

HEC MONTRÉAL

Étude du taux de recouvrement des faillites des particuliers

par

Yannick Assor

**Sciences de la gestion
(Option Finance)**

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)*

Décembre 2018
© Yannick Assor, 2018

Résumé

Alors que l'endettement des ménages et les faillites de particuliers au Canada connaissent une croissance soutenue depuis plusieurs années, il existe toujours peu d'articles portant sur la perte en cas de faillite des individus. Entre autres, on connaît très peu la relation entre le taux de recouvrement des prêts non garantis et les variables pouvant l'influencer. Il serait pertinent de fournir de meilleurs instruments pour prévoir le taux de recouvrement des faillites. À l'aide d'une base de données exclusive sur les faillites canadiennes, nous étudions le taux de recouvrement des consommateurs individuels et proposons aux créanciers non garantis un modèle unique de prédiction du taux de recouvrement lors du dépôt du dossier de faillite. Cet instrument pourrait permettre aux créanciers de calculer leur capital réglementaire en fonction de la prédiction du taux de recouvrement conditionnel à la faillite et non pas simplement au défaut. Nous abordons les problèmes d'endogénéité et de biais de sélection. Nous innovons en corrigeant ce dernier au moyen du modèle en deux parties de Heckman (1977), ce qui à notre connaissance, n'a jamais été appliqué à la perte en cas de défaut. Les résultats les plus probants impliquent que certains mécanismes à la disposition des créanciers ont un effet significatif sur leur taux de recouvrement. Notre analyse contribue aux recherches préalablement effectuées sur la perte en cas de défaut, notamment quant aux variables déterminant le taux de recouvrement.

Mots-clés

Gestion de risque, faillite, particuliers, consommateurs, perte en cas de défaut, taux de recouvrement, modèle de Heckman (1977), modèles de prédiction.

Table des matières

Résumé	i
Liste des tableaux	vi
Liste des figures	viii
Remerciements	ix
Introduction	1
1 Revue de littérature	7
1.1 Contexte	8
1.1.1 Les accords de Bâle	9
1.1.2 Cycles économiques	11
1.1.3 Le défaut et la faillite — définitions et distinctions	12
1.2 Perte en cas de défaut des entreprises	14
1.2.1 Recouvrement d'obligations corporatives	14
1.2.2 Recouvrement de prêts bancaires aux entreprises	18
1.2.3 Calcul du taux de recouvrement des entreprises	20
1.3 Perte en cas de défaut des individus	21
1.3.1 Variables explicatives du taux de recouvrement	22
1.3.2 Prédiction du taux de recouvrement	24
1.3.3 Classification de taux de recouvrement par type de dette	28

1.3.4	Modélisation du taux de recouvrement	29
1.3.5	Insolvabilité et taux de recouvrement	31
1.3.6	Calcul du taux de recouvrement des prêts individuels	35
2	Faillite et règlementation au Canada	41
2.1	L'insolvabilité	41
2.2	Séniorité de la dette	42
2.2.1	Créanciers garantis	42
2.2.2	Créanciers privilégiés	43
2.2.3	Créanciers non garantis	43
2.3	Recours des créanciers avant la faillite	44
2.4	Syndic autorisé en insolvabilité (SAI)	45
2.5	Proposition	45
2.6	Faillite	46
2.6.1	Dépôt de la faillite	47
2.6.2	Administration de l'actif	48
2.6.3	Clôture de la faillite	49
3	Présentation des données	53
3.1	Description de la base de données	53
3.2	Mesures des variables et prévisions	54
3.2.1	Variable dépendante du modèle	54
3.2.2	Variables explicatives du modèle	56
3.3	Nettoyage des données	67
3.3.1	Variable dépendante du modèle	67
3.3.2	Variables explicatives du modèle	68
3.4	Statistiques descriptives	68
3.4.1	Variable dépendante du modèle	69
3.4.2	Variables explicatives du modèle	72

4	Méthodologie	79
4.1	Estimation des paramètres	80
4.1.1	Régression linéaire	80
4.1.2	Validation des modèles	82
4.1.3	Régression avec interactions	85
4.1.4	Modèle de Heckman (1977)	85
4.2	Prédiction du taux de recouvrement	88
4.2.1	Validation croisée	88
4.2.2	Sélection de variables	89
5	Résultats principaux	91
5.1	Régression principale	91
5.1.1	Interprétation des variables continues	93
5.1.2	Interprétation des variables catégorielles	95
5.2	Régression Types de dette	100
5.3	Régression Types d'actif	104
5.4	Coefficient de détermination	106
5.5	Problèmes relatifs à la validité interne du modèle	107
6	Résultats additionnels	111
6.1	Régression sur la rémunération du syndic	111
6.2	Régression avec interactions	113
6.3	Application du modèle de Heckman (1977)	118
6.3.1	Choix de modèle	118
6.3.2	Modèle de Heckman (1977)	120
6.4	Modèles de prédiction	127
6.4.1	Modèle de classification probit	129
6.4.2	Modèle de prédiction linéaire	135
6.4.3	Modèle de prédiction linéaire censuré	136
6.4.4	Modèle de classification lasso	136

Conclusion	141
7.1 Résumé	141
7.2 Résultats	142
7.3 Limites et extensions de la recherche	144
Bibliographie	147
Annexes	i
Normes du surintendant	i
Régressions additionnelles	ii
Étude des résidus	iii
Sélection de variables lasso	iv

Liste des tableaux

1.1	Taux de recouvrement selon les cycles économiques	12
1.2	Taux de recouvrement selon la séniorité de la dette	17
1.3	Comparaison de la qualité des modèles de prédiction de Zhang et Thomas (2012)	25
2.1	Délai d'obtention de la libération automatique	50
3.1	Types de conditions à la libération	65
3.2	Moments du taux de recouvrement	69
3.3	Statistiques descriptives des variables continues et numériques	72
3.4	Statistiques descriptives des variables catégorielles	73
3.5	Statistiques descriptives des Types de dette (TD_j)	74
5.1	Modèle principal	92
5.2	Comparaison des trois modèles de régression	103
6.1	Modèle Rémunération du syndic	112
6.2	Modèle avec interactions	115
6.3	Modèle de sélection	123
6.4	Modèle avec correction Heckman	126
6.5	Matrice de confusion générale	131
6.6	Modèle de classification probit	134
6.7	Modèles de prédiction linéaires	136

6.8	Modèle de classification lasso	138
A1	Normes du surintendant - Revenu excédentaire mensuel total, 2009	i
A2	Régression des Actifs (\$) sur le Sexe	ii
A3	Régression des Recettes totales (\$) sur la Région	ii
A4	Régression de la Rémunération du syndic (\$) sur la Région	iii
A5	Résidus studentisés en fonction du numéro d'observation	iii
A6	Variables explicatives retenues par le lasso	iv

Liste des figures

1.1	Niveau de séniorité en fonction de la structure de capital des entreprises . . .	15
1.2	Distribution beta du taux de recouvrement par séniorité	16
1.3	Distribution du taux de recouvrement de Prívarva et collab. (2014)	27
1.4	Histogramme du taux de recouvrement proposé	33
1.5	Histogramme du taux de recouvrement ultimement reçu	33
3.1	Histogramme - Taux de recouvrement	70
3.2	Histogramme - Rémunération relative du syndic	75
3.3	Diagramme de dispersion - Taux de recouvrement et Rémunération du syndic	77
4.1	Validation croisée en $k = 5$ parties	89

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué à ce mémoire.

Je voudrais dans un premier temps remercier mon directeur de mémoire, Georges Dionne, pour sa disponibilité, sa patience et ses précieux conseils qui reflètent sa compréhension inégalable de la gestion des risques financiers. Je lui exprime ici toute ma gratitude pour la confiance qu'il m'a accordée.

Je remercie du fond du cœur ma mère, mon père et mes soeurs, pour leur soutien inconditionnel. Ils ont su m'encourager, m'aider, et me conseiller tout au long de mes études. Je vous en suis éternellement reconnaissant.

Je remercie finalement mes amis, mon entourage et mes collègues, qui m'inspirent au quotidien et me poussent vers l'excellence.

Sans vous, ce travail n'aurait pas été possible.

Introduction

Dans le contexte de prospérité économique globale d'aujourd'hui, la hausse des taux d'intérêt est une question centrale. Au cœur de cette croissance, l'endettement des ménages est actuellement en pleine expansion. La dette totale des ménages est la somme des dettes hypothécaires et de la dette à la consommation. Cette dernière n'est pas associée qu'à des biens donnés en garantie : il peut s'agir de soldes de cartes de crédit, d'emprunts divers (bancaires ou autres), marges de crédit, factures impayées. Aujourd'hui, l'accès au crédit relevant d'une « norme sociale » et étant de plus en plus courant, les ménages canadiens ont atteint un niveau record d'endettement. En termes absolus, la dette totale des ménages est passée de 147 milliards de dollars en 1982 à 1454 milliards de dollars en 2010¹ (dollars courants). En termes relatifs, le ratio de la dette des ménages sur le revenu personnel disponible était de 66% en 1980. Il s'élevait à 150% en 2011², ce qui signifie que les ménages devaient 1,50 \$ pour chaque dollar de revenu disponible.

L'endettement des ménages est un facteur qui retient la Banque centrale d'augmenter les taux. En effet, les banques sont très exposées au défaut des individus, compte tenu de leurs portefeuilles de prêts hypothécaires et de prêts à la consommation. Par conséquent, les banques et institutions financières auraient tout avantage à disposer d'une estimation précise de la probabilité de défaut et de la perte en cas de défaut des emprunteurs.

Aussi, dans les dernières décennies, l'industrie bancaire a connu un développement

1. R. K. Chawla, 2011, « La répartition de la dette hypothécaire au Canada », L'emploi et le revenu en perspective, été 2011, vol. 23, no 2, no 75-001-XIF au catalogue de Statistique Canada.

2. Statistique Canada, 2011, « Tableau 2 : Indicateurs du secteur des ménages – Données non désaisonnalisées », Comptes du bilan national, Le Quotidien, le 13 décembre, no 11-001-XIF au catalogue de Statistique Canada.

important dans la compréhension du risque de crédit, avec l'introduction des accords de Bâle. Ces accords fournissent un cadre de référence de gestion des risques inhérents à l'entreprise. Bâle II introduit l'approche interne (*Internal Rating Based approach*) qui permet aux banques d'utiliser des estimations du risque de défaut pour calculer leur capital. Les institutions qui utilisent cette approche doivent développer des méthodes pour estimer les composantes clés du risque de défaut. C'est pourquoi la probabilité de défaut et la perte en cas de défaut ont fait l'objet d'une littérature abondante. La probabilité de défaut se définit comme étant la probabilité que le débiteur ne puisse faire face à ses obligations de remboursement. La perte en cas de défaut, quant à elle, mesure la perte qu'une banque est susceptible de subir dans le cas du défaut.

La définition du défaut est large : elle va d'un retard de paiements à la faillite. Schuermann (2004) recommande de modéliser certains événements de recouvrement séparément pour ne pas surestimer ou sous-estimer la perte en cas de défaut. En effet, alors que, dans la littérature, on observe que la distribution de la perte en cas de défaut est bimodale (Schuermann, 2004, Prívará et collab., 2014, Hoechstetter et collab., 2012), la distribution du taux de recouvrement en cas de faillite n'a qu'un seul mode sur les recouvrements nuls (Flynn et collab., 2002). La recommandation de Schuermann d'estimer les événements de défaut séparément s'avère donc particulièrement adaptée puisque ces événements de défaut ont des distributions différentes. Il convient de préciser d'emblée que la faillite est un processus encadré par la loi, peu connu du public, ce qui la différencie des autres défauts.

Toutefois, peu d'auteurs se sont attardés sur le taux de recouvrement en cas de faillite ; et moins encore dans les cas de faillites d'individus. Quelques rares études abordent l'insolvabilité des particuliers : Eraslan, Koşar, Li et Sarte (2017) ne traitent pas à proprement parler de la faillite, mais plutôt du *Chapter 13* du *United States Bankruptcy Code* qui est l'équivalent de la proposition de consommateur au Canada ; ils concluent que certaines variables ont un effet sur les taux de recouvrement, mais ne développent pas un modèle de prédiction du taux de recouvrement. Quant à Flynn, Bermant et Bakewell (2002), ils abordent la question de la faillite personnelle aux États-Unis, mais simplement par des

statistiques descriptives.

Pourtant, en proportion, le nombre de faillites déposées par des consommateurs représente plus de 95% des faillites enregistrées au Canada en 2009 : 116 381 faillites individuelles sur 121 801 faillites totales.

Comme en témoignent les *Statistiques sur l'insolvabilité au Canada* de 2009³, la valeur des passifs attribués aux faillites des consommateurs s'élève à plus de 12 milliards de dollars, tandis que la valeur des passifs déclarés par les entreprises en faillite est de 5,5 milliards de dollars. Les faillites de consommateurs représentent donc plus de 68% des passifs totaux des faillites. D'où l'importance de se pencher sur le recouvrement des faillites individuelles, qui représentent une plus grande exposition au défaut pour les créanciers que les faillites des entreprises.

Il en va de même pour la valeur des actifs. En effet, toujours d'après les *Statistiques sur l'insolvabilité au Canada* de 2009, sur un total de 8 milliards d'actifs déclarés dans les faillites d'entreprises et d'individus, les consommateurs en déclarent près de 7 milliards. Il apparaît donc qu'en moyenne, les individus possèdent un actif pour deux passifs. Ce ratio est moins élevé pour les entreprises qui ont davantage de levier financier. Il est d'autant plus pertinent, donc, de se pencher sur le recouvrement des individus, qui présentent un potentiel de recouvrement largement supérieur; le principe de la faillite étant la liquidation des biens pour le bénéfice de la masse des créanciers.

C'est pour ces raisons que nous avons entrepris d'estimer la perte en cas de faillite d'individus. Alors que les faillites des consommateurs sont de plus en plus nombreuses au Canada, les prêteurs courent plus de risques face à leurs créances. Une gestion adéquate des risques de défaut est souhaitable, notamment une estimation plus précise du taux de recouvrement des faillite de particuliers.

Nous nous proposons dans notre étude de combler ce manque dans la littérature. Puisque le comité de Bâle exige aux institutions que le taux de recouvrement soit soumis à des simulations de récession économique, nous avons utilisé des données extraites d'une période de crise financière. En effet, nos données recensent les faillites de l'année

3. Bureau du surintendant des faillites, *Statistiques sur l'insolvabilité au Canada*, 2009.

2009, année record d'insolvabilité au Canada, avec un taux de 5,8 pour 1000 consommateurs s'étant déclarés sous la protection de la *Loi sur la faillite et l'insolvabilité* (LFI). Wirenhammar (2011) ayant ventilé le taux de recouvrement en catégories de créanciers (garantis ou non garantis), nous nous concentrerons dans notre étude sur les créanciers non garantis au Canada. En effet, les créanciers garantis ayant des actifs pour garantir le remboursement de leurs prêts, il nous a semblé moins pertinent de les inclure dans notre recherche.

L'objectif de notre recherche est donc, dans un premier temps, d'expliquer la variabilité du taux de recouvrement des créanciers non garantis dans les faillites individuelles, en l'estimant grâce à des variables indépendantes, puis, dans un deuxième temps, proposer aux créanciers un modèle de prédiction du taux de recouvrement lors du dépôt du dossier de faillite. En effet, nous pouvons nous demander quelles variables ont un effet statistiquement significatif sur le taux de recouvrement dans un contexte de faillite personnelle, et si ces variables peuvent nous conduire à des modèles de prédiction précis en matière de perte en cas de défaut ?

Notre hypothèse de départ est la suivante : les coûts de la faillite, les types de dette et d'actif, le caractère volontaire ou pas de la faillite et les mécanismes à la disposition des créanciers ont un effet substantiel sur le taux de recouvrement.

Les coûts associés à la faillite des entreprises ont largement été examinés dans le cadre d'études sur la structure de capital des entreprises⁴. Qu'en est-il des coûts de la faillite des individus ? La rémunération des syndics, qui sont les administrateurs des faillites, constitue la part la plus importante du coût d'une faillite. Quel est l'impact de cette rémunération sur le taux de recouvrement des créanciers ? À cet effet, nous nous interrogerons sur l'endogénéité de cette variable explicative, que nous traiterons au Chapitre 6.

4. Modigliani et Miller (1958), "The cost of capital, corporation finance and the theory of investment", *The American Economic Review* 48.3 (1958) : 261-297.
M. J. Gruber et J. B. Warner. "Bankruptcy costs : Some evidence." *The Journal of Finance* 32.2 (1977) : 337-347).

Par ailleurs, Hoehstoetter et collab. (2012) détectent que le recouvrement diffère selon les catégories de dette, qui sont fonction du domaine d'activité ou du type de dette. Nous proposerons un modèle qui vérifiera cette hypothèse. De surcroît, puisque nous avons accès aux actifs des individus en faillite, nous établirons si la présence de certains actifs fait varier le taux de recouvrement.

D'un autre côté, la faillite d'un individu peut survenir de trois manières : soit il déclare volontairement faillite, soit il fait l'objet d'une ordonnance de faillite, soit encore il est présumé en faillite. Gupton et collab. (2000) estiment que le genre de procédure de faillite aux États-Unis conduit à différents recouvrements. D'une manière similaire, nous vérifierons dans notre recherche si le genre de procédure – volontaire ou pas – de la faillite présente des différences dans le taux de recouvrement.

Enfin, parmi les mécanismes à la disposition des créanciers, ces derniers peuvent s'opposer à la libération des faillis. Nous vérifierons dans notre travail si l'opposition à la libération peut avoir un effet significatif sur le taux de recouvrement.

Afin de répondre à ces questions, nous nous inspirerons des modèles et des variables explicatives utilisés par les auteurs en ce qui a trait à la perte de cas de défaut. Les études sur le taux de recouvrement des faillites étant limitées, nous élargirons l'étendue de notre revue de littérature aux entreprises. Nous expliquerons dans un deuxième temps le processus de faillite au Canada. Ces deux étapes nous permettront de construire les variables de nos modèles sur la base de nos connaissances et de la revue de littérature.

Nous vérifierons nos hypothèses grâce aux modèles qui seront présentés dans le Chapitre 5. Nous perfectionnerons ces modèles dans le Chapitre 6 en abordant dans un premier temps l'endogénéité des coûts de la faillite, puis en intégrant des interactions entre certaines variables explicatives, et enfin, en corrigeant le biais de sélection avec le modèle en deux parties de Heckman (1977).

Finalement, nous développerons des modèles de classification de la probabilité d'observer un recouvrement et des modèles de prédiction linéaire du taux de recouvrement par la technique de validation croisée. Nous comparerons nos résultats avec ceux de la littérature et avec des modèles estimés à l'aide de méthodes de sélection automatique de

variables (*data driven*), le lasso.

L'intérêt de notre travail réside dans le fait que, pour améliorer l'estimation du taux de recouvrement, nous corrigeons le biais de sélection en appliquant le modèle en deux parties de Heckman (1977). À notre connaissance, cette démarche n'a jamais été entreprise ni documentée dans la littérature de la perte en cas de défaut. Nous avons également, à partir des documents publics des faillites canadiennes, créé un modèle unique qui permettra aux créanciers de prédire le taux de recouvrement à partir des données dont ils disposent. Alors qu'en général, les études utilisent des données privées et dont les modèles ne sont pas utilisables par d'autres, du fait de leur singularité, notre étude est basée sur des données uniformes et propose un modèle utilisable par n'importe quel créancier d'un individu en faillite au Canada. En effet, puisque la LFI prescrit les documents statutaires devant accompagner le dépôt d'une faillite, n'importe quel créancier a accès à ces documents, qui sont uniformes et identiques pour chaque procédure au Canada.

Outre la pertinence de notre recherche dans le domaine particulier des faillites d'individus, nous croyons que nos résultats pourraient être utilisés plus largement dans le domaine de la gestion des risques des prêts au détail ainsi que par l'organisme réglementaire du système de faillites au Canada, le Bureau du surintendant des faillites.

Chapitre 1

Revue de littérature

Introduction

Dans le contexte large du risque de crédit, notre revue a pour objectif de cerner spécifiquement les recherches traitant de la perte en cas de défaut des individus, particulièrement dans les cas de faillite. Pour ce faire, nous nous sommes concentrés sur la littérature la plus récente ainsi que des articles qui font référence en la matière. Ce domaine de recherche ayant été peu abordé par les auteurs, nous avons dû élargir le spectre de notre étude afin d'être en mesure de faire des parallèles avec des travaux moins spécifiques, dans l'objectif de répondre à nos questions de recherche.

Notre revue de littérature s'articule donc sous un angle thématique : nous abordons d'abord les notions générales de la gestion des risques, puis la perte en cas de défaut des entreprises, en terminant avec celle des individus – la moins documentée. En effet, nous cherchons principalement à comprendre quels modèles sont utilisés pour traiter de la perte en cas de défaut et quelles variables ont un effet statistiquement significatif sur celle-ci afin d'établir un modèle spécifique au recouvrement dans les faillites individuelles au Canada. L'intérêt de notre recherche réside dans le fait que ce type de données n'a jamais fait l'objet d'un traitement exhaustif, à notre connaissance.

1.1 Contexte

Penchons-nous tout d'abord sur la gestion des risques de l'entreprise. L'objectif de celle-ci est de créer un cadre de référence qui permettra aux entreprises de gérer les risques et l'incertitude. Dionne (2017) dénombre cinq principaux types de risques inhérents aux activités financières et économiques des entreprises :

- le risque pur qui peut être assurable ou pas,
- le risque de marché découlant de la variation des prix des marchandises, des taux de change, des taux d'intérêt, etc.,
- le risque de défaut défini par les paramètres de probabilité de défaut, perte en cas de défaut et exposition au défaut,
- le risque opérationnel attribuable à une inadéquation ou une défaillance liée à des processus, des employés, ou des événements extérieurs, et
- le risque de liquidité, soit le risque de manquer de fonds pour respecter ses obligations financières à court terme.

Dans le cadre de notre recherche, nous nous concentrerons sur le risque de défaut. Selon Engelmann et Rauhmeier (2006), le risque de défaut se ventile de la manière suivante :

- probabilité de défaut (PD) : probabilité que l'institution prêteuse subisse le défaut d'un obligataire ou d'une transaction,
- exposition au défaut (EAD) : donne une estimation de la valeur de la dette encourue (*exposure outstanding*) au moment du défaut et indique également la perte maximale sur les produits de crédit correspondants, et
- perte en cas de défaut (PCD) : mesure la perte en pourcentage qu'une banque est susceptible de subir dans le cas du défaut.

Notre étude sera axée sur la perte en cas de défaut. La gestion des risques d'entreprise est encadrée par une convention internationale connue sous le nom des accords de Bâle. En

effet, comme nous le verrons dans la prochaine section, les accords de Bâle permettent d’appréhender le risque de défaut au moyen des trois paramètres susmentionnés afin de conserver un niveau de capital réglementaire approprié, compte tenu du niveau de risque.

1.1.1 Les accords de Bâle

Afin de soutenir le cadre de référence de gestion des risques d’entreprises, le groupe des dix pays les plus industrialisés (G10) signait à Bâle, en 1988, un accord pour définir une réglementation des banques. Cet accord se traduit en Bâle I et entré en vigueur quatre ans plus tard, en 1992. Depuis, plusieurs autres pays ont signé cet accord. Les pays membres peuvent imposer des règles plus strictes à leurs banques, mais ils doivent se conformer aux obligations imposées par Bâle. L’accord exige des banques de détenir un montant de capital réglementaire minimal afin de se protéger contre certains des risques inhérents aux entreprises (Dionne, 2017).

Dans les dernières décennies, l’industrie bancaire a connu un développement important dans la compréhension du risque de crédit. Plusieurs crises ont secoué cette industrie, en particulier la crise des *subprimes* aux États-Unis en 2008. De nouvelles méthodes ont été proposées pour l’estimation des paramètres clés de risque de crédit, en particulier la probabilité de défaut.

En 2004, Bâle II propose une réforme majeure relative aux risques de crédit et opérationnels. Sous Bâle II, le niveau de capital réglementaire dépend des caractéristiques de risque de chaque type de crédit. Cette réglementation entrera en vigueur en 2006, mais son application accusera un certain retard en raison de la crise.

Ainsi, depuis Bâle II, les institutions financières peuvent, dans le cadre du risque de crédit, calculer leur capital réglementaire de deux manières (Schuermann, 2004) :

- l’approche standardisée de 1988 où chaque classe d’actifs est assignée à un risque et pondérée grâce à un poids qui est fixé par l’accord, et
- l’approche interne (*Internal Rating Based approach*), qui donne la possibilité aux banques de calculer les différentes composantes de leur risque de crédit et ainsi,

d'évaluer leur capital réglementaire à l'interne.

La formule du capital réglementaire dans l'approche interne implique, pour la banque, un calcul détaillé de la probabilité de défaut, de la perte en cas de défaut et de l'exposition au défaut.

Il existe une littérature abondante sur la probabilité de défaut. En effet, jusqu'à présent, la littérature du risque de crédit s'est intéressée davantage à l'estimation de la probabilité de défaut, puisque dans le cadre de Bâle II, l'élément sans doute le plus important à déterminer est la probabilité de défaut. De nombreuses recherches ont également été réalisées sur la relation entre la perte en cas de défaut et la probabilité de défaut. En utilisant des données sur des obligations en défaut de 1982 à 2002, Altman, Brady, Resti et Sironi (2005) établissent qu'il y a une relation négative entre la probabilité de défaut et le taux de recouvrement. De leur côté, Hu et Perraudin (2002) observent une relation négative similaire sur les données de Moody's concernant les défauts durant la période de 1971 à 2000.

Il appert que les recherches portant sur l'estimation de la perte en cas de défaut ou du taux de recouvrement sont plus limitées. C'est pourquoi nous nous penchons sur la question. En effet, il est d'autant plus important d'avoir une estimation précise de ce paramètre puisque le facteur de pondération du risque d'une institution financière est linéaire en ce qui concerne la perte en cas de défaut, comme l'illustre la formule de la perte attendue (PA - *expected loss*) :

$$PA = PD \times EAD \times PCD \quad (1.1)$$

Ainsi, la capacité des banques à estimer avec précision la perte en cas de défaut aura un effet direct sur le montant de capital réglementaire requis par Bâle II (Engelmann et Rauhmeier, 2006). Cela explique l'importance d'une estimation précise de la perte en cas de défaut et donc, du taux de recouvrement adéquat. D'après Bastos (2010), les banques peuvent gagner un avantage concurrentiel en améliorant leur modélisation interne de la prévision de la perte en cas de défaut, puisqu'elles seront en mesure d'avoir une allo-

cation optimale entre capital réglementaire et capital économique. C'est pourquoi notre recherche se concentrera sur la perte en cas de défaut et le taux de recouvrement.

Par ailleurs, le comité de Bâle exige un calcul de la perte en cas de défaut qui doit tenir compte des périodes de crise économique. Selon Frye et collab. (2000), « Defending against economic downturns is the principle reason that banks hold capital ». En effet, un des aspects affectant le taux de recouvrement est la présence de cycles économiques.

1.1.2 Cycles économiques

Comme nous venons de le voir, le comité de Bâle exige que le taux de recouvrement soit calculé ou soumis à des simulations de récession économique, ce que la flexibilité des modèles utilisant des variables macroéconomiques permet de faire. Altman et collab. (2005) étudient la relation entre la perte en cas de défaut et la probabilité de défaut en contexte de récession économique. Ils montrent que lorsque les taux de défaut augmentent, les taux de recouvrement diminuent. Dionne, Gauthier, Hammami, Maurice et Simonato (2011), à l'aide de modèles de Markov à changement de régimes, estiment également cette relation négative en tenant compte des cycles économiques. Ces résultats sont corroborés par Hu et Perraudin (2002) qui mettent en évidence que le recouvrement et le défaut aux États-Unis sont négativement corrélés de 20% en moyenne. Lorsque les auteurs considèrent uniquement les queues de la distribution, ils observent une corrélation négative de 30%. Ces études sont le résultat d'une augmentation du taux de défaut et de la dépréciation du taux de recouvrement durant la période de 1999 à 2002. Dans ces travaux, menés dans plusieurs pays, il est mis en évidence que le taux de recouvrement peut être volatile, d'autant plus que celui-ci tend à diminuer lorsque le nombre de défauts augmente, en période de récession économique.

Selon Schuermann (2004), il a été largement démontré que le taux de recouvrement en période de récession est inférieur, voire très inférieur, à ceux des périodes d'expansion économique, comme on peut le voir dans le tableau 1.1.

	Mean	Std. Dev.	25%	50%	75%	N
Recessions	32.07	26.86	10.00	25.00	48.50	322
Expansions	41.39	26.98	19.50	36.00	62.50	1,703
All	39.91	27.17	18.00	34.50	61.37	2,025

TABLE 1.1 – Taux de recouvrement selon les cycles économiques, *Moody's* 1970-2003.

De son côté, Frye et collab. (2000) vérifie, grâce à la base de données *Moody's Default Risk Service*, qu'en période de récession, le taux de recouvrement est trois fois plus petit que le taux de recouvrement en période d'expansion économique.

Carey (1998) ajoute, après avoir examiné des investissements privés dans 13 compagnies d'assurance vie, que la récession a une grande importance pour les queues de la distribution de la perte en cas défaut, particulièrement pour des niveaux inférieurs de séniorité de la dette. En effet, en utilisant des méthodes de simulation, il établit que les taux de recouvrement de créances de notation spéculative dans les queues de la distribution (de 99,90% à 99,95%) du taux de recouvrement en période de récession sont au moins deux fois plus bas qu'en période d'expansion économique. Néanmoins, cette différence s'estompe lorsque la notation de la dette s'améliore. Cela implique que les instruments de dette de notation spéculative sont plus sensibles au risque systématique, entre autres l'état de l'économie. Nous aborderons les principes de séniorité de la dette au sein des entreprises dans la section 2.2.

Dans cette optique, et pour satisfaire aux exigences du Comité de Bâle, nous choisirons comme contexte de notre étude une période de crise économique. Et comme nous le verrons dans la section 3.2, notre étude qui concerne uniquement les individus traite le taux de recouvrement du niveau le plus inférieur de séniorité.

1.1.3 Le défaut et la faillite — définitions et distinctions

Schuermann (2004) utilise la définition du défaut donnée par le comité de Bâle¹.

1. Basel Committee on Banking Supervision (2001c), section III.F, §146.

« By definition, a debt instrument can experience a loss only if there has been a default. However, there is no standard definition of what constitutes a default. Different definitions may be used for different purposes. »

En particulier, on dira qu'il y a défaut lorsque l'une de ces conditions est remplie :

- un prêt est classé comme prêt à intérêts non comptabilisés : on détermine que le débiteur ne pourra probablement pas honorer intégralement sa dette – valeur nominale, intérêts ou frais),
- une radiation a déjà eu lieu : une perte associée au prêt, par exemple une radiation, une provision spécifique ou une restructuration de la dette,
- le débiteur a des retards de paiement de plus de 90 jours, ou
- le débiteur a déclaré faillite.

Comme l'explique Schuermann (2004), plusieurs de ces conditions peuvent mener à un recouvrement total. En effet, la contrepartie pourrait ne pas avoir fait de paiements pendant plus de 90 jours, et ainsi être en défaut selon la définition de Bâle, mais s'acquitter par la suite de l'intégralité de sa dette : le créancier n'aurait ainsi subi aucune perte.

Il n'est donc pas étonnant que la littérature, comme nous le verrons par la suite, relate que la distribution de la perte en cas de défaut est bimodale. Par exemple, Schuermann (2004) distingue deux modes à cette distribution : un premier où le recouvrement est très élevé (entre 70% et 80%) et un autre où le recouvrement est beaucoup plus faible (entre 20% à 30%).

Pour s'assurer de ne pas sous-estimer ou surestimer la perte en cas de défaut, Schuermann (2004) conseillerait de modéliser certains événements de recouvrement séparément. Dans cette optique, nous nous proposons de traiter uniquement la dernière catégorie qui définit le défaut, soit le cas de la faillite.

La faillite se définit comme une procédure juridique qui peut permettre d'alléger le fardeau d'endettement de personnes qui sont dans l'incapacité de rembourser leurs dettes.

C'est une procédure de dernier recours qui est un sous-ensemble du défaut tel que nous venons de le définir. Le prochain chapitre de notre étude y sera consacré.

Après avoir fait le tour des composantes du risque de défaut, nous avons établi que nous nous pencherons sur la perte en cas de défaut. Nous avons déterminé que le défaut pouvait survenir dans le cadre de différents événements, dont la faillite. De plus, la perte en cas de défaut peut impliquer une entreprise ou un individu comme contrepartie. En ce qui nous concerne, notre travail s'articulera uniquement sur le taux de recouvrement lors d'une faillite effectuée par un individu. Or, il existe peu de littérature concernant la perte en cas de défaut lors de prêts à des individus et encore moins uniquement en cas de faillite. De ce fait, notre base de données étant assez unique, nous l'aborderons en nous inspirant de paramètres et d'applications de modèles dans des contextes relativement similaires. Aussi, nous élargirons notre revue de littérature de manière à nous inspirer des modèles et des variables utilisés non seulement pour la perte en cas de défaut des individus, mais aussi celle des entreprises.

1.2 Perte en cas de défaut des entreprises

Il existe, sur le sujet de la perte en cas de défaut, une littérature qui concerne les entreprises (*corporate*) et une autre axée sur les individus (*retail*). La littérature traitant le recouvrement des entreprises peut se décomposer en deux catégories principales : le recouvrement d'obligations corporatives et le recouvrement de prêts corporatifs. Bien que nous étudions le taux de recouvrement de portefeuilles d'individus, nous ferons un bref tour d'horizon de certains points majeurs et prééminents de la littérature concernant le recouvrement dans le cas de défaut d'entreprises.

1.2.1 Recouvrement d'obligations corporatives

La majorité des études traitent de l'estimation du recouvrement d'obligations corporatives. Wirenhammar (2011) confirme que de nombreuses études se sont concentrées

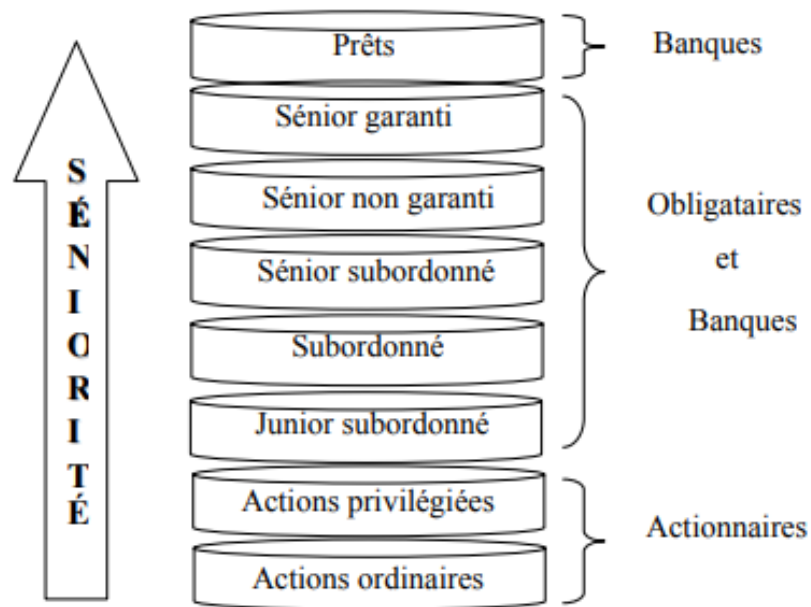


FIGURE 1.1 – Niveau de séniorité en fonction de la structure de capital des entreprises

sur des prêts aux entreprises puisque les données sont plus accessibles que celles des prêts aux consommateurs. De surcroît, Schuermann (2004) ajoute que la plupart des études sont faites sur le taux de recouvrement des obligations étant donnée l'abondance de données dans cette industrie. Il dresse le portrait de la structure de capital générale des entreprises afin d'expliquer les différents niveaux de séniorité dans la figure 1.1.

Il explique qu'aux États-Unis et dans plusieurs autres pays, il existe une règle de priorité très importante – *the Absolute Priority Rule (APR)* – parmi les créanciers de l'entreprise. Cette règle de priorité est fonction de la séniorité des prêts, c'est-à-dire que la tranche de séniorité supérieure sera remboursée intégralement, en cas de défaut, avant le paiement de la tranche subordonnée, et ainsi de suite.

