; $d_2(y,\tau) \equiv (ln(y) - 0.5\sigma^2\tau)/(\sigma\sqrt{\tau})$; $N_1(\cdot)$ is the standard normal c.d.f.; $N_2(\cdot)$ is the standard bivariate normal c.d.f. with correlation $\sqrt{\tau_C/\tau_S}$; P^* is implicitly determined by solving $P^*N_1(d_1(P^*)) - N_1(d_2(P^*)) = K/(HF_0)$. Calculating the simplified real option values of the coal plant and the natural gas plant, we consider a decision that will be made one year from now. That is, $\tau_C = 1$. Because a typical plant life is 30 years, we set $\tau_S - \tau_C$ to be 15 years.

Figure 3.1 depicts the values of two investment opportunities with and without assuming emissions costs. Without assuming the emission cost, the value of a coal plant investment opportunity (\$36.68) is greater than that of a natural gas plant investment opportunity (\$32.87). In contrast, assuming full emission cost of \$35/ton, the value of a natural gas plant investment opportunity (\$30.85/MWh) is greater than that of a coal plant investment opportunity (\$26.21/MWh). It is known that the intrinsic value of option is zero because the all-in cost (\$76.41/MWh) including both the physical capital and generation costs is greater than the electricity price (\$70.00/MWh). Hence, all of \$26.21/MWh is the extrinsic value. According to the same logic, the extrinsic value of natural gas plant investment opportunity is \$30.85/MWh, which is more valuable than that of coal plant. Therefore, a rational firm may change its resource choice depending on the emission cost.

Under the presence of government credibility problem, the relative economics will be between the right and left ends. Therefore, the government credibility problem will decrease the degree that emission permits market encourages firms to invest in a green resource. Specifically, if the emission price is less than the break-even emission price \$2.61/ton, the natural gas plant investment opportunity is less valuable than the coal plant investment opportunity.[10]

The government credibility problem may also increase the time variation of generation cost. Furthermore, it may change $\bar{\sigma}$ in (3.6). Because

$$\bar{\sigma} = \sqrt{((\sigma_E - H\sigma_F)^2 + 2H(1-\rho)\sigma_E\sigma_F)},$$

$\bar{\sigma}$ increases (decreases) in σ_F for a sufficiently large (small) σ_F. Hence, both the mean and standard deviation of SEOs increase (decrease) in σ_F for a sufficiently

[10] The break-even point is dependent on the initial parameter of the problem.

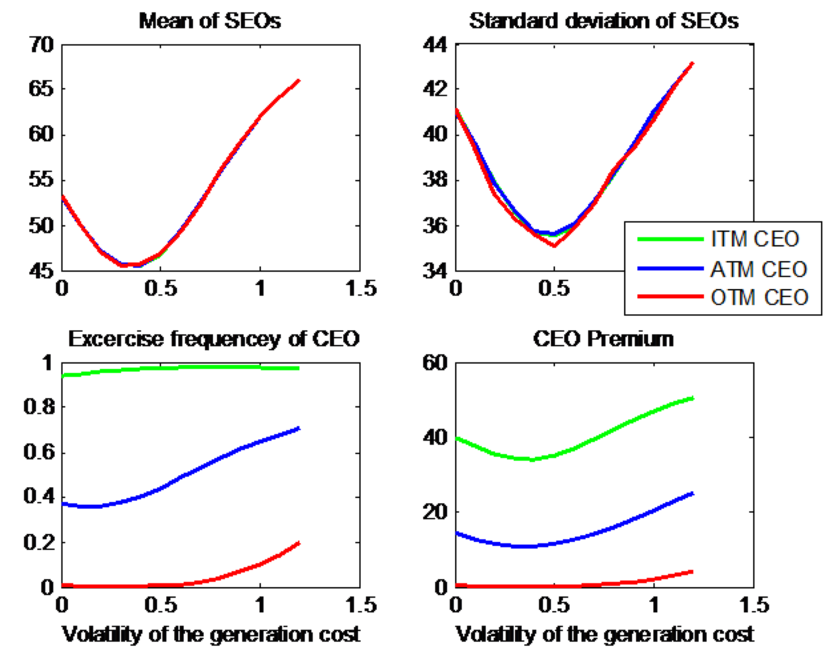

Figure 3.2: **Statistics of the CEO and SEO options as a function of volatility of generation cost.**
This figure depicts the mean and standard deviation of SEOs, the exercise probability of CEO and the CEO premium as a function of the volatility of generation cost. From table 1, the heat rate for coal (natural gas) plant is 8.732 (7.233) mmbtu/MWh. From table 3, the electricity price=\$70/MWh, the coal price=\$0.50/mmbtu, the natural gas price=\$8.50/mmbtu, the electricity volatility=0.60, the natural gas volatility=0.60, the coal volatility=0 and the correlation between the natural gas and electricity=0.6. Finally, from table 2, the capital cost for coal (natural gas) plant is \$36.06/MWh (\$19.18/MWh). To calculate the exercise frequency and CEO premium of ATM CEO, we set K to be the same as the mean of SEOs. To calculate the exercise frequency and CEO premium of ITM (OTM) CEO, we set K to be the mean of SEO divided (multiplied) by three.

