HEC MONTRÉAL

Projet supervisé
06-002-08

Modèle de probabilités de défaut des prêts d'une banque canadienne

Andréanne Simard

11088242

Directeur de projet :

M. Georges Dionne

Lundi, le 20 février 2012

Table des matières

Introduction **3**
Mise en contexte 3
Objectifs *3*
Plan de travail de la recherche *4*

Risque de défaut des banques **4**

Revue de littérature **6**
Revue de littérature – modélisation du risque de défaut des PME 7

Données et statistiques descriptives des variables **9**
Choix des variables 11
Matrice de corrélations *16*

Modèles – Probit et Logit **17**
Modèle probit 18
Modèle logit 18

Résultats **19**
Résultats des régressions (voir simultanément avec annexe D) 19
Signes obtenus et coefficients *22*
Interprétation des coefficients pour le modèle choisi 25
Test de robustesse 25
Tests de stabilité 26

Critiques du modèle **27**

Conclusion **30**
Résumé 30
Limites et recommandations 30

Références **32**

Annexe A **34**

Annexe B **35**

Annexe C **37**
Matrice des corrélations entre les variables 37
Test de multicolinéarité 39

Annexe D **40**

Introduction

Mise en contexte

Dans le cadre d'un stage au sein de l'équipe de crédit d'une banque canadienne, l'idée est venue de reproduire un modèle économétrique d'estimation de probabilités de défaut, similaire à celui utilisé par la Banque tout en utilisant ses données. En tant qu'analyste au sein de l'équipe de rédaction de prêts, je me retrouvais à jongler avec plusieurs informations financières et non financières afin d'évaluer le risque de crédit auquel s'exposait la banque, tout en ne connaissant pas l'apport de chacune de ces variables à l'évaluation de la probabilité de défaut. Je crus donc qu'il était pertinent de tester par moi-même certaines de ces informations puisqu'elles me sont utiles à mon travail d'analyste à la Banque, de même qu'apporter des détails supplémentaires à la Banque concernant ses pratiques d'analyse du risque de défaut.

La banque commerciale où le stage a été effectué est d'origine canadienne et favorise les prêts aux PME ainsi qu'aux entreprises dans le secteur des technologies de l'information. Ma tâche principale au sein de mon équipe à la Banque se résumait ainsi à évaluer le risque des divers prêts acheminés au centre de rédaction dans le but de préparer les dossiers pour le crédit. Pour ce faire, je procédais à différents types d'analyse, soit principalement:

- Analyse financière basée sur divers ratios comptables
- Analyse économique basée sur les industries et secteurs d'activités dans lesquels opèrent les clients, de même que leur position concurrentielle dans leurs marchés respectifs.
- Analyse environnementale, à laquelle est accordée une attention particulière due à la mission et au code éthique de la Banque en question. L'importance de cette analyse provient aussi de son impact sur la valeur attribuée au collatéral offert en garantie sur les prêts.
- Analyse de l'équipe de gestion de la firme, dont :
    - Niveau de scolarité des entrepreneurs et leur expérience appliquée au domaine étudié
    - Scores obtenus grâce à un bureau de crédit (Equifax)

Dans le cadre de ce projet, les différentes informations provenant de cette analyse ont été pour la plupart évaluées, dans l'optique de confirmer leur pertinence dans la prévision du risque de défaut. Ainsi, l'apport de cette recherche ne concerne pas seulement l'estimation d'un nouveau modèle probabiliste, mais réside principalement en l'évaluation du modèle en place ainsi que des variables le constituant.

Objectifs

L'objectif principal de ce projet constitue donc à tester les pratiques de la Banque en termes d'analyse des différents risques des prêts octroyés, en estimant un modèle économétrique d'estimation des probabilités de défaut de ces derniers. Pour ce faire, certaines données rendues disponibles par la Banque seront utilisées afin de déterminer quelles sont les variables

significatives dans la prédiction du défaut. De plus, les méthodes statistiques pour en venir à ces résultats seront étudiées, dont principalement les modèles probit et logit.

Plan de travail de la recherche

Le reste de ce projet sera présenté comme suit. Le risque de défaut (et ainsi de crédit) auquel sont exposées les banques sera défini dans la prochaine section, suivi d'une revue de littérature sur les différents modèles de probabilité de défaut à travers le temps, en élaborant sur certaines recherchent portant sur le défaut des PME. Ensuite, les différentes variables testées seront définies et évaluées. Les modèles économétriques probit et logit seront expliqués dans la section suivante. Les résultats des différentes régressions opérées seront présentés et analysés, suivis par quelques critiques portant sur le modèle. Finalement, une courte conclusion terminera cette recherche. Les annexes peuvent être trouvés à la suite de cette section.

Risque de défaut des banques

Le risque de crédit est devenu une préoccupation importante des banques à la fin des années 1980, et n'a cessé d'être évalué et modélisé depuis. Il présente en fait trois composantes, soit la probabilité de défaut (ou la probabilité que l'emprunteur ne respecte pas ses conditions), le taux de recouvrement au moment du défaut (1- *loss given default)* et l'exposition au risque de crédit au moment du défaut. En fait, le taux de recouvrement peut être estimé en utilisant la séniorité de la dette, où par exemple une dette senior a un ratio de recouvrement plus élevé.

De par cette définition, il est donc clair que le risque de défaut ne correspond pas entièrement au risque de crédit. En fait, selon Longstaff et al. (2011), ce dernier n'y serait composé qu'à 50%. Il est ainsi crucial de faire la distinction entre les deux termes, souvent confondus par plusieurs théoriciens. Le risque de crédit correspond ainsi au risque de perte financière résultant de l'incapacité d'un débiteur à satisfaire ses obligations de dette. Légèrement différent, le risque de défaut est défini comme étant la probabilité qu'une entreprise se retrouve en position de défaut dans un certain horizon temporel[1]. En fait le risque de crédit peut se diviser en quatre catégories, soit le risque de défaut du client, le risque de liquidité, le risque de transaction (et de dégradation du *rating*) et le risque de marché.

Diverses méthodes d'évaluation du risque de crédit sont utilisées au sein des banques, principalement :

1. *Bureaux de crédit* : pour les prêts de plus petite envergure

2. *Cotes d'agences* : pour les firmes publiques

[1] Beaulieu (2003), p.9

Ces dernières regroupent des opinions d'experts sur la nature qualitative et quantitative de leur analyse. Elles sont calculées par plusieurs agences d'importance, dont Moody's, Dun & Bradstreet et Standard and Poor's.

3. *Modèle de Merton (KMV)*

4. *Modèles statistiques* basés sur les états financiers des firmes

Cette méthode est priorisée selon le mandat octroyé au courant de cette recherche.

5. *Credit Metrics*, pour le risque de portefeuille d'une banque

Ce modèle, développé par la banque J.P. Morgan, analyse le risque de crédit d'un portefeuille basé sur la valeur à risque et sur les valeurs au marché (et non pas sur les valeurs comptables historiques). Le modèle de Merton utilise aussi les valeurs au marché des actions de l'entreprise. Pour ces dernières raisons, et puisque les entreprises évaluées dans le contexte actuel sont pour la plupart privées, ces deux méthodes ne sont pas approfondies dans le cours de cette recherche.

Ainsi, la méthode la plus discutée dans la littérature pour évaluer le risque de crédit est appliquée aux obligations et où les cotes d'agences (telles que mentionnées précédemment) sont amplement utilisées. Ces cotes permettent en fait d'évaluer les probabilités de défaut et de migration d'une classe de risque à une autre. Plusieurs recherches ont ainsi démontré que ces cotes ont une corrélation positive avec le taux de recouvrement, et une corrélation négative avec la probabilité de défaut. Cependant, il est à noter que cette méthode suppose que les entreprises au sein d'une même cote sont homogènes, ce qui consiste en une hypothèse très forte en soi.

La gestion du risque de crédit (et par le fait même le risque de défaut) a été marquée par l'introduction en 1988 de l'accord de Bâle, qui impose une réglementation sur le capital réglementaire des banques, ainsi en lien avec leur risque de défaut et faillite (et exclut ainsi le risque économique). Cependant, la Banque étudiée n'y étant pas sujette, les règles maintenues par les accords Bâle ne seront pas définies dans ce contexte.

Dans la gestion du risque de crédit, les banques font face à deux types de risque moral, soit *ex ante* et *ex post*. Le risque moral *ex ante* est relié aux actions (donc prévention) de l'emprunteur avant la réalisation des revenus du projet, et que le prêteur ne peut observer. Ainsi, il est reconnu que les dirigeants peuvent manquer d'incitatifs dans la gestion de leur risque d'entreprise, ce qui les encourage à prendre plus de risque et conserver leur capital bas dans le but d'augmenter leur rendement. À l'opposé, le risque moral *ex post* réfère aux actions (déclaration des vrais revenus) effectuées après la réalisation de ces derniers[2] et subsiste dû au fait que la fonction de paiement au temps de remboursement est fonction d'une variable aléatoire et des vrais revenus non observables. Ainsi, les efforts de l'emprunteur ne peuvent être observés par la banque, ce qui affecte les déclarations de ce premier (enclin à mentir), dû à la présence d'asymétrie d'information. Cette observation est d'autant plus renforcée en ce qui a trait aux firmes lourdement endettées et qui doivent payer une grande proportion de leurs revenus à leur banque en guise de remboursement. Une autre forme du risque moral des entreprises en difficulté

[2] Dionne, G. Notes de cours de *Gestion des risques et assurances*, Chapitre 6.2, p. 59

financière peut aussi se révéler lorsque ces dernières empruntent non pas pour fins d'investissement, mais pour survivre.

Revue de littérature

Deux des pionniers dans l'analyse de la probabilité de défaut à partir d'information spécifique aux états financiers sont Beaver (1967) et Altman (1968). Le second est reconnu pour son modèle basé sur l'information comptable, ou plus particulièrement pour son analyse discriminante multivariée qui a donné naissance au reconnu « Z-*score* ». Il a ainsi soulevé cinq principaux ratios de cinq ordres différents et essentiels à la prédiction de défaut, soit un ratio de liquidité, de profitabilité, d'endettement, de solvabilité ainsi que d'activité :

- Fonds de roulement / actifs totaux
- BNR[3] / actifs totaux
- BAII[4] / actifs totaux
- Valeur de l'équité au marché / passifs totaux
- Ventes nettes / actifs totaux

L'avantage de ce modèle est qu'il ne tient pas compte de la non-linéarité entre les données de défaut et les variables explicatives. Cependant, bien que son apport fût considérable durant ces années, il nous faut garder en mémoire que son modèle balancé n'était basé que sur 33 données de défaut de compagnies publiques seulement (et 33 données de non défaut). Il y faisait ainsi plusieurs hypothèses sur la distribution normale des variables indépendantes, ainsi que sur l'homogénéité de la variance des défauts.

Afin de parer aux hypothèses généralistes d'Altman, Ohlson (1980) fut le premier à proposer le modèle de régression logit, dans lequel il analysa 2163 firmes selon des observations de défaut non balancées (*unbalanced*). Il étudia ainsi quatre principaux facteurs, soit la taille de la firme, la structure du capital, la performance financière ainsi que les liquidités immédiates.

Malgré cet ajustement innovateur au modèle d'Altman, il est impossible d'ignorer l'apport important de Merton (1974) sur l'évaluation de la probabilité de défaut de firmes publiques, de par son hypothèse que l'avoir des actionnaires correspond à une option d'achat sur les actifs de la firme. Ainsi le prix d'exercice est égal à la dette totale (ou partielle) et les prêts à des obligations zéro coupon, déterminés selon la théorie d'efficience des marchés. Cependant, considérant que ce modèle est conçu pour les firmes publiques, aucune attention particulièrement n'y sera accordée dans ce travail.

Le modèle d'analyse discriminante multivariée a par la suite été remis en question par Altman lui-même et Sabato en 2007, selon quoi l'application d'un modèle logistique est préférée à la première approche puisque ce premier prend en considération le problème de multicollinéarité des ratios financiers. Les deux auteurs ont ainsi effectué une régression logit contenant des

[3] Bénéfices non répartis
[4] Bénéfices avant intérêts et impôts

données de 2000 PME américaines sur une période de 1994 à 2002 en y testant dix ratios financiers, desquels cinq sont significatifs selon leur analyse :

- BAIIA[5] / actifs totaux
- Dette à court terme / valeur de l'équité aux livres
- BNR / actifs totaux
- Encaisse / actifs totaux
- BAIIA / dépenses d'intérêts

Le degré de prévisibilité de leur modèle logit est ainsi plus élevé que la cote Z d'Altman, ce qui suggère que ce premier devrait être utilisé aux dépens du modèle initial. En effet, il ne suppose pas de relations linéaires entre les variables ; la variable dépendante, distribuée selon une loi logistique, n'est pas obligatoirement normalement distribuée. Finalement il ne suppose pas l'hypothèse d'homoscédasticité.