Eberhart, Moore et Roenfeldt (1990) définissent cette règle de la manière suivante :

« The absolute priority rule [...] states that a bankrupt firm's value is to be distributed to suppliers of capital such that senior creditors are fully satisfied before any distributions are made to more junior creditors, and junior creditors are paid in full before common

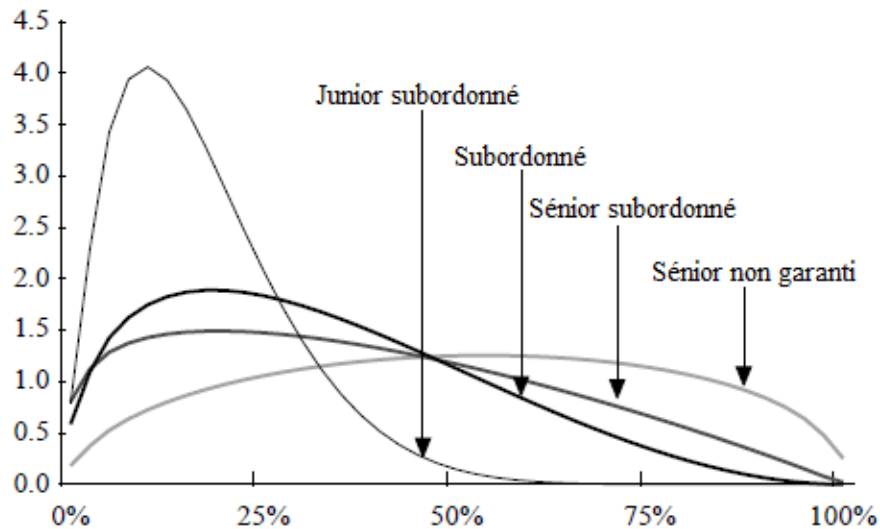


FIGURE 1.2 – Distribution beta du taux de recouvrement par séniorité. En estimant le taux de recouvrement par une distribution beta, la figure illustre les différences entre chacun des niveaux de séniorité des obligations. Il apparaît donc que plus le niveau de séniorité du prêt est prioritaire, plus le taux de recouvrement est élevé.

shareholders »

De son côté, CreditMetrics (1997) se base sur des études sur des obligations ayant fait défaut aux États-Unis après le 1^{er} octobre 1979, la date d'entrée en vigueur du *Bankruptcy Reform Act*. Les auteurs utilisent les prix au marché des obligations un mois après le défaut afin de calculer le taux de recouvrement.

CreditMetrics (1997) se penche sur la distribution du taux de recouvrement pour chaque niveau de séniorité des obligations et fait correspondre une distribution aux taux de recouvrement tout en préservant les bornes requises de 0 à 100%. Ils concluent que la distribution du taux de recouvrement, pour chaque type de séniorité, s'apparente à une distribution beta (voir figure 1.2), puisque celle-ci est bornée entre zéro et un.

Fridson, Garman et Okashima (2000) utilise les données d'Altman (2000)², qui traitent des défauts d'obligations entre 1978 et 1999. Il conclut son étude en expliquant que le re-

2. Altman, Edward I., Naeem Hukkawala et Vellore Kishore. *Defaults and Returns on High Yield Bonds : Analysis through 1999 and Default Outlook for 2000-2002*. New York University Salomon Center, Janvier 2000.

Séniorité	Moyenne	Écart-type	25%	50%	75%	N
Sénior garanti	54,26	25,82	33,00	53,50	75,00	433
Sénior non garanti	38,71	27,80	14,50	30,75	63,00	971
Sénior subordonné	28,51	23,41	10,00	23,00	42,25	260
Subordonné	34,65	22,23	19,50	30,29	45,25	347
Junior subordonné	14,39	8,99	9,13	13,00	19,13	12

TABLE 1.2 – Taux de recouvrement selon la séniorité de la dette

couvrement des obligations est très variable : une compagnie peut faire faillite avec beaucoup d’actifs, alors que dans d’autres cas, il arrive qu’il reste très peu pour les créanciers. Fridson et collab. (2000) démontre que le taux de recouvrement des créanciers diffère selon chaque niveau de séniorité des obligations, et lorsqu’on se penche sur le recouvrement au sein de chaque niveau, il est variable d’année en année.

Schuermann (2004) confirme les résultats de Fridson et collab. (2000) et de CreditMetrics (1997) en rapportant les statistiques descriptives de recouvrement pour des obligations américaines, grâce aux données de *Moody’s Default Risk Service Database*³, 1970-2003, par niveau de séniorité (voir tableau 1.2).

En conclusion, les auteurs semblent être du même avis en ce qui concerne le taux de recouvrement des obligations corporatives, c’est-à-dire que le niveau de séniorité est un déterminant majeur de celui-ci.

Par ailleurs, comme l’illustre la structure de capital présentée dans la figure 1.1, les prêts bancaires aux entreprises ont le niveau de séniorité le plus élevé. On devrait donc s’attendre à ce que le taux de recouvrement de ces prêts soit équivalent, ou même meilleur que celui des obligations. Amihud, Garbade et Kahan (2000) ajoutent que les prêteurs privés – soit les prêts bancaires – « contrôlent mieux les coûts d’agence de la dette grâce à des clauses plus restrictives, de meilleures négociations et une surveillance plus étroite ». Les banques sont ainsi en mesure d’exploiter leur relation avec la contrepartie afin de se hisser au plus haut niveau de la structure de séniorité de la dette, et donc de faire

3. À la fin de l’exercice 2003, la base de données Moody’s Default Risk Service contenait des informations sur 2958 dettes d’entreprises ayant fait défaut avec 2025 taux de recouvrement disponibles. De ces 2025, environ 88% de ces entreprises étaient domiciliées aux États-Unis.

augmenter leur taux de recouvrement.

1.2.2 Recouvrement de prêts bancaires aux entreprises

Plusieurs études portent sur l'estimation du taux de recouvrement des prêts bancaires accordés à des entreprises. Gupton, Gates et Carty (2000) examinent la perte en cas de défaut de prêts bancaires contractés par des entreprises. Ils envisagent de donner une estimation plus précise du taux de recouvrement, mais aussi d'examiner les facteurs qui influencent la perte en cas de défaut. Les données utilisées dans l'étude sont des prêts de banques américaines de type « syndicated loans ». Ceux-ci sont généralement d'un montant plus élevé que les prêts bancaires standards, et permettent de répartir le risque de défaut de l'emprunteur entre plusieurs prêteurs. Selon leur définition, un emprunteur est en défaut lorsqu'il dépose une procédure sous le « Chapter 11 » ou le « prepackaged Chapter 11 » du Bankruptcy Code des États-Unis. Le Chapitre 11 du *Bankruptcy Code* des États-Unis est l'équivalent de la procédure sous le régime de la *Loi sur les arrangements avec les créanciers des compagnies* (LACC) au Canada. La LACC est une loi fédérale autorisant les entreprises insolubles qui doivent plus de 5 millions de dollars à leurs créanciers à présenter un plan de restructuration de leurs affaires et de leurs finances. La base de données de Gupton et collab. (2000) contient un total de 181 prêts octroyés à 121 entreprises qui ont fait défaut de 1989 à 2000. Les auteurs déterminent six facteurs qui influencent le recouvrement de prêts bancaires accordés à des entreprises qui ont fait défaut :

La séniorité de la dette Les auteurs dressent un tableau récapitulatif des statistiques sur le recouvrement des prêts, ventilées par séniorité de la dette. Ces résultats convergent avec ceux des études sur le taux de recouvrement des obligations que nous avons analysées dans la section 1.2.1. En effet, la séniorité de la dette semble avoir un effet sur le taux de recouvrement, que ce soit dans le cas d'obligations corporatives ou de prêts à des entreprises.

Multiple loans Gupton et collab. (2000) arrivent à la conclusion que pour les dettes séniors non garanties, le recouvrement peut être différent en fonction du nombre de prêts octroyés au débiteur. En effet, pour une entreprise qui n'a contracté qu'un seul prêt, le taux de recouvrement moyen est de 63,4%, alors que pour une entreprise ayant contracté plusieurs prêts, le taux de recouvrement est de 36,8%. Ainsi, dans le cas de prêts non garantis, le fait d'avoir plusieurs prêts semble être une variable ayant un impact négatif sur le taux de recouvrement des prêteurs.

Le domaine d'activité de l'entreprise Bien que ce facteur soit considéré comme ayant une influence sur le recouvrement (Borenstein et Rose, 1995⁴), les auteurs ne peuvent conclure, sur la base de leurs données, que les entreprises ont un taux recouvrement différent selon leur domaine d'activité.

La cote de crédit initiale Moody's Les auteurs confirment l'hypothèse selon laquelle lorsque la cote de crédit diminue, le taux de recouvrement diminue généralement aussi.

Biens en garantie L'article conclut en indiquant que la présence d'un bien en garantie augmente de 17,4% le recouvrement. Également, Schuermann (2004) estime que la présence d'un collatéral dans le contrat du prêt, c'est-à-dire de biens en garantie, est l'un des facteurs principaux qui font varier le taux de recouvrement de la dette.

Le genre de dépôt de procédure L'article conclut que le type de dépôt de procédure (« prepackaged Chapter 11 ») a un effet considérable et positif sur le recouvrement ainsi que sur la rapidité de résolution du défaut.

Toujours en ce qui concerne la procédure aux États-Unis en vertu du Chapitre 11, Schuermann (2004) estime que le temps passé en faillite peut réduire considérablement la valeur de recouvrement pour les créanciers. Également, en ce qui a trait au temps passé en

4. S. Borenstein, N. L. Rose, 1995, « Bankruptcy and pricing behavior in US airline markets », *The American Economic Review*, Vol. 85, No 2, p. 397-402.

faillite, Helwege (1999) conclut que la présence de réclamations contingentes et que la taille des entreprises (qui est un instrument (*proxy*) pour la complexité du dossier) tend à faire augmenter la durée de la faillite.

En outre, Caouette, Altman, Narayanan et Nimmo (2011) cite l'étude de Neto de Carvalho et Dermine (2003)⁵, qui présente des données d'une banque portugaise sur des prêts en défaut accordés à 371 PME entre 1985 et 2000. Cette étude arrive aux trois conclusions suivantes :

- la fréquence de la distribution des pertes en cas de défaut est bimodale, avec de nombreux cas où il n'y a aucun recouvrement et d'autres aucune perte,
- la taille du prêt a un impact négatif statistiquement significatif sur le taux de recouvrement, et
- le type de bien en garantie est statistiquement significatif pour déterminer le recouvrement. Cela corrobore les résultats de Gupton et collab. (2000).

Après ce tour d'horizon des taux de recouvrement d'obligations et de prêts aux entreprises ainsi que des variables qui affectent ces taux, nous examinerons la formule utilisée pour calculer le taux de recouvrement.

1.2.3 Calcul du taux de recouvrement des entreprises

Définissons tout d'abord la relation entre la perte en cas de défaut et le taux de recouvrement. Nous utilisons les deux terminologies au cours de notre travail puisque :

$$\text{Perte en cas de défaut} = 1 - \text{Taux de recouvrement.}$$

Dans la littérature consultée pour les fins de notre étude, les chercheurs utilisent généralement les prix du marché des obligations afin de calculer le taux de recouvrement. Par exemple, Khieu, Mullineaux et Yi (2012) démontrent que ces prix au marché constituent de justes approximations du taux de recouvrement. En effet, le prix observé un mois

5. N. De Carvalho et J. Dermine, 2003, « Bank loan losses-given-default-empirical evidence », Tentative Draft.

après le défaut intègre avec précision les valeurs de réalisations futures des créanciers. On part donc de l'hypothèse que le marché est efficient et que les agents qui le composent estiment avec rationalité le recouvrement futur. Gupton et collab. (2000) expliquent que le prix au marché secondaire observé un mois après le défaut représente le « turnover » – soit le prix de transfert entre un investisseur qui est disposé à conserver le prêt et un investisseur disposé à détenir le prêt suite au défaut. Toutefois, cette mesure n'est pas applicable au contexte de notre recherche puisqu'il n'existe pas de marché secondaire pour des prêts aux individus.

Caouette et collab. (2011) confirment que le prix d'un marché secondaire d'obligations ou de prêts bancaires est souvent utilisé pour estimer le taux de recouvrement. Cependant, les auteurs ajoutent que Altman, Haldeman et Narayanan (1977) n'utilisaient pas le prix des marchés secondaires pour estimer le taux de recouvrement, parce que ces marchés n'existaient pas à l'époque. Ils utilisaient donc les sommes ultimement reçues par les créanciers pour établir le taux de recouvrement. De plus, ils n'actualisaient pas ces sommes à la date du défaut. Ce taux de recouvrement, que Altman et collab. (1977) appellent l'*Ultimate Nominal Recovery Rate*, est le ratio du recouvrement post-défaut non actualisé sur l'exposition au moment du défaut. Cette méthode de calcul sera utilisée dans le cadre de notre recherche.

1.3 Perte en cas de défaut des individus

Comme nous l'avons observé dans la section précédente, de nombreuses études ont été effectuées sur la perte en cas de défaut lorsque la contrepartie est une entreprise. Cela est notamment attribuable à la disponibilité des données. En ce qui nous concerne, nous avons choisi d'étudier la perte en cas de faillite des individus parce qu'il existe peu de recherches dans ce domaine et que, d'après les *Statistiques sur l'insolvabilité au Canada* de 2009⁶, le nombre de faillites déposées par des individus représente plus de 95% de toutes les faillites enregistrées au Canada, en 2009. En outre, sur les 17,5 milliards de

6. Bureau du surintendant des faillites, *Statistiques sur l'insolvabilité au Canada*, 2009.

dollars en perte des créanciers que représentent toutes les faillites au Canada (incluant celles des entreprises), 12 milliards sont attribuables aux individus, soit plus de deux tiers des pertes. Il nous apparaît donc d'autant plus pertinent et utile de modéliser adéquatement cette perte ainsi que les variables ayant un effet sur celle-ci.

En outre, sous l'approche interne du calcul du capital réglementaire, le Comité de Bâle (2006) stipule que « pour une exposition de type *retail*, les banques doivent fournir leur propre estimation de la PD, de l'EAD et de la PCD ». Le comité définit plusieurs groupes de type *retail* : celui des prêts à des PME de moins de un million d'euros (cautionnés ou pas par un individu) ; celui des prêts hypothécaires à des particuliers ; et celui de l'exposition à des individus, c'est-à-dire le crédit et les lignes de crédit, mais aussi toute la catégorie des prêts personnels. Dans notre étude, nous nous concentrerons plutôt sur la dernière catégorie, celle du *retail*, et notamment les prêts à des consommateurs individuels.

Cette section traite donc des études sur la perte en cas de défaut d'individus, notre objectif étant de déterminer quelles variables et quels modèles utilisent les chercheurs. Nous commencerons par présenter les variables explicatives utilisées dans la littérature, puis nous aborderons des modèles de prédiction de la perte en cas de défaut d'individus, avec des variables macro et microéconomiques, pour nous concentrer finalement sur des variables spécifiques aux prêts individuels. Nous finirons cette section avec des études décrivant le taux de recouvrement dans un contexte d'insolvabilité individuelle. Ultime-ment, nous serons en mesure de sélectionner les variables les plus pertinentes et d'appliquer les modèles les plus adaptables à notre base de données sur les faillites individuelles au Canada.

1.3.1 Variables explicatives du taux de recouvrement

Comme nous l'avons vu, les accords de Bâle exigent un calcul de la perte en cas de défaut qui tienne compte des périodes de crise économique. À cet effet, Wirenhammar

(2011) étudie les pertes en cas de défaut en période de récession économique. L'étude a été réalisée à partir des données d'une banque (Nordea) sur des portefeuilles de prêts personnels (*retail*) dans les pays nordiques : Finlande, Norvège, Danemark, Suède. Le but de l'étude est de permettre à cette banque de calculer des pertes en cas de défaut en période de récession et de les utiliser à l'interne et dans les calculs d'exigence de capital réglementaire de la banque.

Wirenhammar (2011) utilise des données de 2002 à 2006 pour calculer la valeur du recouvrement réalisé post-défaut, c'est-à-dire, comme nous l'avons vu dans la section 1.2.3, le *Ultimate Recovery Rate*.

Grâce à ces données, l'auteur développe un modèle de régression « hors crise » avec des variables macroéconomiques, et utilise ensuite la valeur de ces variables durant la crise économique des années 1990 afin de prédire une perte en cas de défaut en période de récession économique. Il dresse une liste de variables macroéconomiques pouvant affecter la perte en cas de défaut de prêts aux consommateurs :

- taux de chômage
- ratio de la dette sur le revenu disponible
- PIB (produit intérieur brut)
- revenu disponible
- valeurs immobilières

Outre les variables macroéconomiques, l'auteur propose d'ajouter des variables individuelles qui sont des microfacteurs devant être appliqués au niveau du prêt :

- côte de crédit
- domaine d'activité
- région géographique

L'auteur laisse supposer que la perte en cas de défaut pourrait varier en fonction de la région géographique. En effet, deux pays peuvent avoir des modèles culturels tels que le défaut sur une dette ou la faillite est plus acceptable dans l'un que dans l'autre. L'objectif

de cette variable est de saisir les différences qui pourraient exister entre la culture de la dette et le défaut de paiement, selon les pays.

Wirenhammar (2011) procède à des régressions multiples par pays et par catégorie de prêts à l'aide de différents indicateurs macro-économiques. Il dresse trois catégories de prêts :

- résidentiel : principalement des hypothèques
- autres : autres prêts ayant un actif en garantie
- non garantis : par exemple, cartes de crédit, prêts bancaires

L'auteur développe un modèle linéaire multifacteurs en utilisant des variables macroéconomiques pour expliquer les différences dans le taux de recouvrement des trois catégories de prêt et des variables non macroéconomiques pour expliquer les variations du taux de recouvrement entre chaque prêt. L'article conclut que le modèle le plus prometteur est celui qui utilise les microfacteurs, c'est-à-dire des variables individuelles au niveau du prêt : ces variables permettent une modélisation réaliste, intuitive et mathématiquement simple du taux de recouvrement.

Toutefois, Wirenhammar (2011) démontre que le caractère positif ou négatif de certains coefficients des variables macroéconomiques ne correspond pas à ses attentes et ne suit pas une logique économique acceptable. À la lumière de ces résultats, nous sommes confrontés de ne pas être en mesure d'utiliser des variables macroéconomiques puisque nos données sont sur une période d'un an.

1.3.2 Prédiction du taux de recouvrement

Zhang et Thomas (2012) essaient de prédire le taux de recouvrement en utilisant deux différents types de modèles, d'une part des modèles de régression et d'analyse de survie qui utilisent une seule distribution et, d'autre part, des modèles qui utilisent des combinaisons de distributions afin de se rapprocher de la distribution réelle des pertes en cas de défaut. Par exemple, les modèles peuvent être améliorés en segmentant la population et en

Recovery Rate	R-square	Spearman	MAE	MSE
(1) Linear Regression	0.0904	0.29593	0.3682	0.1675
(2) A – Weibull	0.0598	0.25306	0.3586	0.2042
(3) A – log-logistic	0.0638	0.25990	0.3560	0.2060
(4) A – gamma	0.0527	0.23496	0.3635	0.2015
(5) Cox – including 0's	0.0673	0.27261	0.3546	0.2006
(6) Cox – excluding 0's	0.0609	0.25506	0.3564	0.2072
(7) Linear Regression*	0.0292	0.22837	0.4077	0.2432
(8) A – weibull*	0.0544	0.24410	0.3606	0.2070
(9) A – log-logistic*	0.0591	0.25315	0.3575	0.2077
(10) Cox – including 0's*	0.0425	0.22646	0.3693	0.2216
(11) Cox – excluding 0's*	0.0504	0.23269	0.3624	0.2108

TABLE 1.3 – Comparaison de la qualité des modèles de prédiction de Zhang et Thomas (2012). Les modèles avec le recouvrement en dollars sont notés avec un *. *Mean Absolute Error* (MAE) représente l'erreur absolue moyenne et *Mean Squared Error* (MSE) représente l'erreur quadratique moyenne.

obtenant différentes estimations pour chaque segment, puisque chaque sous-groupe peut avoir des caractéristiques différentes et un taux de recouvrement distribué différemment.

Les données utilisées consistent en plus de 27 000 prêts personnels en défaut accordés par une banque du Royaume-Uni de 1987 à 1999. Les paiements relatifs à ces dettes en défaut ont néanmoins été enregistrés jusqu'à la fin de l'année 2003. Les auteurs utilisent un Ultimate Nominal Recovery Rate, et le taux de recouvrement moyen de leur base de données est de 42%.

Zhang et Thomas (2012) construisent un modèle de régression linéaire multiple avec le taux de recouvrement en pourcentage comme variable dépendante et un autre modèle avec le recouvrement en dollars en tant que variable dépendante (voir le tableau 1.3). Ils concluent que le modèle de prédiction avec le taux de recouvrement en pourcentage donne de meilleurs résultats que celui qui utilise le recouvrement en dollars. Nous nous baserons donc sur ces résultats pour notre recherche et utiliserons le taux de recouvrement en pourcentage en tant que variable dépendante.

Les auteurs établissent que les modèles de prédictions à l'aide de régressions linéaires

sont plus performants que les modèles d'analyse de survie, en se basant sur le R^2 ainsi que le coefficient de rangs de Spearman, et ce, pour des prêts non garantis à des individus. Le R^2 , aussi connu sous le nom de coefficient de détermination, sera défini au Chapitre 4. Celui-ci mesure la qualité de prédiction du modèle puisqu'il représente la proportion de la variance prédite de la variable dépendante à partir des variables explicatives.

Rappelons que le coefficient de corrélation de rang de Spearman est un test non paramétrique qui ne suppose pas la normalité des variables étudiés. Par ailleurs, alors que l'erreur absolue moyenne représente la moyenne sur l'échantillon test des différences absolues entre la prévision et l'observation réelle, l'erreur quadratique moyenne mesure la magnitude moyenne de l'erreur. Ce sont deux scores orientés négativement, ce qui signifie que les valeurs basses sont meilleures.

Zhang et Thomas (2012) présentent finalement les variables significatives qu'ils ont trouvées dans leurs modèles : le ratio du montant en défaut face au montant du prêt total est la variable la plus significative, et a une relation négative avec le taux de recouvrement. La seconde variable la plus significative est le fait d'avoir un deuxième demandeur sur le prêt, ou un cosignataire, ce qui augmente le taux de recouvrement. Le statut d'emploi, le statut résidentiel et l'exposition au défaut sont aussi des variables significatives de leurs modèles.

Prívará, Kolman, Witzany et collab. (2014), tout comme Zhang et Thomas (2012), cherchent à prédire le taux de recouvrement. Leurs données, qui sont des prêts personnels non garantis en défaut, proviennent d'une grande banque tchèque. Les auteurs ont eu accès à 10 000 prêts choisis de manière aléatoire, sur une période de 10 ans, débutant en 2000. Cependant, l'objectif de l'étude n'est pas de trouver un modèle de perte en cas de défaut en période de récession économique, mais plutôt de comparer la qualité de différents modèles de prédiction du taux de recouvrement.

Prívará et collab. (2014) retranscrit la distribution du taux de recouvrement dans la figure 1.3. Comme dans de nombreuses autres études, la distribution est bimodale, avec un poids élevé sur les 0 et les 1. Dans 28% des cas, il n'y a aucun recouvrement tandis que 15% des individus ont repayé intégralement la dette. Les auteurs rapportent un taux

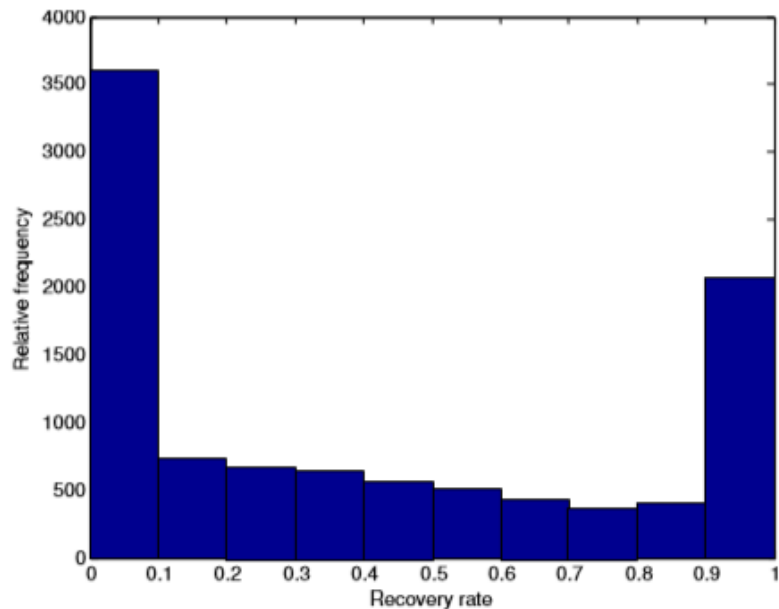


FIGURE 1.3 – Distribution du taux de recouvrement de Prívará et collab. (2014)

de recouvrement moyen de 40%.

Les auteurs comparent la régression linéaire et logistique à l’analyse de survie (*Survival Analysis Models*). L’analyse de survie permet généralement de traiter l’information partielle censurée en utilisant toute l’information à disposition au moment de la modélisation. L’analyse de survie permet de prédire le taux de recouvrement sans pour autant que le défaut soit terminé et que les paiements post-défaut soient observables. En ce qui concerne notre étude, et contrairement aux auteurs, la procédure de défaut est terminée et nous avons accès aux paiements post-défaut.

Tout comme Zhang et Thomas (2012), Prívará et collab. (2014) comparent la qualité des modèles, entre autres, au moyen du R^2 . Généralement, les modèles de prédiction du taux de recouvrement présentent un R carré entre 3% et 12%, comme nous l’avons vu dans Zhang et Thomas (2012). Toutefois, certains modèles de prédiction que Prívará et collab. (2014) obtiennent affichent un R^2 de 15%, notamment les modèles de survie.

Voici les variables explicatives utilisées dans leur modèle :

- éducation

- maison
- emploi
- sexe
- âge
- état civil

Les auteurs constatent que le fait d’être un homme a un effet significativement négatif sur le taux de recouvrement lorsqu’ils emploient la régression linéaire en guise de méthode de prédiction. Ils observent également, en utilisant le modèle de Cox, qui fait partie d’une classe de modèle de survie, que plus le niveau d’éducation est élevé, plus le taux de recouvrement est élevé et que les individus mariés ont un taux de recouvrement plus faible. Puisque nous avons des variables explicatives semblables à Prívará et collab. (2014), nous serons en mesure de comparer nos résultats aux leurs.

Prívará et collab. (2014) concluent que les modèles de survie semblent plus performants que les modèles de régression, contrairement à Zhang et Thomas (2012), qui privilégient les modèles de régression linéaire. Toutefois, la nature de nos données ne nous permet pas d’utiliser des modèles d’analyse de survie.

1.3.3 Classification de taux de recouvrement par type de dette

Hoechstóetter, Nazemi, Rachev et Bozic (2012) cherchent à créer un modèle capable de prédire le taux de recouvrement d’une agence de recouvrement de dette qui est externe à la banque ou au créancier initial. Les auteurs expliquent que la régression logistique est la plus utilisée pour analyser les taux de recouvrement. Comme les études que nous venons de voir, ils vont également utiliser des variables liées au débiteur ou individuelles, sans aucune variable liée à des facteurs macroéconomiques.

Les données que les auteurs utilisent sont constituées d’environ dix millions de dettes non garanties en défaut rachetées entre 2001 et 2010 par Arvato Infoscore, l’une des

plus grandes agences de recouvrement d'Allemagne. Les données sont classées en neuf catégories de prêts, qui sont fonction du domaine d'activité ou du type de dette :

- achat en ligne
- *business-to-business* (B2B)
- énergie et services
- services financiers, incluant les cartes de crédit
- secteur public
- note de débit
- télécommunication
- transport public
- divers, comprenant tout ce qui n'entre pas dans les autres catégories, par exemple, contraventions impayées.

Dans cette étude, chaque dette est unique, c'est-à-dire qu'elle ne peut qu'appartenir à une seule catégorie de prêt; cependant, un individu peut détenir plusieurs dettes différentes. Le taux de recouvrement est ici spécifique à une catégorie de prêt et non à un individu particulier, et est donc ventilé en fonction du type de dette. Les auteurs constatent que le taux de recouvrement fluctue beaucoup d'une catégorie à une autre et qu'il se situe en moyenne entre 9,6% et 38,4%. Alors que le secteur public possède le taux de recouvrement le plus faible, la catégorie des achats en ligne présente le taux de recouvrement le plus élevé des neuf catégories. Les services financiers, quant à eux, ont le montant moyen le plus élevé de dettes (2174\$), tout en ayant l'un des recouvrements les plus faibles (10,99%).

1.3.4 Modélisation du taux de recouvrement

Comme la majorité des auteurs qui étudient la perte en cas de défaut, Hoechstoetter et collab. (2012) font face à une distribution bimodale du taux de recouvrement, avec un recouvrement dans la plupart des cas nul et dans d'autres cas intégral. Pour pallier cette

distribution particulière, les auteurs construisent un modèle en deux parties. La première partie consiste à effectuer une régression logistique sur la variable dépendante, soit le taux de recouvrement, défini de la manière suivante :

$$y = \begin{cases} 1 & \text{si extrême} \\ 0 & \text{sinon} \end{cases} .$$

Lorsque le taux de recouvrement est égal à 0% ou 100%, $y = 1$, et l'on dira que le taux de recouvrement est extrême. Dans les autres cas, soit lorsque le taux de recouvrement est continu et se situe exclusivement entre 0% et 100%, $y = 0$, et l'on dira que le taux de recouvrement est non extrême. Ainsi, ce modèle de régression logistique nous permet d'obtenir la probabilité d'observer un taux de recouvrement extrême.

La première partie de ce modèle propose les variables individuelles suivantes ayant un impact significatif sur la prédiction du taux de recouvrement :

- montant de la dette au moment du défaut
- âge du débiteur
- temps écoulé entre le moment du défaut et le moment du rachat de la dette par l'agence de recouvrement
- côte de crédit (classification de la solvabilité selon une échelle ordinale à sept niveaux)
- adresse (validité de l'adresse du débiteur)
- type de débiteur (homme, femme ou personne morale)

Bien que l'effet – positif ou négatif – de certaines de ces variables sur le taux de recouvrement diffère d'une catégorie de prêt à l'autre, elles ont toutes un impact statistiquement significatif sur le taux de recouvrement.

La deuxième partie du modèle consiste à prédire les taux de recouvrement non extrêmes, lorsque $y = 0$. Pour ce faire, Hoehstoetter et collab. (2012) utilisent un modèle

de régression linéaire avec les mêmes variables explicatives individuelles que celles utilisées dans la première partie du modèle. Les auteurs concluent que ces variables sont toujours statistiquement significatives, et ce, pour tous les types de dette.

De leur côté, Thomas, Matuszyk et Moore (2012) utilisent une technique de modélisation similaire. Ils étudient la perte en cas de défaut d'une agence de recouvrement externe à l'émetteur initial de la dette et rapportent un taux de recouvrement nul dans 83% de leurs observations.

Thomas et collab. (2012) construisent également un modèle en deux parties. En effet, puisque 83% des recouvrements observés sont nuls, les auteurs jugent nécessaire de séparer la base de données en deux groupes. Leur première étape est donc d'estimer la probabilité d'observer un taux de recouvrement non nul au moyen d'une régression logistique. Une fois établie la probabilité d'avoir un recouvrement non nul, la deuxième étape consiste à prédire le taux de recouvrement à condition qu'il soit non nul à l'aide de régressions linéaires et d'autres modèles.

Étant donné que la distribution de nos données est similaire, puisque plus de 30% des taux de recouvrement que nous observons sont nuls, nous appliquerons également un modèle en deux parties, qui sera présenté dans le Chapitre 6 de notre étude.

1.3.5 Insolvabilité et taux de recouvrement

Nous abordons à présent des études traitant non seulement du défaut, mais, plus spécifiquement, des procédures d'insolvabilité des individus, soit la proposition de consommateur et la faillite. Il s'agit précisément du contexte particulier de notre recherche.

Tout d'abord, Eraslan, Koşar, Li et Sarte (2017) se penchent sur les procédures d'insolvabilité en vertu du Chapitre 13 du *United States Bankruptcy Code*. Ce chapitre, aussi connu sous le nom de « plan de salarié », permet aux individus qui ont un revenu d'emploi régulier d'établir un plan afin de payer l'intégralité ou une partie de leur endettement. Il est équivalent à la proposition canadienne en vertu de la LFI, que nous aborderons dans le Chapitre 2 de notre étude. Dans le Chapitre 13 du *United States Bankruptcy Code*, il existe

deux plans permettant aux débiteurs d'effectuer leurs paiements aux créanciers : un plan à court terme de 3 ans et un plan à long terme de 5 ans. Les débiteurs ont le choix entre l'un ou l'autre. Il est également pertinent de mentionner que le débiteur garde le contrôle de ses actifs et propose simplement aux créanciers un remboursement partiel de ses dettes. Cela signifie que ses actifs ne seront pas liquidés pour rembourser les créanciers. Ces derniers peuvent refuser cette proposition.

Eraslan et collab. (2017) ont mené l'étude sur un échantillon final de 821 dossiers sous le Chapitre 13, déposés dans l'État du Delaware sur une période d'un an (d'août 2001 à août 2002). Afin de calculer le taux de recouvrement, Eraslan et collab. (2017) utilisent le ratio des paiements effectués par le débiteur sur sa dette totale au moment du défaut. Ils emploient le même procédé que Altman et collab. (1977), c'est-à-dire le *Ultimate Nominal Recovery Rate*. En moyenne, le taux de recouvrement global pour la masse des créanciers est de 29%. Celui-ci est une moyenne pondérée des créanciers garantis et non garantis. De son côté, le taux de recouvrement moyen des créanciers non garantis est de 17%, avec une médiane de 0%.

Les figures 1.4 et 1.5 illustrent respectivement la distribution du taux de recouvrement du Chapitre 13 en fonction du montant initialement proposé par le débiteur et du montant ultimement reçu par les créanciers. En effet, les auteurs expliquent que la disparité des deux distributions provient essentiellement du fait que plusieurs débiteurs n'exécutent pas intégralement leur plan de salarier en vertu du Chapitre 13 et peuvent, volontairement ou non, faire défaut à leur plan initial. Aussi, si le revenu du débiteur change durant la période d'exécution de son plan, ce dernier doit être revu à la hausse ou à la baisse. Ainsi, le plan exécuté finalement peut considérablement différer du plan initialement proposé.