large (small) σ_F as the upper panels of Figure 3.2 depicts. Furthermore, as the lower right panel depicts, a CEO premium increases (decreases) in σ_F for a sufficiently large (small) σ_F because a CEO premium increases in $\bar{\sigma}$.[11]

Observe, from the lower left panel, that in the case of out-of-the-money (hereafter, OTM) and at-the-money (hereafter, ATM) CEOs, the exercise probability exhibits a U-shape pattern. The interpretation is that as $\bar{\sigma}$ increases, there will

[11] Adapting Carr (1988), we removed his technical condition that strike K is proportional to the second asset price, namely, K=qD where D is the fuel and emission costs. The value of the option with a fixed strike is solved by simulation.

be a higher probability that an OTM or ATM CEO expires in the money.[12] The inverse of such an exercise probability of CEO may serve as a proxy for investment timing. Specifically, the increase in the generation cost volatility may encourage rational firms to invest in an ATM and OTM CEOs earlier.

Carr (1988) has several limitations in our study of the effect of the government credibility problem on the power plant investment decision:

- Because the dispatch decision is made only once during the life of the plant, it fails to address the "time series" nature of a power plant as (3.3) suggests.
- Strictly speaking, this simple CEO model does not directly address the investment timing of three choices: (a) building a coal plant, (b) building a natural gas plant and (c) do nothing and wait.[13]
- Finally it is difficult to formally introduce variables to model the government credibility problem.

Because of these limitations, the next section proposes a more sophisticated model using the LSM. Furthermore, we will analyze rich information that the LSM provides including the investment timing, the investment choice, the emission amount and the profitability.

[12] In contrast, observe that in the case of deeply in-the-money CEOs, the exercise frequency is of a "reversed" U-shape pattern. The interpretation is that as $\bar{\sigma}$ increases, there will be a lower probability that a deeply in-the-money CEO expires in the money.

[13] The early exercise of Carr (1988)'s CEO may be optimal if the convenience yield of electricity is greater than that of fuel. However, the convenience yield is not defined for electricity.

3.6 LSM analyses

To evaluate the dispatch and investment timing optionality, we evaluate the exotic compound exchange option expressed in (3.5) using the LSM method. Working backward from the terminal nodes[14], the algorithm compares non-zero exercise values with discounted continuation values of the next exercise payoffs, estimated from ordinary least squares (OLS) with a set of basis functions. Following Schwartz and Trolle (2007), we choose a constant, an electricity price, a fuel price, the emission permit prices, a latent variable that will be introduced in the next subsection, their square terms, and their cross terms as the basis functions. The complete set of polynomials up to order two are used. Various combinations of higher-order terms are also tested, but the results do not change substantially. The LSM algorithm calculates not only the option value but also a series of auxiliary quantities. It calculates the exercise time of (3.5) and the exercise (dispatch) profile of "plant value" options.

3.6.1 The credibility of government

We take a reduced-form approach to model political pressure to lower the emissions permit prices and the credibility of government as

$$(PRES)_t \equiv W\left(t\right) \, and \, \tau \equiv min\left\{t > 0 | (PRES)_t \geq (CRED)\right\} \tag{3.7}$$

where $W(t)$ is a Brownian motion, $(PRES)_t$ represents political pressure to a government, $(CRED) > 0$ represents resistance of a government to the pressure, τ

[14]Considering a typical operating life of electric power plant and a career length of manager, we assume a thirty years investment horizon.

is the first passage of time representing the time required to change the emission policy and $(PRES)_t$ is reset to zero after time τ because the government can change the policy multiple times. A "good" (credible; time-consistent) government is modeled as one with high $(CRED)$ in (3.7), whereas a "bad" (not credible; time-inconsistent) government is modeled as one with low $(CRED)$. If $(CRED) = \infty$, this extremely credible government never gives in to any political pressure to increase the amount of permits. (CRED) can be easily calibrated from a rational firm's belief on how credible its government is, for example, "with 80% probability, the government will not increase the amount of permits for the next 5 years."

Upon the first passage time, $(PRES)_t$ is reset to zero. The rationale is that once a government increases the amount of permits, the pressure on the government disappears. $(PRES)_t$ then evolves toward the next first passage time. In other words, several submissions of a government are possible during a finite time horizon.

The correlation between an electricity (natural gas) forward price and political pressure on a government is modeled as a positive number. The rationale is that if both electricity forward prices and natural gas forward prices are high, more demand for a coal plant is naturally created, and a government will receive more pressure to increase the amount of permits.

3.6.2 Simulation cases

The following three cases of belief are evaluated by the LSM:

- Case A: A firm does not pay for any CO_2 emission permits.

- Case B: A firm pays \$35/ton for CO_2 emission permits, and a government is time-consistent (good; credible). $(CRED)$ in (3.7) is set to 10,000, our proxy for positive infinity. The quantity of emissions permits practically never increases.