Revue de littérature – modélisation du risque de défaut des PME

En ce qui a trait à la modélisation de la probabilité de défaut propre aux PME, plusieurs articles ont fait d'importants avancements durant ces dernières années, et ont ainsi été utiles à cette recherche portant sur le même sujet. La première étude empirique consacrée à la prédiction du défaut des PME fut complétée par Edminster (1972), dans laquelle il appliqua la cote Z d'Altman sur 43 firmes afin de déterminer les ratios financiers ayant le plus grand pouvoir prévisionnel. Ainsi selon l'auteur, les deux principales catégories de ratios significatifs sont celles de liquidité ainsi que d'endettement.

Plusieurs auteurs ont par la suite adopté la même approche que celle d'Edminster, mais cette fois-ci en y incluant des données qualitatives, plus communément appelées « soft info ». Ainsi Lehman (2003) a démontré que le modèle prévisionnel contenant cette *soft info* était plus précis que celui l'excluant, tandis que Behr et al (2004) a obtenu un modèle logit binaire combinant ratios financiers et information qualitative en se basant sur des données provenant de 485 firmes allemandes. Plus précisément, Lopez (2006) a démontré, en analysant la performance de prêts octroyés au Puerto Rico durant 2004 et 2005, que seules les caractéristiques portant sur le propriétaire de la firme (expérience, scolarité, etc.) étaient pertinentes pour faire la distinction entre les bonnes et mauvaises firmes.

Au niveau des PME canadiennes, Beaulieu (2003) a étudié les défauts sur une base de données comprenant 3332 entreprises au niveau commercial, entre 1997 et 2001. Les sociétés publiques ainsi que les grandes et petites entreprises ont été exclues de l'étude, de même que celles en démarrage et provenant des secteurs soit financier, immobilier et public. L'auteur a ainsi obtenu un modèle contenant sept différents ratios financiers (plus que certaines autres études sur le sujet) en y indiquant clairement les signes obtenus :

- Croissance des ventes (+)
- Croissance de l'avoir net (+)

[5] Bénéfices avant intérêts, impôts et amortissement

- ROA (-)
- Couverture des intérêts (-)
- Comptes recevables en jours divisés par les ventes nettes (+)
- Ratio de liquidité immédiate (-)
- Passifs totaux / actifs totaux (+)

Dans cet ordre d'idées, Bonfim (2009) a réalisé une étude sur plus de 100 000 observations recueillies auprès d'environ 30 000 firmes portugaises entre 1996 et 2002, afin d'y examiner les déterminants de défaut reliés au risque de crédit (donc aux caractéristiques idiosyncratiques de la firme) et à certains facteurs macroéconomiques (donc au risque systématique). Les données macroéconomiques incluent entre autre la croissance du PIB, les taux d'intérêt, la production industrielle ainsi que les écarts de taux obligataires L'auteure a ainsi démontré une tendance à entreprendre des activités plus risquées en périodes de croissance économique, même si cela crée des déséquilibres en ralentissement économique. Malgré tout, ses résultats provenant de régressions probit et logit démontrent que ce sont les ratios propres à la firme qui déterminent principalement le défaut. Ainsi, lorsque la taille de la firme est isolée seul le ratio d'investissement, défini comme le nouvel investissement net de la dépréciation, divisé par les ventes totales, est significatif et négatif. En général, l'auteure soulève cinq principaux ratios utiles au modèle de probabilité de défaut, soit la croissance des ventes, le ROA, le ratio de solvabilité, le taux d'investissement ainsi que le ratio de liquidité. De par cette recherche, Bonfim a entre autre démontré l'importance de la prise en compte du type d'entreprise dans l'analyse de crédit.

Plus récemment, Gunawidjaja et Hermanto (2010) ont estimé deux modèles logistique et probit sur 1 845 PME indonésiennes sur une la période de 2005 à 2007 (similaire à celle évaluée lors de cette recherche). Ils ont eux aussi proposé un modèle corrigeant le problème de multicollinéarité soulevé par Altman et Sabato (2007), entre autre afin de déterminer le coussin de capital requis par les banques selon Bâle et afin d'offrir un outil d'analyse du risque de crédit pour les PME. Leur modèle a soulevé deux points importants : 1) les principaux ratios utiles à la prédiction du risque de défaut sont ceux de liquidité courante, de couverture d'intérêt (confirmant les résultats obtenus par Edminster (1972) et Altman et Sabato (2007)), et de la période de collection moyenne des comptes à recevoir, 2) les modèles logit et probit sont consistants. Finalement, ils ont aussi démontré, étonnamment, que l'utilisation d'états financiers non vérifiés n'avait pas d'impact significatif sur l'évaluation du risque de crédit. Cela pourrait être simplement expliqué par le fait que les états financiers sont moins vérifiés en Indonésie.

L'étude de Fidrmuc et Hainz (2010) porte elle aussi sur la modélisation du risque de défaut des PME, en utilisant cette fois-ci des données sur 700 prêts accordés de 2000 à 2005 en Slovaquie (donc durant une période de crise). Ainsi, bien que cet article porte sur une économie en transition, il est intéressant de comparer son contexte à la présente recherche. Malgré la situation économique fragile, il est cependant à noter que les auteurs ont remarqué une croissance des prêts de 6%, considérée comme étant modérée. De plus, ils n'ont fait l'analyse que de quelques ratios, ce qui encore une fois réfère à l'étude actuelle. Un de leur apport important consiste en l'introduction de variables de contrôle telles que l'industrie, la forme légale ainsi qu'une variable temporelle à leur modèle. Ils ont ainsi observé que les prêts sont robustes pour le secteur de l'agriculture et des services seulement, tandis que la dette n'a pas de relation significative avec le défaut dans l'industrie de la construction. En d'autres mots, les effets sont clairement différents entre les secteurs c'est pourquoi Fidrmuc et Hainz soulignent l'importance de considérer ces

variables. En ce qui a trait aux différences de formes légales entre les firmes (entreprises incorporées, donc à responsabilité limitée, coopératives, sociétés par actions et entreprises privées), leurs résultats démontrent que les entreprises privées ont une plus faible probabilité de faire défauts selon les incitatifs et engagements personnels de leurs propriétaires.

Données et statistiques descriptives des variables[6]

La base de données utilisée correspond à un échantillon représentatif des prêts octroyés par la Banque au Canada et sur une période de deux ans, donc pour 2007 et 2008. Il est cependant à noter que considérant la fin d'année financière de la banque en date du 31 mars, les données en date de 2007 proviennent ainsi principalement de 2006 (donc d'avril 2006 à mars 2007), et celles de 2008 de 2007. Suite à ces années, la Banque a instauré de nouvelles pratiques au sein de ses équipes d'analyse, ce qui rend la comparaison avec les données des années ultérieures plus complexe.

À l'intérieur de cette base de données, les prêts faisant défaut sont définis comme étant ceux: (1) transférés aux comptes spéciaux suite à trois non paiements consécutifs (ou arrérages), (2) pour lesquels le remboursement est improbable, (3) inscrits en perte dans les livres de la Banque. Bien que la nature des entreprises ne soit pas précisément déterminée, la majorité des prêts contenus dans la base de données ont été octroyés à des PME. En effet, la présence d'entreprises de plus grande taille est non significative, tel que le démontre une analyse de stabilité effectuée ultérieurement. Les défauts (ou non défauts) sont évalués en septembre 2010, soit généralement plus de deux ans après l'autorisation du prêt. La durée entre l'autorisation et la date d'observation n'est cependant pas disponible, mais se situe dans un intervalle de 30 à 53 mois.

Ainsi, des 6095 observations retenues (pour lesquelles suffisamment de données étaient disponibles), 694 sont reconnues comme étant en défaut, représentant ainsi 11.4% des prêts de l'échantillon. Selon sa mission, la Banque est consciente que certains des prêts octroyés sont plus risqués que ceux d'une banque à charte, tel que le démontre le taux de défaut beaucoup plus faible de Beaulieu (2003) se situant à 1.22%. Voici quelques statistiques concernant le nombre de prêts ayant fait défaut et certaines variables sur les deux années:

[6] Le lecteur peut se référer aux annexes A et B.

Tableau 1 : Statistiques descriptives des données de défaut

	2007	2008	
Défauts	295	399	
	Obs	Défauts	%
But du prêt:			
Fond de roulement	1472	293	19,90%
Immobilisation	3378	278	8,23%
Refinancement & autre	517	42	8,12%
Changement de propriétaire	728	81	11,13%
Industrie:			
Manufacture	1871	294	15,71%
Agriculture et construction	1250	138	11,04%
Ventes au détail	1391	135	9,71%
Service	1366	112	8,20%
Autre	217	15	6,91%

Dans ce tableau, le nombre de firmes correspond au nombre total de prêts octroyés durant la période étudiée, comprenant ceux en défaut. Le nombre de défauts pour chaque année est chiffré dans le haut du tableau. Finalement, le pourcentage de défaut par catégorie de but du prêt et industrie est indiqué à l'extrême droite.

Même si elles sont indiquées dans ce tableau, les variables « but du prêt – refinancement et autre » et « industrie – autre » seront considérées comme les variables de référence dans les régressions ultérieures.

La date minimale des défauts pourrait survenir trois mois suivant la première autorisation, donc en juillet 2006, et la date maximale en septembre 2010. De plus, il est impossible de savoir si une entreprise a contracté plus d'un prêt durant la période d'étude dû à la nature de la base de données. Cependant, l'hypothèse que les prêts sont octroyés à des compagnies différentes est posée afin de simplifier la recherche et la concentrer sur l'analyse de données transversales recueillies sur deux années seulement. Il En fait, seulement un faible pourcentage des prêts a été réalisé en refinancement de prêts existants, où la probabilité d'avoir affaire à une même entreprise à plus d'une reprise est plus élevée. Finalement, les plus anciens états financiers disponibles lors de l'évaluation du risque à l'autorisation sont datés de février 2004 tandis que les plus récents sont en date de février 2008. La Banque exige généralement des états financiers sur trois ans précédant l'octroi d'un prêt, mais cette règle peut être détournée selon certaines exceptions.

Plusieurs ratios ont été calculés vu les variables contenues dans la base de données, tel qu'il le sera démontré en plus amples détails dans une section ultérieure. Voici une liste exhaustive des principales informations accessibles à la recherche :

- Année d'octroi du prêt
- Dates des deux états financiers finaux précédant l'autorisation du prêt

- Dette long terme contractée auprès de la Banque, de même que toute autre dette long terme
- Capitaux propres à l'autorisation
- Actifs et passifs court terme
- Marge bénéficiaire nette à l'autorisation ainsi qu'à l'année précédente
- Ventes totales à l'autorisation ainsi qu'à l'année précédente
- Fonds disponibles à l'autorisation ainsi qu'à l'année précédente
- Parties courantes existante et additionnelle de la dette à court terme à l'autorisation ainsi qu'à l'année précédente
- Intérêts additionnels à l'autorisation ainsi qu'à l'année précédente
- Ajustements à la partie courante de la dette à court terme à l'autorisation ainsi qu'à l'année précédente (dans le cas de compagnies étroitement liées)
- Ratio de couverture après octroi du prêt à l'autorisation et sur une année rétroactive
- Engagement total à l'autorisation
- Industries – classées en quatre grandes classes distinctes

L'engagement total correspond au montant total des prêts octroyés à la firme et aux compagnies y étant étroitement liées, elles aussi clientes de la Banque. Selon ses pratiques internes au crédit, la banque accorde une attention particulière aux fonds disponibles des firmes, utilisés en remplacement au BAIIA habituel et correspondant aux flux de trésorerie d'exploitation des firmes, corrigés des manipulations comptables et ajustés afin d'inclure les flux des compagnies étroitement liées.

Choix des variables

Deux types de facteurs sont déterminants lors de régression statistique : les facteurs stratégiques et non stratégiques. Dans leur article « On the Determinants of the Implied Default Barrier », Dionne et Lajimi (2010) font référence à trois catégories de facteurs stratégiques : aux coûts de liquidation, au pouvoir de négociation relatif et aux frictions de renégociation. Ceux-ci sont en fait endogènes, donc spécifiques au modèle de défaut considéré. À l'opposé, les facteurs non stratégiques étudiés sont les suivants :

- Endettement de la firme
- Ratio de liquidité
- Coût de la dette
- Taille de la firme
- Volatilité
- Conjoncture économique

Dans cette étude, plusieurs variables explicatives, provenant pour la plupart de la littérature abordée dans la section précédente ainsi que de l'expérience pratique, ont été testées afin de déterminer celles qui ont un effet significatif sur la prévisibilité du défaut au sein de la Banque en question. En effet, bien que certains ratios tels qu'un haut levier financier et une faible liquidité soient corrélés de façon significative avec le défaut[7], il n'est pas aussi évident de déterminer

[7] Beaulieu (2003), p. 32

l'effet des autres variables. Somme toute, la sélection des variables explicatives, qu'elles soient financières ou non, demeure une partie cruciale de la modélisation du risque de défaut.