En ce qui nous concerne, nous avons uniquement accès à des dossiers de faillite, parmi lesquels certains étaient initialement des propositions. Alors que l'étude de Eraslan et collab. (2017) compare les taux de recouvrement perçus à ceux qui étaient initialement proposés, nous vérifierons, quant à nous, s'il existe des différences significatives entre le taux de recouvrement d'une faillite et celui d'un dossier de faillite qui était initialement en proposition. Dans le Chapitre 3 de notre étude, nous créerons une variable *Genre de dépôt*,

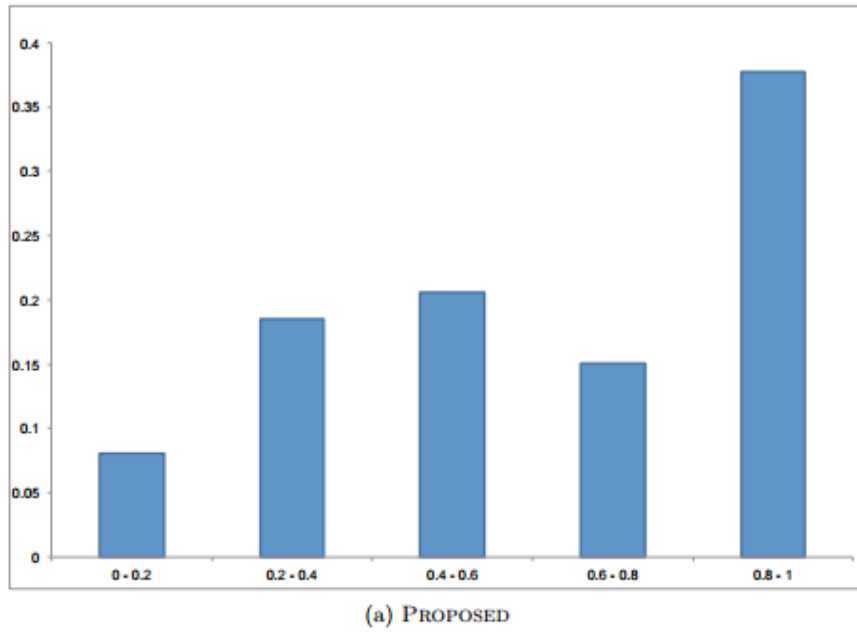


FIGURE 1.4 – Histogramme du taux de recouvrement proposé

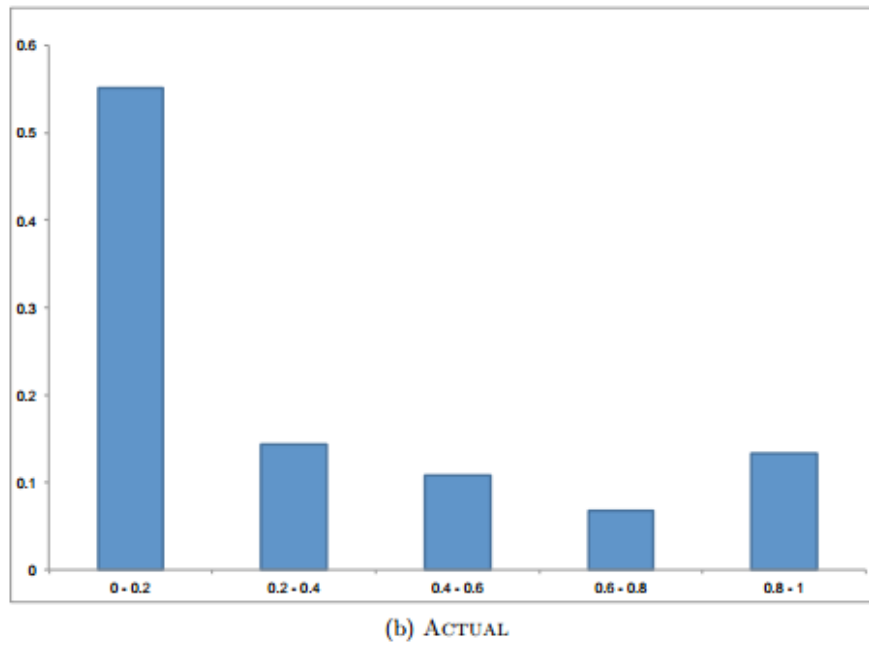


FIGURE 1.5 – Histogramme du taux de recouvrement ultimement reçu

qui nous permettra de tester cette hypothèse.

En ce qui a trait aux variables ayant un effet sur les taux de recouvrement, Eraslan et collab. (2017) concluent que le taux de recouvrement d'un débiteur qui en est à sa première faillite est en moyenne de 33%, contrairement à un récidiviste dont le recouvrement moyen est de 23%. Ils démontrent une relation négative entre le montant d'arriérés relatifs à l'endettement total du débiteur et le recouvrement. Aussi, les débiteurs ayant un ratio de revenu excédentaire relatif à leur dette totale plus élevé présentent un taux de recouvrement plus élevé. Finalement, la durée du plan de remboursement des créanciers semble avoir un effet sur la distribution du taux de recouvrement. En effet, les débiteurs qui proposent des paiements sur cinq ans ont un recouvrement médian (15%) inférieur à ceux qui proposent des paiements sur trois ans (19%), même si le recouvrement moyen est similaire.

Toujours aux États-Unis, le Chapitre 7 du *United States Bankruptcy Code* renvoie à la liquidation des biens du débiteur : celui-ci fait cession de ses biens, ce chapitre étant l'équivalent du processus de faillite au Canada. Dans ce cas-là, et contrairement au Chapitre 13 du *Bankruptcy Code*, le débiteur cède tous ses actifs et en confie la saisine à un fiduciaire – un syndic de faillite – qui les liquidera au profit des créanciers.

Flynn, Bermant et Bakewell (2002) observent que durant l'année terminée le 30 juin 2002, dans près de 96% des procédures déposées sous le Chapitre 7, le taux de recouvrement des créanciers a été nul. En effet, tout comme dans les procédures de faillites sommaires au Canada, il n'y a presque aucun recouvrement : les débiteurs ne possèdent généralement pas d'actifs, ne sont pas employés ou ne présentent aucun revenu excédentaire, c'est-à-dire qu'ils ont un budget très serré. De ce fait, en 2005, le *Bankruptcy Abuse Prevention and Consumer Protection Act* oriente en premier lieu les débiteurs ayant un revenu vers le Chapitre 13, avant qu'ils déposent une faillite en vertu du Chapitre 7. Il en est de même pour la législation canadienne avec le concept de proposition aux créanciers.

En partant de l'hypothèse que le débiteur ne proposera pas moins à ses créanciers que la valeur nette de ses actifs, on devrait s'attendre à ce que le taux de recouvrement des procédures d'insolvabilité aux États-Unis en vertu du Chapitre 13 soit plus élevé que celui

du Chapitre 7. En effet, selon Eraslan et collab. (2017), « *the debtor must propose to pay at least as much as the value of the assets creditors would have otherwise received under Chapter 7* ». Le processus étant similaire au Canada, la situation devrait être équivalente pour le système d'insolvabilité canadien. Aussi pourrait-on s'attendre à ce que notre taux de recouvrement soit inférieur au taux de recouvrement ultimement reçu par les créanciers dans l'étude de Eraslan et collab. (2017).

Après ce tour d'horizon de la littérature sur la perte en cas de défaut, nous nous inspirerons des modèles et des variables explicatives utilisés par les différents auteurs pour orienter notre recherche. Nos choix de variables seront aussi basés sur notre connaissance du domaine de l'insolvabilité.

1.3.6 Calcul du taux de recouvrement des prêts individuels

Comme nous l'avons présenté dans la section 1.2.3, le calcul du taux de recouvrement d'obligations ou de prêts octroyés à des entreprises est similaire à celui des prêts aux individus. Chalupka et Kopecsni (2008) proposent deux manières de calculer les taux de recouvrement : la perte en cas de défaut du marché (*market LGD*) et la perte en cas de défaut ex-post (*workout LGD*).

La perte en cas de défaut du marché est calculée lorsqu'on observe la valeur d'une dette sur le marché, après le défaut, en supposant que le marché reflète le juste prix ou la juste valeur de la dette, en fonction du recouvrement anticipé d'entrées de fonds futures. C'est donc un bon instrument pour estimer le taux de recouvrement. Cependant, la perte en cas de défaut du marché est beaucoup plus utilisée pour le calcul du recouvrement des prêts aux compagnies, surtout dans le cas d'obligations, et avec raison : c'est un marché beaucoup plus transparent, liquide et fiable.

En ce qui concerne les prêts aux individus, il existe peu de marchés pour ce type de dette. Certains portefeuilles de dettes de consommateurs sont cotés sur des marchés secondaires. Par exemple, dans le cas du recouvrement de dettes de cartes de crédit, il est

possible d'utiliser le prix de rachat de la dette, proposé par les agences de recouvrement, comme instrument pour estimer la juste valeur du recouvrement futur.

Toutefois, la majorité des études sur le taux de recouvrement *retail* utilisent la perte en cas de défaut ex-post.

La perte en cas de défaut ex-post, quant à elle, utilise les paiements effectués par le débiteur afin de rembourser sa dette durant la période de défaut. Pour calculer le taux de recouvrement de cette manière, il faut attendre que le défaut soit terminé, et calculer le ratio des paiements observés ex-post sur le solde de la dette au moment du défaut (EAD). Il est aussi possible d'actualiser les paiements effectués par le débiteur jusqu'au moment du défaut.

Prívará et collab. (2014) définissent le taux de recouvrement de la manière suivante :

$$TR = \frac{1}{EAD} \sum_{t=t_0}^{t_{\max}} \frac{CF_t}{(1+r)^t}$$

avec EAD l'exposition au moment du défaut, c'est-à-dire la valeur de la dette au moment du défaut. Les CF sont les entrées de fonds – les paiements du débiteur – après le défaut et actualisés au moment du défaut, avec r le taux d'intérêt approprié à l'actualisation de ces flux.

De leur côté, Zhang et Thomas (2012) et Eraslan et collab. (2017) n'actualisent pas les montants recouverts. Ils utilisent donc la valeur nominale de ces paiements, et leur taux de recouvrement se calcule de la manière suivante :

$$TR = \frac{\text{Montant recouvré}}{EAD}$$

En ce qui concerne notre recherche, nous utiliserons la valeur nominale des paiements pour calculer nos taux de recouvrement.

Analyse des différences

D'après Hoechstetter et collab. (2012), « la définition de la perte en cas de défaut des accords de Bâle soulève un problème ». En effet, certains paramètres ne sont pas précisés, d'où le questionnement des auteurs concernant la manière de calculer le taux de recouvrement. Par exemple, le taux de recouvrement peut être considéré comme le ratio du prix de la dette en défaut si elle était vendue immédiatement après le défaut sur le solde de la dette originale ; ou bien le ratio de la somme des paiements effectués par le débiteur et actualisés sur le solde de la dette originale.

Comme nous l'avons vu dans les sections sur le calcul du taux de recouvrement, nous constatons qu'il existe un prix au marché de la dette tant dans le cas des entreprises que des individus. Cependant, il est clair que dans le cas de prêts à des particuliers, il est plus rare et difficile d'observer des prix de portefeuilles de prêts individuels sur le marché, contrairement aux prix d'obligations ou même de prêts bancaires à de grandes compagnies.

En ce qui a trait au taux de recouvrement ex-post dans le cas d'individus, on parle de *market LGD*. Celui-ci est similaire au taux de recouvrement de Altman et collab. (1977), avec leur calcul du *Ultimate Nominal Recovery Rate* dans le cas de perte en cas de défaut d'entreprises. Le taux de recouvrement des dettes personnelles est le plus souvent calculé une fois la période de défaut terminée, avec la valeur des paiements ex-post. Il existe deux manières de faire :

- *Ultimate Recovery Rate* – comme l'ont préconisé Prívará et collab. (2014), en actualisant, au taux d'intérêt adéquat, les paiements du débiteur jusqu'à la date du défaut,
- *Ultimate Nominal Recovery Rate* – tout comme Zhang et Thomas (2012), qui utilisent la valeur nominale du montant recouvré à l'issue du défaut.

Nous favoriserons l'approche de Zhang et Thomas (2012). En effet, le taux d'actualisation adéquat suscite de nombreuses discussions. De plus, les faillites s'étalent sur de longues périodes (elles durent en moyenne trois ans et demi dans notre échantillon), ce qui est supérieur aux périodes de défaut abordées dans la littérature. Ainsi, la valeur de

ces paiements serait considérablement réduite, ce qui reverrait à la baisse notre taux de recouvrement. Finalement, même si le failli effectue des paiements mensuels au syndic et que le syndic liquide certains biens du failli, le syndic verse dans la majorité des cas un dividende aux créanciers uniquement à la fin de son administration. Par ailleurs, comme il est stipulé à l’alinéa 7(1)i) de l’instruction 5R5 de la *Loi sur la faillite et l’insolvabilité*, le syndic « doit déployer des efforts raisonnables pour obtenir un taux d’intérêt concurrentiel sur tous les comptes en fiducie », où l’argent de l’actif est déposé tout au long de l’administration de la faillite, comme nous verrons au Chapitre 2.

À la lumière des résultats de la littérature et des données disponibles, nous utiliserons le taux de recouvrement nominal reçu par les créanciers (*Ultimate Nominal Recovery Rate*) dans le cadre de notre étude.

Conclusion

Notre travail se concentrera sur le contexte particulier du défaut en situation de faillite. Nous nous intéresserons uniquement aux taux de recouvrement des individus, et non des entreprises. Nous n’utiliserons pas la valeur de la dette sur un marché secondaire, mais plutôt le taux de recouvrement ex-post.

À la lumière des auteurs que nous avons recensés dans notre revue de littérature, nous commencerons par un modèle d’estimation du taux de recouvrement à l’aide d’une régression linéaire multiple. Comme nous l’avons vu dans certaines études, nous proposerons par la suite un modèle en deux parties pour pallier les particularités de la distribution du taux de recouvrement.

Nous pouvons retenir entre autres de cette revue de littérature que la modélisation du taux de recouvrement diffère selon les divers instruments de dette. Nous avons établi que plusieurs variables permettent de l’estimer et de le prédire. Les types de dette peuvent se ventiler en fonction du niveau de séniorité, du type de garantie ou du domaine d’activité.

Comme nous l’avons vu, la distribution du taux de recouvrement diffère à l’échelle internationale – en effet, les auteurs supposent que le taux de recouvrement peut varier

de pays en pays. Par ailleurs, les variables qui l'expliquent diffèrent aussi. Dans cette optique, notre étude proposera un modèle pour le taux de recouvrement des individus applicable au Canada, conditionnel à la faillite. Nous serons en mesure d'expliquer, ex-post, les variations du taux de recouvrement, et proposerons un modèle de prédiction pertinent pour les créanciers non garantis. Ce modèle constitue, à notre avis, l'intérêt et l'originalité de notre travail, en raison du caractère unique de notre base de données.

Étant donné que nous n'avons accès qu'à une année de données, nous ne serons pas en mesure d'incorporer des variables macroéconomiques dans notre modèle, et, de ce fait, de confirmer, à l'instar des auteurs que nous avons étudiés dans notre revue de littérature, les variations du taux de recouvrement par rapport aux cycles économiques. Cependant, le taux de faillite a atteint son pic à 4,5 pour mille canadiens en 2009 – 2008-2009 étant des années de récession économique dans la conjoncture nord-américaine⁷. Ainsi, cette statistique confirme le fait que le nombre de défaut augmente en période de crise économique, et devrait également impliquer que le taux de recouvrement est moindre durant ce creux économique. De concert avec Bâle II, nous estimerons un modèle de perte en cas de défaut sur les données des faillites canadiennes déposées durant l'année 2009.

7. Selon l'encyclopédie canadienne, le sommet mensuel de la récession était en octobre 2008, et le creux mensuel se situait en mai 2009 au Canada

Chapitre 2

Faillite et règlementation au Canada

Cette étude étant centrée sur le recouvrement dans le cadre du système de faillite au Canada, nous avons jugé essentiel de faire le tour des éléments clés et du déroulement du processus de faillite, afin de notamment comprendre certains choix de variables dans notre modèle.

2.1 L'insolvabilité

Comme nous l'avons vu auparavant, il faut faire une distinction entre la faillite et le défaut. En effet, il y a deux mécanismes, régis par la *Loi sur la faillite et l'insolvabilité* (LFI), qui permettent de régler les conflits engendrés lorsqu'une personne est insolvable¹ : la faillite et la proposition. Ces dernières, en vertu de l'article 69.3 LFI, sont

1. Définition d'une personne insolvable, article 2 de la LFI : « Personne qui n'est pas en faillite et qui réside au Canada ou y exerce ses activités ou qui a des biens au Canada, dont les obligations, constituant à l'égard de ses créanciers des réclamations prouvables aux termes de la présente loi, s'élèvent à mille dollars et, selon le cas :

- (a) qui, pour une raison quelconque, est incapable de faire honneur à ses obligations au fur et à mesure de leur échéance ;
- (b) qui a cessé d'acquitter ses obligations courantes dans le cours ordinaire des affaires au fur et à mesure de leur échéance ;
- (c) dont la totalité des biens n'est pas suffisante, d'après une juste estimation, ou ne suffirait pas, s'il en était disposé lors d'une vente bien conduite par autorité de justice, pour permettre l'acquittement de toutes ses obligations échues ou à échoir. »

administrées par un syndic de faillite, et sont sujettes à la suspension des procédures, c'est-à-dire une suspension temporaire du droit de recours habituel des créanciers qui se trouvent dans l'incapacité de recouvrer une réclamation prouvable². Cela signifie, notamment, qu'à compter de la faillite d'un débiteur, les créanciers sont dans l'incapacité de poursuivre le débiteur pour obtenir la somme d'argent qui leur est due, de continuer les poursuites judiciaires déjà intentées et de prendre toute autre mesure visant à recouvrer ce qui leur est dû. Nous y reviendrons en temps et lieu.

2.2 Séniorité de la dette

Comme nous l'avons vu avec CreditMetrics (1997), Schuermann (2004) et l'*Absolute Priority Rule*, il existe une structure de capital par rapport à la séniorité de la dette pour les obligations corporatives. De la même manière, dans la faillite au Canada, les réclamations des créanciers répondent à un ordre de priorité. Commençons par le début.

2.2.1 Créanciers garantis

L'article 2 de la LFI définit le créancier garanti comme suit : « Personne titulaire d'une hypothèque, d'un gage, d'une charge ou d'un privilège sur ou contre les biens du débiteur ou une partie de ses biens, à titre de garantie d'une dette échue ou à échoir [...] ». Autrement dit, dans un cas de défaut du débiteur, le créancier garanti pourrait amorcer des procédures afin de réaliser sa garantie, pour ultimement rembourser la dette échue ou à échoir. Selon Deslauriers (2004), les créanciers garantis peuvent être étrangers à la faillite dans une certaine mesure parce qu'ils ne participent pas à la distribution des dividendes ; ils peuvent réaliser leurs garanties, ou les laisser réaliser par le syndic et être payés, selon leur rang hypothécaire, malgré la faillite. Nous discuterons plus tard des effets d'une procédure d'insolvabilité en vertu de la LFI sur une créance garantie. Notons cependant que notre travail ne visera pas à estimer le taux de recouvrement des créanciers garantis.

2. « Réclamation prouvable en matière de faillite » : il s'agit essentiellement d'une réclamation applicable à une dette non remboursée au créancier, que ce dernier peut présenter au syndic de faillite.

2.2.2 Créanciers privilégiés

Le plan de répartition, conforme à l'ordre de priorité des réclamations, est défini aux articles 136 à 141 de la LFI. En effet, la *Loi* ne définit pas la créance privilégiée, mais en vertu du paragraphe 136(1), sous réserve des droits des créanciers garantis, les montants réalisés provenant des biens d'un failli sont distribués d'après un ordre de priorité de paiement. Cet ordre vise certains montants dus, comme par exemple la rémunération du SAI, le prélèvement du Surintendant, des salaires impayés, des arrérages de loyer, etc.

Le paiement des créanciers privilégiés est un processus en cascade, c'est-à-dire que l'on paie en priorité les premiers créanciers sur la liste, puis une fois ces derniers totalement payés, on passe au paiement des deuxièmes sur la liste, et ainsi de suite, jusqu'aux derniers créanciers privilégiés. À noter que toutes ces catégories de créanciers ne sont pas forcément présentes au sein de la même faillite et qu'il est même fort probable qu'il n'y ait aucun créancier privilégié dans une faillite donnée.

Une fois le paiement de tous les créanciers privilégiés effectué, le reste est distribué au prorata de la réclamation de chaque créancier non garanti.

2.2.3 Créanciers non garantis

Les créanciers non garantis sont ceux qui ne sont pas définis ci-dessus, tel que décrit au paragraphe 136(3) de la LFI : « Tout créancier dont le présent article restreint les droits prend rang comme créancier non garanti, quant à tout solde de réclamation qui lui est dû ».

Ainsi, dans le cas d'une faillite, c'est une fois que la tranche des créanciers privilégiés a été remboursée intégralement que les créanciers non garantis ont droit à un paiement. « Les créanciers non garantis sont payés au prorata de leurs créances, à même les biens compris dans la faillite et après les créanciers privilégiés » (Deslauriers, 2004).

Deslauriers (2004) conclut que le créancier non garanti est celui qui ne détient ni priorité, ni hypothèque, ni priorité constitutive de droit réel, ni droit de rétention, ni faculté

de rachat ou réserve de propriété sur les biens du failli pour garantir le paiement de sa créance.

2.3 Recours des créanciers avant la faillite

Comme nous l'avons vu à la section 1.1.3, il existe une distinction entre le défaut et la faillite, ainsi que les recours à la disposition des créanciers durant la période de défaut du débiteur. Il nous paraît pertinent d'énoncer les mesures que les créanciers peuvent prendre une fois la contrepartie en défaut, mais avant que celle-ci ne fasse faillite, ce qui, comme nous le savons, entraîne la suspension des procédures des créanciers.

En effet, selon Hoehstoetter et collab. (2012), les créanciers ont des recours afin d'augmenter le recouvrement de ce qui leur est dû. Ces mesures sont, entre autres, des rappels, des lettres, des appels téléphoniques et éventuellement des actions en justice telles que des ordonnances judiciaires ou l'envoi d'huissiers. Ces différentes mesures donnent lieu à certaines actions du débiteur, par exemple la négociation, entre les deux parties, d'un plan de remboursement.

Eraslan et collab. (2017) ajoutent que lorsqu'un débiteur fait défaut sur sa dette sans pour autant déclarer faillite, les créanciers garantis – hypothécaires ou de prêts automobiles – saisissent leur bien en garantie afin de récupérer ce qui leur est dû. De leur côté, les créanciers non garantis peuvent prendre les mesures proposées par Hoehstoetter et collab. (2012). Eraslan et collab. (2017) ajoutent que les créanciers non garantis peuvent aussi obtenir des jugements de la cour afin de saisir, à même son employeur, le salaire du débiteur – c'est ce qu'on appelle la saisie de salaire. Les créanciers non garantis ont aussi la possibilité de faire saisir le compte bancaire du débiteur ou d'obtenir une hypothèque légale sur sa maison ou d'autres actifs qu'il possède. Cependant, les lois provinciales ou des différents états américains diffèrent et restreignent à certains degrés les montants et les divers actifs que les créanciers peuvent saisir. C'est pour cette raison que ces derniers processus peuvent être coûteux pour les créanciers non garantis et sont donc moins fréquents. En dernier recours, ces mêmes créanciers non garantis – généralement des émetteurs de

cartes de crédit – vendront leurs dettes à des agences de recouvrement.

Ces facteurs expliquent donc pourquoi le taux de recouvrement observé peut être plus élevé dans le cas de défaut – mais avant que le débiteur ne fasse faillite.

2.4 Syndic autorisé en insolvabilité (SAI)

Le syndic est une personne qui détient une licence et possède des pouvoirs conférés par la loi. Selon Deslauriers (2004), il agit en triple qualité. D'une part, il représente et conseille, d'une certaine façon, le débiteur dont il est le fiduciaire des biens et des droits accessoires. C'est ce que l'on appelle la saisine du syndic, puisqu'en vertu de la LFI, un failli cesse d'être habile à céder ou autrement aliéner ses biens, qui doivent immédiatement être dévolus au syndic. D'autre part, il représente les créanciers, en agissant pour sauvegarder leurs intérêts. Pour ce faire, il doit prendre possession du plus grand nombre de biens possible et les liquider de la façon la plus profitable pour payer les créanciers garantis, privilégiés et non garantis de la faillite. Le SAI cherche ainsi à maximiser la réalisation des biens du failli au bénéfice de la masse des créanciers. Enfin, il est aussi un officier de justice veillant à ce que la loi et l'ordre public soient respectés dans le cadre d'une faillite. Ainsi, en tant que fiduciaire au sens de l'article 2 du *Code criminel*, il est soumis aux codes de déontologie qui régissent sa conduite. À titre d'officier de justice il doit se comporter honnêtement, de façon à maintenir la confiance du public. Un code de déontologie régit l'exercice de la fonction de syndic (règles 34 à 53 de la LFI).

2.5 Proposition

Pour éviter une faillite, un débiteur peut déposer une proposition à ses créanciers, qui doit être acceptée par ceux-ci. La proposition suggère une offre pour rembourser aux créanciers un pourcentage du montant qui leur est dû, ou pour rééchelonner le montant de la dette, voire les deux. Les paiements sont effectués par l'intermédiaire du SAI, et

celui-ci utilise cet argent pour rembourser les créanciers selon les modalités prévues. Il existe deux propositions en vertu de la LFI : celles de la section I, pouvant être déposés par des entreprises et des particuliers ; celles de la section II, plus connues sous le nom de proposition de consommateurs, qui est un processus abrégé réservé aux personnes physiques. Le débiteur conserve la propriété de ses biens, contrairement à la faillite.

2.6 Faillite

Nous décrivons le processus de faillite en tenant compte de sa chronologie. Tout d'abord, il existe deux types d'administration dans la faillite : l'administration sommaire et l'administration ordinaire.

Faillite sommaire : c'est un processus abrégé – seule une personne physique (individu) peut déposer ce genre d'administration. Elle est réservée aux individus ayant, à priori, moins de 15 000 \$ de réalisation nette d'actifs. Le travail de réalisation des biens du SAI ainsi que les étapes du processus de faillite sont écourtés. Ce processus est extrêmement répandu, compte tenu du nombre d'individus faisant faillite, ces derniers ne présentant pas de potentiel de réalisation élevé, souvent parce qu'ils ne possèdent pas d'actifs. Ils représentent la grande majorité des faillites déposées au Canada.

Faillite ordinaire : la faillite d'une personne morale (compagnie) peut uniquement être un dossier d'administration ordinaire. Toutefois, une personne physique peut également déposer un dossier d'administration ordinaire. En effet, lorsque, de l'avis du Bureau du surintendant des faillites ou du syndic, les avoirs réalisables du failli, déduction faite des réclamations des créanciers garantis, dépasseront 15 000 \$, la faillite sera un dossier d'administration ordinaire. Cela signifie essentiellement que les dispositions de la LFI concernant l'administration sommaire ne s'appliqueront pas. Comme nous l'avons vu dans l'étude de Flynn et collab. (2002), les taux des recouvrements reportés dans 96% des cas étaient de zéro. C'est pour cette raison que notre étude se concentrera uniquement

sur les faillites ordinaires, qui ont un potentiel de recouvrement plus élevé que les faillites sommaires, comme nous venons de le définir.

Il est important de saisir la différence majeure entre une proposition et une faillite. La proposition équivaut à un contrat entre le débiteur et la masse des créanciers, exécuté par le syndic, proposant une somme d'argent qui peut être inférieure à la dette globale, tout en la rééchelonnant sur un horizon maximum de cinq ans. Les actifs du débiteur ne sont pas saisis par le syndic : ce n'est pas une liquidation d'actifs, comme l'est une faillite. La faillite, par contre, est une liquidation des actifs du débiteur dans le but de payer les créanciers.

2.6.1 Dépôt de la faillite

Une fois que la faillite ou la cession des biens est déposée auprès du SAI, la suspension des procédures des créanciers entre en vigueur. Le dépôt d'une faillite est accompagné, entre autres, des documents statutaires suivants :

Bilan

La cession doit être accompagnée d'un bilan, ou encore le formulaire 79. Il énumère l'actif et le passif du débiteur, y compris le nom et l'adresse de tous les créanciers et le montant de leur réclamation. Il contient des renseignements concernant le débiteur (comme les renseignements sur son conjoint et son employeur) et les activités qu'il peut avoir exercées au cours des cinq ans précédant l'ouverture de la faillite.

État des revenus et dépenses du failli

La faillite est également accompagnée d'un état mensuel courant des revenus et dépenses, ou encore le formulaire 65. Il inclut tous les revenus, bruts et nets; les détails de toutes les dépenses, y compris les dépenses pour besoins spéciaux, les pensions alimentaires, les paiements de soutien, les frais de transport, les dépenses médicales et les frais de médicaments, appelées dépenses non discrétionnaires. C'est à l'aide de cet état que le SAI

déterminera si le failli se qualifie au revenu excédentaire.

Une fois la faillite déposée, le SAI commence son administration de l'actif. Tout au long du processus, le SAI conservera le produit de la vente des actifs du failli et des paiements de celui-ci dans un compte en fiducie afin de le répartir entre ses créanciers.

2.6.2 Administration de l'actif

Les biens du failli constituent le patrimoine attribué à ses créanciers. À cet effet, le SAI vendra les biens du failli pour en distribuer le produit aux créanciers. Toutefois, comme il a été expliqué auparavant, certains biens sont détenus en garanti par des créanciers. Il y a généralement deux possibilités lorsqu'il s'agit de réaliser un bien donné en garanti :

- le bien en question possède une équité - c'est-à-dire que suite au paiement intégral du créancier garanti, il resterait un surplus d'argent pour rembourser le reste des créanciers non garantis. Dans ce cas, le SAI liquiderait le bien afin d'acquitter la créance garantie et de verser le surplus à la masse des créanciers,
- le bien ne possède pas d'équité - le SAI ne verrait donc pas d'intérêt de réaliser ce bien. Il ferait donc une « mainlevée » en faveur du créancier garanti qui pourrait alors se prévaloir de ses recours de réalisations à ses propres frais.

Autrement dit, les créanciers garantis ne participent que rarement à la distribution du SAI des montants disponibles.

Par ailleurs, en vertu de l'article 67 de la LFI, certains biens sont exclus du patrimoine attribué aux créanciers. Plus particulièrement, des exclusions notables sont en vertu des lois provinciales. En effet, toutes les provinces disposent d'une législation qui détermine quels sont les biens qui ne peuvent faire l'objet d'une mesure d'exécution ou de saisie. Par exemple, la législation provinciale peut rendre insaisissables des meubles, des vêtements ou des outils nécessaires pour exercer un métier jusqu'à concurrence d'un montant précis. Si l'on prend l'exemple du Québec, le *Code de procédure civile* prévoit que les

meubles d'une valeur de 7000\$ qui garnissent la résidence principale sont insaisissables. Toujours au Québec, le véhicule automobile peut être insaisissable s'il est nécessaire au maintien du revenu du travail ou s'il est nécessaire pour assurer la subsistance, les soins de santé ou d'éducation. Cela suggère donc un effet réduisant les taux de recouvrement, et ce, variant parmi les provinces. Nous testerons cet effet dans notre étude. Pour ce faire, nous construirons la variable *Région* dans au Chapitre 3.

En vertu de l'article 68 LFI, le revenu excédentaire est le montant du revenu total du failli supérieur au montant nécessaire pour maintenir un niveau de vie raisonnable, selon les normes établies par le Bureau du surintendant des faillites (présentées en Annexe A1). Si le revenu du failli est supérieur à la norme établie, ce dernier doit effectuer, à même ce revenu excédentaire, des paiements à l'actif de la faillite au bénéfice des créanciers.

À la fin de la procédure, le syndic répartit le montant disponible pour distribution entre les créanciers : c'est ce qu'on appelle, dans le jargon de la faillite, le *dividende*. Il ne faut pas confondre le dividende versé aux créanciers, auquel nous ferons référence tout au long de ce travail, au dividende d'un actionnaire. Notre dividende sera défini un peu plus bas.

2.6.3 Clôture de la faillite

Libération du failli

On dit d'une personne physique qui a fait faillite qu'elle est libérée de ses dettes lorsqu'elle obtient sa libération. Le débiteur sera libéré des dettes contractées avant la date de la faillite ; il ne sera pas libéré des dettes contractées après le jour où il est devenu failli, ces créances n'étant pas admissibles comme des réclamations prouvables dans la faillite.

Il existe deux types de libération : la libération automatique et la libération à la cour. Le tableau 2.1 résume le délai d'obtention de la libération automatique en vertu de l'article 168.1 de la LFI.

	Sans revenu excédentaire	Avec revenu excédentaire
Première faillite	9 mois	21 mois
Deuxième faillite	24 mois	36 mois

TABLE 2.1 – Délai d’obtention de la libération automatique

Cependant, il existe une procédure d’opposition à la libération du failli : le syndic, les créanciers ou le Bureau du surintendant des faillites peuvent s’opposer à la libération du failli dans certains cas (p. ex. si le failli n’a pas respecté les obligations et responsabilités qui lui sont imposées par la *Loi*).

En cas d’opposition à la libération du failli, celui-ci n’est plus admissible à sa libération automatique : il y aura une audition à la cour afin d’obtenir un jugement sur sa libération. Le failli n’a pas non plus droit à sa libération automatique s’il en est à sa troisième faillite ou plus. Pour obtenir sa libération, il devra également se présenter à une audition au tribunal et c’est à ce moment-là que la durée de la faillite sera déterminée.

État des recettes et débours du syndic

C’est en vertu de l’article 151 de la LFI que le syndic prépare l’état des recettes et débours, lorsque celui-ci a réalisé tous les biens du failli et qu’il partage la réalisation de ces biens entre les créanciers qui ont prouvé leur réclamation. Le relevé contient de l’information sur la réalisation de l’actif, sur les honoraires du syndic, sur ses débours administratifs, et sur le dividende versé aux créanciers. Le dividende, dans le jargon de la faillite et de l’insolvabilité, est le paiement effectué aux créanciers (qui ont prouvé leur réclamation) par le SAI à l’issue de son administration. Ce paiement est le produit de la liquidation des biens dont le SAI avait la saisine, distribué aux créanciers, suivant leurs rangs respectifs et au prorata des créanciers de même rang. C’est le dividende versé qui nous permettra de calculer notre taux de recouvrement.

Libération du syndic

Après avoir terminé son administration, le syndic doit demander sa libération. Il l'obtiendra après avoir rendu compte de l'administration et de la liquidation de tous les biens dont il a eu la saisine, et de la distribution du produit de cette liquidation sous forme de dividendes aux créanciers, en vertu de l'article 41(1) de la LFI.

La libération du syndic a lieu en vertu d'une ordonnance et elle relève le syndic de toute responsabilité découlant de l'exercice de ses fonctions à l'égard de ces actifs, sauf dans les cas de fraudes ou de manquements graves.

Il doit également, avant sa libération, faire fixer sa rémunération par les créanciers ou par le tribunal, au moyen du relevé des recettes et débours. Finalement, c'est à la date de libération du syndic que la suspension des procédures pour les créanciers est levée, en vertu de l'article 69.3(1.1) LFI.

Chapitre 3

Présentation des données

3.1 Description de la base de données

L'étude est effectuée grâce aux données du Bureau du surintendant des faillites. La base de données qui nous a été fournie contient de l'information sur les 2227 dossiers de faillites ordinaires d'individus qui ont été déposées entre le 1^{er} janvier 2009 et le 31 décembre 2009, à l'instar de Eraslan et collab. (2017) qui utilisaient une fenêtre d'un an de dépôts. La première partie de cette base de données a été reçue le 16 novembre 2015, suivie de la seconde partie le 15 janvier 2016. Les données étudiées n'ont pas été mises à jour depuis ces dates. L'information contenue dans cette base de données provient principalement de trois formulaires qui sont d'ordre public :

- Bilan (formulaire 79)
- État mensuel des revenus et dépenses du failli (formulaire 65)
- État des recettes et débours du syndic (formulaire 12)

Les deux premiers formulaires, soit le bilan et l'état mensuel des revenus et dépenses du failli, sont remplis au moment du dépôt de la faillite auprès du syndic autorisé en insolvabilité (SAI), sur la base des déclarations du failli, qui sont faites sous serment, au meilleur de ses connaissances. Le bilan contient des informations sur les actifs et les passifs du failli, tandis que l'état mensuel des revenus et dépenses du failli dresse un

portrait de sa situation financière. Quant à l'état des recettes et débours (ERD), il est rempli par le SAI à l'issue de son administration, c'est-à-dire vers la fin de la procédure de faillite, comme nous l'avons vu à la section 2.6.3.

Pour mener notre étude, nous sélectionnons uniquement les dossiers qui se sont soldés par le paiement de dividendes (positifs ou nuls – non vides). En effet, même si le dossier de faillite a été déposé en 2009, il est possible que l'administration du dossier ne soit pas terminée et qu'aucun dividende, au jour de la réception de la base de données, n'ait été versé aux créanciers. Par ailleurs, nous ne retenons que les dossiers dont la réclamation totale a été prouvée non nulle, c'est-à-dire où il y a au moins un créancier non garanti. Cela nous permettra de définir et de construire notre variable d'intérêt dans la section suivante. Notre échantillon est ainsi réduit à 1590 dossiers de faillites individuelles.

3.2 Mesures des variables et prévisions

Nous décrivons dans cette section de quelle manière nous avons construit notre variable d'intérêt et nos variables explicatives. Nous discuterons également de nos prévisions vis-à-vis de l'effet sur le taux de recouvrement que nous anticipons.

3.2.1 Variable dépendante du modèle

Comme l'illustre la revue de littérature, la plupart des études traitent de la perte en cas de défaut, du point de vue du créancier, généralement sur la base des prix au marché des obligations. L'originalité de notre étude est d'avoir accès au dividende versé à la masse des créanciers et ainsi d'obtenir le taux de recouvrement des faillites ordinaires. Comme nous l'avons vu au Chapitre 1, nous utiliserons, comme Altman et collab. (1977), le *Ultimate Nominal Recovery Rate*. Voici la définition de notre variable dépendante

$$y = \text{TR} = \frac{\text{Dividende} - \text{Prélèvement}}{\text{Total des réclamations prouvées}}. \quad (3.1)$$

De manière similaire, Zaniboni, de Araújo et de Avila Montini (2013) utilisent la formule suivante pour estimer le « *Workout LGD* »

$$\text{PCD} = 1 - \frac{\text{Paiements} - \text{Coût}}{\text{EAD}}, \quad (3.2)$$

où PCD représente le taux de perte en cas de défaut, les paiements représentent le montant récupéré à l'issue du défaut (notre dividende), et les coûts sont ceux associés au recouvrement - ils sont déjà pris en compte dans notre dividende, sauf le prélèvement. EAD représente l'exposition au moment du défaut qui, dans notre cas, est l'exposition totale de la masse des créanciers, soit le total des réclamations prouvées.