- Case C: A firm pays \$35/ton for CO_2 emission permits, and a government is time-inconsistent (bad; not credible). In (3.7), $(CRED)$ is set to 1. The probability that the amount of emission permits will increase in the next year is about 16% ($= 1-\phi(1)$, where $\phi(\cdot)$ is the cumulative distribution of a standard normal distribution). It is assumed that whenever $(PRES)$ hits $(CRED)$, in other words, whenever a government gives in the political pressure, the price of an emission permit decreases by 50%, and $(PRES)$ is reset to zero. This is represented by $P_{\tau^k} = \frac{1}{2}P_{\tau^k-}$, where $\tau^k = min\{t > 0|(PRES)_t \geq (CRED)\}$ and $k = 1, 2, 3, ...$

Table 3.III summarizes the parameters of LSM simulations. Each column represents Cases A, B, and C, respectively. The parameters are categorized into three areas: Emission permits price and their evolution, Commodity forward prices and Second moments of commodities. The second and third categories are identical across all cases. In other words, electricity forward prices, coal price forecasts, natural gas forward prices, and the associated volatilities and correlations are identical across Cases A, B, and C. The only difference lies in how emission permits prices evolve over time.

Table 3.III: Inputs to LSM simulations

Category	Parameter	Case A	Case B	Case C
Emission permits prices and their evolution	Emission price (\$/pound)	\$0.0000	\$0.0175/pound indexed by interest rate	\$0.0175/pound indexed by interest rate
	CRED	N/A	10000	1
	The decrease % of emission permit price when PRES hits CRED	N/A	N/A	50%
	Electricity-PRES correlation	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90
	Natural gas-PRES correlation	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90
Commodity forward prices	Electricity forward curve (\$/MWh)	\$70/MWh indexed by interest rate	\$70/MWh indexed by interest rate	\$70/MWh indexed by interest rate
	Coal price forward curve (\$/mmbtu)	\$0.50/mmbtu indexed by interest rate	\$0.50/mmbtu indexed by interest rate	\$0.50/mmbtu indexed by interest rate
	Natural gas forward curve (\$/mmbtu)	\$8.50/mmbtu indexed by interest rate	\$8.50/mmbtu indexed by interest rate	\$8.50/mmbtu indexed by interest rate
	Interest rate	4% flat	4% flat	4% flat
Second moments of commodities	Electricity forward volatility	Decreasing shape from 0.60 to 0.30	Decreasing shape from 0.60 to 0.30	Decreasing shape from 0.60 to 0.30
	Natural gas forward volatility	Decreasing shape from 0.60 to 0.30	Decreasing shape from 0.60 to 0.30	Decreasing shape from 0.60 to 0.30
	Electricity-Natural gas correlation	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90	Increasing shape from 0.60 to 0.90

3.6.3 Numerical results and discussion

Table 3.III reports the outcomes of 10,000 path simulations, with the parameters summarized in Table 3.I and the forward curves evolution documented in Appendix A. Each column represents Cases A, B, and C, respectively. Different rows represent different output variables. These variables are categorized into three areas: Resource choice, Investment timing and Economics.

From the Resource Choice category, observe that in Case A, in which a firm does not pay any emission permit price, a coal plant is chosen with 99.33% probability, but in Case B where a firm pays \$0.0175/pound emission permit price, a coal plant is chosen with only 6.74% probability. As discussed in a previous section, in Case A without emission cost, a coal plant is the obvious choice (99.33%), but in Case B with emission cost, a natural gas plant is the obvious choice (93.20%). The interpretation is that the introduction of CO_2 permits discourages investment in a relatively less "green" coal plant and encourages investment in a relatively "greener" natural gas plant.

Recall from the net present value analysis section that without considering the dispatch and investment timing optionality, the coal plant (\$76.41/MWh cost) is more economical than the natural gas plant (\$95.49/MWh cost). In Case B, which considers both emission cost and "optionality", rational firms chose the natural gas plant (93.20%) more frequently than the coal plant (6.74%). Observe that in Case B, a natural gas plant is chosen with 93.20% probability, but in Case C where a government is time-inconsistent (bad; not credible), a natural gas plant is chosen with 87.00% probability. This 6.2% reduction in the investment choice probability of a natural gas plant is consistent with the first theoretical predictions discussed in the introduction.

Observe that the percentages of not investing in any plant for the next 30 years are relatively small (0.00%, 0.06%, and 0.31% in Cases A, B, and C, respectively). Under the parameters discussed in an earlier section, investing in a plant is a more rational alternative than purchasing all electricity from the wholesale market. From the Investment Timing category, observe that the average years to build a plant are 1.31, 1.40, and 1.54 years in Cases A, B, and C, respectively. These relatively short investment timings reassert the interpretation in the previous paragraph. Moreover, observe that in Case B, the average years to build a plant is 1.40 years, but in Case C, it is 1.54 years. This outcome is consistent with the second theoretical predictions discussed in the introduction.

Higher uncertainty delays the investment, but not substantially. What explains such short investment timing is the optionality of the firm to dispatch the power plant when conditions are favorable or shut down production when they are not. When the option to shut down the power plant is removed (as in Yang et al. (2008)) the average year to invest increases up to 15 years.

From the Economics category, observe that the expected profit of a firm decreases from $1,965.97 in Case A to $698.30 in Case B. The interpretation is that the introduction of carbon emission permits takes away the firm's profit from used-to-be profitable coal plants. Moreover, observe that the expected profit of a firm increases from $698.30 in Case B to $891.60 in Case C. The reasons for this outcome are twofold. First, because of the time-inconsistency of a government, the expected emission permit price decreases. Second, due to the increased income variation caused by policy uncertainty, the extrinsic value of an option increases. The second reason for this outcome is consistent with the third theoretical prediction discussed in the Introduction section.