Ainsi, plusieurs variables et ratios financiers ont été analysés et classés en diverses catégories reconnues : ratios de liquidité (ou trésorerie), ratios de profitabilité, ratios de structure financière (ou endettement), ratios de couverture et ratios de croissance. En plus de ces ratios financiers propres à la firme, plusieurs autres variables ont été testées, correspondant au contexte du prêt (année et but du prêt), tenant compte du secteur économique, et en relation avec les pratiques d'analyse de la Banque. Ainsi, selon la disponibilité des données, seul des facteurs non stratégique ont été testés. Pour terminer, les effets de la corrélation ont été déterminés, puisque leur importance est considérable surtout en contexte multivarié. Les variables retenues suivent ainsi le signe prédit sur la probabilité de défaut entre parenthèses.

1. *Ratios de liquidité*

Ces derniers donnent une image de la solvabilité à court terme de la firme

- Ratio de liquidité courante (ou fonds de roulement) (-)
- Équité / Actifs totaux (-)
- Actifs court terme / Dette totale (-)

Les ratios de liquidité, quoiqu'il soient nombreux, témoignent de la capacité de la firme à répondre à ses engagements à court terme et ce à partir de ses actifs à court terme. Ainsi, plus le ratio de liquidité courante, couramment utilisé en analyse financière, est élevé, plus l'entreprise disposera de liquidités pour subvenir à sa dette. Selon le calcul de la Banque, ces actifs et passifs à court terme comprennent tous les postes à ce nom au bilan, et non pas seulement ceux plus liquides (utilisés dans le calcul du ratio de liquidité immédiate). Ainsi, les actifs à court terme incluent les comptes clients, encaisse, inventaires et titres liquides, tandis que les passifs à court terme sont composés des comptes fournisseurs, revenus perçus d'avance, etc. Le ratio de solvabilité, défini comme l'équité divisée par les actifs totaux, suggère que des firmes avec de meilleures conditions financières sont moins enclines à faire défaut. Finalement, le ratio d'actifs à court terme sur la dette totale témoigne de la facilité (ou non) des firmes à rembourser leurs obligations de dette selon leurs contraintes de liquidité.

2. *Ratios de profitabilité*

Ces derniers renseignent sur le rendement des ventes et du capital utilisé.

- Rendement sur actif moyen (ROA) (-)
- Fonds disponibles (-)

Les ratios de rentabilité consistent en de bons indicateurs prévisionnels des défauts ultérieurs de la firme sur ses obligations où une corrélation négative subsiste entre ces derniers et le défaut Ainsi, le ratio de rendement sur actifs témoigne de l'efficacité de la firme à générer des revenus grâce à ses actifs.

Les fonds disponibles, tel que mentionné antérieurement, sont une mesure développée par la Banque afin de parer aux plusieurs manipulations comptables qui affectent le BAIIA (bénéfice avant intérêts, impôts et amortissement). Ces derniers sont calculés sur deux ans, soit pour

l'année d'autorisation du prêt et pour l'année précédente. Ils sont calculés en débutant avec le bénéfice net et en y rajoutant l'amortissement, les impôts reportés et les prêts d'actionnaires (possibilité de les déduire selon certaines conditions), en y enlevant les dividendes et en y ajoutant les fonds disponibles des compagnies étroitement liées ou garantes. Ainsi, comme le bénéfice net ou le BAIIA, il est espéré qu'une hausse des fonds disponibles est négativement corrélée à la probabilité de défaut.

3. *Ratios d'endettement*

Ces derniers expriment la position de la dette dans la firme, ainsi que de l'utilisation qui en est faite.

- Dette / Équité (+)

Il est reconnu qu'un plus grand ratio de D/E, toutes choses étant égales ailleurs, augmente la probabilité de défaut puisqu'il réduit ainsi la capacité de la firme à répondre à un choc quelconque. De plus, et tel que mentionné en lien avec le risque moral *ex post*, les firmes endettées et devant payer d'importantes redevances à leur banque sont plus enclines à déclarer défaut.

4. *Ratios de couverture*

Ces derniers apportent aussi une appréciation de la position de solvabilité à court terme de la firme.

- Ratio de couverture (-)
- Retard (lag) ratio de couverture (-)
- Différence entre le ratio de couverture et son retard (-)

Il porte à croire que le ratio de couverture est un excellent indicateur de la probabilité de défaut puisqu'il combine des notions d'endettement et de rentabilité. Il témoigne ainsi de la « suffisance » de fonds afin de couvrir la nouvelle dette, tout en poursuivant le cours normal des opérations et subvenir aux dettes à court terme. En d'autres mots, il mesure le degré d'autosuffisance de l'entreprise. Ainsi, un ratio supérieur à 1 indique une couverture suffisante des obligations, tandis qu'un ratio inférieur à 1 témoigne d'un manque de fonds pour couvrir la nouvelle dette.

Le ratio de couverture est calculé en partant des fonds disponibles et en y déduisant la partie courante de la dette actuelle, à laquelle est ajoutée la partie courant de la nouvelle dette et les paiements d'intérêts s'y rattachant. Ce calcul est en fait effectué sur deux années (donc en appliquant la nouvelle dette à l'année précédente) afin de dégager la constance de la couverture malgré le nouveau prêt. La variable « retard ratio de couverture » provient ainsi de ce deuxième calcul. De plus, afin d'éviter une possible corrélation entre les deux variables, la différence entre ces dernières a aussi été testée dans une régression à part.

5. *Ratios de croissance*

Ces derniers démontrent la variation de certains postes du bilan et de l'état des résultats à travers le temps, et renseignent sur le positionnement de la firme.

- Ratio de croissance des ventes (-)
- Ratio de croissance de la marge brute (-)
- Ratio de croissance du bénéfice net (-)

Même si la relation semble assez directe et négative, il n'est pas aussi simple de prédire l'effet des taux de croissance de la firme sur sa probabilité de défaut. En effet, dans le cas d'une croissance rapide et soutenue, il est peu probable que les bénéfices non répartis de la firme soient suffisants pour couvrir les nouveaux besoins ainsi créés, ce qui peut ainsi forcer cette dernière a contracter une dette supplémentaire et être exposée au risque qui y est relié. Il porte ainsi à croire qu'il existe une relation en « U » entre les ratios de croissance et le taux de défaut (Beaulieu, 2003). En général, et bien que trois ratios distincts soient testés dans cette recherche, celui de la croissance des ventes est le plus utilisé de par son degré de significativité et puisqu'il représente bien comment l'entreprise évolue dans son marché.

6. *Autres variables – non financières*

Ces dernières sont spécifiques aux méthodes d'analyse de la banque, ou contrôlent pour des effets autres que financiers et qui peuvent influencer la probabilité de défaut.

- Années (?)
- Taille de la firme (-/?)
- Dette de la Banque / dette total (?)
- Engagement total (?)
- But du prêt :
    - Fonds de roulement (?)
    - Immobilisation (?)
    - Changement de propriétaire (?)
    - Refinancement et autre
- Industrie:
    - Manufacture (?)
    - Agriculture et construction (?)
    - Vente au détail (?)
    - Service (?)
    - Autre

Par rapport à la variable « années », selon Altman et Sabato (2007) les PME ont une certaine immunité face à des événements importants, tel que la crise financière qui a débuté en 2007, dû à leur forme relativement simple. Ainsi, il serait possible de n'obtenir aucune différence significative entre les deux années étudiées. Certaines recherches, dont celle de Bonfim (2009), ont démontré une corrélation positive entre l'âge des firmes et leur propension à faire défaut, due à la plus longue période durant laquelle la firme est exposée au risque. Cependant, rien n'indique que les prêts octroyés à en 2007 (ou 2006) l'ont été à des firmes de plus longue date que ceux en 2008.

En ce qui a trait à la taille et tel que l'ont démontré Bun et Redwood (2003), les firmes de plus petite taille ont une plus forte probabilité de défaut. Cependant, cet effet ne semble pas cohérent à travers d'autres articles de recherche, où la probabilité de défaut serait en fait positivement reliée à la taille de la firme (Benito et al, 2004). Il existe effectivement un effet contre-intuitif par

rapport à la taille, où les petites compagnies, malgré qu'elles présentent habituellement des probabilités de défaut plus élevées, ne sont pas aussi endettées que les plus grosses firmes. À l'opposé, les grandes firmes semblent avoir moins de difficulté à surmonter des problèmes liés au crédit de par leur plus grand pouvoir de négociation avec la Banque. De plus, selon Stulz (1996), elles ont moins tendance à faire défaut puisqu'elles font plus de gestion des risques que les plus petites firmes. Au cours de cette recherche, la taille de la firme a été testée en utilisant deux variables *proxy* séparément, soit les ventes et les actifs de la firme.

La variable « dette de la Banque /dette totale » représente une mesure d'exposition de la Banque au risque de la firme. Ainsi, la proportion de dette de la Banque par rapport à la dette totale de la firme peut augmenter la probabilité de cette dernière à faire défaut vu la plus grande exposition de la Banque au risque de la firme. Cependant, le fait que la Banque exerce ainsi un meilleur contrôle sur la firme lorsqu'elle y détient une majorité de prêts, ce qui est ainsi un négatif au défaut, ne permet pas de déterminer la direction de l'impact du ratio sur le défaut *a priori.*

La variable « but du prêt » est caractérisée par quatre classes, soit les prêts de fonds de roulement (donc injection de liquidités, le plus souvent à des fins de croissance et de développement de marché), les prêts pour achat ou améliorations d'immobilisations, les prêts relatifs à un changement de propriétaire et finalement les prêts de refinancement (de prêts auprès de la Banque ou autre) et autre. Il est aussi à noter que la variable « but du prêt » revêt une certaine importance puisqu'elle sous-entend le déficit de garantie par rapport au collatéral sur le prêt. Ainsi, les prêts « fonds de roulement » ont une plus grande propension à ne pas être couvert (donc déficit de garantie à 100%), tandis que les prêts en immobilisations sont plus souvent qu'autrement garantis à plus de 70%. Cependant, les garanties consenties par les propriétaires et actionnaires des firmes peuvent varier selon les buts de prêts, malgré qu'elles soient plus souvent de mise lors des prêts « fonds de roulement ».

Finalement, tel que l'a prouvé Bonfim (2009), certains facteurs macroéconomiques influençant le défaut sont différents à travers différents secteurs. Ainsi, des variables *dummy* on été utilisées afin de mesurer l'impact des différentes industries (regroupées en cinq classes) dans le modèle. Il est à noter que l'industrie « agriculture et construction » inclut aussi les industries minières, fermières et de pêche, ainsi que les industries reliées au bois et au tourisme. De plus, l'industrie « vente au détail » inclut aussi les industries culturelles, des médias ainsi que les pourvoyeurs de locaux, et l'industrie « service » regroupe aussi la vente en gros, le transport routier de même que l'entreposage. Ces regroupements sont basés sur les codes primaires SCIAN (en se basant sur le premier chiffre) et selon une logique de la Banque.

Il est ainsi intéressant de constater que la plupart des variables testées dans cette recherche sont similaires à celles utilisées par Dionne et Laajimi (2010) quoique certaines sont ajoutées ou modifiées selon les critères de la Banque. Il est à noter que la valeur zéro a été attribuée aux ratios où le dénominateur était égal à zéro, dans l'optique de conserver le plus grand nombre de valeurs possibles. Puisque cette opération n'était pas fréquente dans la base de données, nous ne croyons pas que les résultats seront affectés par cette manipulation. De même, en ce qui a trait aux ratios de croissance, il est à noter que celui-ci est aussi égal à zéro pour les entreprises en démarrage (où les données au temps 0 n'étaient pas disponibles).

Ainsi, les PME inclues dans la base de données présentent en moyenne une excellente santé financière, avec des ratios de croissance élevés (en moyenne 85% pour les ventes), et ce sur deux années seulement, un bon rendement sur actif, un ratio de fonds de roulement en moyenne très élevé (plus de 2) et des ratio de couverture de près de 1.5 en moyenne (donc couvrant la nouvelle dette en excédant de 50%). Cependant, il est à noter que les firmes semblent assez endettées avec un ratio D/E de près de 2, quoique la majorité de leur dette ne soit pas contractée à la Banque en question. Finalement, les prêts en immobilisation et ceux dans le secteur manufacturier représentent la plus grande proportion des prêts étudiés.