En effet, le dividende est le résidu des recettes totales moins les débours dans le dossier, qui comprennent les frais et dépenses administratives relatives à un dossier (telles que les frais de dépôt, les frais judiciaires, les taxes, les frais de consultation, etc.) et la rémunération du syndic. Comme nous l'avons vu au chapitre précédent, le dividende étant versé aux créanciers privilégiés et non garantis, l'étude du taux de recouvrement ne concerne que ces derniers, et exclut les créanciers qui possèdent des garanties sur certains actifs du failli.

Le prélèvement est celui du Bureau du surintendant des faillites. Le taux de prélèvement est prescrit par la Règle 123(1) de la *Loi sur la faillite et l'insolvabilité* (LFI). Celui-ci est fonction du montant total des paiements à la masse des créanciers (c'est-à-dire qu'il n'est pas fonction du montant versé à un créancier en particulier). Le taux est le suivant :

- 5% pour le premier million de dollars ;
- 1,25% pour le montant en sus de 1 000 000 \$, mais ne dépassant pas 2 000 000 \$;
- et
- 0,25% pour le montant en sus de 2 000 000 \$.

En règle générale, le prélèvement est payable à même le dividende versé par le syndic aux créanciers (privilégiés et non garantis). Il convient donc de le soustraire du dividende brut afin d'obtenir le dividende net versé aux créanciers.

Comme nous l'avons vu dans Zhang et Thomas (2012), le taux de recouvrement en pourcentage présente de meilleurs résultats dans le cas de modèles de prédiction. Afin d'obtenir un taux de recouvrement en pourcentage, nous divisons le dividende net versé par le total des réclamations prouvées. En effet, afin d'être en mesure de recevoir un dividende, le créancier doit prouver sa réclamation au SAI, comme nous l'avons vu à la section 2.6.3. Le taux de recouvrement représente donc le ratio entre le montant versé aux créanciers et la dette totale prouvée.

Finalement, notre taux de recouvrement est le taux de recouvrement global de la masse des créanciers privilégiés et non garantis. Nous n'avons pas l'information concernant le taux de recouvrement des créanciers qui possèdent une garantie sur un bien appartenant au failli.

3.2.2 Variables explicatives du modèle

Dionne, Artís et Guillén (1996) définissent trois groupes de variables pour expliquer le nombre de non-paiements d'un prêt dans une banque espagnole. Sur la base de notre compréhension et nos connaissances de la faillite et du recouvrement, nous avons été en mesure de classer, d'une manière similaire, les variables à notre disposition en trois catégories :

- Personnelles (Sexe, État civil, Nombre de membres dans l'unité familiale, Région);
- Financières (Actifs sur passifs, Types de dette et d'actif, Revenu d'emploi disponible);
- Faillite (Durée de la faillite, Genre de dépôt, Conditions à la libération, Rémunération du syndic).

Sexe

Puisque la littérature montre que les hommes ont généralement une probabilité de défaut plus élevée, et du fait de la relation inverse entre la probabilité de défaut et le

recouvrement, nous cherchons à tester si le fait d'être un homme conduit à un taux de recouvrement plus faible. Aussi, comme le montre l'étude de Prívará et collab. (2014), le fait d'être un homme a un effet négatif sur le taux de recouvrement. Nous documentons cette information grâce à la variable dichotomique construite de la manière suivante :

$$Sexe = \begin{cases} 1 & \text{si Homme} \\ 0 & \text{si Femme} \end{cases} .$$

La catégorie en référence (qui sera contenue dans la constante) sera Femme.

État civil

Nous avons accès à la variable catégorielle de l'état civil du failli, qui appartient à l'une des six catégories suivantes :

- Loi commune (*common law*);
- Divorcé;
- Marié;
- Séparé;
- Célibataire;
- Veuf.

Prívará et collab. (2014) observent un recouvrement plus faible chez les individus mariés. En effet, il est possible que les personnes mariées aient de plus en plus d'engagements financiers et soient donc plus endettées. On pourrait s'attendre alors à une dilution du taux de recouvrement puisque le dénominateur augmente. Mais d'un autre côté, on peut présumer que les personnes mariées aient accumulé plus d'actifs, ce qui ferait augmenter le numérateur du taux de recouvrement, la faillite étant la liquidation des actifs. La catégorie en référence (qui sera contenue dans la constante) sera Marié.

Nombre de membres dans l'unité familiale

Cette variable discrète représente le nombre de membres dans l'unité familiale incluant le failli inscrit au bilan. Notre intuition est que plus il y a aura de membres dans l'unité familiale, moins il y aura de recouvrement, puisque les paiements effectués à l'actif risquent d'être réduits. En effet, la norme du surintendant en ce qui a trait au revenu excédentaire, reportée en Annexe A1, se fonde directement sur le nombre de membres dans l'unité familiale. Il apparaît donc raisonnable de penser que l'augmentation de membres dans l'unité familiale diminue le risque du failli de se « qualifier » au revenu excédentaire, ce qui a pour effet de réduire la durée de la faillite, et par conséquent, les paiements faits à l'actif.

Région

Nous avons accès à la variable qualitative qui représente la région du failli. Comme nous l'avons vu dans l'étude de Wirenhammar (2011), nous sommes en mesure de séparer notre échantillon en quatre régions présentant plus ou moins le même nombre de dossiers, soit :

- Alberta ;
- Ontario ;
- Québec ;
- Autre, c'est-à-dire les neufs autres provinces et territoires du Canada.

Nous serons ainsi capables de mesurer le taux de recouvrement par région canadienne. Bien que nos régions soient au sein d'un même pays, et donc que la loi fédérale qu'est la LFI s'applique de la même manière, nous serons en mesure de capter les différences éventuelles concernant la culture de la dette dans les provinces canadiennes. De plus, les méthodes de travail des syndics, leur prise de risque et leur manière d'administrer un dossier pourraient être différentes de région en région. Aussi, comme nous l'avons vu au

Chapitre 2, les normes quant aux biens non saisissables ne sont pas uniformes au sein du Canada : elles varient de province en province. La variable Région s'avère intéressante, puisqu'elle permet peut-être de capter ces aspects des lois provinciales. La catégorie en référence (qui sera contenue dans la constante) sera Autre.

Actifs sur passifs

Les actifs sont définis comme le total estimatif des actifs détenus par le débiteur, tel que ce dernier l'a déclaré au syndic au moment du dépôt, dans son bilan. Ces actifs représentent la valeur des actifs totaux, qu'ils soient exemptés ou pas, et qu'ils soient garantis ou pas. Ils pourraient donc inclure la valeur estimée d'une maison en garantie ou d'une assurance insaisissable.

De la même manière, les passifs sont le total estimatif des passifs dont le débiteur est redevable, tel que ce dernier l'a déclaré au syndic au moment du dépôt. Ces passifs représentent le total des montants non garantis, garantis et privilégiés.

La variable Actifs sur passifs est le ratio entre les actifs totaux et les passifs totaux du failli.

En général, lors de l'évaluation de compagnies, les chercheurs utilisent le ratio d'endettement, soit passifs sur actifs. Celui-ci représente le montant total de dette relatif au montant total d'actif, et permet d'avoir une idée du niveau du levier d'endettement. Il sert donc à comparer le niveau d'endettement de différentes compagnies. Plus ce ratio est élevé, plus le risque financier augmente. Les actifs totaux sont aussi bien tangibles qu'intangibles. C'est la portion des actifs financés par des passifs, ou encore le nombre d'actifs qui pourrait servir à rembourser un passif.

Cependant, dans un contexte de faillite, les passifs sont généralement plus élevés que les actifs du failli, et il arrive fréquemment que celui-ci ne présente aucun actif. C'est pourquoi nous avons choisi l'inverse du ratio d'endettement, le ratio des actifs sur les passifs. Il servira à mesurer combien d'actifs couvrent un passif. Nous estimons que ce

ratio devrait avoir un effet positif sur le taux de recouvrement, c'est-à-dire que si les actifs augmentent ou les passifs diminuent, le taux de recouvrement devrait se retrouver à la hausse.

Types de dette

Comme le montre l'étude de Hoechstoeffer et collab. (2012), le taux de recouvrement varie beaucoup en fonction du type de dette ou de l'industrie qui a consenti le prêt. En ce qui concerne notre étude, lorsque le failli effectue son bilan auprès du syndic, il dresse la liste de ses dettes et détaille ainsi son passif. Chaque dette est ensuite classée dans une catégorie par le SAI. Pour les neuf catégories de types de dette, nous créons une variable dichotomique indiquant si le failli détient ce type de dette ou pas. Un individu peut donc détenir plusieurs types de dette au sein de son passif. Les 9 variables sont décrites dans le tableau 3.5.

Ces catégories de dette peuvent aussi, en quelque sorte, représenter l'industrie du prêt. La catégorie Autre contient, en principe, toute dette qui n'entre pas dans les huit autres catégories. Il peut s'agir, par exemple, d'une facture impayée pour l'électricité (Hydro Québec), pour le téléphone, etc. Cependant, nous ne sommes pas à l'abri d'un certain risque opérationnel entourant la manière dont un SAI classera les dettes du failli.

Types d'actif

Tout comme les types de dette, lorsque le failli dresse son bilan, il détaille les actifs en sa possession au moment du dépôt de la faillite. Nous créons donc des variables dichotomiques indiquant si le failli détient l'actif en question ou pas.

Au premier abord, nous pourrions nous attendre à des coefficients positifs pour certains actifs que le failli détient. En effet, nous pouvons supposer que la présence d'un actif devrait affecter positivement le taux de recouvrement, puisque le principe de la faillite est la liquidation des biens au profit des créanciers. Cependant, rappelons que nous estimons

uniquement le taux de recouvrement des créanciers non garantis, et qu'il arrive fréquemment que les actifs au sein d'une faillite soient en garantie, en particulier dans le cas de maisons.

Revenu d'emploi disponible

Ce champ est issu du formulaire 65. Il constitue le revenu d'emploi mensuel net des dépenses non discrétionnaires du failli. Puisque nous n'avons pas accès à l'emploi du débiteur, il constitue un très bon indicateur de la situation financière du failli et une bonne mesure pour contrôler si le failli a un emploi ou non.

Durée de la faillite

La durée de l'administration du dossier de faillite en années est calculée en soustrayant la date de libération du syndic à la date de dépôt de la procédure de faillite, de la manière suivante :

Durée de la faillite (en année) = Date libération du syndic – Date dépôt de la faillite.

La date de dépôt de la procédure de faillite, comme nous l'avons expliqué dans la section 3.1, se situe entre le 1^{er} janvier 2009 et le 31 décembre 2009. C'est la date où le dossier de faillite ordinaire de l'individu a été ouvert.

La date de libération du syndic, quant à elle, est la date à laquelle le dossier est fermé, puisqu'elle correspond au jour où la suspension des procédures est levée, comme nous l'avons vu dans la section 2.6.3. Elle ne doit pas être confondue avec la date de libération du failli, date à partir de laquelle celui-ci est officiellement libéré de ses dettes (libérables au sens de la LFI). En effet, même après la libération du failli, le syndic peut continuer son administration, déposer de nouveaux États des recettes et débours (formulaire 12) et verser des dividendes aux créanciers. C'est une fois la libération du syndic effective que

le dossier de faillite se dit terminé et fermé. Par ailleurs, il est opportun de noter que le syndic peut obtenir sa libération sans que le failli l'obtienne pour autant. Cependant, le SAI ne peut obtenir sa libération sans émettre un relevé de recettes et débours, dans lequel nous calculons notre variable dépendante.

La durée de l'administration du dossier de faillite en années est ensuite catégorisée en six niveaux. Les catégories sont présentées dans le tableau 3.4.

Ainsi, notre variable qualitative qui désigne la durée de la faillite pourrait permettre d'estimer le taux de recouvrement par rapport à la complexité du dossier, comme le montrent les conclusions de Helwege (1999). En effet, sur la base de notre connaissance du domaine des faillites, il semble qu'un dossier plus complexe peut prendre plus de temps : la réalisation des actifs et les multiples auditions à la cour, la présence de réclamations contingentes, les différentes procédures telles que les oppositions à la libération du failli de la part du créancier, du syndic ou du BSF peuvent facilement s'étaler dans le temps. Ce facteur complexité/temps signifie que plus la faillite est complexe, plus la durée est élevée, et finalement, plus elle devrait impliquer, selon nous, un recouvrement supérieur.

Par ailleurs, en présence de revenu excédentaire, la faillite dure plus longtemps : comme nous l'avons vu dans le tableau 2.1, le délai d'obtention de la libération automatique est supérieur et les paiements effectués à l'actif le sont également.

Toutefois, la durée de la faillite pourrait exercer un contrôle par rapport au débiteur : comme nous l'avons également vu dans le tableau 2.1, une faillite de plus longue durée caractérise peut-être une deuxième faillite. Le cas échéant, Eraslan et collab. (2017) observent que les faillis récidivistes présentent un taux de recouvrement inférieur.

Pour conclure, notre intuition concernant cette variable explicative est que la durée de la faillite sera positivement corrélée avec le taux de recouvrement, contrairement à Schuermann (2004) qui estime que le temps passé en faillite peut réduire considérablement la valeur de recouvrement pour les créanciers.

Genre de dépôt

Selon Gupton et collab. (2000), le genre de dépôt de procédure de faillite aux États-Unis présente des différences dans le taux de recouvrement. En nous inspirant de leur variable, nous sommes en mesure de créer la variable du genre de dépôt de dossier de faillite, ventilée en trois groupes :

- Cession volontaire ;
- Ordonnance de faillite ;
- Cession présumée.

La cession volontaire est caractérisée par le fait que le débiteur a choisi la faillite ordinaire. Le mot « volontaire » signifie que la faillite découle du choix délibéré de l'individu à faire faillite. On dit qu'il est allé volontairement voir un SAI pour faire faillite.

L'ordonnance de faillite, quant à elle, est une procédure moins fréquente. Dans ce cas, la faillite découle du fait qu'un des créanciers a entrepris la démarche, auprès de la cour, pour mettre en faillite le débiteur en question.

Enfin, la cession présumée peut être la conséquence de plusieurs événements indépendants. En effet, elle peut résulter d'une proposition de la section I qui n'a pas été déposée auprès des créanciers, ou d'une proposition qui a été refusée par les créanciers ou par la cour.

- En effet, comme le prescrit la LFI, lorsqu'un débiteur fait une proposition de la section I à ses créanciers, il a un certain délai à respecter pour la déposer. S'il ne respecte pas ce délai, il tombe involontairement en faillite.
- Par ailleurs, les créanciers sont en droit de refuser la proposition. Le cas échéant, le débiteur tombe involontairement en faillite.
- Mais si les créanciers acceptent la proposition, celle-ci, pour être homologuée, doit être présentée à la cour, laquelle peut à son tour refuser la proposition.

C'est ce qu'on appelle la cession présumée. C'est de cette manière que nous testerons s'il existe des différences significatives entre le taux de recouvrement d'une faillite et celui

d'un dossier de faillite qui était initialement en proposition, en nous inspirant de l'étude de de Eraslan et collab. (2017).

Grâce à cette variable, nous sommes en mesure d'estimer l'effet de la « volontarité » de la faillite sur le recouvrement. Nous croyons raisonnable de penser que l'ordonnance de faillite mène à un recouvrement plus élevé que la cession volontaire, puisque c'est le créancier qui actionne le débiteur en faillite. Le même raisonnement s'applique à la cession présumée : le débiteur a choisi une option avant la faillite, qui consiste à faire une proposition. Il doit certainement se trouver dans une meilleure situation financière qu'un débiteur qui a choisi la faillite en premier lieu. La catégorie en référence (qui sera contenue dans la constante) sera la Cession volontaire.

Conditions à la libération

Cette variable dichotomique est construite de la manière suivante :

$$\text{Conditions à la libération} = \begin{cases} 1 & \text{Avec conditions} \\ 0 & \text{Sans conditions} \end{cases} .$$

La catégorie en référence (qui sera contenue dans la constante) sera Sans conditions.

La variable Condition désigne s'il y a eu une condition à la libération du failli, non-obstant le fait qu'il soit actuellement libéré ou pas. En effet, comme nous l'avons vu au Chapitre 2, il est possible que le failli n'obtienne pas une libération automatique de ses dettes. Les types de conditions sont les suivantes :

- Médiation : la médiation est une entente hors cour entre le failli et son SAI lorsqu'il y a une mésentente entre les deux parties. À l'issue de la médiation, le failli aura en quelque sorte des conditions à respecter afin d'obtenir sa libération.
- Suspension de la libération du failli : le failli est contraint de rester en faillite pendant une période de temps, sans conditions financières ; une fois ladite période écoulée, le failli sera automatiquement libéré de ses dettes.

- Conditions à la libération du failli : il s’agit souvent d’une condition financière, soit un paiement jugé nécessaire à l’actif de sa faillite. La condition peut également contenir une suspension pendant une période de temps. Parfois, les conditions peuvent également exiger que le failli fournisse certains documents ou informations au SAI.
- Libération refusée : le refus de la libération du failli est la condition la plus sévère rendue par la cour. Il ne sera jamais libéré de ses dettes, et pourra peut-être se réadresser à la cour après une certaine période de temps pour demander de réévaluer son dossier. Ces conditions sont plus rares.

Nous pensons que le fait d’avoir eu une condition à la libération aura un effet positif sur le recouvrement, puisqu’elle est souvent d’ordre financier et pourrait ainsi faire augmenter le recouvrement du dossier de l’actif en question. Cependant, une condition pourrait aussi avoir un effet contraire, puisqu’elle implique en quelque sorte une « délinquance » de la part du failli, des paiements non effectués auparavant, un manque de collaboration de sa part, ou même des actes frauduleux. Nous détaillons la variable Conditions à la libération dans le tableau 3.1.

Valeur de la variable	Statut de la libération du failli	Fréquence	Pourcentage
Condition = 0	Audition fixée	6	0,38
Condition = 0	Libération absolue	235	14,91
Condition = 0	Audition ajournée	63	4,00
Condition = 0	Failli décédé	6	0,38
Condition = 0	Libération automatique	622	39,47
Condition = 1	Libération suspendue	147	9,33
Condition = 1	Libération conditionnelle	36	2,28
Condition = 1	Libération refusée	7	0,44
Condition = 1	Libération absolue - Conditions du syndic remplies	101	6,41
Condition = 1	Libération absolue - Entente de médiation respectée	10	0,63
Condition = 1	Libération absolue - Conditions de la cour remplies	343	21,76

N = 1576

TABLE 3.1 – Types de conditions à la libération

Rémunération du syndic

La rémunération relative du syndic est le ratio de ses honoraires sur les recettes totales du dossier :

$$\text{Rémunération du syndic (\%)} = \frac{\text{Honoraires du syndic (\$)}}{\text{Recettes totales (\$)}}. \quad (3.3)$$

Les honoraires du syndic sont le montant que celui-ci se verse pour le travail qu'il a effectué durant l'administration du dossier de faillite. En règle générale, dans le cas des faillites ordinaires, le syndic facture son temps et le temps de ses employés. En effet, la Règle 58(1) de la LFI prévoit ce qui suit : « Sauf ordonnance contraire du tribunal, la rémunération du syndic est censée englober tous les services rendus par lui, ses associés et ses employés. »

Les recettes totales, quant à elles, sont les sommes totales perçues par le syndic au moyen de la réalisation de l'actif. Les recettes totales sont donc en quelque sorte le recouvrement brut du dossier de faillite, avant les débours du SAI.

Ce ratio nous donne une idée du pourcentage de rémunération du syndic par rapport aux recettes dans le dossier. C'est la portion de l'actif que le syndic se verse, compte tenu de ses services et de son travail.

Par ailleurs, il convient d'ajouter qu'en vertu de l'article 39(1) de la LFI, « la rémunération du syndic est celle que lui vote une assemblée de créanciers ». Si sa rémunération n'est pas votée, alors « le syndic peut inclure dans son état final et retenir pour sa rémunération, sauf l'augmentation ou la réduction ci-après prévue, une somme ne dépassant pas sept et demi pour cent du montant qui subsiste de la réalisation des biens du débiteur après que les réclamations des créanciers garantis ont été payées ou acquittées ». En d'autres mots, la rémunération du syndic, relative aux recettes dans le dossier en question, ne devrait pas dépasser 7,5%. En pratique – dans notre échantillon, sa rémunération relative est en moyenne de 45% des recettes totales des dossiers. Cela signifie qu'en moyenne, la rémunération d'un SAI représente près de la moitié des sommes totales perçues lors de la

réalisation de l'actif, et ce, sans compter les frais qui sont déboursés à même l'actif. Ces frais comprennent les dépenses administratives relatives à un dossier, telles que les frais de dépôt, les frais judiciaires, les taxes, les frais de consultation et les sommes forfaitaires.

3.3 Nettoyage des données

Comme nous l'avons expliqué dans la section 3.1, notre base de données brute contenait 2227 dossiers de faillite ordinaire d'individus qui ont été déposés entre le 1^{er} janvier 2009 et le 31 décembre 2009.

3.3.1 Variable dépendante du modèle

Nous avons tout d'abord retiré, afin de construire notre taux de recouvrement, les dossiers dont la dette prouvée n'était pas reportée, soit 475 faillites.

Ensuite, nous retirons les dossiers ayant une dette prouvée de 0, puisque c'est le dénominateur du ratio du taux de recouvrement. En tout, 127 dossiers avec une dette prouvée nulle ont été retirés.

Nous retirons par la suite les dossiers n'ayant pas de dividende reporté. Il nous reste alors 1615 observations.

Cependant, 25 dossiers contenaient des taux de recouvrement supérieurs à 1. Rappelons que le taux de recouvrement est le dividende versé à la masse des créanciers, divisé par les réclamations prouvées des créanciers. D'un point de vue économique, le taux de recouvrement ne devrait pas être supérieur à 1, puisque les créanciers ne peuvent percevoir de montants supérieurs à leur réclamation. Nous retirons donc ces dossiers.

L'échantillon final contient 1590 observations après l'élimination d'individus ayant des valeurs incomplètes ou incohérentes.

3.3.2 Variables explicatives du modèle

Il y a 35 dossiers où la rémunération relative du syndic était supérieure à 1. Étant donné que ce ratio représente la rémunération du syndic par rapport aux recettes totales, le syndic ne peut percevoir un montant supérieur aux recettes du dossier. Nous retirons donc ces observations de l'échantillon.

Nous avons également retiré d'autres observations de variables explicatives du modèle. Notre démarche a consisté à retirer les valeurs extrêmes, afin de garder uniquement les valeurs dans un intervalle de confiance à 99,9% défini de la manière suivante :

$$[\bar{X} - z_{(\frac{\alpha}{2})}s_X; \bar{X} + z_{(\frac{\alpha}{2})}s_X],$$

avec $z_{(\frac{\alpha}{2})} = 3,291$, \bar{X} la moyenne de l'échantillon tel que $\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$, et s_X est l'écart type de l'échantillon tel que $s_X = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}}$.

Voici les variables explicatives qui ont été touchées par notre processus :

Pour *Revenu d'emploi disponible*, nous retirons les valeurs au-dessus du 99,9^e percentile, soit au-dessus de 13 354 \$. Nous perdons ainsi deux observations pour cette variable.

Actifs sur passifs a également été affectée par notre processus. En effet, nous retirons les observations supérieures au 99,9^e percentile, soit supérieures à 5,39.

Nous estimons que cela n'aura pas d'effet sur notre modèle puisque seulement deux observations ont été retirées sur la base de ce critère d'élimination de valeurs extrêmes.

Une fois le nettoyage des données terminé, nous sommes en mesure de présenter les statistiques descriptives de nos variables.

3.4 Statistiques descriptives

Après avoir couvert les variables à notre disposition, nous pouvons présenter les statistiques descriptives de ces dernières.

3.4.1 Variable dépendante du modèle

Nous présentons la matrice des moments du taux de recouvrement dans le tableau 3.2.

Statistiques descriptives du taux de recouvrement	
N	1590
Moyenne	0,1209
Écart-type	0,1783
Asymétrie	2,2593
Kurtose	5,7205

TABLE 3.2 – Moments du taux de recouvrement

Le taux de recouvrement moyen, pour les faillites ordinaires individuelles, est de 12,09%. Il diffère de celui rapporté dans la littérature sur la perte en cas de défaut, qui varie selon les études, mais qui est pratiquement toujours plus élevé. À cet effet, le taux de recouvrement de 17% pour les créanciers non garantis, recensé dans l'étude de Eraslan et collab. (2017), est celui qui se rapproche le plus du nôtre.

Il n'est pas étonnant que la littérature rapporte un recouvrement plus élevé, puisque nous nous situons dans un environnement de faillite, qui constitue un cas particulier du défaut. Nous pouvons expliquer cet écart par la rapidité d'agissement des créanciers au moment du défaut de l'individu, avant le dépôt de la procédure de faillite. En effet, comme nous l'avons vu dans l'étude de Hoehstoetter et collab. (2012), les créanciers peuvent prendre plusieurs mesures de recouvrement lors du défaut de l'individu, telles que la réalisation d'actifs ou la prise de garantie sur des biens du débiteur, et ce, avant la suspension des procédures lors du dépôt de la faillite. Ainsi, le recouvrement lors du défaut serait considérablement supérieur à celui de la faillite.

Par ailleurs, comme le montre le tableau 3.2, la distribution de notre taux de recouvrement n'est pas tout à fait normale, ce qui correspond à l'étude de Gupton et collab. (2000). En effet, les coefficients d'asymétrie et de kurtose retranscrits laissent croire que le taux de recouvrement n'est pas normalement distribué.

L'histogramme 3.1 illustre la distribution du taux de recouvrement. Dans notre échantillon, plus de 30% des taux de recouvrement observés sont nuls. Cela concorde avec

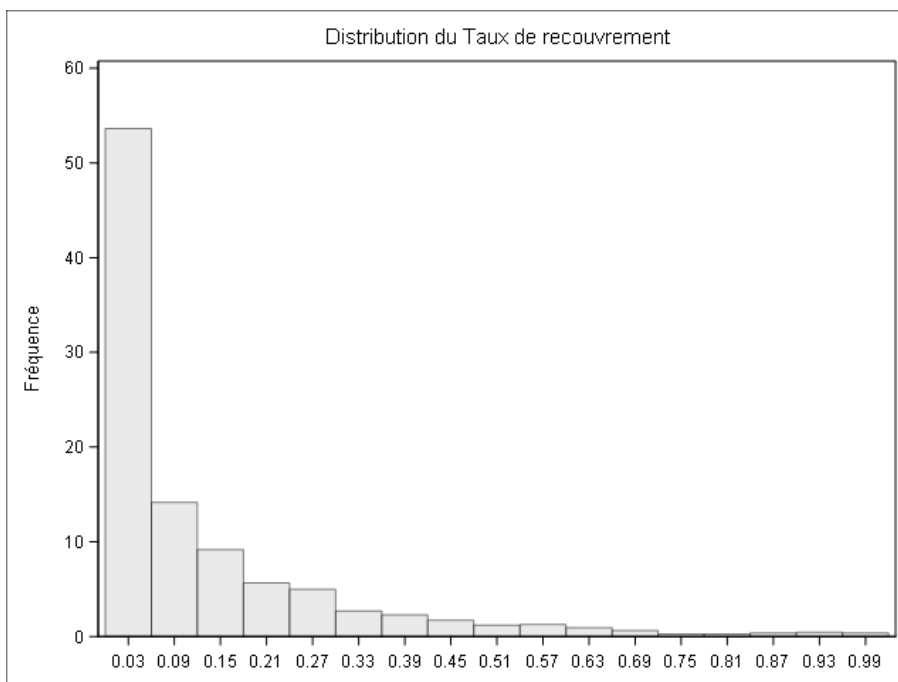


FIGURE 3.1 – Histogramme - Taux de recouvrement

l'étude de Eraslan et collab. (2017) : la distribution du recouvrement ultimement reçu (figure 1.5) s'apparente beaucoup à la nôtre, avec cependant un poids plus fort sur les recouvrements très élevés (de 80% à 100%), puisque, comme l'expliquent les auteurs, environ 20% des débiteurs proposent de rembourser la totalité de leur dette. Notre distribution est également similaire à Prívará et collab. (2014), dans les cas où il n'y a pas de recouvrement. Cependant, nous n'avons presque aucun recouvrement intégral. En effet, le 95^e et le 99^e percentile de notre distribution présentent respectivement un taux de recouvrement de 50,76% et 86,89%. Cette différence s'explique par le fait que nos recouvrements sont conditionnels à la faillite, alors que ceux des études susmentionnées sont conditionnels au défaut, lequel est un ensemble beaucoup plus large que la faillite, comme nous l'avons établi dans la section 1.1.3. Des quatre événements donnant lieu à un défaut, la faillite est considérée comme étant le dernier recours. Aussi, par opposition à un individu ayant fait défaut sans pour autant faire faillite, il apparaît logique, sur le plan économique, qu'un individu qui en vient à la faillite ne rembourse pas intégralement ses

créanciers ou qu'il ne possède pas assez d'actifs pour le faire.

Considérant les particularités de notre distribution, qui présente une grande proportion de recouvrements nuls et presque aucun recouvrement intégral, contrairement à la distribution bimodale généralement recensée dans la littérature, nous proposerons une méthodologie différente afin de trouver des modèles adaptés à notre distribution, soit le modèle en deux parties de Heckman (1977).

Discussion

Au vu des similarités et des différences que nous avons exposées entre le taux de recouvrement dans la littérature et celui de notre recherche, il est sans doute bon de rappeler que dans le cadre de la perte en cas défaut, le créancier se trouve face à un éventail d'individus en défaut – quelle que soit la raison du défaut – et que son taux de recouvrement sera déterminé par sa capacité à recouvrer la dette auprès de chacun des débiteurs. Par opposition, dans le contexte particulier de la faillite qu'est celui de notre étude, c'est l'individu en défaut qui fait face à un éventail de créanciers, et le taux de recouvrement n'est plus déterminé par la capacité du créancier à recouvrer sa dette, mais par un prorata, comme nous l'avons vu à la section 2.6.3.

Ainsi, les moyens financiers d'un créancier et sa capacité à recouvrer intégralement sa dette ne sont pas représentatifs de ceux de la masse des créanciers en cas de faillite d'un particulier. Il est possible que la littérature ne prenne pas en compte ce biais et qu'elle présente une certaine disparité dans le taux de recouvrement des créanciers, disparité que nous ne retrouverons pas dans notre étude.

Notre modèle a l'originalité de nous permettre d'isoler les caractéristiques du débiteur sans être biaisés par certains facteurs propres au créancier, comme pourrait l'être les études de Prívará et collab. (2014), Zhang et Thomas (2012), Wirenhammar (2011), qui étudient le recouvrement d'un créancier en particulier. Notre modèle propose une estimation du taux de recouvrement qui n'est pas fonction de caractéristiques spécifiques à un créancier.

3.4.2 Variables explicatives du modèle

Nous présentons dans les tableaux 3.3 et 3.4 les statistiques descriptives des variables indépendantes du modèle.

Variabiles	N	Minimum	Moyenne	Ecart-type	Maximum
Rémunération relative du syndic	1546	0,00	0,455	0,241	0,95
Actifs sur passifs	1586	0,00	0,660	0,467	5,52
Durée de la faillite (années)	1471	0,64	3,536	1,427	6,84
Nombre de membres dans l'unité familiale	1469	1,00	2,339	1,402	8,00
Actifs (\$)	1587	0,00	240 790	264 147	4 424 000
Passifs (\$)	1588	1 000	471 573	1 286 798	32 520 000

TABLE 3.3 – Statistiques descriptives des variables continues et numériques

Nous analyserons ces tableaux en reprenant nos catégories de variables.

Personnelles

Du côté des variables personnelles, nous constatons que notre échantillon contient plus d'hommes que de femmes, soit deux hommes pour une femme qui déposent une faillite ordinaire. Comme nous l'avons vu dans la revue de littérature, il semble que la fréquence d'hommes en faillite soit effectivement plus élevée que celle des femmes.

En ce qui concerne la taille des ménages des faillis, ces derniers sont en moyenne constitués de 2,3 personnes, et ce, en incluant le failli. Les catégories d'état civil des faillis semblent homogènes, sauf les veufs qui sont moins nombreux et les faillis mariés qui constituent 46,61% de notre échantillon.

Par ailleurs, comme nous l'avons décrit dans la section précédente, les groupes de régions sont homogènes.

Financières

Le ratio Actifs sur passifs est en moyenne de 0,66 : ainsi, en moyenne, les faillis présentent 2 \$ d'actif pour 3 \$ de passif. En situation de faillite, il apparaît normal que

Variables catégorielles	N	Proportion
Sexe		
Femme	500	0,3165
Homme	1080	0,6835
Conditions à la libération		
Sans conditions	946	0,5950
Avec conditions	644	0,4050
Genre de dépôt		
Cession volontaire	1378	0,8667
Ordonnance de faillite	35	0,0220
Cession présumée	177	0,1113
Durée de la faillite		
Moins de 2 ans	225	0,1415
2 à 3 ans	386	0,2428
3 à 4 ans	339	0,2132
4 à 5 ans	236	0,1484
Plus de 5 ans	285	0,1792
Syndic non libéré	119	0,0748
Région		
Autres	326	0,2050
Alberta	466	0,2931
Ontario	478	0,3006
Québec	320	0,2013
État civil		
Marié	681	0,4661
Conjoint de fait	144	0,0986
Divorcé	173	0,1184
Célibataire	192	0,1314
Séparé	228	0,1561
Veuf	43	0,0294

TABLE 3.4 – Statistiques descriptives des variables catégorielles

les faillis possèdent moins d'actifs que de passifs. En effet, on observe que ce ratio est au maximum à 5,52.

En ce qui concerne la variable Types de dette, puisqu'un individu peut détenir plusieurs dettes, nous présentons des statistiques descriptives séparément dans le tableau 3.5. Ce tableau résume l'information relative aux types de dette, qu'elles soient privilégiées ou non garanties.

Variables	Pourcentage	Fréquence
Hypothèques sur les biens immeubles	0,0356	183
Prêts bancaires (excepté hypothèques sur les biens immeubles)	0,1861	972
Prêts des sociétés de crédit	0,0518	497
Cartes de crédit émises par les banques et les sociétés de fiducie	0,2154	1222
Autres émetteurs de cartes de crédit	0,0973	936
Impôts fédéral/provincial et taxes municipales	0,1774	824
Prêts étudiants	0,0069	50
Prêts de particuliers	0,0238	36
Autres	0,2056	1087
N = 1 469		

TABLE 3.5 – Statistiques descriptives des Types de dette (TD_j). Puisque nous avons accès au montant en dollars de chaque types de dette, la deuxième colonne représente le pourcentage entre la valeur du type de dette j et la valeur totale de tous les types de dette, tel que $\text{Pourcentage} = \frac{TD_j(\$)}{\sum_{i=1}^n TD_i(\$)}$, pour $j = 1, \dots, 9$. La troisième colonne représente la Fréquence lorsque $TD_j = 1$, c'est-à-dire lorsque la dette est présente au sein du passif de l'individu i .

Tout d'abord, il appert que les dettes les plus fréquentes au sein des faillites canadiennes sont les dettes de cartes de crédit (émises par les banques et les autres émetteurs), celles des prêts bancaires, des impôts et des autres types de dette. Ce sont également ces types de dette qui représentent les plus grandes proportions en dollars.

Il n'est pas étonnant que les hypothèques sur les biens immeubles soient aussi faibles, que ce soit en dollars ou en fréquence : en effet, ce type de dette est très fréquemment une dette qui possède une garantie, et n'est donc pas comprise dans notre étude.

Aussi, la valeur totale en dollars des neuf types de dette, en ce qui concerne les créanciers privilégiés, représente uniquement de 0,0028 du total des réclamations prouvées, soit celles des créanciers privilégiés et des créanciers non garantis combinées. Comme nous l'avons expliqué au Chapitre 2, les créanciers privilégiés sont assez peu fréquents dans les faillites ordinaires, et ne représentent qu'une infime partie du passif des faillis.