Still from the Economics category, observe that CO_2 emission decreases from 47,910 pounds in Case A to 8,697 pounds in Case B. The interpretation is that the introduction of CO_2 emission permits reduces CO_2 emission by 82%. In addition, observe that CO_2 emission increases from 8,698 pounds in Case B to 12,847 pounds in Case C. The interpretation is that the time-inconsistency of a government may increase the amount of CO_2 emission by 48%, everything else being equal. Even though the (belief in a) government time-inconsistency problem may not increase the investment timing, it provides incentive for rational private companies to emit a significantly larger amount of CO_2 emission, by 48%; in other words, the system fails to efficiently internalize pollution.

From Case B to Case C, the coal consumption increases from 10.8 mmbtu to 25.8 mmbtu by 138%, whereas the CO_2 emission increases from 8,697 pounds to 12,847 pounds, a 47% increase. It should be noted that the relative magnitude of the change in coal consumption is not necessarily the same as that of CO_2 emissions because CO_2 emissions are from both coal and natural gas. In addition, the investment in coal plants increases from 6.80% (=100%-93.20%) to 13.00% (=100%-87.00%), which is translated to a 91% increase. A reader may be puzzled by the difference between the 91% increase in a coal plant investment choice and the 138% increase in coal consumption. However, this difference is reasonable because of a non-linearity in the sense that the submission of a government to strong political pressure tends to occur when both electricity and natural gas prices are skyrocketing and coal plants are very much needed by private firms.

Table 3.IV: Outputs of LSM simulations

Category	Variables	Case A	Case B	Case C
Resource choice	% choose a coal plant	99.33%	6.74%	12.69%
	% chose a natural gas plant	0.67%	93.20%	87.00%
	% do not choose any plant for the next 30 years	0.00%	0.00%	0.31%
Investment timing	Expected years to exercise "built-or-no-build" option in the case that any plant is built	1.31	1.40	1.54
Economics	CO_2 emission (pound)	47,910	8,697	12,847
	Coal consumption (mmbtu)	232.7	10.8	25.8
	Natural gas Consumption (mmbtu)	1.0	54.9	64.0
	Fuel coast (Present value)	$121.59	$649.23	$585.85
	Power cost (Present value)	$152.08	$1,419.74	$1,226.45
	Expected profit of a firm (Present value)	$1,965.97	$698.30	$891.60

3.6.4 Sensitivity test

To assess the robustness of our explanations, we report two cases of robustness tests among numerous tests we have performed:

- Cases A1, B1 and C1: These cases are identical to Cases A, B and C except that both capital costs are multiplied by three. Table 3.V reports numerical outputs. The goal of these cases is to reduce the extrinsic value and increase the volatility of the "build-or-no-build" option. As expected, the invest timing of Case B1 (2.24) is greater than that in Case B (1.40). Similarly, that of Case C1 (2.68) is greater than that of Case C (1.54). As the standard real option theory suggests, the more volatile the investment project is, the more likely a rational firm will defer the investment project.

- Cases A2, B2 and C2: These cases are identical to Cases A, B and C except that the emission price is multiplied by three. TableTable 3.VI reports numerical output. The goal of these cases is to see the sensitivity of a "green" resource choice in the emission cost. As expected, more rational firms choose the natural gas plant in Case C2 (96.69%) than in Case C (87.00%).

3.7 Conclusion

This paper proposes a couple of real option approaches to address an optimal real option decision of a physical capital investment under policy uncertainty or the time-inconsistency problem. The first one is to use a simple CEO and the second is a more advanced LSM method combining a first passage time to a compound

Table 3.V: Sensitivity test (high capital cost)

Category	Variables	Case A	Case B	Case C
Resource choice	% choose a coal plant	9.90%	2.45%	5.71%
	% chose a natural gas plant	84.56%	86.97%	82.31%
	% do not choose any plant for the next 30 years	5.54%	10.58%	11.98%
Investment timing	Expected years to exercise "built-or-no-build" option in the case that any plant is built	1.99	2.24	2.68
Economics	CO_2 emission (pound)	16,906	6,934	9,70
	Coal consumption (mmbtu)	21.9	3.2	9.7
	Natural gas Consumption (mmbtu)	105.1	53.2	63.4
	Fuel coast (Present value)	$611.00	$626.32	$582.75
	Power cost (Present value)	$1,159.69	$1,501.00	$1,501.00
	Expected profit of a firm (Present value)	$958.36	$617.04	$755.17

Table 3.VI: Sensitivity test (high emission cost)

Category	Variables	Case A	Case B	Case C
Resource choice	% choose a coal plant	99.33%	0.05%	3.05%
	% chose a natural gas plant	0.67%	99.95%	96.69%
	% do not choose any plant for the next 30 years	0.00%	0.00%	0.26%
Investment timing	Expected years to exercise “built-or-no-build” option in the case that any plant is built	1.31	1.29	1.51
Economics	CO_2 emission (pound)	47,910	3,327	6,690
	Coal consumption (mmbtu)	232.7	0.0	5.4
	Natural gas Consumption (mmbtu)	1.0	28.1	47.3
	Fuel coast (Present value)	$121.59	$570.96	$605.04
	Power cost (Present value)	$152.59	$1,698.56	$1,497.20
	Expected profit of a firm (Present value)	$1,965.97	$419.48	$620.84

option. These approaches may be applicable to various areas, such as banking, R&D management and international finance, as the belief of rational firms about a policy risk or time-consistency problem may prevent a government from achieving its policy goal. A practitioner may use our approaches to assist with real-world capital investment decisions by comparing several government credibility scenarios.