Matrice de corrélations

Nous avons par la suite calculé les corrélations entre les variables, afin d'identifier celles significativement corrélées avec les fréquences de défaut, de même que pour éviter tout problème de multicolinéarité (le résultat est présenté dans l'annexe C). Ainsi, il est intéressant de constater que les variables « fonds de roulement » et « immobilisation » (toutes deux dans « but du prêt ») sont respectivement positivement et négativement corrélées de façon significative avec le défaut. Ce résultat c'est pas surprenant en ce qui a trait aux prêts en fonds de roulement puisqu'ils sont souvent octroyés sans collatéral important. Il est aussi possible que les prêts en immobilisation soient corrélés au défaut (quoique l'effet est de faible magnitude), bien qu'ils offrent une garantie considérable à la Banque en cas de liquidation. En effet, la nature des opérations des firmes ayant contracté ces prêts peut être risquée et augmenter la probabilité de défaut. De même, les variables sur l'industrie semblent toutes négativement corrélées (à faible pourcentage) entre elles. La variable « actifs à court terme / dette totale » est aussi significativement et positivement corrélée avec la variable « équité / actifs totaux », tandis que les « fonds disponibles » sont significativement et positivement corrélés avec les ventes comme *proxy* pour la taille, ce qui n'est pas surprenant considérant que les fonds disponibles sont une mesure similaire au BAIIA. La variable « dette / équité » est elle aussi corrélée, cette fois-ci négativement et significativement, avec celle de l'équité sur les actifs totaux.

De plus, la variable « engagement total des entreprises étroitement liées » semble être corrélée de façon significative avec plusieurs variables dont la taille (avec les actifs comme *proxy*) et « fonds de roulement » et « immobilisation » comme buts du prêt. Ainsi, les entreprises de plus grande taille sembleraient plus endettées (prêts plus élevés). Au sein de ce groupe (soit « but du prêt »), il est aussi intéressant de constater certaines corrélations négatives entre les variables « immobilisation » et « fonds de roulement » (significative et surprenante intuitivement), de même que « changement de propriétaire » et « immobilisation », et « refinancement et autre » et « immobilisation ». En fait, les prêts en « immobilisation » semblent corrélés avec les autres catégories de prêts, ce qui peut être expliqué par leur plus forte densité dans l'échantillon. Finalement, il n'est pas surprenant de constater une corrélation positive et significative entre les variables de croissance du bénéfice net et de la marge brute, de même qu'entre les variables de ratio de couverture (ainsi que le retard) et la différence entre les deux.

Modèles – Probit et Logit

Dans le cas où la variable dépendante d'intérêt est dichotomique (donc non continue), les modèles non structurels reconnus et favorisés des économistes consistent en les modèles probit et logit. Au cours de cette recherche, le second servira de test de référence (ou *benchmark*) au premier. En fait, les tests effectués démontrent que les deux modèles génèrent des résultats très similaires[8]. Selon leur définition, les modèles étudiés sont qualifiés de non structurels (ou de forme réduite) puisqu'ils se concentrent sur l'analyse statistique de l'impact de certaines variables sur un événement quelconque (ici le défaut)[9].

En fait, dans le cas où la variable dépendante est discrète et non pas continue, il y a de fortes chances que l'application d'un modèle général linéaire génère des estimés biaisés. En régression linéaire, une des hypothèses de base correspond à l'exogénéité des variables explicatives:

$$E(\varepsilon_i \mid x_{j1}, x_{j2}, \ldots, x_{jk}) = 0$$

où les données $x_{ji}, x_{j2}, \ldots, x_{jk}$ sont générées de façon exogène, soit indépendamment de ε_i.
Cependant, cette condition n'est pas respectée en régression logistique (ou probit) dû à la présence d'erreurs binaires, où la variance des résidus est fonction de la moyenne conditionnelle :

$$\text{var}(y_j \mid x_j) = E(y_j \mid x_j) * [1 - E(y_j \mid x_j)]$$

Ainsi, si un modèle linéaire avait été utilisé afin de modéliser la probabilité de défaut au cours de cette recherche, force est de croire qu'il y aurait eu présence d'hétéroscédasticité dans les erreurs de la régression. En fait, cette dernière est même corrigée dans certaines des régressions effectuées. Finalement, il aurait aussi été possible que la prédiction de la variable dépendante ne soit pas contrainte aux deux seuls résultats escomptés, soit entre 0 et 1, c'est-à-dire qu'elle soit plus faible que zéro ou plus élevée que 1.

Il est important de considérer que le modèle probit utilise en fait la distribution normale standard, tandis que le modèle logit est basé sur une distribution logistique. Il est reconnu que dans le cas où la distribution a des queues plus importantes (et épaisses), ce qui reflète ainsi une plus grande concentration des données aux extrêmes, le modèle logit est plus approprié. Dans les deux types de modèles, la probabilité résultante (ici le défaut) est fonction d'une combinaison linéaire de paramètres (β) :

$$P(y = 1 \mid x) = G(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k) = G(\beta_0 + x\beta)$$

où *G* est une fonction prenant strictement les valeurs entre 0 et 1. En fait, les modèles probit et logit sont dérivés d'un modèle avec une variable latente qui introduit la fonction *indicatrice*, qui ne prend que des valeurs de 0 ou 1.

[8] Le lecteur peut se référer à l'annexe D.
[9] Beaulieu (2003), p.17

Ces deux modèles supposent quelques hypothèses, soit que les résidus sont indépendants des variables explicatrices, et qu'ils suivent une distribution normale standard ou logistique standard. Dans les deux cas, nous supposons que les résidus sont symétriquement distribués autour de zéro. En fait, les économistes semblent pencher en faveur de l'hypothèse de normalité des erreurs, ce qui explique pourquoi le modèle probit est habituellement plus populaire que le logit.

Modèle probit

Dans le modèle probit, *G* correspond à la fonction de distribution cumulative normale standard, exprimée par l'intégral :

$$G(z) = \Phi(z) = \int_{-\infty}^{z} \phi(v)dv$$

où $\phi(z)$ est la fonction de densité normale standard exprimée par :

$$(2\pi)^{-1/2} e^{(-z^2/2)}.$$

Dans cette recherche, supposons les fonctions de probabilité suivantes :

$$\mathrm{Prob}(y_k = 1) = G(x_k^{'}\beta)$$
$$\mathrm{Prob}(y_k = 0) = 1 - G(x_k^{'}\beta)$$

où x_k correspond au vecteur des variables explicatives de la firme *k* et β au vecteur des coefficients. Ainsi le modèle linéaire est le suivant :

$$y_k = x_k^{'}\beta + \varepsilon_k.$$

Il est ainsi important que le modèle réponde aux deux conditions suivantes :

$$\lim_{x_k^{'}\beta \to \infty} \mathrm{Prob}(y_k = 1) = 1$$
$$\lim_{x_k^{'}\beta \to -\infty} \mathrm{Prob}(y_k = 1) = 0$$

Ainsi, en utilisant la distribution normale, le modèle probit déduit est le suivant :

$$\mathrm{Prob}(y_k = 1) = \int_{-\infty}^{x_k^{'}\beta} \phi(v)dv$$

Modèle logit

Dans le modèle logit, *G* correspond à la fonction logistique. Il peut ainsi être exprimé comme :

$$\mathrm{Prob}(y_k = 1) = \frac{e^{x_k^{'}\beta}}{1 + e^{x_k^{'}\beta}}$$

Si le modèle probit donne sensiblement un meilleur pourvoir prédictif, alors il porte à croire que les variables indépendantes sont asymptotiquement proches d'une distribution normale dû au fait que l'échantillon est considérablement large[10] (et que la loi des grands nombres s'applique).

L'analyse économétrique a été réalisée avec le logiciel statistique *Stata*, grâce aux commandes *.probit* et *.logit* conçues pour les modèles binaires du même nom. *Stata* utilise ainsi les données individuelles des différents prêts afin d'ajuster les modèles, et calcule par la suite les statistiques de diagnostic par un modèle covariable[11], en conservant la taille de l'échantillon initial. Ainsi, les sujets (ou prêts) ont des résultats différents malgré qu'ils aient les mêmes valeurs covariables et ajustées, de même que les mêmes statistiques de diagnostic.

Afin d'estimer ces modèles binaires non linéaires, la méthode par maximum de vraisemblance est utilisée. Ainsi, le modèle avec la probabilité de défaut $G(x_k^{'}\beta)$ mène à la probabilité jointe (ou fonction de vraisemblance) suivante :

$$L = \prod_{y_k=0}(1 - G(x_k^{'}\beta))\prod_{y_k=1}(G(x_k^{'}\beta))$$

où L dénote la fonction de vraisemblance. Ainsi, la vraisemblance du modèle est le produit des vraisemblances individuelles des différents sujets (ou prêts). Cette fonction a pour but de maximiser le produit en procédant par itérations.

Résultats

Résultats des régressions (voir simultanément avec annexe D)

Dans le cadre de cette recherche basée sur l'évaluation du modèle de la Banque, plusieurs régressions ont été effectuées afin d'isoler les effets des variables explicatives dans un contexte multivarié. Tel que mentionné antérieurement, le modèle probit fut celui testé, tandis que le modèle logit n'a servi qu'à des fins de validation.

À la lumière de ces régressions, nous pensons qu'un modèle comprenant un ensemble de 10 ratios est bien spécifié (comme c'est le cas du modèle RiskCalc Plus™ de Moody's[12]) et tel que l'ont démontré plusieurs auteurs. Dans cette recherche, et selon les tests de robustesse effectués, l'ajout de variables additionnelles n'apporte pas nécessairement de pouvoir explicatif supplémentaire au modèle « épuré » de huit variables. En d'autres termes, le ratio de vraisemblance ne s'améliore pas. Le premier modèle testé (1), en appliquant la même méthodologie de Bonfim (2009), ne contient que les données financières contenues dans la base de données. Les autres régressions probit (2) à (4), représentent les différentes étapes qui ont

[10] Gunawidjaja et Hermanto (2010), p.8
[11] Modèle qui utilise une variable explicative (covariable) pour laquelle la variation est étudiée comme facteur causant la variable principale (ici le défaut).
[12] www.moodysanalytics.com

permis de dériver le modèle final. Finalement, plusieurs tests on été effectués (incluant la modélisation d'un modèle logit en (5)) afin de valider le modèle ainsi obtenu.

RÉGRESSION (1) – 9 variables (robust)

Cette régression ne contient que les variables financières testées dans le modèle, en omettant les variables étant liées à la politique de crédit de la Banque de même qu'à l'industrie et au but du prêt. Il est à noter que les variables « croissance du bénéfice net », « croissance de la marge brute » et « retard ratio de couverture » ont été omises de toutes les régressions puisqu'elles étaient corrélées avec des variables incluses au modèle et étaient largement non significatives[13].

- D/E (-)
- E / A tot (-)
- Fonds de roulement (-)
- A CT / D tot (+)*
- Croissance ventes (+)**
- ROA (-)*
- Fonds disponibles (-)
- Ratio de couverture (-)**

Nous pouvons ainsi observer que des variables provenant de quatre des cinq catégories de ratios financiers mentionnés précédemment sont significatives, dont les ratios de liquidité, de rentabilité, de croissance ainsi que de couverture. Les résultats obtenus sont similaires à ceux de Bonfim (2009). Il est cependant surprenant qu'aucune variable d'endettement ne soit significative dans ce modèle entièrement financier. De plus, il est intéressant de constater que lorsque la régression est opérée sans la commande *robust* (qui corrige pour l'hétéroscédasticité des erreurs selon le test de White), la variable « fonds disponibles » devient significative à 10%. Il porte ainsi à croire qu'il y a effectivement présence d'hétéroscédasticité dans les résidus, c'est pourquoi la fonction *robust* (pour l'estimation robuste à l'hétéroscédasticité de la matrice de variance-covariance, White (1982)) sera utilisée pour toutes les régressions subséquentes.

RÉGRESSION (2) – 22 variables (robust)

En second lieu, le modèle probit comprenant toutes les variables estimées a été testé. La variable de la taille de l'entreprise estimée par les actifs a aussi été testée, mais n'était pas significative. Ainsi, les ventes ont été retenues comme meilleur *proxy* de la taille de la firme.

- Année (-)
- Taille estimée par les ventes (-) (* pour le modèle non *robust*)
- D/E (-)
- D Banque / D tot (-)
- E / A tot (-)*
- Fonds de roulement (-)
- A CT / D tot (+)

[13] *, **, *** significatives aux degrés 10%, 5% et 1% respectivement

- Croissance du bénéfice net (+)
- Croissance de la marge brute (-)
- Croissance ventes (+)**
- ROA (-)
- Fonds disponibles (+) (mais effet marginal (coefficient) très petit)
- Retard ratio de couverture (+)
- Ratio de couverture (-)*
- Engagement total (-)
- But du prêt – fonds de roulement (+)***
- But du prêt – immobilisation (+)
- But du prêt – changement de propriétaire (+)*
- Industrie – manufacture (+)***
- Industrie – agriculture et construction (+)
- Industrie – vente au détail (+)
- Industrie – service (+)

Il est à noter que les variables « but du prêt – refinancement et autre » de même que « industrie – autre » n'ont pas été incluses dans la régression (elles sont en fait dans la constante), puisqu'elles constituent les variables de références et puisqu'elles rassemblaient le plus petit nombre d'observations. Les résultats obtenus pour les catégories de buts de prêts et d'industries doivent ainsi être interprétés par rapport aux variables de référence. Les signes obtenus dans ce modèle sont tous identiques à ceux de la régression précédente, et le demeurent dans la totalité des régressions ultérieures, ce qui est très satisfaisant.