Faillites

En revenant sur le tableau 3.4, nous observons que les six groupes de Durée de la

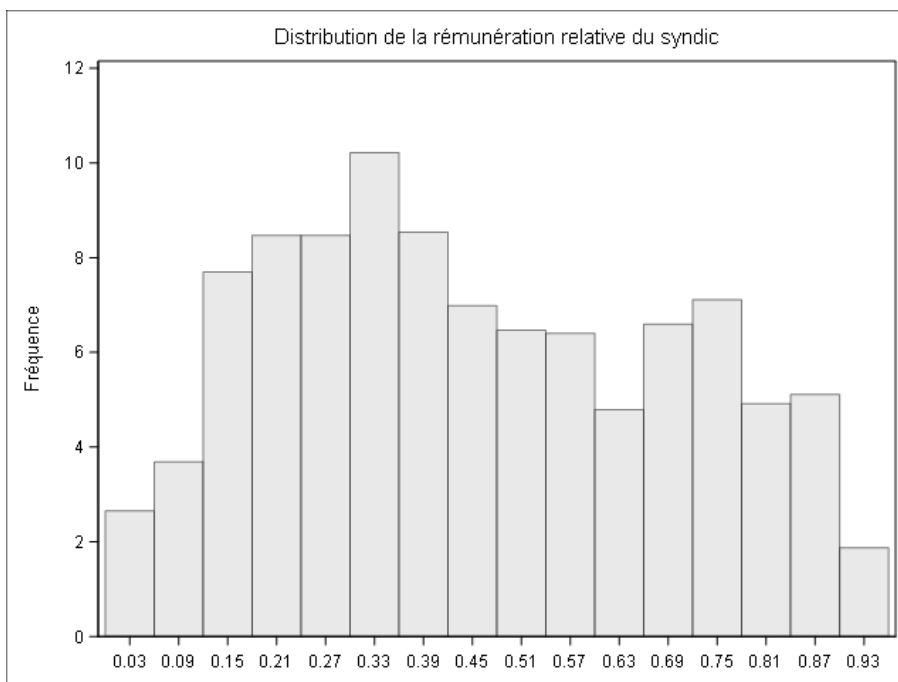


FIGURE 3.2 – Histogramme - Rémunération relative du syndic

faillite sont relativement homogènes. Cependant, les genres de dépôts diffèrent : en effet, la cession volontaire est de loin la procédure la plus fréquente au sein de notre échantillon. La plupart des débiteurs font volontairement faillite. L'échantillon contient un petit groupe de cessions présumées, et très peu d'ordonnances de faillites, qui sont effectivement plus rares dans le cas de faillites d'individus que des faillites d'entreprises.

Environ 40% des individus ont obtenu une condition à leur libération, qu'ils soient actuellement libérés ou non.

Enfin, la rémunération relative du syndic autorisé en insolvabilité (SAI), comme nous l'avons vu au chapitre précédent, est votée par une assemblée des créanciers ; si elle n'est pas votée, la rémunération du SAI relative aux recettes dans le dossier ne devrait pas dépasser 7,5%. Comme on peut le voir dans notre échantillon, elle est en moyenne de 45%. La figure 3.2 représente la distribution de la rémunération relative du syndic.

Si nous revenons à la définition de la rémunération relative du syndic, ce sont les honoraires du syndic divisés par les recettes totales, pour un dossier donné. Dans notre

échantillon, qui représente toutes les faillites ordinaires de l'année 2009 à travers le Canada, les honoraires totaux des syndics s'élèvent à un montant cumulé de 15 954 289 \$. Les recettes totales cumulées, quant à elles, sont de 53 344 160 \$. Ces recettes cumulées représentent le potentiel de recouvrement que les créanciers auraient pu obtenir, sans les frais et coûts associés à la faillite.

Enfin, le dividende versé total cumulé en dollars s'élève à 26 778 327 \$. Cela nous renseigne sur les coûts engendrés par la faillite, lesquels sont inévitables et payés à même l'actif. Les coûts de la faillite à un niveau agrégé s'illustrent ainsi : pour deux dollars de recettes totales brutes dans les dossiers, le créancier en reçoit un.

Pour terminer, nous présentons la figure 3.3, illustrant la relation entre le taux de recouvrement et la rémunération relative du syndic. Nous remarquons rapidement la relation inverse entre ces deux variables : en effet, il semble cohérent que plus le syndic se rémunère à même les recettes disponibles dans le dossier, plus le taux de recouvrement des créanciers est faible.

En suivant l'axe des abscisses, nous constatons de multiples observations où le taux de recouvrement des créanciers est nul, mais où le syndic se rémunère. La concentration d'observations s'accroît lorsque la rémunération relative du syndic dépasse 68%. Nous observons également qu'à partir de ce niveau de rémunération, le taux de recouvrement des créanciers, à l'exception d'un seul dossier, ne dépasse plus 10%.

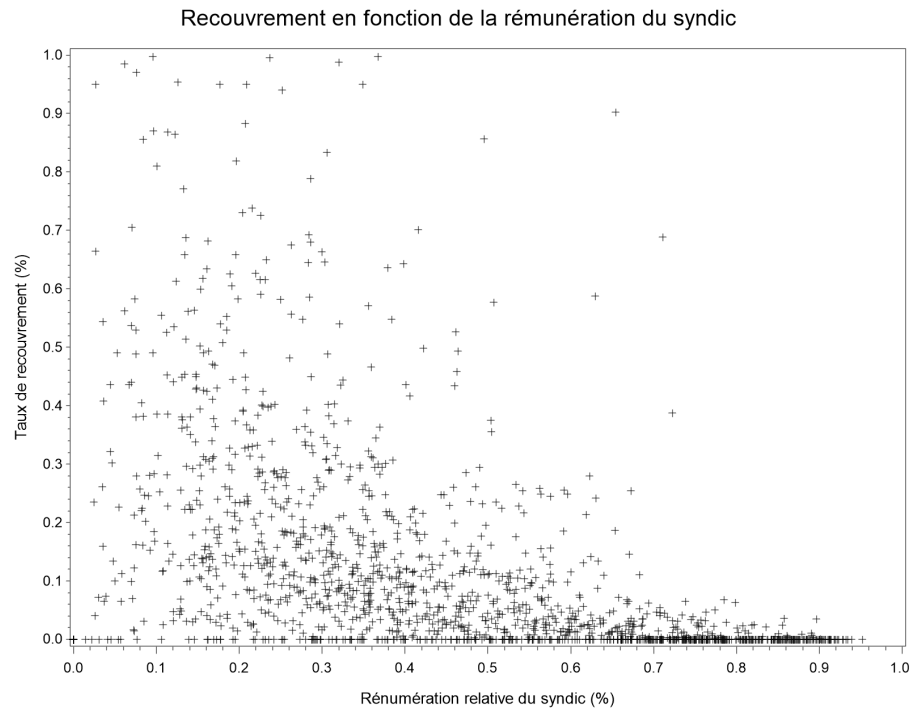


FIGURE 3.3 – Diagramme de dispersion - Taux de recouvrement et Rémunération du syndic

En revanche, il n’y a aucun dossier où un dividende a été versé sans que le SAI se soit rémunéré, ce qui semble cohérent au niveau économique. Cela concorde également avec l’ordre de priorité des paiements de la faillite, présenté à la section 2.2, et en sachant que les professionnels en insolvabilité - syndics - ne travaillent pas gratuitement.

Chapitre 4

Méthodologie

Dans le Chapitre 1, nous avons présenté la revue de littérature détaillant à la fois les modèles statistiques et les variables explicatives liés à la perte en cas de défaut. Nous avons par la suite décrit le processus de la faillite au Canada au Chapitre 2. Finalement, dans le Chapitre 3, nous avons présenté les statistiques descriptives du taux de recouvrement, la variable d'intérêt de notre étude ainsi que celles des variables potentiellement explicatives du taux de recouvrement. L'objectif du présent chapitre est de détailler la méthodologie de notre analyse statistique et économique du taux de recouvrement.

Les mêmes modèles statistiques peuvent servir autant à l'estimation qu'à la prédiction d'une variable dépendante. En effet, il est tout à fait possible d'estimer les paramètres d'un modèle (β) sans biais, et de prédire la réponse (y) en minimisant une distance entre les valeurs observées et les valeurs prédites par le modèle. Le défi entre en jeu lorsque le modèle est uniquement destiné à l'estimation ou uniquement à la prédiction. Lorsque nous cherchons à estimer l'effet de variables explicatives sur une variable dépendante, le modèle se veut tout d'abord facile d'interprétation. Il s'agit alors de se concentrer sur l'inférence statistique, à savoir si les paramètres sont significatifs, sans biais, et ce, sans se préoccuper forcément des ajustements nécessaires afin d'utiliser ce même modèle pour prédire la variable d'intérêt pour des observations futures.

La prédiction, quant à elle, nécessite bien évidemment l'estimation des paramètres,

mais ne se restreint pas seulement aux modèles interprétables. Les méthodes peuvent alors s'étendre à des modèles non paramétriques, par exemple les arbres de décisions ou de régressions (Breiman, 2017), ou même à des techniques d'apprentissage automatique (Friedman et collab., 2001) (*machine learning* en anglais), ou bien d'apprentissage profond (Goodfellow et collab., 2016) (*deep learning* en anglais) tels que les réseaux de neurones. Ces méthodes, qui peuvent produire de meilleurs résultats en matière de prédiction sur un nouveau jeu de données, font perdre l'interprétation économique.

La différence entre l'estimation et la prédiction réside donc dans le niveau d'interprétation souhaité, qui lui, dépend du domaine d'application. Dans cette optique, notre travail se concentrera dans un premier temps sur l'estimation des effets des variables explicatives du taux de recouvrement, puis sur l'établissement d'un modèle de prédiction du taux de recouvrement. Dans ce qui suit, nous présentons en ordre chronologique les modèles qui seront utilisés au long de notre étude.

4.1 Estimation des paramètres

Nous chercherons à expliquer la variabilité du taux de recouvrement et à estimer l'effet des variables indépendantes à notre disposition. Nous présentons les résultats dans le Chapitre 5.

4.1.1 Régression linéaire

Nous commencerons notre étude en utilisant un modèle de régression linéaire par les moindres carrés ordinaires. En effet, d'après Zhang et Thomas (2012), la régression linéaire est le modèle le plus adapté pour la modélisation du taux de recouvrement. Il est également utilisé comme modèle prédictif dans d'autres domaines financiers. Le modèle de régression linéaire exprime une variable dépendante, y_i , en fonction de variables explicatives $x_{1i}, x_{2i}, \dots, x_{pi}$ et de paramètres $\beta_0, \beta_1, \dots, \beta_p$. Le modèle de régression linéaire de manière générale s'exprime ainsi

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i \quad (4.1)$$

pour $i = 1, \dots, n$ contrats individuels, où

- y_i représente le taux de recouvrement de l'individu i ,
- $\beta_0, \beta_1, \dots, \beta_p$ sont des paramètres inconnus à estimer,
- $x_{1i}, x_{2i}, \dots, x_{pi}$ sont des variables indépendantes, connues, servant à décrire le taux de recouvrement, et
- ε_i est un terme d'erreur suivant une loi Normale de moyenne 0 et de variance σ^2 , soit $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$.

Sous forme matricielle, le modèle s'écrit

$$y_{n \times 1} = X_{n \times (p+1)} \beta_{(p+1) \times 1} + \varepsilon_{n \times 1} \quad (4.2)$$

où

$$X_{n \times (p+1)} = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{pmatrix} \text{ et } \beta_{(p+1)} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}. \quad (4.3)$$

L'objectif est d'estimer les paramètres inconnus du vecteur β . L'estimation consiste à trouver les valeurs des paramètres par minimisation de la somme des carrés des écarts entre les taux de recouvrements observés et prédits par le modèle. Ceci est résolu par la méthode des moindres carrés ordinaires (MCO). Le critère des moindres carrés ordinaires s'écrit

$$\operatorname{argmin}_{\beta} \sum_{i=1}^n (y_i - X_i \beta)^2. \quad (4.4)$$

Nous trouvons la valeur des paramètres qui minimisent cette fonction : ce sont les réalisations des estimateurs $\hat{\beta}$. Dans le cadre de notre étude, l'estimation des paramètres des modèles de régression linéaire est effectuée avec la fonction *lm* du logiciel R.

Nous présentons 3 modèles de régression linéaire multiple. Le premier, que nous nommons *Modèle principal*, ajuste le taux de recouvrement en fonction des variables suivantes : *Rémunération du syndic, Actifs sur passifs, Nombre de membres dans l'unité familiale, Région, Sexe, Genre de dépôt, Durée de la faillite, État civil et Conditions à la libération*. Le deuxième modèle, *Modèle Types de dette*, reprend le *Modèle principal* auquel nous ajoutons les variables indicatrices des types de dette. Enfin, le troisième modèle, *Modèle Types d'actif* ajoute en plus, les variables indicatrices des types d'actif.

4.1.2 Validation des modèles

Les résidus sont définis, pour $i = 1, \dots, n$, par la différence suivante

$$\hat{y}_i = X_i \hat{\beta}, \quad (4.5)$$

$$\hat{\varepsilon}_i = y_i - \hat{y}_i. \quad (4.6)$$

La variable aléatoire ε_i suit une loi Normale de moyenne 0 et de variance σ^2 , soit $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. Ainsi, lorsque nous utilisons un modèle de régression linéaire, nous supposons pour $i = 1, \dots, n$ que

1. $E(\varepsilon_i | x_{1i}, x_{2i}, \dots, x_{pi}) = 0$: la moyenne des erreurs est nulle,
2. $\text{Var}(\varepsilon_i) = \sigma^2$: la variance des erreurs est constante,
3. $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$: les observations i et j sont indépendantes,
4. $\text{Cov}(\varepsilon_i, x_{1i}) = \text{Cov}(\varepsilon_i, x_{2i}) = \dots = \text{Cov}(\varepsilon_i, x_{pi})$: les régresseurs sont non endogènes,
et
5. il y a absence de colinéarité entre les régresseurs.

En pratique, nous vérifions la validité des hypothèses visuellement, au moyen d'un graphique des résidus (4.6) en fonction des valeurs ajustées par le modèle (4.5).

Notons aussi d'autres menaces possibles à la validité interne du modèle de régression, soit

- biais de variables omises,
- mauvaise forme fonctionnelle,
- biais dû à l'entrée des données (*Errors-in-variables bias*),
- biais de sélection, et
- biais de causalité simultanée.

Elles impliquent que les résidus ne soient pas centrés en 0. Dans ce cas, les estimateurs des moindres carrés sont biaisés et non consistants. Nous essayons de remédier au biais de causalité simultanée, mais nous ne trouvons pas de variables instrumentales adéquates dans le jeu de données disponible. Ainsi, nous nous concentrons seulement sur le biais de sélection.

L'inférence statistique, dans le contexte d'une régression linéaire, se fait sur les paramètres du modèle $\beta_0, \beta_1, \dots, \beta_p$. L'idée est de tester si ces paramètres sont *significativement* différents de zéro. Les valeurs estimées $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ sont spécifiques à un échantillon. En effet, avec un échantillon différent, les valeurs estimées sont différentes. Alors, une mesure de leur précision est indispensable. Sans cette mesure, aucune conclusion ne peut être tirée. La précision d'une valeur estimée est donnée par son erreur-type. L'erreur-type indique quelle est, en moyenne, la distance entre les valeurs estimées $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ et les "vraies" valeurs $\beta_0, \beta_1, \dots, \beta_p$.

Lorsque nous parlons de *significativité*, nous faisons allusion à un test d'hypothèse. Pour ce qui est des paramètres de régression, nous devons tester les hypothèses suivantes

$$H_0 : \beta_k = 0 \text{ contre } H_1 : \beta_k \neq 0$$

pour $k = 0, 1, \dots, p$. Le test utilisé est un test- t et la statistique de test est définie par

$$t_0 = \frac{\hat{\beta}_k}{\hat{\sigma}_{\beta_k}}. \quad (4.7)$$

Sous l'hypothèse nulle H_0 , la statistique de test t_0 suit une distribution de Student à $n - p$ degrés de liberté, T_{n-p} .

L'hypothèse nulle est rejetée lorsque la valeur du paramètre estimé divisée par son écart-type est supérieure en valeur absolue à la statistique de Student pour un niveau de confiance donné α , généralement égal à 95%. La règle de décision peut s'écrire de la façon suivante : nous rejetons $H_0 : \beta_k = 0$ si $|t_0| > t_{n-p; 1-\alpha/2}$. Ainsi, lorsque l'hypothèse nulle est rejetée pour le paramètre β_k , nous dirons que la variable explicative X_k a un effet statistiquement significatif sur la variable dépendante y , puisque le paramètre estimé de ladite variable est significativement différent de zéro.

Le coefficient de détermination (R^2) mesure la qualité d'ajustement d'un modèle de régression linéaire. Pour un échantillon donné, le R^2 est défini par

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (4.8)$$

où $\bar{y} = n^{-1} \sum_{i=1}^n y_i$. Le R^2 représente la proportion de la variance de y expliquée par les variables indépendantes du modèle, x_1, x_2, \dots, x_p . Toutefois, l'ajout de plusieurs variables explicatives pourrait artificiellement augmenter le R^2 , qui est fonction du nombre de variables. Il convient d'utiliser le R^2 ajusté lorsque l'on compare des modèles n'ayant pas le même nombre de variables explicatives, puisque la valeur de R^2 ajusté intègre le nombre de prédicteurs dans le modèle. Le R^2 ajusté se calcule comme suit

$$R^2_{\text{ajusté}} = 1 - (1 - R^2) \frac{n - 1}{n - p - 1}, \quad (4.9)$$

avec p le nombre total de variables explicatives dans le modèle (sans compter la constante) et n la taille de l'échantillon. Nous utiliserons donc cette statistique afin de comparer les différents modèles ajustés à nos données.

4.1.3 Régression avec interactions

Nous intégrerons par la suite des modèles de régression linéaire avec interactions entre une ou plusieurs variables explicatives du modèle, dans le but de fournir une meilleure explication du taux de recouvrement. Nous définissons donc un nouveau *Modèle d'interaction* auquel nous ajoutons deux interactions, soit, la région avec *Actifs sur passifs* ainsi que la région avec la *Rémunération du syndic*.

Prenons un exemple simple avec deux variables explicatives que nous avons définies au Chapitre 3 : le ratio d'actifs sur passifs (AP_i) et la région (R_i). Le modèle de régression linéaire s'écrit

$$y_i = \beta_0 + \beta_1 R_i + \beta_2 AP_i + \varepsilon_i.$$

Ce modèle permet d'estimer l'effet β_2 du ratio d'actifs sur passifs sur le taux de recouvrement, en gardant constant la région. De la même manière, β_1 est l'effet de la région sur le taux de recouvrement, en fixant la valeur du ratio d'actifs sur passifs. Toutefois, ce modèle a d'importantes limites. En effet, dans cette spécification, l'effet du ratio d'actifs sur passifs est interprété indépendamment des régions canadiennes que nous avons définies. Il n'y a pourtant aucune raison de penser que la taille des faillites est homogène au Canada. Nous voudrions, par exemple, estimer l'effet de la taille des faillites, et ce, par région. Nous introduirons donc un terme d'interaction dans notre modèle de la manière suivante

$$y_i = \beta_0 + \beta_1 R_i + \beta_2 AP_i + \beta_3 (R_i \times AP_i) + \varepsilon_i. \quad (4.10)$$

Ainsi, le nouveau régresseur ($R_i \times AP_i$) est une variable d'interaction et a une interprétation particulière. Nous la décrirons lorsque nous utiliserons le modèle d'interaction dans notre étude au Chapitre 6.

4.1.4 Modèle de Heckman (1977)

Le modèle présenté par la suite utilisera la méthode de Heckman (1977). En effet, comme l'expliquaient Zhang et Thomas (2012), la distribution du taux de recouvrement a

une forme en « u ». Toutefois, comme nous l'avons vu dans l'histogramme 3.1 du chapitre précédent, la distribution du taux de recouvrement des faillites ne contient presque aucun recouvrement intégral ; il contient néanmoins une forte concentration d'observations nulles. Un modèle linéaire n'estime peut-être pas cette distribution correctement, au risque d'obtenir des estimateurs biaisés.

Nous proposons donc d'estimer les taux de recouvrement strictement positifs, c'est-à-dire, $y_i > 0$. Nous pourrions ainsi n'utiliser qu'un sous-ensemble de notre échantillon afin de modéliser le taux de recouvrement. Toutefois, ceci pourrait engendrer un biais de sélection.

L'idée principale du modèle de Heckman (1977) est de corriger le biais de sélection. En effet, selon cette méthode, le taux de recouvrement n'est observable que selon certains critères. Le modèle est ajusté aux données en deux étapes : d'abord, la modélisation d'une variable binaire z permettant de discriminer les taux de recouvrement positifs ($z = 1$, observés) des taux de recouvrements nuls ($z = 0$, censurés ou non observés) ; ensuite, dans un deuxième temps l'estimation des taux de recouvrements, y , conditionnels au fait qu'ils soient observés ($z = 1$).

Le premier modèle est défini par

$$z_i^* = W_i a + e_i, \quad (4.11)$$

où W_i représente un vecteur de prédicteurs, a les paramètres qui lui sont associés, $e_i \sim \mathcal{N}(0, \sigma_e^2)$ l'erreur aléatoire et

$$z_i = \begin{cases} 1 & \text{si } z_i^* > 0 \\ 0 & \text{sinon} \end{cases} .$$

Cette première étape, communément appelée le modèle de sélection, est estimée par une régression probit. Nous présentons les résultats du *Modèle de sélection* au Chapitre 6. La deuxième étape consiste à examiner les effets des variables indépendantes sur la variable expliquée, en l'occurrence le taux de recouvrement, à condition que celui-ci soit observable, lorsque $z = 1$. Supposons le modèle latent suivant

$$y_i^* = X_i b + u_i, \quad (4.12)$$

où X_i représente un vecteur de prédicteurs, b les paramètres qui lui sont associés, $u_i \sim \mathcal{N}(0, \sigma_u^2)$ l'erreur aléatoire et

$$y_i = \begin{cases} y_i^* & \text{si } z_i = 1 \\ \text{non observé} & \text{si } z_i = 0 \end{cases}.$$

Alors, puisque y_i^* est observable uniquement si l'individu i satisfait un critère de sélection, il peut y avoir un biais de sélection parmi les y_i^* observés, si $E(u_i | W_i, z_i = 1) \neq 0$.

En effet, le fait de présenter un taux de recouvrement positif n'est pas forcément aléatoire. Bien que nous ayons bâti un *Modèle de sélection* afin d'expliquer ce phénomène, il existe peut-être des effets ou facteurs non mesurables. Ces effets se retrouvent alors dans le terme d'erreur e de l'équation de sélection. Comme l'expliquent Dahen et Dionne (2010), z est la réalisation de la variable latente z^* . On observe y lorsque $z = 1$, à savoir quand le taux de recouvrement est strictement positif. Ainsi, y est la réalisation d'une autre variable latente y^* . Les deux termes d'erreur ont une corrélation ρ . Si $\rho \neq 0$, alors les résultats du modèle de régression sur les taux de recouvrement strictement positifs seront biaisés. Heckman propose donc, à partir de l'estimation des paramètres a de l'équation de sélection, de calculer l'inverse du ratio de Mills de la manière suivante

$$\lambda_i = \frac{\phi(W_i \hat{a})}{1 - \Phi(W_i \hat{a})} \quad (4.13)$$

avec $\phi(\cdot)$ et $\Phi(\cdot)$ les fonctions de densité et de répartition de la loi normale centrée réduite. Une fois le ratio de Mills estimé, nous avons donc pour les individus $i = 1, \dots, n$, une nouvelle variable $\hat{\lambda}_i$. Cette estimation est ajoutée comme variable explicative dans la deuxième partie du modèle de régression sur les taux de recouvrement observés. Alors, si le coefficient estimé $\hat{\beta}_\lambda$ de l'inverse du ratio de Mills est statistiquement significatif, cela signifie qu'il y a un biais de sélection qui a été corrigé par le régresseur λ . Autrement, nous n'aurions pas eu une estimation biaisée des valeurs de recouvrement observés. Nous présentons les résultats du *Modèle avec correction Heckman* au Chapitre 6.

4.2 Prédiction du taux de recouvrement

Une fois l'estimation achevée, le dernier objectif de l'étude sera d'établir des modèles de prédiction du taux de recouvrement. Nous comparons différentes méthodes au Chapitre 6.

4.2.1 Validation croisée

Afin de quantifier le pouvoir de prédiction d'un modèle statistique, il est courant d'utiliser un jeu de données non utilisé afin d'estimer les paramètres du dit modèle. En pratique, nous avons un nombre limité d'observations. Ainsi, nous utilisons toutes les observations pour estimer les paramètres du modèle. Toutefois, une méthode populaire qui permet d'utiliser un seul échantillon afin d'estimer les paramètres et d'évaluer la qualité de prédiction du modèle est la validation croisée en k parties (Friedman et collab., 2001) ou *k-fold cross-validation*. Il s'agit de segmenter notre échantillon aléatoirement en k sous-échantillons de tailles égales. Nous utilisons $k - 1$ sous-échantillons pour l'estimation des paramètres (apprentissage ou entraînement du modèle) et un sous-échantillon pour la prédiction (validation du modèle).

Toutefois, si nous effectuons une seule fois cette technique, que l'échantillon soit segmenté aléatoirement ou non, les résultats pourraient ne pas être robustes. La technique de validation croisée en k parties consiste à répéter ce processus k fois (itérations), en faisant varier le sous-échantillon de validation de manière à ce que l'entièreté de notre échantillon ait servi en tant que sous-échantillon de validation. Considérant le nombre d'observations à notre disposition, nous segmentons notre échantillon aléatoirement en $k = 10$ parties. La figure 4.1 illustre le processus de validation croisée en $k = 5$ parties.

Ainsi, pour chaque itération, nous calculons une erreur de prédiction telle que l'erreur quadratique moyenne (pour un y continu) ou le taux de bonne classification (pour un y binaire) par exemple. Nous aurons donc $k = 10$ erreurs. Nous serons donc en mesure de présenter la moyenne des erreurs de prédiction, une métrique qui sera davantage robuste pour estimer l'erreur de prédiction d'un échantillon futur (non observé).



FIGURE 4.1 – Validation croisée en $k = 5$ parties. Tour à tour, chacune des 5 parties est utilisée comme sous-échantillon pour tester le modèle prédictif. L’union des 4 autres parties est utilisée comme sous-échantillon pour entraîner le modèle.

Pour un y binaire, nous présenterons un *Modèle de classification probit* au Chapitre 6. Pour un y continu, nous comparerons les résultats du *Modèle de prédiction linéaire* et du *Modèle de prédiction linéaire censuré* au Chapitre 6.

4.2.2 Sélection de variables

Notre dernier modèle de prédiction est effectué en utilisant *Least Absolute Shrinkage and Selection Operator* ou le lasso (Tibshirani, 1996), une régularisation qui permet de faire de la sélection de variables lorsque nous ajustons un modèle de régression.

Tout comme nos modèles de régression linéaire par les moindres carrés ordinaires, le modèle du lasso minimise la somme des carrés des résidus (4.4), mais est assujéti à une contrainte supplémentaire, qui contracte la valeur de certains coefficients à zéro. Les coefficients du modèle du lasso sont la solution au problème d’optimisation suivant

$$\begin{aligned} \min_{\beta} \sum_{i=1}^n (y_i - X_i \beta)^2, \\ \text{s.c. } \sum_{j=1}^p |\beta_j| \leq t. \end{aligned} \quad (4.14)$$

Le tout peut s'écrire sous la forme plus connue de la somme des carrés des résidus pénalisés (SCR) utilisant le Lagrangien

$$\min_{\beta} \sum_{i=1}^n (y_i - X_i \beta)^2 + \lambda \sum_{j=1}^p |\beta_j|. \quad (4.15)$$

Ainsi, nous obtenons plusieurs solutions en fonction de λ , la constante de régularisation, qui contrôle le montant de pénalisation infligé aux paramètres β_j . Si $\lambda = 0$, nous retrouvons les estimateurs des moindres carrés ordinaires. Plus λ augmente (t diminue), plus la contrainte sera forte et plus il y aura de coefficients qui seront amenés à zéro. Le lasso opère alors comme un modèle de sélection de variables. En pratique, nous utilisons la validation croisée en k parties afin d'estimer λ . Nous présenterons un *Modèle de classification lasso* que nous comparerons au *Modèle de classification probit* au Chapitre 6. En effet, le lasso sera utilisé pour son processus de sélection de variables : les variables retenues par le lasso seront intégrées dans un modèle de classification probit.

Chapitre 5

Résultats principaux

Ce chapitre présente les résultats économétriques de l'estimation du modèle. Ces résultats permettent de se prononcer sur les hypothèses émises aux chapitres précédents.

5.1 Régression principale

Dans le but de répondre à notre objectif principal, nous commençons par présenter les résultats du modèle de régression linéaire principal estimé par la méthode des moindres carrés ordinaires. Notre modèle est basé à la fois sur nos connaissances dans le domaine de l'insolvabilité et sur la littérature sur la perte en cas de défaut. Ainsi, nous nous efforcerons de comparer nos résultats à ceux d'autres chercheurs ainsi qu'à nos prévisions de la section 3.2.

L'intérêt de la régression linéaire est d'obtenir une estimation simple des variables faisant varier le taux de recouvrement. Puisque la méthode des moindres carrés ordinaires est facilement interprétable et largement utilisée en ce qui a trait à l'estimation du taux de recouvrement, c'est le premier modèle que nous emploierons, l'objectif étant une modélisation réaliste et intuitive. Le modèle principal est présenté dans le tableau 5.1.

En premier lieu, il est possible d'interpréter la constante de notre modèle de régression. C'est la valeur du taux de recouvrement lorsque toutes les variables explicatives sont nulles. Ainsi, pour une femme mariée, résidant dans le reste du Canada (hors Québec, Al-

	Coefficients (écarts-type)
Intercept	0.2643 (0.0200)***
Rémunération du syndic	-0.4078 (0.0174)***
Actifs sur passifs	0.0569 (0.0088)***
Nombre de membres dans l'unité familiale	-0.0060 (0.0036)*
Région (Ref = Autre)	
Alberta	0.0302 (0.0114)***
Ontario	-0.0181 (0.0113)
Québec	-0.0487 (0.0129)***
Sexe (Ref = Femme)	
Homme	-0.0180 (0.0086)**
Genre de dépôt (Ref = Cession volontaire)	
Ordonnance de faillite	-0.0337 (0.0397)
Cession présumée	-0.0246 (0.0134)*
Durée de la faillite (Ref = Moins de 2 ans)	
2 à 3 ans	0.0088 (0.0130)
3 à 4 ans	0.0215 (0.0136)
4 à 5 ans	0.0180 (0.0148)
Plus de 5 ans	0.0393 (0.0144)***
Syndic non libéré	0.0303 (0.0185)
État civil (Ref = Marié)	
État civil : Conjoint de fait	0.0348 (0.0138)**
État civil : Divorcé	0.0160 (0.0144)
État civil : Séparé	0.0220 (0.0126)*
État civil : Célibataire	0.0365 (0.0139)***
État civil : Veuf	0.0461 (0.0245)*
Conditions à la libération (Ref = Sans conditions)	
Avec conditions	0.0220 (0.0084)***
R ²	0.3461
Adj. R ²	0.3367
Num. obs.	1412
RMSE	0.1457

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 5.1 – Modèle principal - Résultats du modèle de régression linéaire principal estimant le taux de recouvrement par les moindres carrés ordinaires (MCO) en fonction des groupes de variables personnelles, financières et de faillite. Les Types de dette et d'actif ne sont pas inclus dans ce modèle.

berta et Ontario), ne possédant aucun actif et n'ayant personne dans son unité familiale, ayant fait une cession volontaire qui a duré moins de deux ans, sans conditions à sa libération et sans que le syndic se soit rémunéré, le taux de recouvrement serait, en moyenne, de 26,43%. Ce résultat ne peut avoir de signification économique considérant que l'interprétation de certains coefficients est aberrante. En effet, il serait irrationnel de supposer que le syndic maximise sa richesse en travaillant sans rémunération ou encore que l'unité familiale ne compte pas au moins une personne, soit la débitrice elle-même.

Cela dit, la régression linéaire multiple nous permet d'estimer l'effet de chaque variable explicative, tout en maintenant constantes – toutes choses étant égales par ailleurs – les autres facteurs (personnels, financiers et sur la faillite).

5.1.1 Interprétation des variables continues

Rémunération du syndic

Examinons tout d'abord l'effet du ratio de la rémunération relative du syndic sur le taux de recouvrement. L'interprétation du coefficient estimé $\hat{\beta}$ de la rémunération relative du syndic est la suivante : lorsque celui-ci augmente de 1%, le taux de recouvrement des créanciers diminue en moyenne de 0,4078%, toutes choses étant égales par ailleurs. Notons que ce coefficient est statistiquement significatif, c'est-à-dire différent de 0, puisque nous rejetons l'hypothèse nulle H_0 à un niveau de confiance de 99%, à l'aide du test d'hypothèse (4.7) établi dans le chapitre précédent.

Cela nous indique, bien évidemment, que la rémunération relative du syndic a un effet significativement négatif sur le taux de recouvrement des créanciers. On pourrait s'attendre à ce que chaque dollar marginal accordé en rémunération au syndic prive les créanciers du même dollar. En effet, comme nous l'avons vu dans le calcul de la perte en cas de défaut de Zaniboni et collab. (2013), on soustrait des paiements aux créanciers les coûts associés au recouvrement.

Or, il convient de rappeler que la rémunération relative du syndic ne représente pas

tous les coûts de l'administration d'une faillite et qu'il existe d'autres coûts liés à la réalisation des actifs de la faillite. Toutefois, la rémunération des syndics représente le coût le plus important associé au recouvrement.

Notre hypothèse sous entendait un effet de la rémunération des syndics sur le taux de recouvrement des créanciers. Alors que nous concluions dans le Chapitre 3 que les syndics recevaient des honoraires, en moyenne, représentant 45,5% des recettes totales, nous avons maintenant une estimation précise de l'impact de cette appropriation sur la variation du taux de recouvrement des créanciers.

Actifs sur passifs

L'interprétation du ratio Actifs sur passifs est similaire. Lorsque celui-ci augmente de 1%, le taux de recouvrement augmente en moyenne de 0,0569%, en maintenant constantes les autres variables explicatives du modèle. Le coefficient est aussi statistiquement significatif à un niveau de confiance de 99%. On peut aussi interpréter le coefficient de la manière suivante : lorsque le ratio Actifs sur passifs augmente d'une unité, c'est-à-dire de 100%, le taux de recouvrement augmente de 5,69%. Économiquement, cela signifie que si ce ratio double, le taux de recouvrement augmente, en moyenne, de 5,69%. Cette interprétation est pertinente dans la mesure où le ratio Actifs sur passifs pourrait se voir doubler alors que la rémunération relative du syndic est bornée à 1, comme nous l'avons expliqué lors du nettoyage des données présenté dans la section 3.3.

Comme nous l'avons vu dans Hoechstetter et collab. (2012) et dans Zhang et Thomas (2012), le montant de la dette au moment du défaut a un effet sur le taux de recouvrement. Il est équivalent à la valeur du passif que nous reportons dans notre ratio. Par ailleurs, dans Caouette et collab. (2011), nous avons vu que la taille du prêt avait un impact négatif sur le taux de recouvrement. Ainsi, lorsque notre ratio Actifs sur passifs augmente, cela est dû à l'augmentation des actifs ou à la diminution des passifs. Les théories que nous avons vues expliquent donc bien l'effet du ratio Actifs sur passifs sur le taux de recouvrement, et celui-ci correspond à nos attentes.

Nombre de membres dans l'unité familiale

Finalement, lorsque le Nombre de membres dans l'unité familiale augmente d'une unité, soit d'un individu, le taux de recouvrement diminue, en moyenne, de 0,60%, toutes choses étant égales par ailleurs. Nous expliquons le signe de ce coefficient par le fait que plus la taille du ménage de l'individu en faillite augmente, plus les dépenses du débiteur risquent d'être élevées. Par ailleurs, comme nous l'avons vu à la section 3.2, le nombre de membres dans l'unité familiale a un effet direct sur le revenu excédentaire du failli. En effet, plus la taille de l'unité familiale augmente, plus les exigences de revenu mensuel du débiteur sont élevées pour que celui-ci se « qualifie » au revenu excédentaire. Ainsi, plus l'unité familiale compte de membres, moins le débiteur risque de verser du revenu excédentaire à l'actif de sa faillite. Le taux de recouvrement des créanciers s'en trouve donc réduit.

En conclusion, cela confirme notre hypothèse de la section 3.2 concernant l'effet de cette variable sur le taux de recouvrement. Par ailleurs, nous apportons un nouveau résultat à la littérature, soit que la taille du ménage a tendance à réduire le taux de recouvrement lors d'une faillite.