Studying the real option of capital investment decision under the government credibility problem, we found that the increase in volatility of commodities and government uncertainty has two opposite effects on the investment timing. First, the higher volatility leads to higher volatility of SEOs, the underlying of the CEO, which delays the optimal time to invest; this view would be in line with the conventional wisdom. Second, the higher volatility leads to a higher expected value of SEOs, which decreases the time to invest. To our surprise, we find that the second effect is dominant over the first effect. This apparent paradox is easily attributed to the nature of compound exchange options.

The current research opens several opportunities for future research. First, an empiricist may investigate a testable prediction: the lower the (belief in the) credibility of a government, the less investment there will be in "green" end-user technology. For example, someone may add to the literature by analyzing European Union countries, which may be heterogeneous in government credibility and the investment in "green" physical capital.

Second, this paper focuses on the optimal decision of a private firm and the effect of heterogeneous (CRED)s on investment timing and the pollution amount given a advanced American compound real option. However, an equally interesting question would be to ask what the determinants of this heterogeneity in (CRED)s are. We leave this question to future research.

Appendix

Forward curves evolution

The evolution of whole forward curves, as opposed to a spot price, of electricity prices and fuel prices over time should be modeled because a rational firm makes an irreversible investment decision by evaluating (5). Motivated by Schwartz (1997), we model continuous time stochastic processes of electricity forward prices, fuel forward prices, and a latent variable for political pressure at time t as:

$$\mathrm{d}\log E\left(t,T\right) = -\left[\sigma_E^2\left(T-t\right)/2\right]\mathrm{d}t + \sigma_E\left(T-t\right)\mathrm{d}Z_E\left(t,T^{'}\right) \quad for\, T \geq t$$

$$\mathrm{d}\log F\left(t,T\right) = -\left[\sigma_F^2\left(T-t\right)/2\right]\mathrm{d}t + \sigma_F\left(T-t\right)\mathrm{d}Z_F\left(t,T^{'}\right) \quad for\, T \geq t$$

$$\mathrm{d}\left(PRES\right)_t = \mathrm{d}Z_{PRES}\left(t,T^{'}\right)$$

$$\begin{bmatrix} \mathrm{d}Z_E\left(t,T^{'}\right) \\ \mathrm{d}Z_F\left(t,T^{'}\right) \\ \mathrm{d}Z_{PRES}\left(t,T^{'}\right) \end{bmatrix} = \begin{bmatrix} 1 & \rho_{E,F}\left(T^{'}-t\right) & \rho_{E,PRES}\left(T^{'}-t\right) \\ \rho_{E,F}\left(T^{'}-t\right) & 1 & \rho_{F,PRES}\left(T^{'}-t\right) \\ \rho_{E,PRES}\left(T^{'}-t\right) & \rho_{F,PRES}\left(T^{'}-t\right) & 1 \end{bmatrix}^{1/2} \begin{bmatrix} \mathrm{d}W_1\left(t\right) \\ \mathrm{d}W_2\left(t\right) \\ \mathrm{d}W_3\left(t\right) \end{bmatrix} \tag{3.8}$$

where t represents a trading time; $T^{'}$ represents a delivery time of a commodity; $\sigma_E\left(\tau^{'}\right)$ and $\sigma_F\left(\tau^{'}\right)$ is a forward instantaneous volatility with a tenor $\tau^{'}$ of electricity and fuel, respectively; $\rho_{E,F}\left(\tau^{'}\right)$, $\rho_{E,PRES}\left(\tau^{'}\right)$, and $\rho_{F,PRES}\left(\tau^{'}\right)$ is a forward instantaneous correlation with a tenor $\tau^{'}$ between an electricity price and a fuel price, between an electricity price and the latent variable, and between a fuel price and the latent variable, respectively; $W_1\left(t\right)$, $W_2\left(t\right)$ and $W_3\left(t\right)$ are independent Brownian motions. In (3.8), whole forward curves of electricity prices move upward or downward together and so do whole forward curves of fuel prices. The greater a $\rho_{E,F}\left(\tau^{'}\right)$, the more likely that electricity forward prices move in the same direction as fuel prices. As $\sigma_E\left(\tau^{'}\right)$ and $\sigma_F\left(\tau^{'}\right)$ are a function of $\tau^{'}$, the term structure of volatility can be incorporated into the model. A typical pattern of $\sigma_E\left(\tau^{'}\right)$ and $\sigma_F\left(\tau^{'}\right)$ is decreasing in $\tau^{'}$. Specifically, if $\sigma_E\left(\tau^{'}\right)$ and $\sigma_F\left(\tau^{'}\right)$ is a decreasing exponential function, it is identical to Schwartz (1997)'s model. As $\rho_{E,F}\left(\tau^{'}\right)$ is also a function of $\tau^{'}$, the term structure of correlation can be explicitly modeled. A typical pattern of $\rho_{E,F}\left(\tau^{'}\right)$ is increasing in $\tau^{'}$. Figure 3.3 depicts the positive and negative one standard deviation of perturbation in the forward curve. Observe that with these volatility and correlation functions, successive upward (downward) movements lead to backwardation (contango).