Contrairement au modèle précédent, le ratio financier « équité / actifs totaux » est devenu significatif, maintenant que le ratio « actifs à court terme / dette totale » ne l'est plus. Il est surprenant de remarquer des effets contraires en termes de significativité entre les variables « équité / actifs totaux » et « actifs à court terme / dette totale » puisqu'elles sont positivement corrélées à plus de 50%. Ainsi, il porte à croire selon les résultats que seule une des deux variables n'est significative à la fois, ce qui pourrait s'expliquer par le fait que son effet « contient » une part de l'effet de la variable lui étant corrélée. Afin de tester pour un potentiel problème de colinéarité, la commande Stata *.vif* (basée sur le *Variance Inflation Factor*) a été opérée[14]. Selon le test, si le degré de tolérance (représenté par 1/*VIF*) est inférieur à 10%, la variable concernée est possiblement une combinaison linéaire d'autres variables explicatives contenues dans le modèle. Dans les modèles probit et logit étudiés, il est important de noter que les variables « retard ratio de couverture », « but du prêt – changement de propriétaire » de même que « industrie – autre » ont été omises due à une présence de colinéarité. En ce qui a trait aux autres variables, tous les niveaux de tolérance obtenus sont inférieurs au seuil de 10%, ce qui confirme l'absence de colinéarité au sein de ces variables. Il est aussi intéressant de constater que malgré que la variable « fonds disponibles » ne soit pas significative, le ratio de couverture l'est même si son calcul soit basé sur ceux-ci. Il n'est cependant pas surprenant d'observer que le ratio de couverture est significatif contrairement à la disponibilité des fonds dans la détermination de la probabilité de défaut, puisqu'il couvre deux notions soit celle de liquidité et celle d'endettement en y appliquant l'effet du nouveau prêt.

[14] Le lecteur peut se référer à l'annexe C.

De plus, seules deux variables sont significatives pour la catégorie « but du prêt » et une seule pour « industrie », soit « fonds de roulement », « changement de propriétaire » et « manufacture » respectivement. Cela peut être expliqué en consultant la matrice des corrélations, où il est évident que les différents buts de prêts sont significativement négativement corrélés, notamment entre « immobilisation » et « fonds de roulement » (-62.92%) et « immobilisation » et « changement de propriétaire » (-41.07%). L'effet est aussi le même en ce qui a trait aux industries, qui sont elles aussi toutes négativement corrélées, spécifiquement pour « manufacture » et toutes les autres classes d'industries. Cet effet peut aussi être dû aux regroupements effectués au sein des différents buts de prêt et industries, afin de n'en conserver que quelques classes par catégorie. Ainsi, il est possible que les regroupements s'entrecroisent et que des types de prêts ou de firmes aient été mal classés.

Aussi, le but de prêt « fonds de roulement » est positivement corrélé à 15% avec le défaut, ce qui expliquerait pourquoi la variable est toujours significative à 1% dans tous les modèles étudiés. Les adeptes du ratio D/E seront surpris de constater que celui-ci n'est pas significatif dans cette recherche, ce qui remet ainsi en doute son pouvoir prévisionnel du défaut. En effet, un nouveau courant encourage les analystes financiers à se concentrer sur les flux de trésorerie d'exploitation ajustés pour les manipulations comptables, de même que sur différents indicateurs de productivité, et non plus sur ce ratio d'endettement.

Signes obtenus et coefficients

À moins d'indication contraire, les signes obtenus (inscrits à la droite des variables évaluées) sont les mêmes pour tous les modèles testés.

Ainsi, les ***ratios financiers de liquidité*** (« fonds de roulement » et « équité / actifs totaux ») suivent les signes espérés négatifs ce qui confirme que la liquidité a un effet contraire au risque de défaut. En ce qui a trait aux ***ratios de profitabilité***, la variable « ROA » suit le signe négatif estimé. Il est surprenant de constater que les fonds disponibles démontrent un coefficient positif, ce qui signifierait alors, contre toute logique, qu'une firme ayant des fonds disponibles plus élevés ait une plus grande probabilité de défaut. La variable n'est cependant pas significative et son signe semble « corrigé » dans le modèle logit donc ce résultat est non concluant. Le ***ratio d'endettement*** « dette / équité » présente lui aussi un résultat imprévu de par son coefficient négatif, signifiant ainsi qu'un plus grand ratio d'endettement diminuerait la probabilité de défaut.

En ce qui a trait aux ***ratios de couverture***, les résultats obtenus sont similaires à ceux escomptés, excepté pour la variable « lretard ratio de couverture ». Cependant, puisque cette variable n'est pas significative et est même exclue du modèle significatif (3), son effet est ignoré. Les ***ratios de croissance*** présentent eux aussi des résultats surprenants, où les variables de croissance des ventes et du bénéfice net ont des coefficients positifs et ainsi contraires à ceux estimés. Ce résultat peut être expliqué par l'hypothèse selon laquelle une croissance non soutenue des ventes peut augmenter la probabilité de défaut. En effet, la croissance moyenne des ventes des firmes contenues dans la base de données étant d'environ 85% par année, elle vient appuyer cette hypothèse.

Les ***variables non financières*** ont présenté des coefficients positifs (donc augmentant la probabilité de défaut) pour toutes les classes de but de prêt et d'industrie, tandis que la variable « année » est négative. La taille des firmes, estimée par les ventes, démontre un signe négatif, ce qui confirme que des firmes d'une plus petite taille ont une plus grande probabilité de faire défaut. L'effet est le même en ce qui a trait à l'engagement total des firmes, où le signe négatif démontre que les firmes ayant un engagement plus important auprès de la Banque ont une plus faible probabilité de défaut. Finalement, le ratio de « dette de la Banque / dette totale » démontre un signe lui aussi négatif, ce qui confirme l'hypothèse que le fait que la Banque détient un plus grand contrôle sur les prêts des firmes en question diminue leur probabilité de défaut. En effet, plus ce montant est large, plus la banque semble préoccupée par les pertes possibles.

RÉGRESSION (3) – 17 variables

Tel que le démontre la régression (2), les variables de croissance du bénéfice net et de la marge brute, de même que le retard du ratio de couverture ne sont pas significatives et ont été ainsi retirées du modèle lors de cette étape. Il est en aussi de même pour les variables « but du prêt – immobilisation » et « industrie – service », les moins significatives parmi ces catégories.

- Année (-)
- Taille estimée par les ventes (-)(* pour le modèle non *robust*)
- D/E (-)
- D Banque / D tot (-)
- E / A tot (-)*
- Fonds de roulement (-)
- A CT / D tot (+)
- Croissance ventes (+)**
- ROA (-)
- Fonds disponibles (+)
- Ratio de couverture (-)*
- Engagement total (-)
- But du prêt – fonds de roulement (+)***
- But du prêt – changement de propriétaire (+)**
- Industrie – manufacture (+)***
- Industrie – agriculture et construction (+)***
- Industrie – vente au détail (+)

Les résultats de cette régression sont concluants, soit que les variables significatives dans le modèle précédent le sont demeurées dans ce régression, malgré que certaines variables y aient été supprimées (voir les tests de robustesse pour la démonstration). Il est aussi intéressant de constater que la variable « industrie – agriculture et construction » et maintenant significative à 1%, similaire au résultat obtenu par Fidrmuc et Hainz (2010). Il semble ainsi que cette variable soit corrélée à une autre variable non significative.

RÉGRESSION (4) – 8 variables

Afin de tester la stabilité du modèle précédent, la même régression a été effectuée en ne conservant que les variables significatives. La variable de taille a été incluse au modèle puisqu'elle était significative dans certains modèles non robustes.

- Taille estimée par les ventes (-)*
- E / A tot (-)
- Croissance ventes (+)**
- Ratio de couverture (-)* (presque **)
- But du prêt – fonds de roulement (+)***
- But du prêt – changement de propriétaire (+)**
- Industrie – manufacture (+)***
- Industrie – agriculture et construction (+)***

Toutes les variables, excepté pour « équité / actifs totaux » demeurent significatives dans le dernier modèle estimé, ce qui correspond somme toute à des résultats concluants.

La même régression a été performée en y ajoutant les variables financières « actifs à court terme / dette totale » et « ROA », puisqu'elles étaient significatives dans la régression (1). Cependant, ces deux variables ne sont pas significatives au sein du modèle (4). Les résultats des autres variables estimées de ne changent pas.

Finalement, afin de démontrer la faible incidence résultante du fait d'utiliser un modèle probit ou logit, les variables des modèles (3) et (4) ont été modélisées dans un modèle logit (régression (5))[15]. Cette étape teste aussi la pertinence des modèles étudiés. Dans ces deux régressions, les résultats sont similaires et les variables significatives demeurent les mêmes que pour le modèle probit, voire même que la taille l'est encore plus (à 5%) de même que la croissance des ventes (à 1%) pour le modèle (4). De plus, les signes des coefficients (donc la direction des effets marginaux) demeurent les mêmes que pour la régression probit. Il est donc possible de conclure que le choix de modèle n'a pas d'importance significative, confirmant ainsi ce qui est mentionné dans la littérature.

Ainsi le modèle (3) obtient des résultats intéressants puisqu'il tient compte de l'évolution des ventes, de la structure financière et de la position en liquidités de la firme, tout en contrôlant pour le secteur économique et pour le but du prêt. De par cette recherche, il est force de croire que le but du prêt de même que l'industrie ont un pouvoir prévisionnel dans la probabilité de défaut. Cependant, dû aux multiples corrélations entres les différentes classes de ces deux catégories, il est difficile de cibler quels types de prêt et industries auront une plus grande propension au défaut relatif.

[15] Le lecteur peut se référer à la régression (5) de l'annexe D.

Interprétation des coefficients pour le modèle choisi

L'interprétation des coefficients pour les modèles probit et logit ne se fait pas de façon aussi simple et directe que dans le cas d'une régression linéaire normale. Les coefficients obtenus donnent ainsi les signes des effets partiels de chaque variable explicative sur la probabilité de défaut, tandis que leurs niveaux de significativité statistique sont déterminés selon s'il est possible ou non de rejeter l'hypothèse nulle que le coefficient concerné est égal à 0 à un degré de confiance déterminé. Ainsi, dans le cas d'une variable indépendante *X*, le coefficient représente la variation par unité de l'exponentiel de *X* dans la moyenne de la distribution de probabilité de *Y* et en maintenant toutes les autres variables constantes, ou par rapport à une variable *dummy* de comparaison (donc omise de la régression). Les effets marginaux (représentés par les coefficients) des variables explicatives sur la probabilité de défaut on été estimés par la commande *.mfx compute* de Stata.

En lien avec le modèle (3)[16], il est ainsi possible d'observer qu'une augmentation des liquidités représentées par le ratio « équité / actifs totaux » fait diminuer la probabilité de défaut de 7.4%, tandis que l'effet du ratio de couverture est encore plus faible pour faire diminuer la probabilité de défaut. Il semblerait aussi que la croissance des ventes ait un effet marginal positif sur la probabilité de défaut (tel que discuté antérieurement), d'une magnitude de 0.3% en maintenant les autres variables constantes. En ce qui a trait aux variables de « but du prêt » et « d'industrie », il semblerait que les prêts en fonds de roulement de même que pour changement de propriétaire aient un effet marginal d'augmentation de la probabilité de défaut par rapport aux prêts de refinancement et autre, de magnitudes de 11.4% et 3.2% respectivement. Finalement, les industries de manufacture et d'agriculture et construction ont elles aussi un effet marginal d'augmentation de la probabilité de défaut par rapport à l'industrie « autre », de magnitudes 7.5% et 3.8% respectivement. Ainsi, les prêts pour fonds de roulement, de même que l'industrie manufacturière, semblent les plus risqués.

Aucune attention particulière n'a été portée aux pseudo R-carré de McFadden (1974) obtenus dans les régressions, puisque cette mesure peut varier de par l'ajout de variables à un modèle « emboîté ». Ainsi, malgré que le pseudo R-carré puisse augmenter dû à l'intégration de nouvelles variables, cela ne signifie pas nécessairement que le modèle résultant sera modifié de façon significative.

Test de robustesse

Afin de valider les modèles (3) et (4) obtenus à partir du modèle de base comprenant toutes les variables (2), le test de vraisemblance (ou *log likelihood ratio test*) a été effectué. Le test LR est basé sur le même concept que le test F dans un modèle linéaire. En fait, ce test correspond à la différence entre les fonctions de vraisemblance (ou *log pseudolikelihood* dans Stata) des modèles avec et sans restrictions, où les modèles sans restrictions sont ceux contenant le plus grand nombre de variables. Ainsi, puisque l'estimation par maximum de vraisemblance maximise la fonction *log likelihood*, l'effet d'imposer certaines restrictions au modèle initial (donc d'y enlever

[16] Le lecteur peut se référer à la régression (3) de l'annexe D dans la partie « Interprétation des résultats ».

une ou des variables) peut générer une statistique *log pseudolikelihood* plus faible ou égale. Le but d'un tel test est ainsi de vérifier si l'omission de certaines variables dans le modèle de base (ici (2)) créera une diminution du *log pseudolikelihood* assez importante pour conclure que ces variables omises sont importantes au modèle[17].