5.1.2 Interprétation des variables catégorielles

Région

Dans son étude, Wirenhammar (2011) présentait la région géographique comme un microfacteur ayant un effet sur la perte en cas de défaut. Nous avons alors décidé d'examiner les variations du taux de recouvrement par région. Comme nous l'avons expliqué dans le Chapitre 3, la variable incluse dans la constante est la catégorie Autre. Voici l'interprétation des coefficients de régions géographiques statistiquement significatifs.

Le fait d'habiter en Alberta par rapport au reste du Canada fait augmenter le taux de recouvrement en moyenne de 3,02%, toutes choses étant égales par ailleurs. Les résidents

du Québec ont un taux de recouvrement inférieur d'en moyenne 4,87% aux résidents du reste du Canada, toutes choses étant égales par ailleurs. Cela confirme une certaine hétérogénéité parmi les provinces canadiennes, en ce qui a trait aux individus en faillite. Les normes du surintendant présentées en Annexe A1 sont appliquées de la même manière à l'échelle nationale, alors qu'il existe, d'après Statistique Canada¹, une disparité documentée entre les salaires des individus dans les provinces.

De surcroît, les normes quant aux biens non saisissables ne sont pas uniformes au sein du Canada : elles varient de province en province. La variable Région s'avère donc intéressante, puisqu'elle permet sans doute de cerner cette particularité des lois provinciales.

Toutefois, bien que lors de l'estimation de la variable Région, nous maintenions constante la rémunération relative des syndics, il convient de nous questionner sur les comportements et les méthodes de travail des syndics dans ces régions. Nous nous pencherons sur la question au chapitre suivant.

Sexe

En ce qui concerne le sexe de l'individu, notre modèle estime qu'un homme qui fait faillite présente, en moyenne, un taux de recouvrement inférieur de 1,80% à celui d'une femme, toutes choses étant égales par ailleurs. Le coefficient est statistiquement significatif à 95%, ce qui confirme les résultats de l'étude de Prívará et collab. (2014). Intuitivement, les hommes sont peut-être en moyenne plus endettés, ayant moins d'aversion pour le risque. En effet, les hommes de notre échantillon présentent un passif significativement supérieur à celui des femmes alors qu'il n'y a pas de différence significative au niveau de l'actif (voir Annexe A2).

1. Enquête canadienne sur le revenu, 2012 :
<https://www150.statcan.gc.ca/n1/daily-quotidien/141210/dq141210a-fra.htm?cmp=mstatcan>

Genre de dépôt

Rappelons que dans notre modèle, la catégorie en référence pour le genre de dépôt est la faillite volontaire. Les résultats accusent une différence significative entre les débiteurs ayant fait une cession volontaire et ceux en cession présumée, à un niveau de confiance de 90%. En effet, les individus ayant au préalable effectué une proposition et étant tombés en faillite présumée ont un taux de recouvrement inférieur d'en moyenne 2,46% à ceux qui ont fait une faillite volontaire, en maintenant constants tous les autres facteurs de notre modèle. Tout comme Gupton et collab. (2000), nous notons que le taux de recouvrement varie d'une procédure à l'autre. Il n'y a toutefois pas de différence significative entre la cession volontaire et l'ordonnance de faillite. Ces résultats nous permettent donc de confirmer notre hypothèse de départ, quant à l'effet du caractère volontaire ou pas de la faillite sur le taux de recouvrement des créanciers.

Nous pouvons supposer que le débiteur qui fait initialement une proposition est probablement dans une meilleure situation financière qu'un débiteur qui choisit de faire faillite en premier lieu, et donc, qu'il accuserait un taux de recouvrement supérieur. Or, nous observons l'effet inverse. Cette différence dans le taux de recouvrement des créanciers peut sans doute s'expliquer par les coûts inhérents à la cession présumée. En effet, une cession présumée implique deux administrations : d'abord celle de la proposition, puis celle de la faillite. Une double administration implique plus de coûts et de frais, puisque les syndics se rémunèrent dans le cadre de chaque administration et que chaque administration engendre aussi ses propres frais. Ces résultats signifient, pour ce qui concerne le taux de recouvrement des créanciers, que les doubles administrations sont moins payantes que les administrations simples, soit les faillites volontaires. Notons que pour un individu présent dans notre échantillon, nous observons uniquement le taux de recouvrement et la rémunération relative du syndic dans le cadre de la faillite. Notre modèle ne considère donc pas les effets de la rémunération relative du syndic de la procédure antérieure.

Par ailleurs, nos observations se limitent à la comparaison, conditionnelle à la faillite, du taux de recouvrement entre une faillite volontaire et une faillite présumée. Nous ne

sommes pas en mesure de comparer le taux de recouvrement d'une faillite présumée au taux de recouvrement que le créancier aurait eu s'il n'avait pas refusé la proposition. Comme nous l'avons expliqué dans la section 3.2, la cession présumée résulte du non-dépôt de la proposition par le débiteur, ou du refus de celle-ci par les créanciers ou la cour. Dans notre échantillon, 86% des cessions présumées sont dues au fait que les créanciers ont refusé la proposition. On peut supposer que si, d'un point de vue financier, les créanciers refusent un montant d'argent proposé par le débiteur, c'est qu'ils sont persuadés qu'ils auraient avantage à ce que ce dernier soit en faillite. Pareillement, la cour refuserait potentiellement une proposition si elle juge que celle-ci n'est pas équitable pour la masse des créanciers. Cette comparaison dépassant le cadre de notre travail, il serait intéressant de l'étudier dans de futures recherches.

Durée de la faillite

Alors que Eraslan et collab. (2017) n'observaient pas de différences significatives des taux de recouvrement en fonction de la durée du plan de remboursement, nous rapportons que plus la durée de la faillite augmente, plus le taux de recouvrement est élevé. En effet, la catégorie de référence étant les faillites ayant duré moins de deux ans, les résultats impliquent que les faillites ayant duré plus de cinq ans ont un recouvrement supérieur à la catégorie de référence d'en moyenne 3,93%.

Cela corrobore nos attentes : en effet, comme nous l'avons souligné au Chapitre 3, en présence de revenu excédentaire, les faillites durent plus longtemps. De même, nous avons émis l'hypothèse qu'un dossier plus complexe était d'une plus longue durée, selon la réflexion de Helwege (1999). Nos résultats corroborent cette hypothèse puisque la durée de la faillite traduit un dossier présentant un taux de recouvrement supérieur.

État civil

En ce qui concerne l'état civil des faillis, la catégorie en référence est celle des faillis

mariés. Alors que Caouette et collab. (2011) rapportaient une cote de crédit supérieure pour les individus mariés, notre modèle indique qu'un failli marié présente un taux de recouvrement inférieur à presque toutes les autres catégories d'état civil. Parmi les catégories où les différences sont les plus significatives, on observe que les faillis en union libre - conjoints de fait - présentent un taux de recouvrement supérieur, en moyenne, de 3,48% à celui des faillis mariés. Il en va de même pour les faillis célibataires, supérieur de 3,65%.

On pourrait comprendre les différences entre le taux de recouvrement des individus célibataires et celui des individus mariés. En effet, en général, les célibataires représentent de plus petits ménages, ont moins de dépenses que les individus mariés, n'ont pas d'enfant à charge et il est donc plus probable qu'ils aient à verser du revenu excédentaire. Mais en ce qui concerne l'écart entre les conjoints de fait et les personnes mariées, on peut se questionner sur les facteurs expliquant un tel résultat.

Ce que le mariage signifie généralement – santé financière, aboutissement personnel, équilibre, etc. – ne s'applique peut-être plus dans l'univers des faillites. Sans doute notre résultat est-il attribuable à la notion de partage du patrimoine familial, qui n'existe pas dans le cas des faillis célibataires ou en union libre. Ainsi certains actifs (résidences familiales, voitures, meubles, régimes de retraite) n'appartiennent qu'à 50% au failli. Les conjoints mariés sont donc généralement copropriétaires indivis d'une résidence familiale. En pratique, lorsqu'un des conjoints dépose une cession de faillite, il faut alors déterminer si le syndic peut exiger la vente de la résidence familiale pour recueillir le produit de la vente de la part indivise du failli. En pareil cas, il faut tenir compte de la législation provinciale régissant le droit matrimonial. En fait, la loi provinciale en vigueur pourrait empêcher le syndic de réaliser une vente forcée de la résidence familiale sans avoir obtenu l'approbation du conjoint qui n'est pas en faillite. On trouve ce genre de restrictions notamment dans le *Code civil du Québec* et la *Loi sur le droit de la famille* de l'Ontario.

Nos résultats correspondent aux conclusions de Prívará et collab. (2014), qui ont observé que les individus mariés avaient un taux de recouvrement plus faible.

Conditions à la libération

La condition est le résultat d'une opposition à la libération du failli. L'opposition provient généralement du syndic ou des créanciers, voire du Bureau du surintendant des faillites. Elle implique forcément le passage du failli devant les tribunaux, excepté dans les rares cas de médiation (uniquement 10 dossiers de notre échantillon). La libération de la faillite est obtenue uniquement si le failli remplit les conditions qui lui ont été imposées.

Notre modèle estime que le fait d'avoir une condition à la libération aboutit à un recouvrement, qui, en moyenne, est plus élevé de 2,20% que celui des faillis n'ayant pas obtenu de condition à leur libération, toutes choses étant égales par ailleurs. Ainsi, cela confirme les prédictions présentées dans la section 3.2 : la condition est souvent d'ordre financier et peut ainsi faire augmenter le recouvrement du dossier de faillite. Nous supposons dans cette section que cette variable impliquait sans doute deux effets contraires : un effet financier, la condition étant, dans la majorité des cas, d'ordre financier ; et un effet de « délinquance », impliquant probablement un manque de collaboration de la part du failli, des paiements non effectués auparavant ou même des actes frauduleux. Or, nos résultats montrent que l'effet financier de la condition dépasse l'effet de « délinquance » du failli.

Ces résultats probants ont une implication très forte pour les créanciers, puisqu'ils ont à leur disposition le mécanisme d'opposition à la libération au failli. Notre hypothèse de départ concernant les recours des créanciers pour augmenter leur recouvrement est ainsi vérifiée. L'opposition est une action que ces derniers contrôlent et peuvent entreprendre contre le débiteur dans le but d'obtenir des conditions favorables à leur recouvrement.

5.2 Régression Types de dette

À cette étape de notre travail, nous avons estimé l'effet de certaines variables sur le taux de recouvrement. Nous ajoutons maintenant à notre modèle d'autres variables afin de vérifier si certains types de dette mènent à des variations significatives dans le taux de recouvrement, ce qui constitue une de nos hypothèses initiales. En effet, en nous

inspirant de Hoechstoeffer et collab. (2012), nous séparons les types de dette en catégories. Rappelons que dans l'étude de de Hoechstoeffer et collab. (2012), le taux de recouvrement était spécifique à une catégorie de prêt et non à un individu particulier, et pouvait donc être ventilé en fonction du type de dette. Considérant la nature de nos données, nous sommes uniquement en mesure de rajouter à notre modèle principal des variables dichotomiques indiquant si le failli détient ou non le type de dette en question, puisque nous n'avons pas accès au taux de recouvrement par types de dette. Notons aussi qu'un failli peut être détenteur de plusieurs types de dette à la fois. Finalement, il est particulièrement important de souligner que les variables sur les types de dette contiennent uniquement les dettes sans garantie, puisque nous estimons que le taux de recouvrement des créanciers non garantis. Nous retrouvons le *Modèle Types de dette* dans le tableau 5.2.

Lorsqu'on ajoute les types de dette à la régression principale, les coefficients des variables initiales demeurent stables et certaines variables semblent même gagner en significativité, comme par exemple le Genre de dépôt avec la cession présumée.

Nos résultats montrent que les faillis qui possèdent une dette de prêt bancaire ont un recouvrement inférieur d'en moyenne 1,74% à ceux qui n'ont pas ce type de dette, toutes choses étant égales par ailleurs. Le résultat est significatif à un niveau de confiance de 95%. Cette catégorie de dette représente les prêts bancaires autres que les hypothèques sur les biens immeubles, c'est-à-dire des prêts bancaires sans biens en garantie. Ce coefficient négatif pourrait s'expliquer par la voracité des banques. En effet, nous croyons que les institutions financières prennent action avant que le débiteur ne fasse faillite. Comme nous l'avons vu à la section 2.3, les créanciers – et plus particulièrement les banques – ont plusieurs recours. Ces dernières peuvent saisir les biens du débiteur, enregistrer des hypothèques légales ou faire des saisies de salaire lors du défaut, mais avant la suspension des procédures qu'entraîne la *Loi sur la faillite et l'insolvabilité*. Cela pourrait avoir comme résultat une réduction de l'actif à réaliser du failli et, conséquemment, un taux de recouvrement plus faible pour les individus qui détiennent ce type de prêt.

	Modèle principal	Modèle Types de dette	Modèle Types d'actif
Intercept	0.2643 (0.0200)***	0.3020 (0.0238)***	0.2842 (0.0273)***
Rémunération du syndic	-0.4078 (0.0174)***	-0.4104 (0.0174)***	-0.4090 (0.0174)***
Actifs sur passifs	0.0569 (0.0088)***	0.0551 (0.0089)***	0.0562 (0.0089)***
Nombre de membres dans l'unité familiale	-0.0060 (0.0036)*	-0.0055 (0.0036)	-0.0049 (0.0036)
Région (Ref = Autre)			
Alberta	0.0302 (0.0114)***	0.0288 (0.0114)**	0.0261 (0.0121)**
Ontario	-0.0181 (0.0113)	-0.0146 (0.0114)	-0.0211 (0.0123)*
Québec	-0.0487 (0.0129)***	-0.0506 (0.0132)***	-0.0397 (0.0144)***
Sexe (Ref = Femme)			
Homme	-0.0180 (0.0086)**	-0.0157 (0.0086)*	-0.0126 (0.0088)
Genre de dépôt (Ref = Cession volontaire)			
Ordonnance de faillite	-0.0337 (0.0397)	-0.0284 (0.0397)	-0.0302 (0.0398)
Cession présumée	-0.0246 (0.0134)*	-0.0280 (0.0136)**	-0.0274 (0.0137)**
Durée de la faillite (Ref = Moins de 2 ans)			
2 à 3 ans	0.0088 (0.0130)	0.0068 (0.0130)	0.0069 (0.0130)
3 à 4 ans	0.0215 (0.0136)	0.0211 (0.0136)	0.0187 (0.0136)
4 à 5 ans	0.0180 (0.0148)	0.0168 (0.0148)	0.0166 (0.0148)
Plus de 5 ans	0.0393 (0.0144)***	0.0389 (0.0144)***	0.0385 (0.0144)***
Syndic non libéré	0.0303 (0.0185)	0.0295 (0.0185)	0.0306 (0.0184)*
État civil (Ref = Marié)			
Conjoint de fait	0.0348 (0.0138)**	0.0338 (0.0138)**	0.0356 (0.0138)**
Divorcé	0.0160 (0.0144)	0.0196 (0.0145)	0.0216 (0.0145)
Séparé	0.0220 (0.0126)*	0.0242 (0.0126)*	0.0233 (0.0127)*
Célibataire	0.0365 (0.0139)***	0.0349 (0.0140)**	0.0389 (0.0140)***
Veuf	0.0461 (0.0245)*	0.0434 (0.0245)*	0.0446 (0.0243)*
Conditions à la libération (Ref = Sans conditions)			
Avec conditions	0.0220 (0.0084)***	0.0208 (0.0084)**	0.0202 (0.0084)**
Types de dette			
Hypothèques sur les biens immeubles		0.0089 (0.0120)	0.0126 (0.0120)
Prêts bancaires - excepté hypothèques		-0.0174 (0.0088)**	-0.0174 (0.0088)**
Prêts des sociétés de crédit		0.0081 (0.0083)	0.0067 (0.0084)
Cartes de crédit émises par banques et autres		-0.0085 (0.0110)	-0.0057 (0.0109)

	Modèle principal	Modèle Types de dette	Modèle Types d'actif
Autres émetteurs de cartes de crédit		-0.0031 (0.0084)	-0.0027 (0.0085)
Impôts et taxes municipales		-0.0154 (0.0080)*	-0.0103 (0.0081)
Prêts étudiants		0.0276 (0.0218)	0.0061 (0.0220)
Prêts de particuliers		-0.0241 (0.0135)*	-0.0284 (0.0135)**
Autre		-0.0138 (0.0090)	-0.0111 (0.0090)
Types d'actif			
Équipement de loisirs			-0.0030 (0.0132)
Valeur de rachat			0.0021 (0.0079)
Chalet			0.0032 (0.0303)
Remboursement d'impôt estimé			0.0047 (0.0259)
Effets personnels			0.0083 (0.0100)
Motocyclette			0.0061 (0.0180)
Valeurs mobilières			-0.0048 (0.0104)
Ameublement			0.0231 (0.0126)*
Automobile			-0.0062 (0.0099)
Espèces en main			-0.0070 (0.0105)
Motoneige			0.0051 (0.0306)
Terrain			-0.0104 (0.0170)
Autre éléments d'actif			-0.0233 (0.0081)***
Autre véhicules à moteur			-0.0035 (0.0149)
R ²	0.3461	0.3544	0.3633
Adj. R ²	0.3367	0.3409	0.3432
Num. obs.	1412	1412	1409
RMSE	0.1457	0.1452	0.1440

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 5.2 – Comparaison des trois modèles de régression - Nous ajoutons au modèle de régression linéaire principal, qui estimait le taux de recouvrement par les MCO, un groupe de variables indicatrices sur les types de dette, puis un deuxième groupe de variables indicatrices sur les types d'actif.

Deux autres coefficients sont statistiquement significatifs à 90% : la catégorie Impôts

fédéral/provincial et des taxes municipales ainsi que la catégorie Prêts de particuliers. Le fait de détenir une dette d'impôt a un effet négatif sur le taux de recouvrement. Il en est de même pour les Prêts de particuliers : en ce qui concerne les individus qui ont une créance auprès d'un particulier, nous établissons que l'effet négatif sur le taux de recouvrement est probablement attribuable au phénomène même qui s'applique aux Prêts bancaires, à un degré différent. En effet, les particuliers sont généralement le groupe le plus lésé lors de faillites : ce ne sont pas des prêteurs institutionnels, qui détiennent des portefeuilles de prêts diversifiés. Ils ne gèrent donc pas le risque de défaut d'une manière globale comme le font les autres créanciers et, donc, ne récupèrent pas leurs pertes ailleurs (autres prêts ou autres activités). De ce fait, il nous apparaît logique que les prêteurs particuliers se montrent voraces avant la faillite et tentent de recouvrer les montants qui leur sont dûs, privant ainsi la masse des créanciers d'une partie du recouvrement lors de la faillite.

Nous pouvons alors vérifier notre hypothèse de départ concernant les types de dette : les individus qui détiennent des dettes de prêts bancaires, des dettes d'impôts et de taxes et des dettes de prêt de particuliers présentent en moyenne un taux de recouvrement significativement inférieur à ceux qui n'ont pas ces types de dette.

5.3 Régression Types d'actif

Comme nous l'avons fait pour les types de dette, nous rajoutons à nos modèles les types d'actif que le failli détient. Les types d'actif sont des variables dichotomiques indiquant si l'individu détient ou non le bien en question. Précisons qu'un failli peut être détenteur de plusieurs actifs à la fois. Il est aussi pertinent d'indiquer que, contrairement aux types de dette, les types d'actif ne nous permettent pas d'exclure les actifs qui garantissent le prêt d'un créancier. Aussi, nous décidons de ne pas inclure dans notre modèle la variable dichotomique Maison². En effet, la maison est, parmi les actifs, celui qui a le

2. Nous avons testé le modèle incluant la variable Maison et obtenu des coefficients instables. Comme nous le savons, lorsqu'on ajoute à un modèle une nouvelle variable explicative qui a une corrélation importante avec une autre variable, la part d'explication de l'une et de l'autre se chevauche, ce qui peut mener

plus de valeur, sauf exception rare. Toutefois, le passif des individus en faillite n'est pas uniquement constitué de l'hypothèque associée à la maison, mais d'autres dettes généralement de nature non garantie. Cela explique que cette variable soit très corrélée avec le ratio Actifs sur passifs, puisque, lorsque l'individu possède une maison, son ratio d'actifs sur passifs est susceptible d'augmenter. Ainsi, le ratio Actifs sur passifs prend probablement en considération le fait d'être propriétaire d'une maison.

Conserver cette variable dans notre modèle nous aurait confronté à un problème de multi colinéarité, avec pour conséquences une variance exagérée des estimateurs, des signes contradictoires des coefficients et des résultats instables. Le test de corrélation de Pearson entre le ratio Actifs sur passifs et le fait d'avoir une maison révèle que la corrélation entre les deux variables est supérieure à 0,5.

Cela étant dit, nous croyons important d'indiquer que lors de notre analyse des types de dette, nous avons conservé la variable Hypothèque sur les biens immeubles – qui est une hypothèque non garantie³ –, celle-ci ne présentant pas une corrélation problématique avec la variable Actifs sur passifs, pour la simple et bonne raison que les types de dette excluaient les dettes garanties.

Les résultats du *Modèle Types d'actif* sont également présentés dans le tableau 5.2.

La variable Ameublement indique si le débiteur possède du mobilier. Notons que les meubles, qui garnissent la résidence principale, servent à l'usage du ménage et sont nécessaires à la vie de celui-ci, sont insaisissables jusqu'à concurrence d'une valeur marchande de 7 000 \$. Notre modèle estime que le fait d'avoir de l'ameublement fait augmenter le taux de recouvrement en moyenne de 2,31%. Ce coefficient ne peut pas être interprété comme un lien de causalité entre le fait de posséder des meubles et de présenter un taux

à des variations considérables des paramètres estimés. Nous avons soupçonné qu'il y avait une interaction entre ces deux variables. Nous avons donc développé un modèle incluant l'interaction entre la variable Maison et le ratio Actifs sur passifs, avec des résultats toujours instables. Pour alléger notre étude, nous ne présentons pas les différents modèles testés.

3. Cette situation peut survenir si le créancier hypothécaire renonce à sa garantie – ce qui serait contre-intuitif. Cela peut aussi arriver lorsque la valeur de l'hypothèque est supérieure à la valeur du bien en garantie. Alors, le mécanisme de la faillite permet au créancier hypothécaire d'être créancier garanti à hauteur de la valeur de la maison, et de prouver une réclamation non garantie pour son solde hypothécaire, soit le manque à gagner. Il participera, pour sa portion non garantie, au taux de recouvrement que nous étudions.

de recouvrement supérieur : en effet, nous savons par expérience que le mobilier est rarement saisi et réalisé par le syndic. Dans notre échantillon, plus de 85% des individus déclarent posséder de l'ameublement. Or, nous estimons que presque tous les individus, à quelques exceptions près, possèdent du mobilier. Nous pouvons supposer que cette variable procède de facteurs endogènes à l'individu (honnêteté, anxiété, etc.) qui ont un effet sur ses déclarations à propos de ses actifs.

En ce qui concerne la variable Autres éléments d'actif, le modèle indique que le fait de posséder d'autres éléments d'actif fait diminuer, en moyenne, le taux de recouvrement de 2,33%, et ce, à un niveau de confiance de 99%. Les autres éléments d'actif ne sont pas clairement définis par notre dictionnaire de données. L'interprétation de cette variable est compromise par le fait qu'elle est constituée d'éléments épars et sans rapport les uns avec les autres. Nous pourrions toutefois l'inclure dans nos modèles de prédiction du Chapitre 6, ceux-ci n'ayant plus d'objectif d'estimation et d'inférence statistique.

Concernant les types d'actif, nous pouvons alors vérifier notre hypothèse de départ : les individus qui possèdent de l'ameublement présentent en moyenne un taux de recouvrement significativement supérieur à ceux qui n'en déclarent pas. Par contre, les individus qui ont d'autres éléments d'actifs présentent un taux de recouvrement inférieur. Notons finalement que le type de dette Impôts fédéral/provincial et taxes municipales n'est plus significatif dans le modèle.

5.4 Coefficient de détermination

Le coefficient de détermination, ou la statistique du R^2 , est calculée selon la formule 4.8.

Tel que présenté dans le tableau 5.2, nos modèles présentent, respectivement, un R^2 de 34,61%, de 35,44%, et de 36,33%, ce qui, à notre connaissance, est largement supérieur au R^2 recensé par la littérature de la perte en cas de défaut. En effet, Zhang et Thomas (2012) rapportaient que le modèle linéaire était celui qui expliquait le mieux la variabilité du taux de recouvrement avec une statistique R^2 de 9,04%. De leur côté, Prívará et collab. (2014), grâce à leurs modèles de survie, atteignaient 15% d'explication.

Notre *Modèle Types de dette* et *Modèle Types d'actif* améliorent l'explication de notre variable dépendante, si l'on se base sur le R^2 . Toutefois l'ajout de plusieurs variables explicatives pourrait artificiellement augmenter le R^2 , qui est fonction du nombre de variables. Il convient d'utiliser le R^2 ajusté lorsque l'on compare des modèles n'ayant pas le même nombre de variables explicatives, puisque la valeur de R^2 ajusté intègre le nombre de prédicteurs dans le modèle. Le R^2 ajusté se calcule selon la formule 4.9.

Lorsqu'on compare le R carré ajusté des deux modèles, la différence entre les trois modèles est mince : 33,67% contre 34,09% et 34,32%. Ainsi, avec l'ajout des types de dette et d'actif, nous gagnons presque un point de base sur l'explication des variations du taux de recouvrement.

5.5 Problèmes relatifs à la validité interne du modèle

Nous nous penchons sur deux menaces potentielles à la validité de notre modèle de régression :

- l'endogénéité de certaines variables explicatives, en particulier la rémunération relative du syndic,
- un biais de sélection en rapport avec notre variable dépendante, le taux de recouvrement.

Rappelons qu'il existe plusieurs biais d'endogénéité, entre autres, le biais de variables omises et le biais de causalité simultanée. Dans un modèle de régression, le terme d'erreur ε_i contient tout ce qui n'est pas compris dans le modèle. De ce fait, si on omet d'intégrer dans le modèle une variable explicative corrélée à la fois avec la variable dépendante et avec une autre variable explicative X_i , alors X_i sera corrélé avec le terme d'erreur. Cela cause un biais que l'on nomme biais de variables omises puisque H1 sera violée : $E(\varepsilon_i | x_{1i}, x_{2i}, \dots, x_{pi}) = 0$.

Nous soupçonnons plutôt qu'un biais de causalité simultanée entre le taux de recouvrement et la rémunération relative du syndic affecte notre modèle. Alors que la rému-

nération du syndic a un effet direct et relativement linéaire sur le taux de recouvrement – puisque le syndic se rémunère à même les recettes du dossier et verse par la suite, le cas échéant, un dividende aux créanciers –, nous nous questionnons sur la possibilité que le taux de recouvrement ait également un effet causal sur la rémunération du syndic. En effet, on peut se demander si les dossiers qui présentent un taux de recouvrement non nul sont ceux dans lesquels le syndic a reçu moins de rémunération ; ou encore si les dossiers qui n’ont aucun recouvrement sont ceux dans lesquels le syndic a reçu une rémunération plus élevée.

L’éventualité de ce biais signifierait que la rémunération relative du syndic serait corrélée avec le terme d’erreur ε_i de la régression principale. La rémunération relative du syndic serait donc endogène au modèle, ce qui pourrait avoir deux conséquences : les paramètres de notre modèle seraient estimés moins précisément (écarts types plus larges, ou, bien plus grave, notre estimateur des moindres carrés serait biaisé.

Nous vérifierons, dans le prochain chapitre, si la rémunération relative du syndic est corrélée avec d’autres variables du modèle.

D’un autre côté, le biais de sélection, contrairement au biais de variables omises, résulte du processus selon lequel les données sur la variable dépendante ont été récoltées ; il ne dépend pas des variables explicatives incluses dans notre modèle. Ainsi, si des observations du taux de recouvrement sont manquantes du fait de notre méthode de sélection non aléatoire des données, alors ce biais pourrait introduire une corrélation entre l’erreur de notre modèle et les variables explicatives.

Notre échantillon n’est pas sélectionné aléatoirement. Dans l’univers des débiteurs, certains sont en faillite et d’autres ne le sont pas. Le biais peut alors provenir de la sélection non aléatoire de dossiers d’individus uniquement en faillite. Dans notre étude, nous n’observons que le taux de recouvrement pour des individus en faillite. Toutefois, il peut y avoir de l’auto-sélection (*self-selection*) parmi les individus étudiés. En effet, ces derniers ont, dans plus de 85% des cas, fait faillite volontairement. La volonté de faire faillite implique sans doute des raisons particulières, sachant qu’il existe plusieurs moyens de remédier au défaut avant d’en venir à la faillite. Le choix de la faillite constitue une dé-

cision dont les raisons pourraient être endogènes à notre échantillon – p. ex. : les autres solutions au défaut offraient moins d'utilité à l'individu que la faillite, l'individu a subi des pressions de créanciers, la faillite a constitué un moyen vexatoire contre les créanciers, l'individu a fait l'objet de saisies de salaire ou de procédures judiciaires contre lui et voulait utiliser la faillite comme moyen de suspension des procédures.

Selon ce raisonnement, notre fonction du taux de recouvrement est peut-être uniquement adaptée à notre échantillon et il est possible de croire qu'elle n'estimerait pas correctement le taux de recouvrement d'un individu de la population générale sélectionné aléatoirement et assujetti à la faillite.

Heckman (1977) propose un modèle en deux parties pour corriger ce biais. Toutefois, pour l'appliquer, il faut avoir accès aux données à la fois des individus en défaut et des individus en faillite, dans le but de créer un modèle de sélection. Ce modèle mesure la probabilité d'appartenir au groupe de la faillite notamment grâce à des facteurs qui tiennent compte des raisons menant à la faillite. N'ayant pas accès aux données de la population générale en défaut, nous ne serons pas en mesure d'aspirer à corriger ce biais.

Conclusion

Jusqu'à présent, nous avons estimé le taux de recouvrement à l'aide de régressions linéaires. Nos résultats ont montré que certaines variables ont un effet statistiquement significatif sur les variations du taux de recouvrement. Toutefois, les résidus de notre modèle ne semblaient pas être distribués normalement (voir Annexe A5). En effet, une des difficultés auxquelles nous nous heurtons est le poids important sur les taux de recouvrement nul – soit plus de 30% de notre échantillon. Cela peut engendrer des erreurs dans notre estimation linéaire du taux de recouvrement.

Comme nous en avons discuté dans la revue de littérature, plusieurs méthodes permettent de pallier les problèmes qu'entraîne la distribution bimodale du taux de recouvrement. Nous retenons les études de Hoechstetter et collab. (2012) et de Thomas et collab. (2012), qui emploient un modèle en deux parties. Ce type de modélisation semble plus

robuste pour la distribution spécifique au taux de recouvrement : il consiste à séparer l'échantillon en deux sous-groupes, de manière à estimer le groupe qui nous intéresse, dans notre cas celui ayant des taux de recouvrement non nuls. Nous notons toutefois que ce type de modélisation peut mener à un biais de sélection. En effet, en nous inspirant de Dahen et Dionne (2010), étudier uniquement les taux de recouvrement positifs sans estimer la probabilité qu'un individu en faillite ait un recouvrement positif pourrait conduire à des résultats biaisés.

Dans le prochain chapitre, lors de notre estimation du taux de recouvrement non nul, nous prendrons en compte ce biais de sélection potentiel grâce à l'application de Heckman (1977). Nous sommes, à notre connaissance, les premiers auteurs à appliquer la méthodologie développée par Heckman à l'étude du taux de recouvrement.

Chapitre 6

Résultats additionnels

Nous présentons dans ce chapitre les résultats additionnels de notre recherche.

6.1 Régression sur la rémunération du syndic

Nous nous questionnons sur l'endogénéité de la rémunération du syndic. En effet, la rémunération du syndic représente un coût pour le créancier, comme nous l'avons défini avec la formule 3.2 du *Workout LGD* de Zaniboni et collab. (2013).

Notre approche consiste à estimer les facteurs ayant un effet sur le coût principal de la faillite : la rémunération relative du syndic. Est-elle homogène au Canada ? Y a-t-il certaines administrations plus rentables pour les syndics ? Leur rémunération est-elle fonction de la durée de la faillite ? La présence de certaines dettes dans le passif de l'individu a-t-elle une incidence sur la rémunération du syndic ? Les résultats du *Modèle Rémunération du syndic* sont présentés dans le tableau 6.1.

Comme nous l'avons conjecturé, l'hypothèse que les méthodes de travail des syndics, leur prise de risque, leur manière d'administrer un dossier et de se rémunérer pourraient être différentes d'une région à l'autre est ici vérifiée. En effet, les syndics en Alberta ont une rémunération relative aux recettes des dossiers plus élevée de 16,93% que celle des syndics des autres provinces. La différence existe à un degré moindre entre les syndics du Québec et ceux des autres provinces. On peut se demander si, en Alberta et au Québec, la

Modèle Rémunération du syndic	
Intercept	0.3205 (0.0286)***
Région (Ref = Autre)	
Alberta	0.1694 (0.0169)***
Ontario	-0.0093 (0.0174)
Québec	0.0403 (0.0198)**
Genre de dépôt (Ref = Cession volontaire)	
Ordonnance de faillite	0.1068 (0.0556)*
Cession présumée	0.0207 (0.0206)
Durée de la faillite (Ref = Moins de 2 ans)	
2 à 3 ans	0.0470 (0.0201)**
3 à 4 ans	0.0146 (0.0208)
4 à 5 ans	0.0511 (0.0227)**
Plus de 5 ans	0.0623 (0.0220)***
Syndic non libéré	0.0979 (0.0279)***
Types de dette	
Hypothèques sur les biens immeubles	0.0151 (0.0183)
Prêts bancaires - excepté hypothèques	-0.0102 (0.0135)
Prêts des sociétés de crédit	0.0058 (0.0128)
Cartes de crédit émises par banques et autres	0.0379 (0.0169)**
Autres émetteurs de cartes de crédit	-0.0034 (0.0130)
Impôts et taxes municipales	-0.0101 (0.0122)
Prêts étudiants	0.1278 (0.0334)***
Prêts de particuliers	-0.0060 (0.0207)
Autre	0.0121 (0.0138)
R ²	0.1283
Adj. R ²	0.1165
Num. obs.	1428
RMSE	0.2258

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 6.1 – Modèle Rémunération du syndic - Résultats du modèle de régression linéaire estimant la rémunération relative du syndic par les MCO en fonction de la région, du genre de dépôt, de la durée de la faillite et des types de dettes.

rémunération des syndics est supérieure ou si les recettes sont inférieures. Pour ce qui est de la catégorie Autres provinces, les recettes totales des dossiers sont en moyenne de 39 192 \$ alors qu'en Alberta, elles sont significativement inférieures de 13 408 \$. Cependant, il n'existe pas de différence statistiquement significative entre la rémunération moyenne des syndics dans les dossiers du Québec et celle des Autres provinces (voir respectivement l'Annexe A3 et l'Annexe A4).

En ce qui concerne l'administration, nos résultats indiquent que l'ordonnance de faillite, par comparaison à la cession volontaire, mène à une rémunération relative du syndic supérieure de 10,67% à un niveau de confiance de près de 95%. Les résultats témoignent également de l'impact de la durée de la faillite sur la rémunération relative du syndic : plus la durée augmente, plus la rémunération du syndic augmente. Finalement, nos résultats révèlent que certains types de dette, à savoir les Cartes de crédit émises par les banques et les sociétés de fiducie et les Prêts étudiants, lorsqu'ils sont présents dans le passif du débiteur, font significativement augmenter la rémunération relative du syndic.

Bien que nous soyons en mesure de montrer que certaines variables explicatives de notre modèle principal ont un effet causal sur la rémunération relative, nous ne sommes pas en mesure de trouver un instrument valide susceptible de remplacer la rémunération relative du syndic, pouvant expliquer à la fois cette dernière et n'étant pas corrélé avec les erreurs de notre modèle principal.

6.2 Régression avec interactions

Il est naturel de penser que dans nombre de problèmes, des interactions existent entre les variables explicatives. Par exemple, notre *Modèle principal* expliquant le taux de recouvrement des créanciers nous mène à penser que la rémunération relative du syndic et la région interagissent, ou encore que la taille des faillites interfère avec les régions. Pour modéliser ces interactions, nous écrivons un modèle avec un produit entre les variables explicatives qui interagissent. Nous avons décrit le modèle simple dans l'équation 4.10 de la section 4.1.3.