The correlation between an electricity price (a fuel price) and political pressure on a government is explicitly incorporated into $\rho_{E,PRES}\left(\tau^{'}\right)$ and $\rho_{F,PRES}\left(\tau^{'}\right)$. This is an important feature because it is plausible that the higher the electricity forward price is, the higher the political pressure ($PRES$) is on a government. In this paper, $(PRES)_t$ is the single determinant for an emission price. Additional sources of emissions price uncertainty as well as the price discovery from the emissions futures market (e.g., Carmona and Hinz (2011)) are ignored because such an incremental production cost variation may not change the main conclusion of this paper.

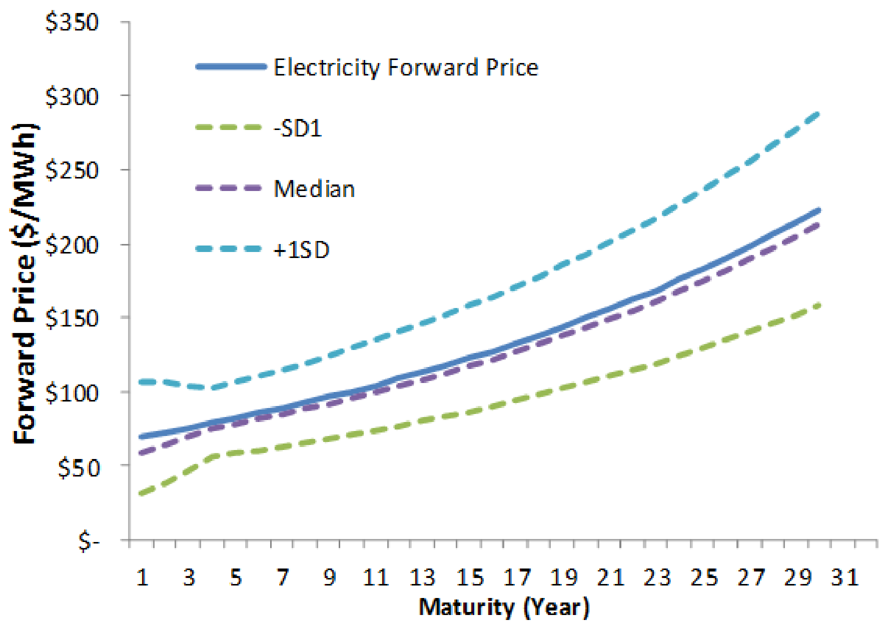

Figure 3.3: **Electricity forward curve movement.**
This figure depicts plus/minus one standard deviation movement of electricity forward curve. First, observe that the whole forward curve is moved in the same direction because of the 1-factor nature of our approach. Second, because of the Samuelson's effect, the movement of relatively near future years is greater than that of relatively far future years. Third, observe that a contango (backwardated) curve may change to a backwardated (contango) curve. The movement of natural gas forward curve also exhibits the same pattern.

Bibliographie

Ait-Sahalia, Y. and J. Duarte (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics 116*(1-2), 9–47.

Areal, N., A. Rodrigues, and M. R. Armada (2008). On Improving the Least Squares Monte Carlo Option Valuation Method. *Review of Derivatives Research 11*(1-2), 119–151.

Ayres, R. U. and J. Walter (1991). The greenhouse effect: damages, costs and abatement. *Environmental and Resource Economics 1*(3), 237–270.

Bacinello, A. R., E. Biffis, and P. Millossovich (2010). Regression-based Algorithms for Life Insurance Contracts with Surrender Guarantees. *Quantitative Finance 10*(9), 1077–1090.

Barraquand, J. and D. Martineau (1995). Numerical Valuation of High Dimensional Multivariate American Securities. *Journal of Financial and Quantitative Analysis 30*(3), 383–405.

Beresteanu, A. (2007). Nonparametric Estimation of Regression Functions under Restrictions on Partial Derivatives. *Working Paper, Duke University*.

Black, F. (1976). The pricing of commodity contracts. *Journal of financial economics 3*(1-2), 167–179.

Bollen, N. P. B. and R. E. Whaley (2004). Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance 59*(2), 711–753.

Boudjellaba, H., J.-M. Dufour, and R. Roy (1994). Simplified conditions for non-causality between vectors in multivariate arma models. *Journal of Econometrics 63*(1), 271 – 287.

Bowman, A. and A. Azzalini (1997). *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*. Oxford University Press, USA.

Brigo, D. and F. Mercurio (2006). *Interest rate models: theory and practice: with smile, inflation, and credit*. Springer Verlag.

Broadie, M. and P. Glasserman (1997). Pricing American-style Securities using Simulation. *Journal of Economic Dynamics and Control 21*(8-9), 1323–1352.

Broadie, M. and P. Glasserman (2004). A Stochastic Mesh Method for Pricing High-Dimensional American Options. *Journal of Computational Finance 7*(4), 35–72.

Carmona, R. and J. Hinz (2011). Risk-neutral models for emission allowance prices and option valuation. *Management Science 57*(8), 1453–1468.

Carr, P. (1988). The valuation of sequential exchange opportunities. *The Journal of Finance 43*(5), 1235–1256.

Carriere, J. (1996). Valuation of the Early-exercise Price for Options using Simulations and Nonparametric Regression. *Insurance: Mathematics and Economics 19*(1), 19–30.