Le ratio de vraisemblance correspond ainsi à LR = 2($log\ likelihood_{sans\ restriction} - log\ likelihood_{avec\ restriction}$). La statistique LR ainsi obtenue est comparée à une chi-carré selon le nombre de degrés de liberté (contenus entre parenthèse suivant chi2).

- Résultats obtenus entre les modèles (2) et (3) – 5 variables omises :

LR chi2(5) = 1.62
Prob > chi2 = 0.8989

L'ajout de variables au modèle obtenu (3) n'est ainsi pas significatif, ce qui signifie que le modèle (3) est une meilleure estimation du modèle (2) qui contient un plus grand nombre de variables. En effet, les variables omises dans le modèle (3) étaient toutes non significatives dans le deuxième modèle.

- Résultats obtenus entre les modèles (3) et (4) – 9 variables omises :

LR chi2(9) = 11.09
Prob > chi2 = 0.2695

Il semblerait ainsi que le modèle (4) consiste en une estimation légèrement meilleure du modèle (3).

Tests de stabilité[18]

Afin de s'assurer que le modèle obtenu n'est pas faussé par des données extrêmes souvent erronées, deux nouvelles régressions (probit) ont été effectuées en éliminant les percentiles extrêmes à 1% et 5% (donc les 1er et 99e percentiles, ainsi que les 5e et 95e respectivement) en se basant sur les ventes. Les valeurs extrêmes n'ont pas été exclues sur la base de l'inférence statistique (en termes d'écarts types) puisque la distribution des variables ne suit habituellement pas une loi normale dans un tel modèle non linéaire[19]. Il est à noter que les premières étoiles correspondent aux résultats obtenus dans la régression (3), tandis que les deuxièmes représentent l'élimination des 1er et 99e percentiles, et les troisièmes étoiles l'élimination des 5e et 95e percentiles.

RÉGRESSIONS (6) et (7)

- Année (-) ˉ/ˉ/*
- Taille estimée par les ventes (-)

[17] Woolridge, Chapitre 17, p. 535
[18] Le lecteur peut se référer aux régressions (6) et (7) de l'annexe D.
[19] Beaulieu (2003), p.28

- D/E (-)
- D Banque / D tot (-)
- E / A tot (-) */*/**
- Fonds de roulement (-)
- A CT / D tot (+)
- Croissance ventes (+) **/**/**
- ROA (-) ¯/¯/**
- Fonds disponibles (+), mais (-) pour régression (6)
- Ratio de couverture (-) */*/ ¯
- Engagement total (-)
- But du prêt – fonds de roulement (+) ***/***/***
- But du prêt – changement de propriétaire **/**/**
- Industrie – manufacture (+) ***/***/***
- Industrie – agriculture et construction (+) ***/***/***
- Industrie – vente au détail (+) ¯/¯/**

Il est ainsi possible d'observer que les résultats varient légèrement entre les deux tests. Cela indique ainsi qu'il y a quelques, voire même plusieurs, données de ventes aberrantes, tel que le démontre l'écart-type (voir annexe A) de 1.21e+08. Le test des premiers percentiles extrêmes n'apporte pas de changements significatifs par rapport au modèle testé précédemment, ce qui est satisfaisant en terme de stabilité. La seule différence survient au niveau de la direction (ou signe) de la variable « fonds disponibles », qui devient négative dans la régression (6). Cependant puisque cette variable n'est pas significative, cet effet est ignoré.

Cependant, la situation est différente en ce qui a trait au test des cinquièmes percentiles extrêmes. En effet, le nouvel écart-type des ventes obtenu de par ce test correspond à 1.47e+07, ce qui demeure très large mais tout de même mieux que dans la régression initiale. Dans ce cas-ci, la variable « année » est maintenant significative à 10%, portant ainsi à croire que les prêts octroyés durant la deuxième année (2008) ont une plus faible probabilité de défaut que ceux en 2007. L'hypothèse de durée serait ainsi confirmée, où un prêt plus « âgé » a de plus fortes chances de faire défaut. Cependant, il serait aussi possible que cette observation soit le résultat de facteurs de la conjoncture économique, n'étant pas nécessairement en lien avec les firmes spécifiquement. De plus, le ratio de couverture n'est plus significatif, étant remplacé par le ratio de profitabilité « ROA » qui lui l'est à 5%, démontrant ainsi que la variable de profitabilité aurait un meilleur pouvoir prévisionnel que celle de couverture de la dette totale par les fonds disponibles. Finalement, toutes les classes d'industrie sont significatives par rapport à « industrie – autre », ce qui indique que les secteurs économiques dans lesquels évoluent les firmes ont une influence sur la probabilité de défaut. Les tests effectués n'offrent aucune explication par rapport aux cycles de ces secteurs, cependant.

Critiques du modèle

Malgré le niveau de significativité satisfaisant obtenu dans le modèle (3), plusieurs critiques peuvent être adressées à cette recherche.

1. *Nombre limité de données disponibles*

La plus importante limite de ce modèle correspond aux dates d'autorisation des prêts, qui n'étaient pas disponibles. En effet, l'utilisation d'un modèle de *durée* tel que proposé par Shumway (2001) aurait certainement été d'un apport considérable dans les tests effectués sur le pouvoir prévisionnel des modèles probit ou logit étudiés[20]. L'auteur a critiqué la nature statique des données, utilisées sur un horizon d'un an seulement. La recherche de Dionne et al. (1996), qui utilise des données provenant d'une banque espagnole, démontre aussi l'importance de la durée dans l'évaluation du risque de défaut. Il faut cependant tenir compte que les différents types de prêts à la Banque ont différentes maturités, soit un maximum d'environ 30 et 5 ans pour les prêts en immobilisation et en fonds de roulement respectivement. Les prêts en immobilisation étant généralement de plus grande envergure, leur maturité est ainsi souvent allongée.

Il aurait été aussi intéressant d'avoir le montant des prêts (et non pas seulement l'engagement total considérant que le défaut résulte principalement non pas de la probabilité de défaut, mais de la dette à risque, c'est-à-dire la probabilité de défaut multipliée par le montant du prêt), de même que des données sur plus de deux périodes (données panel) pour ainsi évaluer leur degré de significativité.

En ce qui a trait aux données financières, il aurait été approprié de tester l'importance du niveau d'encaisse sur la probabilité de défaut, de même que les délais de recouvrement des comptes clients et la rotation des inventaires. Au niveau des variables non financières et ainsi plus subjectives (ou comprenant de la *soft information* tel que l'ont démontré Lehmann (2003) et Behr et al. (2004)), les types d'états financiers utilisés, le nombre d'employés au sein des firmes, la forme légale de ces dernières et leur cycles de vie (démarrage, expansion ou maturité), de même que certains critères sur la qualité du management tels que l'expérience et la scolarité ou autres critères significatifs selon Lopez (2006), et sur la position concurrentielles des firmes sont toutes des données analysées par la Banque mais qui n'ont pas été disponibles dans le cours de cette recherche. En effet, les entreprises en démarrage sont reconnues comme étant plus risquées, surtout durant leurs deux premières années d'existence tel que l'ont confirmé plusieurs spécialistes. Il aurait donc été pertinent de tester toutes ces variables afin de valider leur pouvoir explicatif du défaut et leur présence dans les modèles de probabilité de défaut de la Banque.

Le défaut des garanties en termes du collatéral garantissant le nouveau prêt est une autre variable qu'il aurait été intéressant d'inclure au modèle, surtout que la Banque s'en sert pour calculer le taux de perte prévue. Cette variable est cependant plus reliée au risque de recouvrement qu'au risque de défaut.

Finalement, il aurait été intéressant de tester certains des ratios de liquidité, profitabilité et endettement d'Altman, tous divisés par les actifs totaux de la firme afin de tenir compte de la taille. Cependant, il a été préféré, au cours de cette recherche, d'ajouter directement une variable « taille ».

[20] Dans le cadre de cette recherche, la durée correspond à l'écart de temps entre la date d'autorisation et la date d'observation (en septembre 2010).

2. Biais de sélection

Il est possible de supposer qu'il y a présence d'un biais de sélection au sein des données recueillies par la Banque puisque les caractéristiques des entreprises pour lesquelles les projets furent déclinés ne sont pas enregistrées dans les systèmes. Ainsi, il est impossible de déterminer si les bonnes décisions ont été prises lors de l'évaluation du risque de défaut et en ce qui a trait au risque des firmes. Ce biais est en temps normal inévitable, considérant l'asymétrie d'information entre les emprunteurs et la banque. Ainsi, la base de données pourrait être biaisée par la présence de « bonnes » compagnies seulement. Il aurait été possible de tester ce biais de par l'estimation en deux étapes d'Heckman (1979).

Afin de parer à ce problème d'antisélection, où le prêteur ne connaît pas le risque individuel de l'emprunteur, les banques ont diverses options, soit l'autosélection des risques (où les emprunteurs acceptant un collatéral plus élevé démontrent un « bon » niveau de risque ou le *scoring* bancaire (ou modèle théorique), tel qu'il l'est effectué au sein de la Banque en question[21].

3. Nature des PME

Les règles imputées au PME en ce qui a trait à la qualité de leurs états financiers sont beaucoup moins strictes, voir même qu'elles permettent des incohérences au niveau des pratiques comptables (Carmines et Zeller, 1979). Ainsi, les états financiers non audités (mission d'examen) sont permis lors de l'analyse du crédit au sein de la banque, ce qui laisse place à un manque de fiabilité des données financières. De plus, la nature des PME (étant de plus petite taille selon leur définition), influence directement leur risque de crédit, qui peut être amplifié par le manque d'expérience des dirigeants de même que par le collatéral souvent insuffisant de cette catégorie d'entreprises (Bartels, 2002).

Finalement, il aurait aussi été intéressant de contrôler plus spécifiquement pour le type d'entreprises, tel que l'a démontré Bonfim (2009). En effet, l'utilisation d'un seul modèle pour couvrir diverses catégories d'entreprises et plusieurs secteurs peut s'avérer inadaptée selon leurs caractéristiques parfois considérablement différentes. Par exemple, la réussite des petites entreprises est souvent liée au crédit de leurs propriétaires (évalué de par la valeur nette personnelle), tandis que les plus grandes entreprises sont traitées de façon différente et selon une analyse financière plus assidue. Aussi, l'importance des états financiers pour les entreprises dans le secteur immobilier peut être remise en question, considérant que la réussite (ou l'échec) de ces dernières est souvent liée à un événement particulier[22]. Ainsi, l'analyse de données financières telles qu'elles sont utilisées dans le modèle étudié ne semble pas aussi pertinente pour ce type de prêts.

4. Regroupements au sein des industries et des buts du prêt

Tel que mentionné précédemment, le fait de regrouper certaines industries et buts de prêt en certaines classes peut avoir apporté certains biais statistiques. De plus, dans le cas où un prêt est réalisé dans le but de couvrir différents objectifs (et buts), le but de prêt représentant le plus grand

[21] Dionne, G. Notes de cours *Gestion des risques et assurances*, Chapitre 8.1, p.87
[22] Beaulieu (2003), p.27

montant relatif au prêt total sera inscrit aux livres. Cependant, dans le cas où une minorité du montant correspond à un but jugé plus risqué, il se peut que la partie minoritaire l'emporte et fasse défaut malgré qu'elle ne fût pas inscrite au modèle d'analyse de la Banque.

Conclusion

Résumé

Le modèle développé au cours de cette recherche relie son résultat à une probabilité de défaut (et non une cote de risque telle qu'évaluée par la banque et par différentes agences de notation), ce qui permet de discriminer les firmes entre elles. Tel que mentionné dans l'introduction, l'objectif de ce travail ne consistait pas à développer un nouveau modèle de probabilité de défaut, mais plutôt de faire ressortir et ainsi valider les variables financières et non financières ayant un pouvoir explicatif du défaut. Il prend ainsi la forme d'un complément au jugement d'analystes financiers et de rédacteurs au crédit. Les signes (ou directions) des variables significatives sont demeurés constants à travers toutes les régressions effectuées, ce qui est concluant aux fins d'analyse.

Cette recherche propose une évaluation du risque de défaut des prêts d'une banque canadienne, en se basant sur certaines des données, pour la plupart financière, dont se sert cette dernière. Des données en lien avec le but du prêt de même que l'industrie ont été ajoutées au modèle, et se sont révélées significatives dans la prévision du défaut. Les autres variables déterminantes et significatives aux fins de cette recherche correspondent à la taille de la firme, estimée par les ventes, la croissance des ventes, le ratio de couverture, ainsi que l'équité / actifs totaux (dans le modèle général). Il a en fait été intéressant de remettre en question les ratios financiers traditionnellement utilisés, dont le ratio « dette / équité » qui n'est pas significatif dans cette étude. Il est aussi surprenant de constater que la variable « fonds disponibles », une adaptation au BAIIA, n'est pas significative. Elle est néanmoins reflétée dans le ratio de couverture.