Nous cherchons une explication au fait que certaines régions présentent des taux de recouvrement statistiquement différents. Deux hypothèses nous viennent à l'esprit :

- les syndicats travaillent et se rémunèrent d'une manière différente de région en région (Rémunération du syndic \times Région), ce qui a des conséquences sur le taux de recouvrement. Nous créons donc la variable croisée entre la rémunération relative du syndic et les régions,
- la taille des faillites n'est pas homogène au Canada (Actifs sur passifs \times Région), ce qui devrait également affecter le taux de recouvrement. Nous créons donc la variable croisée entre le ratio Actifs sur passifs et les régions.

Nous ajoutons ces deux variables d'interaction au modèle. Si nos hypothèses sont correctes, alors nos variables d'interaction seront statistiquement significatives. Nous présentons les résultats du *Modèle avec interactions* dans le tableau 6.2.

Lorsque l'interaction est entre une variable continue et une variable catégorielle, les coefficients prennent un sens précis.

	<i>Modèle principal sans interactions</i>	<i>Modèle principal avec interactions</i>
Intercept	0.2643 (0.0200)***	0.2584 (0.0266)***
Rémunération du syndic	-0.4078 (0.0174)***	-0.4469 (0.0373)***
Actifs sur passifs	0.0569 (0.0088)***	0.0841 (0.0166)***
Nombre de membres dans l'unité familiale	-0.0060 (0.0036)*	-0.0061 (0.0036)*
Région (Ref = Autre)		
Alberta	0.0302 (0.0114)***	0.0692 (0.0298)**
Ontario	-0.0181 (0.0113)	-0.0238 (0.0308)
Québec	-0.0487 (0.0129)***	-0.1136 (0.0336)***
Sexe (Ref = Femme)		
Homme	-0.0180 (0.0086)**	-0.0172 (0.0085)**
Genre de dépôt (Ref = Cession volontaire)		
Ordonnance de faillite	-0.0337 (0.0397)	-0.0245 (0.0396)
Cession présumée	-0.0246 (0.0134)*	-0.0255 (0.0135)*
Durée de la faillite (Ref = Moins de 2 ans)		
2 à 3 ans	0.0088 (0.0130)	0.0101 (0.0130)
3 à 4 ans	0.0215 (0.0136)	0.0245 (0.0135)*
4 à 5 ans	0.0180 (0.0148)	0.0216 (0.0148)
Plus de 5 ans	0.0393 (0.0144)***	0.0426 (0.0144)***
Syndic non libéré	0.0303 (0.0185)	0.0345 (0.0184)*
État civil (Ref = Marié)		
Conjoint de fait	0.0348 (0.0138)**	0.0316 (0.0138)**
Divorcé	0.0160 (0.0144)	0.0177 (0.0144)
Séparé	0.0220 (0.0126)*	0.0220 (0.0125)*
Célibataire	0.0365 (0.0139)***	0.0324 (0.0139)**
Veuf	0.0461 (0.0245)*	0.0431 (0.0243)*
Conditions à la libération (Ref = Sans conditions)		
Avec conditions	0.0220 (0.0084)***	0.0215 (0.0083)**
Région (Ref = Autre) * Actifs sur passifs		
Alberta * Actifs sur passifs		-0.0565 (0.0213)***
Ontario * Actifs sur passifs		0.0018 (0.0264)
Québec * Actifs sur passifs		-0.0281 (0.0278)
Région (Ref = Autre) * Rémunération du syndic		
Alberta * Rémunération du syndic		0.0118 (0.0472)
Ontario * Rémunération du syndic		0.0125 (0.0498)
Québec * Rémunération du syndic		0.2032 (0.0566)***
R ²	0.3461	0.3591
Adj. R ²	0.3367	0.3471
Num. obs.	1412	1412
RMSE	0.1457	0.1445

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 6.2 – Modèle avec interactions - Nous comparons le modèle de régression linéaire principal estimant le taux de recouvrement par les MCO au modèle incluant l'interaction entre la région et le ratio d'actifs sur passifs et l'interaction entre la région et la rémunération du syndic.

Rémunération du syndic × Région

Concentrons-nous tout d'abord sur l'interaction entre la rémunération du syndic et la région. Nos résultats démontrent que dans le reste du Canada, lorsque la rémunération relative du syndic augmente de 1%, le taux de recouvrement des créanciers diminue en moyenne de 0,4469%, toutes choses étant égales par ailleurs. La différence statistiquement significative dans les termes d'interaction s'interprète comme suit :

La différence de l'effet de la rémunération du syndic sur le taux de recouvrement entre le reste du Canada et le Québec est de 20,32%. Pour le Québec, l'effet de la rémunération du syndic sur le taux de recouvrement est de $0,2032 - 0,4469 = -0,2437$, c'est-à-dire que lorsque la rémunération relative du syndic augmente de 1%, le taux de recouvrement diminue en moyenne d'uniquement 0,2437%. En d'autres mots, cela signifie qu'au Québec, l'effet de la rémunération relative du syndic est presque deux fois plus faible que dans le reste du Canada.

Il nous est difficile, dans le cadre de notre travail, de justifier ce résultat, en raison des nombreux paramètres pouvant expliquer cet écart. Par exemple, les différences de comportements et de méthodes de travail des syndics entre le Québec et le reste du Canada, les législations provinciales, etc. Dans le tableau 5.1, nous observons dans notre *Modèle principal* que les résidents du Québec avaient en moyenne un taux de recouvrement inférieur de 4,87% à ceux du reste du Canada. Nous expliquions cette différence, entre autres, par les disparités entre les salaires des individus dans les provinces. Le même raisonnement peut s'appliquer ici : s'il s'avérait que les honoraires professionnels de certains intervenants – syndics, avocats, comptables, etc. – étaient moins élevés au Québec, les coûts de la faillite seraient réduits et auraient donc moins d'effet sur le taux de recouvrement au Québec.

Nous pourrions faire le parallèle avec le concept d'élasticité en économie, qui mesure la sensibilité des acheteurs et des vendeurs à une variation dans les conditions du marché. C'est la variation en pourcentage d'une variable à la suite d'une variation de 1% d'une autre variable. Appliqué à notre sujet, on pourrait dire que le taux de recouvrement au

Québec est moins élastique aux coûts de la faillite que le taux de recouvrement des autres régions.

Notre résultat a une forte implication pour les créanciers du Québec, qui voient un effet réduit de la rémunération relative du syndic sur leur taux de recouvrement. Même si le Québec recense un taux de recouvrement inférieur au reste du Canada, le coût majeur et inévitable de la rémunération du syndic prive presque deux fois moins les créanciers québécois de recouvrement.

Actifs sur passifs × Région

En ce qui a trait à l'interaction entre la taille des faillites et la région, dans le reste du Canada, lorsque le ratio Actifs sur passifs augmente de 1%, le taux de recouvrement des créanciers augmente en moyenne de 0,0841%. La différence statistiquement significative dans les termes d'interaction s'interprète comme suit :

La différence de l'effet du ratio Actifs sur passifs sur le taux de recouvrement entre le reste du Canada et l'Alberta est de 5,65%. Pour l'Alberta, l'effet du ratio Actifs sur passifs sur le taux de recouvrement est de $-0,0565 + 0,0841 = 0,0276$, c'est-à-dire que lorsque le ratio Actifs sur passifs augmente de 1%, le taux de recouvrement augmente en moyenne uniquement de 0,0276%. En d'autres mots, cela signifie qu'en Alberta, le taux de recouvrement augmente trois fois moins que dans le reste du Canada quand le ratio Actifs sur passifs augmente, c'est-à-dire lorsque soit les actifs augmentent soit les passifs diminuent. En Alberta, le taux de recouvrement des créanciers est donc moins sensible aux actifs et aux passifs des débiteurs.

Nous pouvons donc conclure à propos des différences observées grâce à nos interactions que les deux hypothèses énoncées dans cette section sont vérifiées. En effet, les syndics travaillent et se rémunèrent d'une manière différente de région en région, et la taille des faillites n'est pas homogène au Canada. Cela a un effet particulier sur le taux de recouvrement dans certaines régions, notamment au Québec et en Alberta.

6.3 Application du modèle de Heckman (1977)

Nous poursuivons l'objectif principal de notre étude, c'est-à-dire une estimation non biaisée des paramètres affectant le taux de recouvrement. Comme nous l'avons constaté plus tôt, la distribution de notre variable dépendante contient un nombre élevé d'observations nulles. Puisqu'il y a une concentration de valeurs de la variable dépendante à une certaine limite (ici à zéro), l'estimation par les moindres carrés ordinaires peut être biaisée puisque la variable dépendante n'est pas continue et est bornée.

6.3.1 Choix de modèle

Certaines techniques sont utilisées quand un nombre élevé d'observations contient des zéros : c'est le cas des *zero inflated models* qui utilisent comme distribution une loi de Poisson ou une loi binomiale négative. Cependant, ces méthodes nécessitent des données de comptage, c'est-à-dire que la variable dépendante prend uniquement des valeurs entières non négatives et provient d'un processus de classement. Ces méthodes ne peuvent pas s'appliquer lorsque la variable dépendante est continue comme le taux de recouvrement de notre travail.

Dans son étude, Tobin (1958) se heurte à un problème de variable dépendante limitée, c'est-à-dire où la variable dépendante n'est observable que sur un certain intervalle. Les modèles à variable dépendante limitée dérivent des modèles à variables qualitatives, dans le sens où l'on doit modéliser la probabilité que la variable dépendante appartienne à l'intervalle pour lequel elle est observable. Tobin (1958) cherche à expliquer la relation entre les dépenses en biens durables et le revenu des ménages. Toutefois, plusieurs observations concernant le montant de dépenses sont nulles, alors qu'il dispose de l'information concernant le revenu, ce qui caractérise un échantillon censuré. Tobin (1958) propose alors un modèle dit de censure, qu'on a subséquemment nommé modèle de Tobit, puisqu'il ne pouvait pas utiliser une densité continue pour expliquer la distribution conditionnelle des dépenses par rapport au revenu.

Cependant, d'après Smith et Brame (2003), une importante limitation du modèle de

Tobit est l'hypothèse voulant que l'effet des variables indépendantes est constant pour la procédure de sélection et celle du phénomène étudié. Les mêmes variables indépendantes déterminent la probabilité qu'une observation soit censurée (égale à zéro) et la valeur de la variable dépendante lorsqu'elle est observable (supérieure à zéro). Cela nous empêche de proprement justifier, d'un point de vue théorique, pourquoi les observations qui sont censurées le sont (sont-ils de vrais zéros?).

De leur côté, Hoehstoetter et collab. (2012) et Thomas et collab. (2012) segmentent l'estimation du taux de recouvrement, en séparant l'échantillon en deux parties. Cela est communément appelé modèle en deux parties ou *hurdle model*. Les *hurdle models* sont similaires aux modèles de Tobit, puisqu'ils concernent une variable dépendante bornée. Ils sont généralement exprimés sous forme de deux équations : la première détermine si la variable dépendante dépasse le seuil ou la borne choisie ; la deuxième estime la valeur de la variable dépendante à condition qu'elle soit supérieure au seuil. Les *hurdle models* se distinguent toutefois des modèles de Tobit dans la mesure où les variables explicatives de la première équation peuvent différer de celles de la deuxième équation.

Thomas et collab. (2012) construisent un *hurdle model* en segmentant leur échantillon : les taux de recouvrement nuls et les taux de recouvrement strictement positifs. La première partie estime la probabilité d'observer un recouvrement non nul à l'aide d'une régression logistique. La deuxième partie du modèle est l'estimation des taux de recouvrement non nuls avec un modèle de régression linéaire par les moindres carrés ordinaires. Selon les auteurs, la segmentation peut améliorer la précision de la prédiction, ce qui a déjà été démontré par Matuszyk et collab. (2010). En effet, une des raisons de modéliser le taux de recouvrement comme tel réside dans la disparité des mentalités des individus en ce qui concerne le remboursement de dettes. Hoehstoetter et collab. (2012) utilisent la même démarche : ils estiment en premier lieu la probabilité d'observer un taux de recouvrement extrême – nul ou intégral –, puis se concentrent sur la modélisation des taux de recouvrement non extrêmes, entre autres à l'aide de régressions linéaires.

Toutefois, les *hurdle models* reposent sur l'hypothèse que les résidus de la première équation ne sont pas corrélés avec ceux de la deuxième équation du modèle. Cette hy-

pothèse implique que ce sont différentes personnes qui choisissent parmi les différentes possibilités, mais qu'il n'y a pas de facteurs endogènes pouvant expliquer les différences entre les deux groupes. Les auteurs ne tiennent donc pas compte du biais de sélection dans la deuxième partie de leur modèle, ce qui risque de mener à une estimation biaisée des paramètres.

C'est pour cette raison que nous privilégions les modèles de sélection d'échantillon. Ces derniers répondent aux lacunes du modèle de Tobit puisqu'un groupe de variables indépendantes déterminent la probabilité de censure, et un groupe de variables distinct peut estimer la valeur de la variable dépendante lorsqu'elle n'est pas censurée. Ces variables explicatives peuvent se chevaucher ou être complètement différentes. Nous y reviendrons dans la section 6.3.2.

De plus, les modèles de sélection d'échantillon, contrairement aux *hurdle models*, nous permettent de prendre en compte le processus de censure, puisqu'on suppose que les résidus de l'équation de sélection ne sont pas indépendants des résidus du modèle non censuré. C'est dans cette optique que nous utiliserons le modèle de Heckman (1977) pour corriger le biais de sélection potentiel, comme nous l'avons décrit à la section 4.1.4.

6.3.2 Modèle de Heckman (1977)

Notre objectif est d'améliorer notre modèle initial, qui incluait les taux de recouvrement nuls, en estimant uniquement l'effet des variables indépendantes sur les taux de recouvrement strictement positifs. Notre modèle d'estimation devient alors l'étude d'une variable dépendante censurée – puisqu'elle est assujettie à la condition d'avoir un recouvrement – et il peut y avoir un biais de sélection. En effet, la question est de savoir si le fait d'avoir un recouvrement à zéro s'explique vraiment par les variables indépendantes du modèle – dans ce cas, ce sont de vrais zéros – ou bien s'il y a des variables endogènes propres à l'individu ou au dossier qui font en sorte que celui-ci a un recouvrement de zéro – par exemple, s'il choisit de ne pas travailler, ou s'il décide de ne pas déclarer

tous ses actifs ou de s'en débarrasser avant la faillite. Les faux zéros se retrouvent dans les erreurs de la première partie du modèle et viennent ainsi biaiser les estimateurs de la deuxième partie du modèle, soit l'estimation du taux de recouvrement conditionnel à un recouvrement non nul. Cela impliquerait qu'il n'est pas aléatoire de présenter un taux de recouvrement positif, et qu'il existe des facteurs pouvant expliquer les différences entre les recouvrements nuls et non nuls.

Nous voulons ainsi corriger le biais de sélection, le cas échéant, en appliquant Heckman (1977). Il s'agit d'un aspect qui n'a pas été abordé par Thomas et collab. (2012), ce qui a rendu leur démarche incomplète. La méthodologie de Heckman se décompose en deux parties.

Modèle de sélection

La première étape de cette procédure consiste à déterminer les sources de différenciation existant entre deux populations, dans notre cas, les individus dont la valeur de la variable dépendante est strictement supérieure à zéro, et les individus dont la valeur prise par cette variable est égale à zéro. À l'aide d'un modèle probit, il s'agit dans un premier temps d'estimer les effets de différentes variables explicatives sur la variable dépendante – équation de sélection – afin de pouvoir contrôler ces mêmes effets lors de l'estimation du modèle – seconde étape. Au-delà de ces facteurs de biais mesurables, la méthode de Heckman autorise l'évaluation des effets non mesurés par les variables prises en compte sur la variable dépendante au cours de l'équation de sélection. Cette information se trouve dans le terme d'erreur contenant les variations résiduelles non expliquées par les variables indépendantes du modèle. Ces variations proviennent de caractéristiques propres aux individus ou aux dossiers étudiés, qui ne sont pas prises en compte dans les variables explicatives. La méthode de Heckman permet précisément de contrôler ces effets. Les résidus de l'équation de sélection correspondant aux effets non mesurés servent à construire un facteur de contrôle du biais de sélection, nommé λ et correspondant à l'inverse du ratio de Mills, tel que décrit dans l'équation 4.13.

Ajoutons finalement que la méthode de correction de Heckman est supérieure en pré-

sence de restrictions d'exclusion (*exclusion restrictions*) : ce sont des variables qui affectent le processus de sélection sans pour autant avoir un effet sur la variable dépendante non censurée. Avec des restrictions d'exclusion, le modèle se prête à une approche causale du problème du biais de sélection. Cela réduit également la problématique de corrélation introduite par le facteur de correction λ dans la deuxième étape du modèle de Heckman, lorsqu'on introduit l'inverse du ratio de Mills en tant que régresseur dans l'équation 4.12. Avec des restrictions d'exclusion valides, l'inverse du ratio de Mills et le vecteur de régresseurs X_i seront moins corrélés, ce qui réduira le problème de multicollinéarité entre les variables explicatives et avec le terme d'erreur. C'est également en présence de restrictions d'exclusion que les résultats du modèle de Heckman (1977) sont différents de ceux du modèle de Tobin (1958).

Le choix des variables explicatives pour le modèle de sélection probit est donc critique. Il faut inclure dans ce modèle les variables qui discriminent notre échantillon en deux groupes : les recouvrements strictement positifs et les recouvrements nuls. Nous ne pensons pas que les variables doivent être entièrement différentes de celles du modèle linéaire. Nous ajoutons toutefois trois variables explicatives au modèle de sélection : le revenu d'emploi disponible du failli, la variable binaire sur le rachat d'actif, et la variable binaire sur l'ameublement. Ces variables aident à expliquer les différences entre nos deux groupes, et donc à prédire la probabilité d'observer un taux de recouvrement strictement positif, sans pour autant avoir un effet sur les taux de recouvrement.

Nous présentons les résultats de l'équation de sélection dans le tableau 6.3 et discutons subséquemment du choix de variables explicatives dans le *Modèle de sélection*.

La première partie du modèle est estimée avec une régression probit, qui suppose que le terme d'erreur est indépendant et normalement distribué. Cela nous permet de calculer l'inverse du ratio de Mills, qui doit être généré à partir d'une loi normale, comme nous l'avons défini à la section 4.1.4. Notre régression probit présente comme coefficients les valeurs z d'une distribution normale, qui n'ont pas d'interprétation économique en tant que telle. Toutefois, le signe des coefficients et leur significativité nous informent sur la

	Modèle de sélection
Intercept	1.8571 (0.2717)***
Rémunération du syndic	-3.7142 (0.2253)***
Actifs sur passifs	0.1047 (0.1052)
Nombre de membres dans l'unité familiale	-0.0775 (0.0427)*
Revenu d'emploi disponible	0.0001 (0.0000)***
Ameublement	0.3810 (0.1347)***
Rachat d'actif (Ref = Sans)	
Avec rachat	0.4175 (0.1415)***
Région (Ref = Autre)	
Alberta	-0.0087 (0.1413)
Ontario	-0.1020 (0.1425)
Québec	-0.6216 (0.1597)***
Sexe (Ref = Femme)	
Homme	0.1891 (0.1000)*
Genre de dépôt (Ref = Cession volontaire)	
Ordonnance de faillite	-1.0034 (0.9138)
Cession présumée	-0.1455 (0.6178)
Durée de la faillite (Ref = Moins de 2 ans)	
2 à 3 ans	0.0451 (0.1549)
3 à 4 ans	0.1649 (0.1636)
4 à 5 ans	0.1335 (0.1764)
Plus de 5 ans	-0.1111 (0.1709)
Syndic non libéré	0.4365 (0.2327)*
État civil (Ref = Marié)	
Conjoint de fait	-0.0033 (0.1615)
Divorcé	-0.1253 (0.1641)
Séparé	-0.0932 (0.1502)
Célibataire	0.1365 (0.1735)
Veuf	-0.0510 (0.2780)
Conditions à la libération (Ref = Sans conditions)	
Avec conditions	0.0610 (0.1014)
Num. obs.	1188

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 6.3 – Modèle de sélection - Résultats du modèle de sélection de Heckman pour l'estimation par régression probit de la probabilité d'observer un taux de recouvrement strictement positif. Nous notons les variables explicatives statistiquement significatives suivantes : la rémunération du syndic, le revenu d'emploi disponible du failli, la variable binaire sur le rachat d'actif, la variable binaire sur l'ameublement et la région du Québec.

probabilité d'observer un recouvrement non nul. En effet, pour des variables ayant des coefficients négatifs, la probabilité d'observer un taux de recouvrement non nul diminue et, inversement, pour des variables ayant des coefficients positifs, la probabilité d'observer un taux de recouvrement non nul augmente.

En ce qui concerne les restrictions d'exclusion, la variable Revenu d'emploi disponible représente le revenu d'emploi net des dépenses non discrétionnaires de l'individu en faillite. Ce facteur indique donc si le failli travaille ou pas, mais plus encore, nous renseigne sur son salaire. Ainsi, plus celui-ci augmente, plus la probabilité d'observer un taux de recouvrement non nul augmente.

De la même manière, la variable Rachat d'actif nous indique si le failli a décidé de racheter un de ses actifs au syndic. Comme nous le savons, lors de la faillite, le débiteur cède ses actifs au syndic. Il incombe alors à ce dernier de réaliser l'équité sur les biens, le cas échéant, en procédant à leur vente. Cependant, la liquidation d'actifs peut être coûteuse pour le syndic. Il va sans dire que les coûts associés à la vente du bien découragent parfois les syndics. C'est pourquoi la *Loi sur la faillite et l'insolvabilité* prévoit que le failli peut lui-même racheter ses actifs au syndic. Il s'engage donc à verser l'équité sur le bien et en conserve l'usage. Il s'avère donc manifeste qu'en cas de rachat d'actif dans la faillite, la probabilité d'observer un taux de recouvrement augmente. Le concept de rachat d'actif sensibilise peut-être les faillis à déclarer leurs actifs. Dans cette optique, notre variable nous permet de différencier les vrais des faux zéros, soit nos deux sous-groupes de recouvrement nul et non nul.

Finalement, comme nous l'avons expliqué dans la section 5.3, nous supposons pour la variable Ameublement que cette dernière procédait de facteurs endogènes à l'individu (honnêteté, anxiété, etc.) qui pouvaient contenir de l'information concernant les déclarations du failli à l'égard de ses actifs. C'est pour cette raison que nous l'ajoutons au *Modèle de sélection*, ce qui vient renforcer la validité des restrictions d'exclusion, considérant que le coefficient de cette variable est également significatif à 99 %.

Modèle de correction

Une fois l'équation de sélection effectuée, nous sommes en mesure de calculer l'inverse du ratio de Mills pour chaque observation. La seconde étape de l'analyse est réalisée à l'aide d'une régression par les moindres carrés ordinaires. Le facteur de contrôle du biais de sélection λ est introduit dans la régression en tant que variable explicative additionnelle. Le coefficient estimé par l'inverse du ratio de Mills procure ainsi la part des effets des caractéristiques non observées corrélée avec notre variable dépendante. On obtient donc pour les autres variables explicatives des coefficients non biaisés par les caractéristiques non prises en compte dans le modèle initial. Les résultats du *Modèle avec correction Heckman* sont présentés dans le tableau 6.4.

Nous nous penchons tout d'abord sur le paramètre λ . Les deux étapes ayant des résidus pour chaque observation, nous devons examiner la relation entre ces résidus pour tester le biais de sélection. Rappelons que les facteurs non observables se retrouvent dans les résidus. Si les facteurs non observables du modèle de sélection sont corrélés avec les facteurs non observables de la deuxième étape, alors nos estimateurs de la deuxième étape seront biaisés. Cela signifierait que le fait d'avoir un taux de recouvrement strictement positif n'est pas aléatoire et, donc, que des facteurs non observables affectent le fait d'avoir un recouvrement positif ou non. L'inverse du ratio de Mills nous permet de contrôler ce biais.

Dans le *Modèle avec correction Heckman*, le coefficient de l'inverse du ratio de Mills n'est pas statistiquement significatif. Cela signifie qu'à un niveau de confiance suffisant, celui-ci n'est pas différent de zéro. Cela implique que les coefficients des variables explicatives de notre modèle sans la correction de Heckman ne sont pas biaisés. En effet, nous pouvons croire qu'il n'y a pas de biais de sélection entre les deux groupes étudiés. Aussi, nous pensons que les variables explicatives choisies à l'étape de l'équation de sélection ont laissé peu d'effets non mesurés qui affectent notre variable dépendante.

L'interprétation de ce modèle est similaire à notre modèle de régression 5.2 qui estimait au préalable tous les recouvrements. Toutefois, dans ce modèle d'estimation des taux

	Modèle avec correction Heckman
Intercept	0.2934 (0.0378)***
Rémunération du syndic	-0.5938 (0.0780)***
Actifs sur passifs	0.1091 (0.0227)***
Nombre de membres dans l'unité familiale	0.0018 (0.0051)
Prêts bancaires - excepté hypothèques	-0.0254 (0.0119)**
Prêt de particuliers	-0.0510 (0.0183)***
Autre éléments d'actif	-0.0316 (0.0109)***
Région (Ref = Autre)	
Alberta	0.0902 (0.0398)**
Ontario	-0.0222 (0.0433)
Québec	-0.1009 (0.0535)*
Sexe (Ref = Femme)	
Homme	-0.0300 (0.0122)**
Genre de dépôt (Ref = Cession volontaire)	
Ordonnance de faillite	0.0449 (0.1567)
Cession présumée	-0.0802 (0.0699)
Durée de la faillite (Ref = Moins de 2 ans)	
2 à 3 ans	0.0166 (0.0183)
3 à 4 ans	0.0411 (0.0188)**
4 à 5 ans	0.0524 (0.0207)**
Plus de 5 ans	0.0663 (0.0204)***
Syndic non libéré	0.0549 (0.0261)**
État civil (Ref = Marié)	
Conjoint de fait	0.0462 (0.0198)**
Divorcé	0.0394 (0.0204)*
Séparé	0.0603 (0.0175)***
Célibataire	0.0678 (0.0193)***
Veuf	0.0544 (0.0338)
Conditions à la libération (Ref = Sans conditions)	
Avec conditions	0.0251 (0.0119)**
Région (Ref = Autre) * Actifs sur passifs	
Alberta * Actifs sur passifs	-0.0764 (0.0281)***
Ontario * Actifs sur passifs	0.0050 (0.0368)
Québec * Actifs sur passifs	-0.0347 (0.0500)
Région (Ref = Autre) * Rémunération du syndic	
Alberta * Rémunération du syndic	0.0387 (0.0780)
Ontario * Rémunération du syndic	-0.0001 (0.0836)
Québec * Rémunération du syndic	0.2426 (0.1039)**
Inverse du ratio de Mills	0.0287 (0.0412)
R ²	0.3565
Adj. R ²	0.3327
Num. obs.	1188
Censored	345
Observed	843

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE 6.4 – Modèle avec correction Heckman - Résultats du modèle de correction de Heckman pour l'estimation des taux de recouvrement strictement positifs par régression linéaire. On ajoute aux variables explicatives le facteur de contrôle du biais de sélection λ , soit l'inverse du ratio de Mills. Les écarts-types des estimateurs sont corrigés.

de recouvrement non nuls, on constate que :

- la rémunération du syndic est d'autant plus importante lorsque le taux de recouvrement des individus est strictement positif. Cela signifie que le syndic absorbe une plus grande partie du taux de recouvrement lorsque celui-ci est strictement positif ;
- la durée de la faillite est d'autant plus significative : les catégories de durée ont un meilleur pouvoir explicatif lorsque les dossiers ont un recouvrement non nul,
- le genre de dépôt n'a plus d'effet significatif sur les taux de recouvrement non nuls.

Alors que nous estimons dans le *Modèle avec correction Heckman* uniquement les dossiers qui présentent un taux de recouvrement non nul, nous observons que les facteurs propres aux individus – soit les variables personnelles et financières – restent relativement inchangés face à la régression sur les taux de recouvrement qui incluaient les zéros. Toutefois, nous observons des différences dans les paramètres des variables de la faillite lorsqu'on estime les taux de recouvrement strictement positifs. En somme, en se basant sur le R^2 ajusté, nos variables indépendantes expliquent plus de 33% des variations du taux de recouvrement non nul, ce qui reste inférieur à l'estimation du modèle linéaire global non censuré.

Finalement, puisque le coefficient de l'inverse du ratio de Mills n'est pas significatif dans le *Modèle avec correction Heckman*, nous serons en mesure, dans la prochaine section de l'étude, d'élaborer des modèles de prédiction sur les taux de recouvrement non nuls sans se soucier du biais de sélection.

6.4 Modèles de prédiction

Jusqu'à présent, nous avons estimé les variables ayant un effet sur le taux de recouvrement, en nous appuyant sur nos connaissances du domaine de l'insolvabilité et sur la revue de littérature. Nous avons bâti un modèle expliquant près de 35% de la variabilité du taux de recouvrement.

Nous avons par la suite abordé les menaces à la validité de notre modèle, plus particulièrement en ce qui concerne la distribution du taux de recouvrement, ainsi que la possibilité d'un biais de sélection. Nous avons donc présenté un modèle estimant uniquement les taux de recouvrement positifs, en corrigeant le biais de sélection, grâce à la méthode de Heckman (1977). Or, nos résultats permettent de croire qu'il n'existe pas de biais de sélection dans notre modèle. Par conséquent, les variables affectant les taux de recouvrement positifs étaient, à une exception près, les mêmes que celles de notre modèle sans correction.

Nous présentons, dans cette dernière section des résultats additionnels, différents modèles de prédiction du taux de recouvrement. Tous les modèles que nous présentons reposent sur un objectif commun : proposer aux créanciers un instrument permettant de prédire le taux de recouvrement lors du dépôt du dossier de faillite. Cela leur permettrait, entre autres, de calculer leur capital réglementaire en fonction de la prédiction du taux de recouvrement conditionnel à la faillite et non pas simplement au défaut.

Dans cette perspective, les variables sur la faillite ne feront plus partie de notre modèle, étant donné que ces données sont uniquement disponibles à l'issue de la faillite. Nous n'utiliserons donc que les variables qui sont à disposition de tous les créanciers au moment du dépôt de la faillite, soit celles du formulaire 78 (Bilan) et du formulaire 65 (État des revenus et dépenses du failli).

Nous proposerons trois modèles de prédiction, dont nous comparerons la qualité mesurée par l'erreur de prédiction. Les trois modèles utilisent la même technique de validation croisée en 10 échantillons, décrite dans la section 4.2.1. Nous entraînons les modèles sur 90% de l'échantillon et effectuons des prédictions sur les 10% de l'échantillon restant, ce qui revient à faire de la prédiction hors échantillon (*out sample*). Nous répétons la procédure 10 fois. Voici la description des trois modèles :

- *Modèle de classification probit* : notre premier modèle a pour objectif de prédire la probabilité d'observer un recouvrement non nul avec un modèle de classification probit,

- *Modèle de prédiction linéaire* : notre deuxième modèle de prédiction est effectué à l'aide d'une régression linéaire, sur tous les taux de recouvrement (incluant les zéros),
- *Modèle de prédiction linéaire censuré* : notre troisième modèle de prédiction est effectué à l'aide d'une régression linéaire, uniquement sur les taux de recouvrement strictement positifs. Nous n'avons pas besoin d'inclure le terme de correction de Heckman, puisque nous avons montré dans la dernière section qu'il n'y avait pas de biais de sélection parmi nos deux sous-groupes.

Le premier modèle est le seul modèle qui emploie une régression probit et qui prédit donc une probabilité. Ce modèle est le seul pour lequel nous effectuerons de la classification. Les deux autres modèles utilisant des régressions linéaires, ils seront présentés séparément et comparés entre eux.

6.4.1 Modèle de classification probit

Le premier modèle permet de prédire la probabilité d'observer un taux de recouvrement non nul. Du point de vue des créanciers, il semble particulièrement pertinent de savoir si l'individu présente un potentiel de recouvrement ou pas. Cela peut conduire à une plus grande implication du créancier dans certains dossiers d'intérêt, puisque, par exemple, comme nous l'avons conclu dans le chapitre précédent, l'opposition à la libération du failli fait significativement augmenter le taux de recouvrement des créanciers. De surcroît, l'opposition a un effet supérieur sur le taux de recouvrement dans le modèle étudiant les individus qui présentent un recouvrement non nul.

Les résultats du modèle de prédiction probit nous permettent de classer les recouvrements nuls et non nuls déterminés au moyen d'un seuil préétabli. Si la probabilité prédite d'observer un taux de recouvrement est supérieure au point de coupure, alors nous assignons cet individu au groupe des taux de recouvrement positifs. Inversement, si la probabilité prédite d'observer un taux de recouvrement est inférieure au point de cou-

pure, alors nous assignons cet individu au groupe des taux de recouvrement nuls. Nous classifions alors les observations parmi les quatre catégories suivantes :

- Vrais positifs (VP) = le nombre d'individus correctement identifiés à un taux de recouvrement positif,
- Faux positifs (FP) = le nombre d'individus incorrectement identifiés à un taux de recouvrement positif,
- Vrais négatifs (VN) = le nombre d'individus correctement identifiés à un taux de recouvrement nul,
- Faux négatifs (FN) = le nombre d'individus incorrectement identifiés à un taux de recouvrement nul.

Nous calculons alors la statistique qui représente la proportion de bonnes classifications, soit la somme des vrais positifs et des vrais négatifs parmi toutes nos observations prédites

$$\text{Proportion de bonnes classifications} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}}. \quad (6.1)$$

De plus, grâce aux classifications, nous sommes en mesure de présenter la matrice de confusion. Cette dernière résume la capacité d'un modèle à prédire les taux de recouvrement nuls et non nuls. La matrice de confusion est un tableau qui fournit le type de classification en suivant les quatre catégories que nous venons de définir.

Si on définit l'hypothèse nulle $H_0 : y_i = 0$ (pas de recouvrement), les mauvaises classifications se traduisent ainsi :

- une erreur de type I se produit lorsque l'hypothèse nulle H_0 est vraie, mais est rejetée – cela revient à affirmer quelque chose qui est absent, soit un faux positif,
- une erreur de type II se produit lorsque l'hypothèse nulle H_0 est fausse, mais n'est pas rejetée – cela revient à ne pas parvenir à affirmer ce qui est présent, soit un faux négatif.

Le tableau 6.5 correspond à une matrice de confusion générale.

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	Vrais négatifs	Faux négatifs - Erreur de type II
	Recouvrements	Faux positifs - Erreur de type I	Vrais positifs

TABLE 6.5 – Matrice de confusion générale

Choix du point de coupure

Ces classifications sont très sensibles au degré de conservatisme du modèle, soit le point de coupure que l'on choisit. Le choix du point de coupure dépend généralement du domaine d'application du modèle de prédiction. Par exemple, en médecine, il est préférable de diagnostiquer un cancer à un patient qui, en réalité, n'est pas atteint du cancer, au risque de ne pas lui diagnostiquer le cancer alors qu'il est en réalité atteint. Dans ce cas particulier, l'objectif est donc de minimiser les erreurs de type II.

Dans le cadre de notre étude, nous souhaitons minimiser les Faux positifs. En effet, le temps d'un créancier étant précieux, nous ne voulons attirer leur attention que sur les dossiers présentant un réel potentiel de recouvrement, au risque de ne pas indiquer un dossier qui présentera un taux de recouvrement non nul, mais qui sera de toute manière à l'avantage du créancier. Nous cherchons donc à minimiser les erreurs de type I.

Nous utilisons donc trois points de coupure pour classifier nos prédictions : 0,5 ; la proportion des taux de recouvrement non nuls ; et enfin, l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen.

Prendre 0,5 comme point de coupure revient à assigner l'observation à la catégorie la plus probable, ce qui est raisonnable. Mais nous savons que les conséquences de classifier faussement une observation à 0 lorsqu'elle vaut 1 en réalité et que les conséquences de classifier faussement une observation à 1 lorsqu'elle vaut 0 en réalité, ne sont pas les mêmes. Il peut donc être avantageux d'utiliser un autre point de coupure. En effet, un point de coupure élevé est plus conservateur pour classifier une observation comme étant positive : il est donc moins probable d'observer des Faux positifs, mais plus probable de ne pas indiquer des dossiers qui s'avèrent, ultimement, positifs. Nous irons donc dans

cette direction.