Casassus, J., P. Collin-Dufresne, and B. Goldstein (2005). Unspanned stochastic volatility and fixed income derivatives pricing. *Journal of Banking & Finance 29*(11), 2723–2749.

Chan, K. C., L. T. W. Cheng, and P. P. Lung (2004). Net buying pressure, volatility smile, and abnormal profit of hang seng index options. *The Journal of Futures Markets 24*(12), 1165–1194.

Collin-Dufresne, P. and R. Goldstein (2002). Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *The Journal of Finance 57*(4), 1685–1730.

Collin-Dufresne, P., R. Goldstein, and C. Jones (2009). Can interest rate volatility be extracted from the cross section of bond yields? *Journal of Financial Economics 94*(1), 47–.

Connor, G. and R. Korajczyk (1993). A test for the number of factors in an approximate factor model. *Journal of Finance*, 1263–1291.

Cont, R. and J. Da Fonseca (2002). Dynamics of implied volatility surfaces. *Quantitative finance 2*(1), 45–60.

Deuskar, P., A. Gupta, and M. Subrahmanyam (2008). The economic determinants of interest rate option smiles. *Journal of Banking & Finance 32*(5), 714–728.

Driessen, J., P. Klaassen, and B. Melenberg (2009). The performance of multi-factor term structure models for pricing and hedging caps and swaptions. *Journal of Financial and Quantitative Analysis 38*(03), 635–672.

Duffie, D. (1996). *Dynamic Asset Pricing Theory.* Princeton, New Jersey: Princeton University Press.

Dufour, J. and D. Pelletier (2008). Practical methods for modelling weak varma processes: identification, estimation and specification with a macroeconomic application.

Fan, R., A. Gupta, and P. Ritchken (2003). Hedging in the possible presence of unspanned stochastic volatility: Evidence from swaption markets. *The Journal of Finance 58*(5), 2219–2248.

Feinerman, R. P. and D. J. Newman (1973). *Polynomial Approximation.* Baltimore, MD.: Williams and Wilkins.

Fengler, M. R. (2009). Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance 9*(4), 417–.

Gamba, A. (2002). Real Options Valuation: A Monte Carlo Approach. *Mimeo, University of Verona, Italy*.

Gupta, A. and M. Subrahmanyam (2005). Pricing and hedging interest rate options: Evidence from cap-floor markets. *Journal of Banking & Finance 29*(3), 701–733.

Hagan, P. (2006). Libor market model with sabr style stochastic volatility. , Working paper.

Hagan, P. and M. Konikov (2004). Interest rate volatility cube: Construction and use. Technical report, Bloomberg technical report.

Hardle, W. (1990). *Applied Nonparametric Regression.* Cambridge University Press Cambridge.

Harrison, J. and D. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of economic theory 20*(3), 381–408.

Harrison, J. M. and S. R. Pliska (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications 11*(3), 215 – 260.

Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica (1986-1998) 60*(1), 77–.

Heidari, M. and L. Wu (2003). Are interest rate derivatives spanned by the term structure of interest rates? *The Journal of Fixed Income 13*(1), 75–75–86.

Helm, D., C. Hepburn, and R. Mash (2003). Credible carbon policy. *Oxford Review of Economic Policy 19*(3), 438–450.

Juneja, S. and H. Kalra (2009). Variance Reduction Techniques for Pricing American Options Using Function Approximations. *Journal of Computational Finance 12*(3), 79–102.

Kan, K. F., R. M. Reesor, T. Whitehead, and M. Davison (2009). Correcting the Bias in Monte Carlo Estimators of American-style Option Values. In P. L'Ecuyer and A. Owen (Eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2008*, pp. 439–454. Springer-Verlag.

Kang, S. B. and M. Klein (2005). Understanding Sam. Technical report.

Lemieux, C. and J. La (2005). A Study of Variance Reduction Techniques for American Option Pricing. *Proceedings of the 2005 Winter Simulation Conference*, 1884–1891.

León, A. and A. Vaello-Sebastià (2009). American GARCH Employee Stock Option Valuation. *Journal of Banking and Finance 33*, 1129–1143.

Li, H. and F. Zhao (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. *The Journal of Finance 61*(1), 341–378.

Li, H. and F. Zhao (2009). Nonparametric estimation of state-price densities implicit in interest rate cap prices. *Review of Financial Studies 22*(11), 4335 – 4376.

Liew, C. (1976). Inequality Constrained Least-Squares Estimation. *Journal of the American Statistical Association 71*(355), 746–751.

Litterman, R. and J. Scheinkman (1991). Common factors affecting bond returns. *The Journal of Fixed Income 1*(1), 54–61.

Litterman, R., J. Scheinkman, and L. Weiss (1991). Volatility and the yield curve. *The Journal of Fixed Income 1*(1), 49–53.

Longstaff, F. and E. Schwartz (2001). Valuing American Options by Simulation: A Simple Least-Squares Approach. *Review of Financial Studies 14*(1), 113.

Longstaff, F. A. (2005). Borrower Credit and the Valuation of Mortgage-Backed Securities. *Real Estate Economics 33*, 619–661.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Cambridge Univ Press.