Le modèle probit a été utilisé (et estimé par maximum de vraisemblance) afin d'estimer les variables explicatives ayant un pouvoir de détermination du défaut, validé par l'estimation d'un modèle logit. Ainsi, nous en concluons que l'un ou l'autre de ces modèles aurait pu être utilisé au cours de cette recherche puisqu'ils génèrent des résultats similaires.

Limites et recommandations

Il est à noter que l'évaluation des prêts s'est faite sur une période marquée par une crise économique de même qu'une récession, qui se faisaient encore ressentir en septembre 2010, ce qui a pu causer un nombre de défauts inhabituellement élevé au sein de la Banque. Ainsi, il semble que le modèle (3) étudié de par cette recherche semble s'appliquer exclusivement à une situation de crise financière, puisque les données analysées en ont nécessairement été affectées. En fait, il a été démontré que la majorité des crises financières est précédée par des périodes de croissance, et de santé et libéralisation financière (Kaminsky et Reinhard, 1999), ce qui explique

le haut taux de croissance et la bonne santé financière en moyenne des firmes évaluées en 2007 et 2008 (donc en 2006 et 2007 principalement).

Finalement, il aurait aussi été intéressant d'inclure la dimension temporelle au modèle économétrique, tel que démontré par Bonfim (2009), afin de non seulement tenter de prévoir le défaut, mais aussi le moment auquel celui-ci serait le plus probable. De plus, il aurait été pertinent d'avoir des informations sur l'état de la marge de crédit (s'il y a) des firmes ayant contracté des prêts au sein de la base de données puisqu'un signal de risque de défaut est lancé à la banque lorsque celle-ci est utilisée à plus de 50%. À l'inverse, l'utilisation d'une marge de crédit peut se révéler être bénéfique pour une firme en croissance et ayant besoin de liquidités courantes. La cote de crédit que calcule la Banque peut être référée à une cote de délinquance, qui ne va habituellement pas chercher le potentiel de remboursement des firmes. Ce qui importe réellement est ainsi la capacité de paiement, ce que le calcul des fonds disponibles vise à évaluer. À titre de recommandation, et puisque la variable « fonds disponibles » n'est pas significative dans cette recherche, la capacité de paiement peut être calculée autrement, soit en déduisant les dépenses courantes mensuelles estimées d'une firme de son revenu mensuel, et en y rajoutant la valeur de ses actifs tangibles divisés par le terme du prêt.

Malgré tout, nous sommes conscients qu'il n'est pas toujours évident d'obtenir des données pertinentes ainsi que des échantillons de taille suffisante afin d'isoler des effets statistiquement significatifs[23].

La présente recherche ayant été un mémoire, il aurait alors été important de procéder à des tests de prédiction « in-sample » et « out-of-sample » en divisant la base de données en deux groupe, et ce dans l'optique de valider les résultats obtenus. Par exemple, 70% des observations auraient pu être modélisées (« in-sample ») afin de prédire les 30% restantes (« out-of-sample »). D'autres tests de validation auraient aussi pu être effectués en faisant varier ces pourcentages. De plus, ces mêmes tests auraient pu être réalisés en utilisant les données de 2007 comme données « in-sample » afin de prédire celles de 2008 (« out-of-sample »). Cependant, puisque le présent travail prend principalement la forme d'un rapport et non pas d'un mémoire, cette extension au modèle n'a pas été testée.

[23] Dionne, G. Note de cours *Gestion des risques et assurances*, Chapitre 4.2, p.37

Références

Notes de cours de M. Georges Dionne pour le cours *Gestion des risque et assurances*, HEC Montréal, hiver 2011.

Altman, E.I., 1968, « Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy », *Journal of Finance*, Vol. 23, No. 4, p. 589-609.

Altman, E.I., Sabato, G., 2007, « Modelling Credit Risk for SMEs : Evidence from the US Market », *Abacus*, Vol. 43, No. 3, p. 332-357.

Bartels, J.C., 2002, « Basel II and the Survival of the SME : Are Lenders and Borrowers Ready to Comply with Basel II ? », *Business Credit*, 1-2.

Beaulieu, S., 2003, « Modèle d'évaluation des probabilités de défaut un an des firmes privées canadiennes : approche probit non-structurelle », *Mémoire en Sciences de la gestion*, 68 p.

Beaver, W.H., 1967, « Financial Ratios as Predictors of Failure », *Journal of Accounting Research*, Vol. 4, p. 71-111.

Behr, P., Gûttler, A., Plattner, D., 2004, « Credit Scoring and Relationship Lending: The Case of German SME », *JEL*, Ver. 16, p. 1-31.

Benito et al., 2004, « A Synthethic Indicator of Financial Pressure for Spanish Firms », *Banco de España Working Paper*, No. 411.

Bonfim, D., 2009, « Credit Risk Drivers : Evaluating the Contribution of Firm Level Information and Macroeconomic Dynamics », *Journal of Banking and Finance*, Vol. 33, No. 2.

Bunn, P., Redwood, V., 2003, « Company Accounts Based Modelling of Business Failures and the Implications for Financial Stability », *Bank of England Working Paper*, No. 210.

Carmines, E.G., Zeller, R.A., 1979, « Reliability and Validity Assessment », *Sage Publications*, Californie.

Dionne, G., Artis, M., Guillen, M., 1996, « Count Data Models for a Credit Scoring System », *Journal of Empirical Finance*, Vol. 3, p. 303-325.

Dionne, G., Laajimi, S., 2010, « On the Determinants of the Implied Default Barrier », *Canada Research Chair in Risk Management, HEC Montréal*, 49 p.

Dionne, G., Laajimi, S., Mejri, S., Petrescu, M., 2008, « Estimation of the Default Risk of Publicly Traded Companies : Evidence from Canadian Data », *Canadian Journal of Administrative Sciences / Revue Canadienne des Sciences de l'Administration*, Vol. 25, No. 2, p. 134-152.

Edmister, R., 1972, « An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction », *Journal of Financial and Quantitative Analysis*, Vol. 7, No. 2, p. 1477-1493.

Fidrmuc, J., Hainz, C., 2010, « Default Rates in the Loan Market for SMEs : Evidence from Slovakia », *Economic Systems*, Vol. 24, No. 2.

Giesecke, K., Longstaff, F., Schaefer, S., Strebulaev, I., 2011, « Corporate Bond Default Risk : A 150-Year Perspective », *Journal of Financial Economics*, Vol. 102, No. 2, p. 233-250.

Gunawidjaja, S., Hermanto, B., 2010, « Default Prediction Model for SME's : Evidence from Indonesian Market Using Financial Ratios », p. 1-20.

Heckman, J.J., 1979, « Sample Selection Bias as a Specification Error », *Econometrica*, Vol. 47, No. 1, p. 53-161.

Kaminskiy, G.L., Reinhart, C.M., 1999, « The Twin Crises : The Causes of Banking and Balance-of-Payments Problems », *American Economic Review*, Vol. 89, No. 3, p. 473-500.

Lehman, B., 2003, « Is it Worth the While ? The Relevance of Qualitative Information in Credit Rating », *EFMA 2003 Helsinki Meetings*, Helsinki, Finlande, p.1-25

Lopez, A.S., 2006, « Improving Access to Credit of SME's in Puerto Rico : Exploring Variables to Forecast Small Business Loan Events », *Doctoral Program in Entrepreneurship and Small Business Management.* Autonomous University of Barcelona, Department of Business Economics.

Merton, R., 1974, « On the pricing of corporate debt: the risk structure of interest rates », *The Journal of Finance*, Vol. 29, No. 2, p. 449-470.

Ohlson, J.S., 1980, « Financial Ratios and the Probabilistic Prediction of Bankruptcy », *Journal of Accounting Research*, Vol. 18, No. 1, p. 109-131.

Shumway, T., 2001, « Forecasting Bankruptcy More Accurately : A Simple Hazard Model », *The Journal of Business*, Vol. 74, No. 1, p. 101-124.

Stulz, R.M., 1996, « Rethinking Risk Management », *Journal of Applied Corporate Finance*, Vol. 9, No. 3, p. 8-24.

White, H., 1982, « Maximum Likelihood Estimation of Misspecified Models », *Econometrica*, Vol. 50, p.1-25.

Woolridge, J.M., 2002, « Introductory Econometrics : A Modern Approach », *South-Western Colelge Pub*, 2nd Revised édition, 896 p.

Annexe A

Statistiques descriptives des variables testées

Variables	Obs	Mean	Std. Dev.	Variance	Skewness	Kurtosis	Min	Max
Défaut	6095	0.11	0.32	0.10	2.43	6.91	0	1
Année	6095	0.59	0.49	0.24	-0.36	1.13	0	1
Taille - ventes	6095	2.43e+07	1.21e+08	1.46e+16	19.86	578.05	0	4.77e+09
D/E	6095	1.89	2.82	7.93	30.01	1371.55	0	150.04
D Banque / D tot	6095	0.81	0.24	0.06	-1.29	3.74	0	1
E / A tot	6095	0.41	0.16	0.03	0.58	3.02	0	1
F/R	6095	2.68	45.77	2094.82	75.25	5787.39	0	3529.41
A CT / D tot	6095	1.34	2.11	4.46	6.31	77.83	0	43.82
Croissance BN	6095	3.96	56.95	3243.02	39.39	2009.73	-529.39	3291.78
Croissance MB	6095	3.53	55.12	3037.90	30.71	1120.35	-404.15	2358.45
Croissance ventes	6095	0.85	39.62	1569.46	77	5978.63	-1	3078.56
ROA	6095	0.40	1.35	1.82	9.34	140.80	-6.77	31.43
Fonds disponibles	6095	1344818	5666998	3.21e+13	15.79	367.82	-1.36e+07	1.85e+08
Lag ratio couverture	6095	1.10	3.83	14.71	-4.44	375.32	-100	100
Ratio couverture	6095	1.46	2.52	6.33	8.79	575.47	-73.33	100
Engagement total	6095	1144234	1371934	1.88e+12	2.67	11.89	15000	1.00e+07
But du prêt:								
Fond de roulement	6095	0.24	0.43	0.18	1.21	2.46	0	1
Immobilisation	6095	0.55	0.50	0.25	-0.22	1.05	0	1
Refinancement & autre	6095	0.08	0.28	0.08	2.98	9.88	0	1
Changement de propriétaire	6095	0.12	0.32	0.11	2.35	6.51	0	1
Industrie:								
Manufacture	6095	0.31	0.46	0.21	0.84	1.70	0	1
Agriculture et construction	6095	0.21	0.40	0.16	1.46	3.13	0	1
Ventes au détail	6095	0.23	0.42	0.18	1.30	2.68	0	1
Service	6095	0.22	0.42	0.17	1.32	2.75	0	1
Autre	6095	0.04	0.19	0.03	5.01	26.12	0	1
Taille - actifs	6095	2988825	9526862	9.08e+13	32.95	1623.14	116760	5.28e+08
Différence ratio couverture	6095	0.36	4.06	16.52	4.55	296.58	-95.82	101.25

Annexe B

Calculs des ratios financiers provenant de la base de données

1. *Ratios de liquidité*

- Ratio de liquidité courante (ou fonds de roulement)
  Actifs à court terme / passifs à court terme

- Équité / Actifs totaux
  Équité totale / actifs totaux

- Actifs court terme / Dette totale
  Actifs à court terme / dette totale

2. *Ratios de profitabilité*

- Rendement sur actif moyen (ROA)
  Bénéfice net / actifs totaux

- Fonds disponibles
  Bénéfice net + (amortissement +/- prêts des actionnaires + impôts reportés – dividendes +/- fonds disponibles des cautions et compagnies garantes)

3. *Ratios d'endettement*

- Dette / Équité
  Dette totale / Équité totale

- Dette de la Banque / Dette totale
  Dette contractée au sein de la Banque en question / Dette totale

4. *Ratios de couverture*

- Ratio de couverture
  Fonds disponibles / (PCDT existance + PDCT additionnelle + intérêts additionnels) au moment de l'autorisation

- Retard (lag) ratio de couverture
  Fonds disponibles / (PCDT existance + PDCT additionnelle + intérêts additionnels) à l'année précédente

- Différence entre le ratio de couverture et son retard (lag)
  Ratio de couverture – retard ratio de couverture

5. *Ratios de croissance*

- Ratio de croissance des ventes
  (Ventes nettes (t) – ventes nettes (t-1)) / ventes nettes (t-1)

- Ratio de croissance de la marge brute
  (Marge brute (t) – marge brute (t-1)) / abs(marge brute (t-1))

  où la marge brute correspond à *ventes – coûts de la marchandise vendue*.