Les deux autres points de coupure, soit la proportion des taux de recouvrement non nuls et l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen, sont estimés uniquement sur les échantillons entraînant nos modèles. Puisque nous effectuons 10 itérations pour notre validation croisée, les points de coupure varieront durant chaque itération, en respectant néanmoins la règle établie. Notons que certaines variables sont des variables catégorielles. Ainsi, pour l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen, nous utilisons le mode statistique pour les variables catégorielles, ce qui consiste à prendre la catégorie la plus fréquente. Nous présenterons ultimement la moyenne de ces deux points de coupure sur les 10 itérations.

Pour vérifier la précision de notre modèle, nous utilisons la métrique de la proportion de bonnes classifications définie plus haut et le AUC (*Area Under the Curve*), l'aire sous la courbe ROC. La courbe ROC est la courbe qui fait varier la sensibilité en fonction de $1 - \text{spécificité}$.

La sensibilité représente la probabilité qu'une observation appartenant à la catégorie 1 soit bien classifiée

$$\text{Sensibilité} = P(\hat{y}_i = 1 \mid y_i = 1).$$

Ainsi, la sensibilité mesure à quel point notre modèle est performant pour détecter la catégorie 1, soit les Vrais positifs.

La spécificité représente la probabilité qu'une observation appartenant à la catégorie 0 soit bien classifiée

$$\text{Spécificité} = P(\hat{y}_i = 0 \mid y_i = 0).$$

La spécificité mesure donc à quel point notre modèle est performant pour détecter la catégorie 0, soit les Vrais négatifs.

Un modèle parfait aurait $\text{Sensibilité} = \text{Spécificité} = 1$. Ainsi, plus le couple $(1 - \text{Spécificité}, \text{Sensibilité})$ est près du point $(0, 1)$ de la courbe ROC, meilleur est le modèle. En effet, plus le point de coupure augmente, plus la sensibilité et le taux de Faux

positifs diminuent, mais plus la spécificité et le taux de Faux négatifs augmentent. Alors, la quantité AUC est souvent utilisée, car elle donne l'aire sous la courbe ROC. Plus cette valeur est élevée, mieux c'est. La valeur maximale est 1.

Résultats

Nous présentons les résultats de notre classification grâce au modèle probit dans les matrices de confusion en fonction des différents points de coupure dans le tableau 6.6. Les variables explicatives utilisées pour le modèle de classification sont les mêmes que celles du *Modèle de sélection* de Heckman 6.3, sans les variables qui sont uniquement disponibles à l'issue de la procédure, soit la rémunération du syndic, la durée de la faillite, et la libération conditionnelle.

Notons tout d'abord que les matrices présentées sont les moyennes des prédictions sur les 10 itérations. C'est pour cette raison que les nombres ne sont pas entiers. Les proportions de bonnes classifications sont respectivement de 69,63%, 62,17% et 47,84% pour les trois modèles. plus le point de coupure augmente, plus la proportion de bonnes classifications diminue. Toutefois, comme nous l'avions conjecturé, l'augmentation du point de coupure fait diminuer le nombre d'erreurs de type I, soit les Faux positifs. C'est pourquoi le modèle utilisant l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen comme point de coupure est celui qui minimise, parmi nos modèles, les erreurs de type I.

Alors que le modèle ayant comme point de coupure 0,5 maximise, parmi nos modèles, la proportion de bonnes classifications, le modèle utilisant la proportion des taux de recouvrement non nuls comme point de coupure est celui qui maximise l'aire sous la courbe ROC.

Il convient ici de choisir le point de coupure qui est en adéquation avec l'objectif du modèle. Dans notre cas, nous expliquions plus tôt que le temps d'un créancier étant précieux, nous ne voulions attirer leur attention que sur les dossiers présentant un réel potentiel de recouvrement, ce qui équivaut à se tromper le moins souvent lorsque nous prédisons un recouvrement non nul.

Point de coupure = 0,5

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	5.1	5.2
	Recouvrements	31.5	79.1
			<i>Moy. auc = 0.5394</i>
			<i>Moy. prop. bonnes classifications = 0.6963</i>

Point de coupure = Proportion des taux de recouvrement non nuls

Moyenne = 0.6818

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	19.9	29.1
	Recouvrements	16.7	55.2
			<i>Moy. auc = 0.5998</i>
			<i>Moy. prop. bonnes classifications = 0.6217</i>

Point de coupure = Estimation probabilité d'un recouvrement non nul pour l'individu moyen

Moyenne = 0.7881

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	30.5	57.1
	Recouvrements	6.1	27.2
			<i>Moy. auc = 0.5787</i>
			<i>Moy. prop. bonnes classifications = 0.4784</i>

TABLE 6.6 – Modèle de classification probit en faisant varier les points de coupures

Ainsi, le choix du point de coupure est endogène à l'utilisateur du modèle, c'est-à-dire qu'il doit tenir compte de sa volonté de s'impliquer dans un dossier donné et des ressources dont il dispose. Le point de coupure est donc adaptable au créancier. Nous suggérons toutefois, comme point de coupure, l'emploi des modèles utilisant la proportion des taux de recouvrement non nuls ou utilisant l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen.

6.4.2 Modèle de prédiction linéaire

Une fois effectuée la prédiction sur la probabilité d’observer un taux de recouvrement strictement positif, attardons-nous à prédire le taux de recouvrement en utilisant un modèle de régression linéaire. Comme nous l’avons vu dans Zhang et Thomas (2012), les meilleurs résultats en ce qui a trait aux modèles de prédiction étaient atteints lorsque le taux de recouvrement était en pourcentage et non en dollars. Selon ce résultat, nos modèles de prédiction seront estimés grâce à notre variable dépendante en pourcentage. Par ailleurs, puisque notre variable dépendante est continue et se situe entre 0 et 1, nous utiliserons la statistique de l’erreur quadratique moyenne (MSE), qui mesure la magnitude moyenne de l’erreur de la manière suivante

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (6.2)$$

les y_i étant les valeurs de la variable dépendante observées et les \hat{y}_i étant les valeurs de la variable dépendante prédites par notre modèle. Plus l’erreur quadratique moyenne sera proche de zéro, plus la précision de notre modèle sera élevée. Lorsqu’il ne risque pas d’y avoir de valeurs extrêmes, le MSE est une méthode plus robuste que l’erreur absolue moyenne (MAE), qui représente la moyenne sur l’échantillon test des différences absolues entre la prévision et l’observation réelle.

Nous comparons donc nos deux modèles de prédiction linéaire sur la base du MSE à l’aide du tableau 6.7. Les variables explicatives utilisées pour les deux modèles linéaires sont les mêmes que celles du *Modèle principal*, en excluant les variables qui sont uniquement disponibles à l’issue de la procédure de faillite : la rémunération du syndic, la durée de la faillite, et la libération conditionnelle. Nous ajoutons dans un deuxième temps les variables du *Modèle Types de dette et d’actif* qui étaient significatives dans le tableau 5.2 et les interactions du *Modèle avec interaction* présentées dans le tableau 6.2.

Notre premier modèle de prédiction linéaire permet de prédire le taux de recouvrement. Nous utilisons ici la population entière, soit les individus ayant un recouvrement nul et non nul. La moyenne de la métrique MSE sur les 10 itérations effectuées est de

		Modèle de prédiction linéaire	Modèle de prédiction linéaire censuré
MSE moyen	Variables <i>Modèle Principal</i>	0.0293	0.0339
	Variables + <i>Dettes, Actif, Interaction</i>	0.0287	0.0327

TABLE 6.7 – Modèles de prédiction linéaires

0,0293. Nos résultats surpassent largement ceux de Zhang et Thomas (2012), qui obtenaient un MSE de 0,1675 sur leur échantillon test. Également, nous présentons une erreur inférieure à Prívará et collab. (2014), qui obtenaient un MSE de 0,137. Lorsque nous ajoutons les variables des types de dette, d’actif et d’interaction qui avaient un effet significatif sur le taux de recouvrement, le MSE est en moyenne légèrement inférieur à 0,0287.

6.4.3 Modèle de prédiction linéaire censuré

Lorsque nous effectuons des prédictions uniquement pour les taux de recouvrement positifs, le modèle présente une erreur quadratique moyenne de 0,0354, supérieure à celle du modèle de prédiction linéaire sur toute la population. Ainsi, le modèle de prédiction sur les recouvrements strictement positifs a un pouvoir prédictif inférieur à celui du modèle linéaire avec les zéros. Lorsque nous ajoutons les variables des types de dette, d’actif et d’interaction susmentionnées, le MSE est légèrement inférieur à 0,0327.

En somme, la prédiction sur les recouvrements strictement positifs n’améliore pas la statistique MSE. Au contraire, elle est moins précise que la prédiction sur le modèle incluant tous les individus. Nous pouvons conclure que la segmentation de notre échantillon en deux sous-groupes de recouvrement nul et non nul, n’est pas bénéfique en ce qui a trait à la modélisation du taux de recouvrement des faillites. De surcroît, comme nous l’avons vu dans la section précédente, l’estimation du modèle linéaire des taux de recouvrement non nuls présentait un R^2 ajusté inférieur à celui du modèle incluant les zéros.

6.4.4 Modèle de classification lasso

Nous avons achevé nos modèles de prédiction linéaires et comparé l’erreur de prédiction de chacun. Nous avons conclu que notre modèle de prédiction linéaire sur le taux de

recouvrement en pourcentage a surpassé ceux de la littérature.

Notre objectif ultime consiste à améliorer notre premier modèle probit, qui prédit la probabilité d'observer un taux de recouvrement non nul. Les prédicteurs de nos modèles précédents étant établis sur la base de nos connaissances et de la littérature, nous proposons finalement une méthodologie alternative, basée sur la qualité de prédiction, qui effectue une sélection de variables selon un critère donné. Ainsi, nous pourrions comparer notre modèle basé sur les connaissances à ce dernier modèle orienté sur les données.

Ainsi, en vue d'améliorer la proportion de bonnes classifications du *Modèle de classification probit*, notre dernier modèle de prédiction est effectué en utilisant *Least Absolute Shrinkage and Selection Operator* ou le lasso, une technique d'apprentissage machine supervisée. Nous fournissons au modèle l'ensemble des variables qui sont à la disposition des créanciers lors du dépôt de la faillite. Tout comme nos modèles de régression linéaire par les moindres carrés ordinaires, le lasso minimise la somme des carrés des résidus, mais est assujéti à une contrainte supplémentaire, qui contracte la valeur de certains coefficients à zéro. Cette contraction des coefficients nous préserve d'une augmentation de la variance des coefficients estimés, comme cela aurait été le cas avec un modèle de régression linéaire. Le modèle ne sera donc pas surestimé (*overfitted*), puisque le coefficient des variables non importantes est réduit à zéro. Ainsi, la contraction des coefficients à zéro exerce un rôle de sélection automatique de variables. Nous pouvons donc fournir au modèle un nombre élevé de variables, c'est-à-dire dans notre cas toutes les variables disponibles au sein des formulaires 65 et 78.

En effet, comme nous l'avons vu dans la formule 4.15, le lasso détermine et sélectionne les variables qui ont un pouvoir prédictif sur le taux de recouvrement, en fonction du λ choisi. Notre objectif étant d'améliorer le modèle de classification probit basé sur nos connaissances, nous choisissons λ pour le lasso qui maximise la métrique AUC. Les variables retenues par le lasso sont présentées en Annexe A6. Les résultats du *Modèle de classification lasso*, estimés par une régression probit, avec les variables explicatives retenues par le lasso, sont présentés dans le tableau 6.8.

Notons que nous ne présentons que deux points de coupure. En effet, selon les va-

riables retenues par le lasso, l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen est en moyenne de 0,6758, alors que la proportion des taux de recouvrement non nuls est en moyenne de 0,6818 ; les deux points de coupures étant similaires, nous présenterons uniquement les résultats du point de coupure avec la proportion des taux de recouvrement non nuls.

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	4.4	2.7
	Recouvrements	32.5	82.2
		<i>Moy. auc</i> = 0.5437	
		<i>Moy. prop. bonnes classifications</i> = 0.7108	

		Observés	
		Recouvrements nuls	Recouvrements
Prédits	Recouvrements nuls	23	35.1
	Recouvrements	13.9	49.8
		<i>Moy. auc</i> = 0.6046	
		<i>Moy. prop. bonnes classifications</i> = 0.5979	

TABLE 6.8 – Modèle de classification lasso en faisant varier les points de coupures

Avec un point de coupure calculé comme étant la proportion des taux de recouvrement non nuls, les résultats du *Modèle de classification lasso* sont inférieurs à ceux du *Modèle de classification probit* basé sur nos connaissances. En effet, le *Modèle de classification lasso* présente un AUC très légèrement supérieur mais une proportion de bonnes classifications inférieure à notre modèle de classification basé sur les connaissances.

Toutefois, lorsque le point de coupure est à 0,5, le *Modèle de classification lasso* présente un couple de bonnes classifications et de AUC qui sont tous les deux supérieur à notre modèle basé sur les connaissances. Cependant, il y a davantage d'erreurs de type I dans les prédictions du modèle Lasso.

Alors, compte tenu de nos résultats, nous ne pouvons conclure que le *Modèle de classification lasso* produit une meilleure classification que le modèle basé sur nos connais-

sances. En effet, compte tenu des seuils que nous avons retenu dans la section précédente, entre autre la proportion des taux de recouvrement non nuls, nous recommanderions l'utilisation du modèle basé sur nos connaissances.

Notons qu'un des problèmes du lasso survient lorsqu'il y a des variables explicatives corrélés entre elles. Le lasso ne retient qu'une variable et fixe les paramètres des autres variables corrélés à zéro. Cela entraîne possiblement en une perte d'information qui résulte en un modèle avec une proportion de bonnes classifications plus faible.

Ainsi, alors que nous choisissons nos variables compte tenu de nos connaissances du domaine de l'insolvabilité et de notre étude de la littérature entourant le taux de recouvrement, la sélection de variables du lasso n'est pas en mesure de choisir des variables plus pertinentes à la prédiction de la probabilité d'observer un taux de recouvrement.

Conclusion

7.1 Résumé

Face au nombre croissant de faillites personnelles au Canada, dans un contexte où les prêteurs courent de plus en plus de risques face à leurs créances, nous avons entrepris, dans notre travail, d'estimer la perte en cas de faillite d'individus. Peu d'études ont été menées dans ce domaine particulier et notre recherche ambitionnait de combler ce manque dans la littérature financière portant sur cette composante principale du risque de défaut.

L'objectif de ce mémoire était d'utiliser les variables à notre disposition pour estimer, dans un premier temps, le taux de recouvrement des créanciers non garantis dans les faillites individuelles, puis, dans un deuxième temps, proposer aux créanciers un modèle de prédiction du taux de recouvrement. Notre travail était innovant dans la mesure où, contrairement aux auteurs recensés (Thomas et collab., 2012 et Hoehstoetter et collab., 2012), nous avons corrigé le biais de sélection en appliquant le modèle en deux parties de Heckman (1977). Nous avons également, à partir des documents publics des faillites canadiennes, créé un modèle unique qui permettra aux créanciers d'un individu en faillite de prédire le taux de recouvrement à partir des données uniformes dont il dispose. Ce modèle a été estimé par validation croisée sur notre échantillon.

Les résultats de nos analyses ont permis de valider notre hypothèse de départ, à savoir que les coûts de la faillite, les types de dette et d'actif, le caractère volontaire ou pas de la faillite et les mécanismes à la disposition des créanciers ont un effet significatif sur le taux de recouvrement. Les modèles que nous avons élaborés étayent notre hypothèse et ont des

implications directes pour les créanciers. Notre instrument de prédiction leur permettrait, entre autres, de calculer leur capital réglementaire en fonction de la prédiction du taux de recouvrement conditionnel à la faillite et non pas simplement au défaut.

7.2 Résultats

Nous avons établi les variables explicatives ayant un effet statistiquement significatif sur le taux de recouvrement, ce qui répond à notre principale question de recherche. Le tableau 5.2 présente de manière exhaustive les résultats de nos trois modèles de régression linéaire incorporant ces variables explicatives. En résumé :

- Les coûts de la faillite, qui sont constitués principalement de la rémunération des syndics, est la variable ayant le plus d'impact sur le taux de recouvrement des créanciers, plus particulièrement lorsqu'elle interagit avec la région. Nous avons montré qu'au Québec, l'effet de la rémunération relative du syndic est presque deux fois plus faible que dans le reste du Canada. Par ailleurs, nous avons examiné la possibilité d'un biais d'endogénéité inhérent à cette variable explicative. Bien que nous ne soyons pas convaincus qu'il existe un biais de causalité simultanée entre cette variable et le taux de recouvrement, nous n'avons pas été en mesure de trouver un instrument valide susceptible de remplacer la variable Rémunération du syndic.
- En ce qui concerne les types de dette et d'actif, nous avons montré que lorsque l'individu détient une dette de prêt bancaire ou de prêt de particulier, il présente un taux de recouvrement inférieur. Lorsque nous intégrons les types d'actif à notre modèle, nous montrons que les individus qui déclarent de l'ameublement présentent un taux de recouvrement supérieur. Enfin, lorsqu'un individu possède d'Autres éléments d'actifs, son taux de recouvrement est inférieur.
- En ce qui a trait au genre de dépôt de faillite (volontaire ou pas), nous sommes arrivés à la conclusion qu'il n'existe pas de différence significative dans le recouvrement entre un individu ayant fait l'objet d'une ordonnance de faillite et un individu ayant

fait une cession volontaire. Toutefois, il existe une différence significative entre les cessions présumées et volontaires : nous avons conclu que les doubles administrations aboutissaient à un recouvrement inférieur à celui des faillites volontaires.

- Enfin, pour ce qui touche aux mécanismes à la disposition des créanciers, nous avons montré qu'un individu ayant eu des conditions à sa libération présente un taux de recouvrement significativement supérieur. Ces résultats sont probants puisque les créanciers ont à leur disposition le mécanisme d'opposition à la libération du failli.

Considérant la distribution particulière du taux de recouvrement qui contient un nombre élevé d'observations nulles, nous avons appliqué la méthode de Heckman (1977) afin d'estimer les taux de recouvrement censurés des zéros tout en contrôlant le biais de sélection, ce qui, à notre connaissance, n'a jamais été appliqué au domaine de la perte en cas de défaut. Nos résultats indiquent que notre échantillon ne présentait pas de biais de sélection, c'est-à-dire qu'il n'y avait pas de facteur endogène qui discriminait les individus présentant un taux de recouvrement nul de ceux dont le recouvrement était positif.

Nous avons réussi à obtenir un modèle d'estimation du taux de recouvrement au moyen de variables indépendantes expliquant plus de 35% de la variabilité du taux de recouvrement, ce qui surpasse largement les résultats rapportés dans la littérature, où les R^2 ne dépassaient pas, à notre connaissance, 15% (Thomas et collab., 2012 et Prívará et collab., 2014). Aussi, nous pouvons conclure que le modèle estimant uniquement les taux de recouvrement non nuls expliquait moins bien la variabilité du taux de recouvrement.

Le deuxième volet de notre objectif de recherche visait à proposer aux créanciers un instrument de prédiction du taux de recouvrement lors du dépôt du dossier de faillite. Nous avons développé des modèles de classification de la probabilité d'observer un recouvrement et des modèles de prédiction linéaire du taux de recouvrement par la technique de validation croisée.

En ce qui concerne les modèles de classification, nous avons fait varier les points de coupure afin d'obtenir la proportion de bonne classification et la métrique AUC en fonc-

tion des points de coupure. Nous avons conclu, considérant l'objectif de notre modèle, qu'il était préférable pour le créancier d'utiliser un modèle ayant un point de coupure minimisant les erreurs de type I, soit la proportion des taux de recouvrement non nuls ou l'estimation de la probabilité d'observer un recouvrement non nul pour l'individu moyen. En effet du point de vue des créanciers, il semble particulièrement pertinent de savoir si l'individu présente un potentiel de recouvrement ou pas. Cela peut conduire à une plus grande implication du créancier dans certains dossiers d'intérêt, puisque, par exemple, l'opposition à la libération du failli fait significativement augmenter le taux de recouvrement des créanciers. Nous avons cherché à améliorer ce premier modèle de classification établi sur la base de nos connaissances et de la littérature en proposant un modèle d'apprentissage machine supervisé, le lasso. Ce modèle opère en faisant une sélection automatique de variables et propose donc un modèle différent du nôtre, basé sur les données (data driven). Nous avons comparé la proportion de bonne classification et la métrique AUC en adéquation avec notre choix de point de coupure : le modèle basé sur les connaissances produit un couple de bonne classification et d'AUC supérieur au modèle basé sur les données.

Quant aux modèles de prédiction linéaire du taux de recouvrement, un premier modèle est estimé à partir de tout notre échantillon (incluant les recouvrements nuls); le second est estimé uniquement à partir d'un échantillon censuré des recouvrements nuls. Nous avons comparé les modèles au moyen de la statistique de l'erreur quadratique moyenne (MSE). Nous concluons que nos modèles présentent une erreur de prédiction largement inférieure aux modèles recensés dans la littérature (Thomas et collab., 2012 et Prívará et collab., 2014). Par ailleurs, nos résultats indiquent que la prédiction sur les recouvrements strictement positifs est moins précise que celle sur l'échantillon non censuré.

7.3 Limites et extensions de la recherche

Les données ont constitué la principale limite de notre travail, réduisant le spectre de notre étude. Tout d'abord, nous n'avons pas été en mesure d'obtenir des données de panel.

Il aurait été pertinent de comparer le taux de recouvrement en période de crise et hors crise afin de vérifier si, dans le contexte particulier de la faillite, les taux de recouvrement varient en fonction des cycles économiques, comme le rapporte la littérature de la perte en cas de défaut (Schuermann, 2004).

Également, nous avons été contraints d'utiliser des modèles qui analysent le taux de recouvrement de manière statique et n'avons pas pu étendre notre étude à des modèles qui nécessitent des données de panel. Par exemple, Prívará et collab. (2014) utilisent des modèles de survie qui performant mieux que les modèles de régression linéaire. Notre analyse aurait peut-être gagné en performance si nous avions eu accès à des données longitudinales.

Au cours de notre recherche, nous avons soupçonné un biais de causalité simultanée entre le taux de recouvrement et la rémunération relative du syndic. Nous avons donc cherché à instrumentaliser la variable sur la rémunération du syndic, mais en vain. En effet, comme nous n'avons pas suffisamment de variables, nous n'avons pas trouvé d'instrument pertinent pour corriger ce biais. Des données de panel nous auraient peut-être permis d'obtenir un facteur macro-économique pouvant servir à l'instrumentalisation. Ou encore, l'accès à des variables telles que la taille des firmes de syndics ou le nombre d'années d'expérience nous auraient permis d'atteindre cet objectif.

En ce qui concerne notre hypothèse sur les mécanismes à la disposition des créanciers, nous n'avons accès qu'aux données relatives au processus d'opposition à la libération. Toutefois, nous savons pertinemment qu'il existe d'autres actions que les créanciers peuvent entreprendre durant les faillites, par exemple demander la tenue d'un interrogatoire, d'une assemblée des créanciers, ou faire diverses requêtes au tribunal. Il aurait été intéressant de vérifier si ces actions ont un effet sur le taux de recouvrement.

Notre recherche a porté exclusivement sur le taux de recouvrement en cas de faillite, et plus particulièrement sur les individus en faillite. Ce domaine d'intérêt a fait l'objet de rares études à ce jour, alors que l'endettement des ménages canadiens est en forte croissance, tout comme les procédures de faillite et d'insolvabilité. Dans ce contexte, nous pensons que notre étude est pertinente et que la recherche devrait davantage se pencher

sur cette question. Aussi, les techniques d'apprentissage automatique et d'apprentissage profond étant de plus en plus puissantes – par exemple les réseaux de neurones –, elles pourraient permettre, dans le domaine particulier de la faillite et de l'insolvabilité des débiteurs, une meilleure modélisation du taux de recouvrement.

Nous avons jusqu'ici tenté de convaincre notre lecteur de la portée de l'étude du taux de recouvrement dans les faillites des particuliers, laquelle, selon nous, gagnerait à être explorée davantage.

Au reste, William Shakespeare, en son siècle reculé, ne craignait pas d'incertitude en pratiquant ce qu'il ne pensait être que de la littérature – il s'agissait pourtant de gestion de risque :

« Neither a borrower nor a lender be »
(Ne sois ni emprunteur ni prêteur).

Bibliographie

- Altman, E. I., B. Brady, A. Resti et A. Sironi. 2005, «The link between default and recovery rates : Theory, empirical evidence, and implications», *The Journal of Business*, vol. 78, n° 6, p. 2203–2228.
- Altman, E. I., R. G. Haldeman et P. Narayanan. 1977, «Zetatm analysis a new model to identify bankruptcy risk of corporations», *Journal of Banking & Finance*, vol. 1, n° 1, p. 29–54.
- Amihud, Y., K. Garbade et M. Kahan. 2000, «An institutional innovation to reduce the agency costs of public corporate bonds», *Journal of Applied Corporate Finance*, vol. 13, n° 1, p. 114–121.
- Basel Committee on Banking Supervision. 2006, *International convergence of capital measurement and capital standards : A revised framework - Comprehensive version*, Bank for International Settlements.
- Bastos, J. A. 2010, «Forecasting bank loans loss-given-default», *Journal of Banking & Finance*, vol. 34, n° 10, p. 2510–2517.
- Breiman, L. 2017, *Classification and regression trees*, Routledge.
- Caouette, J. B., E. I. Altman, P. Narayanan et R. Nimmo. 2011, *Managing credit risk : The great challenge for global financial markets*, vol. 401, John Wiley & Sons.
- Carey, M. 1998, «Credit risk in private debt portfolios», *The Journal of Finance*, vol. 53, n° 4, p. 1363–1387.

- Chalupka, R. et J. Kopecsni. 2008, «Modelling bank loan lgd of corporate and sme segments : A case study», cahier de recherche, IES Working Paper.
- CreditMetrics. 1997, «Creditmetrics - technical document», *JP Morgan, New York*.
- Dahen, H. et G. Dionne. 2010, «Scaling models for the severity and frequency of external operational loss data», *Journal of Banking & Finance*, vol. 34, n° 7, p. 1484–1496.
- Deslauriers, J. 2004, *La faillite et l'insolvabilité au Québec*, Montréal : Wilson & Lafleur.
- Dionne, G. 2017, *Gestion des risques : théories et applications*, Economica.
- Dionne, G., M. Artís et M. Guillén. 1996, «Count data models for a credit scoring system», *Journal of Empirical Finance*, vol. 3, n° 3, p. 303–325.
- Dionne, G., G. Gauthier, K. Hammami, M. Maurice et J.-G. Simonato. 2011, «A reduced form model of default spreads with markov-switching macroeconomic factors», *Journal of Banking & Finance*, vol. 35, n° 8, p. 1984–2000.
- Eberhart, A. C., W. T. Moore et R. L. Roenfeldt. 1990, «Security pricing and deviations from the absolute priority rule in bankruptcy proceedings», *The Journal of Finance*, vol. 45, n° 5, p. 1457–1469.
- Engelmann, B. et R. Rauhmeier. 2006, *The Basel II risk parameters : Estimation, validation, and stress testing*, Springer Science & Business Media.
- Eraslan, H., G. Koşar, W. Li et P.-D. Sarte. 2017, «An anatomy of us personal bankruptcy under chapter 13», *International Economic Review*, vol. 58, n° 3, p. 671–702.
- Flynn, E., G. Bermant et K. Bakewell. 2002, «Bankruptcy by the numbers», *American Bankruptcy Institute Journal*, vol. 21, n° 8, p. 20.
- Fridson, M. S., C. M. Garman et K. Okashima. 2000, «Recovery rates : the search for meaning», *Merrill Lynch High Yield, March*.

- Friedman, J., T. Hastie et R. Tibshirani. 2001, *The elements of statistical learning*, vol. 1, Springer series in statistics New York, NY, USA.
- Frye, J. et collab.. 2000, «Depressing recoveries», *Risk-London Risk Magazine Limited*, vol. 13, n° 11, p. 108–111.
- Goodfellow, I., Y. Bengio et A. Courville. 2016, *Deep learning*, vol. 1, MIT press Cambridge.
- Gupton, G. M., D. Gates et L. V. Carty. 2000, «Bank loan loss given default», *Moody's Investors Service, Global Credit Research, November*.
- Heckman, J. J. 1977, «Sample selection bias as a specification error (with an application to the estimation of labor supply functions)», *National Bureau of Economic Research Cambridge, Mass., USA*.
- Helwege, J. 1999, «How long do junk bonds spend in default?», *The Journal of Finance*, vol. 54, n° 1, p. 341–357.
- Hochstoetter, M., A. Nazemi, S. T. Rachev et C. Bozic. 2012, «Recovery rate modelling of non-performing consumer credit using data mining algorithms», cahier de recherche, Citeseer.
- Hu, Y.-T. et W. Perraudin. 2002, «The dependence of recovery rates and defaults», cahier de recherche, working paper, Birkbeck College.
- Khieu, H. D., D. J. Mullineaux et H.-C. Yi. 2012, «The determinants of bank loan recovery rates», *Journal of Banking & Finance*, vol. 36, n° 4, p. 923–933.
- Matuszyk, A., C. Mues et L. C. Thomas. 2010, «Modelling lgd for unsecured personal loans : Decision tree approach», *Journal of the Operational Research Society*, vol. 61, n° 3, p. 393–398.

- Prívará, S., M. Kolman, J. Witzany et collab.. 2014, «Recovery rates in consumer lending : Empirical evidence and the model comparison», *Bulletin of the Czech Econometric Society*, vol. 21.
- Schuermann, T. 2004, «What do we know about loss given default?», *Wharton Financial Institutions Center Working Paper*, vol. 04-01.
- Smith, D. A. et R. Brame. 2003, «Tobit models in social science research : Some limitations and a more general alternative», *Sociological Methods & Research*, vol. 31, n° 3, p. 364–388.
- Thomas, L., A. Matuszyk et A. Moore. 2012, «Comparing debt characteristics and lgd models for different collections policies», *International Journal of Forecasting*, vol. 28, n° 1, p. 196–203.
- Tibshirani, R. 1996, «Regression shrinkage and selection via the lasso», *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 267–288.
- Tobin, J. 1958, «Estimation of relationships for limited dependent variables», *Econometrica : journal of the Econometric Society*, p. 24–36.
- Wirenhammar, A. 2011, *Modeling Downturn LGD for a Retail Portofolio*.
- Zaniboni, N. C., A. C. de Araújo et A. deAvila Montini. 2013, «Factors that influence lgd for retail loans in financial institutions», *Sixth Brazilian Conference on Statistical Modelling in Insurance and Finance*.
- Zhang, J. et L. C. Thomas. 2012, «Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling lgd», *International Journal of Forecasting*, vol. 28, n° 1, p. 204–215.

Annexes

Normes du surintendant

Personnes	N	Revenu mensuel disponible de l'unité familiale																
		1970	2 070	2 170	2 270	2 370	2 570	2 770	2 970	3 170	3 370	3 570	3 770	3 970	4 170	4 470	4 770	5 070
1	1870	100	200	300	400	500	700	900	1 100	1 300	1 500	1 700	1 900	2 100	2 300	2 600	2 900	3 200
2	2328	0	0	0	0	0	242	442	642	842	1 042	1 242	1 442	1 642	1 842	2 142	2 442	2 742
3	2 862	0	0	0	0	0	0	0	108	308	508	708	908	1108	1 308	1 608	1 908	2 208
4	3474	0	0	0	0	0	0	0	0	0	0	0	296	496	696	996	1 296	1 596
5	3941	0	0	0	0	0	0	0	0	0	0	0	0	0	229	529	829	1 129
6	4444	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	326	626
7+	4948	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	122

TABLE A1 – Normes du surintendant - Revenu excédentaire mensuel total, 2009. Les montants indiqués ci-dessus représentent le revenu excédentaire mensuel total du failli qui dépasse les normes à partir desquelles devrait être calculé le paiement du revenu excédentaire.

Les normes du surintendant (N) sont dérivées des seuils de faible revenu (SFR) publiés par Statistique Canada. Le surintendant utilise les SFR avant impôt pour les régions urbaines de 500 000 habitants et plus. Les normes de 2009 sont mises à jour en ajoutant aux SFR de 2007 l'Indice des prix à la consommation (IPC) de 2008 (2,3%), plus un ajustement de 1,2% qui représente l'IPC anticipé pour 2009.

Régressions additionnelles

	Coefficients (écart-type)
Intercept	243062 (11848)***
Sexe (Ref = Femme)	
Homme	-2532 (14337)
R ²	0
Adj. R ²	-0
Num. obs.	1577
RMSE	264948

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE A2 – Régression des Actifs (\$) sur le Sexe

	Coefficients (écart-type)
Intercept	39191 (2738)***
Région (Ref = Autre)	
Alberta	-13408 (3569)***
Ontario	-2257 (3552)
Québec	-5020 (3890)
R ²	0
Adj. R ²	0
Num. obs.	1589
RMSE	49428

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE A3 – Régression des Recettes totales (\$) sur la Région

	Coefficients (écart-type)
Intercept	9597 (587) ^{***}
Région (Ref = Autre)	
Alberta	896 (765)
Ontario	1387 (761) [*]
Québec	-1201 (833)
R ²	0
Adj. R ²	0
Num. obs.	1590
RMSE	10592

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TABLE A4 – Régression de la Rémunération du syndic (\$) sur la Région

Étude des résidus

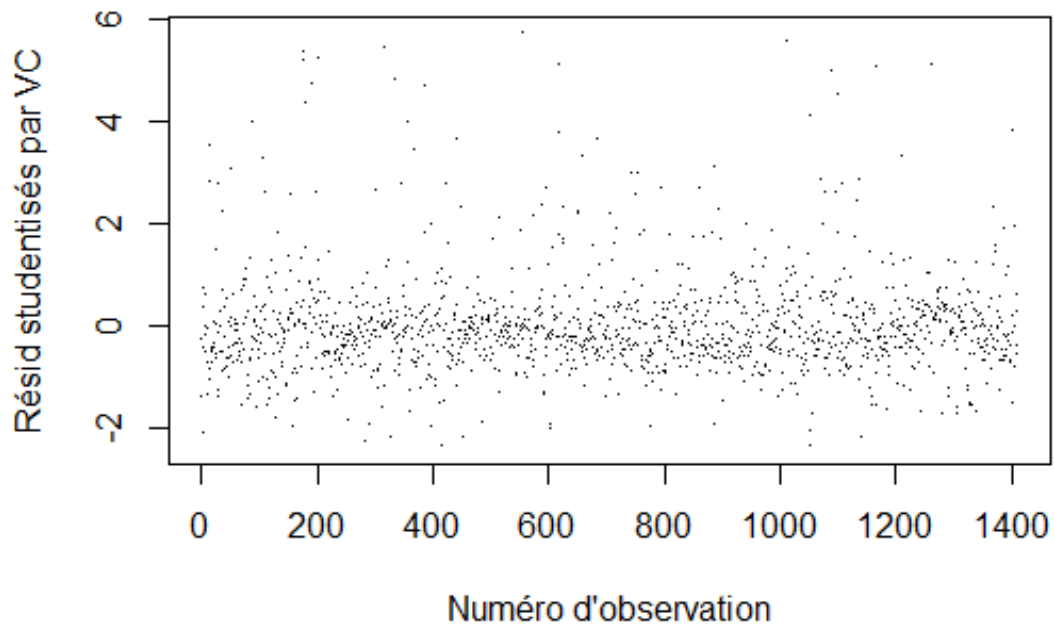


TABLE A5 – Résidus studentisés en fonction du numéro d'observation

D'après la table A5, il semble que la variance soit constante - homogène $\forall i$, c'est-à-dire que $\text{Var}(\varepsilon_i) = \sigma^2$. La forme linéaire du modèle semble donc adéquate. Toutefois, les résidus ne semblent pas être centrés en zéro, c'est-à-dire que $E(\varepsilon_i | x_{1i}, x_{2i}, \dots, x_{pi}) \neq 0$. Il y a peut-être des valeurs extrêmes au sein des y_i .

Sélection de variables lasso

Variables retenues par le lasso
Maison vendue dans les 5 dernières années - Vrai
Revenu disponible du failli
Dette (\$) privilégiée - Prêt de particulier
Valeur (\$) du Type d'actif - Équipement de loisirs
Effets personnels
Valeur (\$) du Type d'actif - Valeurs mobilières
Type d'actif - Espèces en main
Administration antérieure - Proposition
Région - Ontario
Sexe - Homme
Montant (\$) pour le rachat d'actif
Dette (\$) non garantie et privilégiée totale
Passifs moins actifs
Type de dette - Autre
Faillite commerciale
Paiements (\$) mensuels au syndic

TABLE A6 – Variables explicatives retenues par le lasso