Majd, S. and R. S. Pindyck (1987). Time to build, option value, and investment decisions. *Journal of financial Economics 18* (1), 7–27.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *The Journal of Finance 33* (1), 177–186.

McDonald, R. and D. Siegel (1986). The value of waiting to invest. *The Quarterly Journal of Economics 101* (4), 707–727.

Moreno, M. and J. Navas (2003). On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Derivatives. *Review of Derivatives Research 6* (2), 107–128.

Morini, M. and F. Mercurio (2007). No-Arbitrage Dynamics for a Tractable SABR Term Structure Libor Model. *SSRN eLibrary*.

Nawalkha, S. (2009). The libor/sabr market models: A critical review. , Working paper.

Nawalkha, S., N. Beliaeva, and G. Soto (2007). *Dynamic term structure modeling: the fixed income valuation course*. Wiley.

Nawalkha, S. K. and R. Rebonato (2011). What interest rate models to use? buy side versus sell side. *Journal of Investment Management : JOIM 9* (3), 1–.

Newey, W. K. (1997). Convergence Rates and Asymptotic Normality for Series Estimators. *Journal of Econometrics 79* (1), 147–168.

Nordhaus, W. D. (2007). A review of the" stern review on the economics of climate change". *Journal of Economic Literature*, 686–702.

PacifiCorp (2007). Integrated Resource Plan. pp. 93–96.

Paolella, M. S. and L. Taschini (2008). An econometric analysis of emission allowance prices. *Journal of Banking & Finance 32* (10), 2022–2032.

Pawlina, G. and P. M. Kort (2005). Investment under uncertainty and policy change. *Journal of Economic Dynamics and Control 29* (7), 1193–1209.

Paxson, D. A. (2007). Sequential american exchange property options. *The Journal of Real Estate Finance and Economics 34* (1), 135–157.

Pena, I., G. Rubio, and G. Serna (1999). Why do we smile? on the determinants of the implied volatility function. *Journal of Banking & Finance 23* (8), 1151–1179.

Pindyck, R. S. (1991). Irreversibility, uncertainty, and investment. Technical report, National Bureau of Economic Research.

Rasmussen, N. (2005). Control Variates for Monte Carlo Valuation of American Options. *Journal of Computational Finance 9*(1), 2–12.

Rebonato, R. (2007). A time homogeneous, sabr-consistent extension of the lmm. *Risk 20*, 92–97.

Rebonato, R., K. McKay, and R. White (2010). *The SABR/LIBOR market model: pricing, calibration and hedging for complex interest-rate derivatives*. Wiley.

Rogers, L. C. G. and M. R. Tehranchi (2010). Can the implied volatility surface move by parallel shifts? *Finance and Stochastics 14*(2), 235–248.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance 52*(3), 923–973.

Schwartz, E. S. and A. B. Trolle (2007, November). Pricing expropriation risk in natural resource contracts–a real options approach. Harvard University.

Shiu, Y.-M., G.-G. Pan, S.-H. Lin, and T.-C. Wu (2010). Impact of net buying pressure on changes in implied volatility: Before and after the onset of the subprime crisis. *Journal of Derivatives 17*(4), 54–66,5.

Shreve, S. E. (2004). *Stochastic calculus for finance I: The binomial asset pricing model*. Springer.

Stentoft, L. (2004a). Assessing the Least Squares Monte-Carlo Approach to American Option Valuation. *Review of Derivatives research 7*(2), 129–168.

Stentoft, L. (2004b). Convergence of the Least Squares Monte Carlo Approach to American Option Valuation. *Management Science 50*(9), 1193–1203.

Stentoft, L. (2012a). American Option Pricing using Simulation with Application to the GARCH Model. In C. B. Adrian R. Bell and M. Prokopczuk (Eds.), *Handbook of Research Methods and Applications in Empirical Finance*, Chapter 5, pp. 114–147.

Stentoft, L. (2012b). Value Function Approximation or Stopping Time Approximation: A Comparison of Two Recent Numerical Methods for American Option Pricing Using Simulation and Regression. *forthcoming in Journal of Computational Finance*.

Svensson, L. (1995). Estimating forward interest rates with the extended nelson & siegel method. *Quarterly Review, Sveriges Riksbank 3*, 13–26.

Tilley, J. (1993). Valuing American Options in a Path Simulation Model. *Transactions of the Society of Actuaries 45*(83), 499–520.

Tsitsiklis, J. and B. Van Roy (2001). Regression Methods for Pricing Complex American-style Options. *IEEE Transactions on Neural Networks 12*(4), 694–703.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*(2), 177 – 188.

Wang, Y. and R. Caflisch (2010). Pricing and Hedging American-Style Options: A simple Simulation-Based Approach. *Journal of Computational Finance 13*(4), 95–125.

Wills, A. and W. Heath (2002). Using a Modified Predictor-Corrector Algorithm for Model Predictive Control. In *Proceedings of the 15th IFAC World Congress on Automatic Control. Barcelona, Spain*, Volume 21, pp. 26.

Yang, M., W. Blyth, R. Bradley, D. Bunn, C. Clarke, and T. Wilson (2008). Evaluating the power investment options with uncertainty in climate policy. *Energy Economics 30*(4), 1933–1950.

Yüceer, U. (2002). Discrete Convexity: Convexity for Functions Defined on Discrete Spaces. *Discrete Applied Mathematics 119*(3), 297–304.