- Ratio de croissance du bénéfice net
  (bénéfice net (t) – bénéfice net (t-1)) / abs(bénéfice net (t-1))

  où le bénéfice net est calculé après amortissement, intérêts et impôts.

Annexe C

Matrice des corrélations entre les variables[24]

	dfaut	anne	taill~es	de	dbanqu~t	eatot	fr
dfaut	1.0000						
anne	-0.0095	1.0000					
tailleave~es	-0.0258*	-0.0461*	1.0000				
de	-0.0139	0.0179	-0.0144	1.0000			
dbanquedtot	-0.0318*	0.0255*	-0.1254*	-0.0293*	1.0000		
eatot	0.0075	-0.0433*	0.0525*	-0.4527*	0.0794*	1.0000	
fr	-0.0067	0.0127	-0.0046	-0.0054	0.0193	0.0049	1.0000
actdtot	0.0186	-0.0220*	0.1218*	-0.1651*	0.0270*	0.5150*	-0.0090
croissancebn	-0.0087	-0.0040	-0.0028	0.0007	0.0045	-0.0070	-0.0011
croissancemb	-0.0122	0.0013	-0.0035	-0.0039	0.0067	0.0066	0.0003
croissance~s	0.0401*	0.0110	-0.0011	0.0008	-0.0101	-0.0083	-0.0007
roa	-0.0311*	-0.0865*	0.1900*	-0.0156	-0.0077	0.0256*	-0.0016
fondsdispo~s	-0.0259*	-0.0543*	0.6693*	-0.0217*	-0.1464*	0.0406*	-0.0032
lagratioco~e	-0.0092	-0.0103	0.0267*	-0.0257*	-0.0021	0.0832*	-0.0022
ratiocouve~e	-0.0304*	0.0029	0.0321*	-0.0713*	0.0499*	0.1930*	-0.0053
engagement~l	-0.0205	0.0096	0.1231*	0.0106	0.0088	-0.0874*	0.0020
fondderoul~t	0.1513*	-0.0028	-0.0090	-0.0579*	-0.0718*	0.1588*	-0.0109
immo	-0.1108*	-0.0284*	0.0045	0.0316*	0.0443*	-0.1019*	-0.0119
refinancem~e	-0.0313*	0.0179	0.0190	0.0192	0.0048	-0.0057	-0.0060
changement~o	-0.0030	0.0318*	-0.0114	0.0114	0.0227*	-0.0485*	0.0378*
manufacture	0.0907*	-0.0117	0.0264*	-0.0546*	-0.0609*	0.0935*	-0.0104
agricultur~n	-0.0055	0.0157	-0.0197	0.0330*	0.0014	-0.0499*	0.0195
venteaudtail	-0.0288*	-0.0142	0.0615*	-0.0528*	-0.0080	0.0857*	-0.0057
service	-0.0539*	0.0078	-0.0606*	0.0770*	0.0595*	-0.1297*	-0.0015
autre	-0.0271*	0.0097	-0.0255*	0.0102	0.0328*	-0.0261*	-0.0006
tailleave~fs	-0.0097	0.0005	0.2006*	-0.0180	-0.1679*	0.0464*	-0.0016
difference~e	-0.0101	0.0115	-0.0053	-0.0199	0.0329*	0.0410*	-0.0012

[24] Les étoiles représentent les corrélations significatives à un niveau de 10% ou mieux.

	actdtot	croiss~n	croiss~b	croiss~s	roa	fondsd~s	lagrat~e
actdtot	1.0000						
croissancebn	0.0018	1.0000					
croissancemb	0.0009	0.8342*	1.0000				
croissance~s	-0.0017	0.0029	-0.0010	1.0000			
roa	0.0820*	0.0155	0.0135	0.0007	1.0000		
fondsdispo~s	0.0337*	0.0015	0.0004	-0.0014	0.2604*	1.0000	
lagratioco~e	0.0870*	-0.0151	-0.0133	-0.0071	0.0238*	0.0244*	1.0000
ratiocouve~e	0.1802*	0.0144	0.0140	0.0007	0.1129*	0.0858*	0.2340*
engagement~l	-0.0694*	-0.0208	-0.0205	-0.0084	-0.0757*	0.1692*	0.0180
fondderoul~t	0.1753*	0.0177	0.0212*	-0.0045	-0.0077	-0.0382*	0.0050
immo	-0.1118*	-0.0092	-0.0112	0.0096	-0.0155	0.0424*	-0.0070
refinancem~e	-0.0284*	-0.0112	-0.0114	-0.0045	-0.0245*	0.0042	0.0242*
changement~o	-0.0356*	0.0003	-0.0011	-0.0049	0.0549*	-0.0182	-0.0167
manufacture	0.0377*	-0.0114	-0.0167	-0.0100	-0.0403*	0.0336*	-0.0078
agricultur~n	-0.0977*	0.0081	0.0089	0.0282*	-0.0051	-0.0107	0.0069
venteaudtail	0.1913*	0.0159	0.0200	-0.0088	-0.0182	0.0222*	0.0238*
service	-0.1091*	-0.0091	-0.0074	-0.0066	0.0725*	-0.0373*	0.0007
autre	-0.0688*	-0.0049	-0.0066	-0.0020	-0.0105	-0.0267*	-0.0511*
tailleave~fs	0.0094	-0.0096	-0.0103	-0.0012	-0.0366*	0.2091*	0.0039
difference~e	0.0294*	0.0232*	0.0212*	0.0071	0.0475*	0.0301*	-0.7985*

	ratioc~e	engage~l	fondde~t	immo	refina~e	change~o	manufa~e
ratiocouve~e	1.0000						
engagement~l	0.0003	1.0000					
fondderoul~t	-0.0068	-0.1199*	1.0000				
immo	-0.0031	0.1364*	-0.6292*	1.0000			
refinancem~e	0.0557*	0.0210	-0.1718*	-0.3395*	1.0000		
changement~o	-0.0341*	-0.0690*	-0.2078*	-0.4107*	-0.1121*	1.0000	
manufacture	-0.0243*	0.0948*	0.0591*	-0.0236*	-0.0111	-0.0323*	1.0000
agricultur~n	-0.0163	-0.0477*	-0.0749*	-0.0113	0.0612*	0.0635*	-0.3381*
venteaudtail	0.0205	0.0414*	0.0457*	-0.0125	-0.0056	-0.0363*	-0.3619*
service	0.0174	-0.0839*	-0.0165	0.0411*	-0.0394*	-0.0075	-0.3577*
autre	0.0104	-0.0370*	-0.0505*	0.0191	-0.0045	0.0412*	-0.1279*
tailleave~fs	0.0147	0.2936*	-0.0216*	0.0314*	0.0200	-0.0368*	0.0616*
difference~e	0.3984*	-0.0168	-0.0090	0.0047	0.0117	-0.0054	-0.0077

	agricu~n	ventea~l	service	autre	taill~fs	differ~e
agricultur~n	1.0000					
venteaudtail	-0.2762*	1.0000				
service	-0.2730*	-0.2923*	1.0000			
autre	-0.0976*	-0.1045*	-0.1033*	1.0000		
tailleave~fs	-0.0264*	-0.0018	-0.0302*	-0.0236*	1.0000	
difference~e	-0.0167	-0.0097	0.0101	0.0547*	0.0054	1.0000

Test de multicolinéarité[25]

Variable	VIF	1/VIF
manufacture	6.81	0.146884
venteaudtail	5.88	0.169945
service	5.71	0.175099
agricultur~n	5.41	0.184771
immo	3.39	0.294746
croissancemb	3.29	0.303777
croissancebn	3.29	0.303792
fondderoul~t	3.01	0.332184
changement~o	2.15	0.464962
fondsdispo~s	1.97	0.507547
tailleave~es	1.88	0.533194
eatot	1.77	0.564806
actdtot	1.49	0.671063
ratiocouve~e	1.28	0.783066
de	1.27	0.785762
difference~e	1.20	0.835104
engagement~l	1.19	0.842354
tailleave~fs	1.18	0.849460
roa	1.13	0.885576
dbanquedtot	1.08	0.925013
anne	1.01	0.986010
fr	1.00	0.997547
croissance~s	1.00	0.998588
Mean VIF	2.50	

[25] Test VIF – *Variation Inflation Factor*

Annexe D

Régressions effectuées

RÉGRESSION (1) : Modèle probit contenant les variables financières seulement (robust)

```
Probit regression                                 Number of obs   =       6095
                                                  Wald chi2(8)    =      22.21
                                                  Prob > chi2     =     0.0045
Log pseudolikelihood =  -2142.655                 Pseudo R2       =     0.0084

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          de |  -.0242494    .023269    -1.04   0.297    -.0698559    .0213571
       eatot |  -.1276171   .2333404    -0.55   0.584    -.5849559    .3297216
          fr |  -.0061779   .0072467    -0.85   0.394    -.0203811    .0080253
     actdtot |   .0198941   .0110067     1.81   0.071    -.0016787    .0414669
croissance~s |   .0177885   .0069922     2.54   0.011     .0040841    .0314929
         roa |  -.0416882   .0242188    -1.72   0.085    -.0891561    .0057798
fondsdispo~s |  -7.38e-09   7.89e-09    -0.94   0.349    -2.28e-08    8.08e-09
ratiocouve~e |  -.0255636   .0119414    -2.14   0.032    -.0489682   -.0021589
       _cons |  -1.072863   .1268128    -8.46   0.000    -1.321411   -.8243144
------------------------------------------------------------------------------
```

RÉGRESSION (2) : Modèle probit contenant toutes les variables étudiées (robust)

```
Probit regression                               Number of obs   =       6095
                                                Wald chi2(22)   =     202.05
                                                Prob > chi2     =     0.0000
Log pseudolikelihood = -2051.7013               Pseudo R2       =     0.0505

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anne |   -.052272   .0440359    -1.19   0.235    -.1385807    .0340367
tailleave~es |  -9.53e-10   5.88e-10    -1.62   0.105    -2.11e-09    1.99e-10
          de |  -.0206257    .018564    -1.11   0.267    -.0570105    .0157591
 dbanquedtot |  -.0938707   .0941332    -1.00   0.319    -.2783683    .0906269
       eatot |  -.4078518   .2203588    -1.85   0.064    -.8397471    .0240434
          fr |  -.0033828   .0065559    -0.52   0.606    -.0162322    .0094666
     actdtot |   .0125672   .0126429     0.99   0.320    -.0122126    .0373469
croissancebn |   .0001668   .0016098     0.10   0.917    -.0029884     .003322
croissancemb |  -.0009082   .0014309    -0.63   0.526    -.0037127    .0018963
croissance~s |   .0164494   .0069921     2.35   0.019     .0027452    .0301537
         roa |  -.0356763    .024388    -1.46   0.144     -.083476    .0121234
fondsdispo~s |   3.26e-09   9.97e-09     0.33   0.744    -1.63e-08    2.28e-08
lagratioco~e |   .0009733    .007032     0.14   0.890    -.0128092    .0147558
ratiocouve~e |   -.021146   .0118284    -1.79   0.074    -.0443292    .0020372
engagement~l |  -9.53e-09   1.69e-08    -0.56   0.572    -4.26e-08    2.35e-08
fondderoul~t |   .5391723   .0898323     6.00   0.000     .3631043    .7152403
        immo |   .0055114   .0870466     0.06   0.950    -.1650968    .1761196
changement~o |   .1715271   .1028818     1.67   0.095    -.0301176    .3731718
 manufacture |   .4046787   .1382633     2.93   0.003     .1336877    .6756697
agricultur~n |   .2230403   .1414881     1.58   0.115    -.0542713    .5003519
venteaudtail |    .110867   .1420492     0.78   0.435    -.1675443    .3892783
     service |   .0313709   .1424343     0.22   0.826    -.2477952    .3105371
       _cons |  -1.244806    .206231    -6.04   0.000    -1.649011   -.8406003
------------------------------------------------------------------------------
```

RÉGRESSION (3) : Modèle probit en excluant certaines variables non significatives (robust)

```
Probit regression                                 Number of obs   =       6095
                                                  Wald chi2(17)   =     198.68
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -2052.5109                 Pseudo R2       =     0.0501

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anne |  -.0523577   .0440417    -1.19   0.235    -.1386778    .0339624
tailleave~es |  -9.47e-10   5.84e-10    -1.62   0.105    -2.09e-09    1.98e-10
          de |    -.02049   .0185283    -1.11   0.269    -.0568048    .0158248
 dbanquedtot |  -.0954121   .0943933    -1.01   0.312    -.2804194    .0895953
       eatot |  -.4065016   .2199031    -1.85   0.065    -.8375036    .0245005
          fr |  -.0034264   .0066269    -0.52   0.605     -.016415    .0095621
     actdtot |   .0127987   .0126149     1.01   0.310    -.0119261    .0375235
croissance~s |   .0164658   .0069706     2.36   0.018     .0028037     .030128
         roa |  -.0360915    .024406    -1.48   0.139    -.0839264    .0117434
fondsdispo~s |   3.22e-09   9.94e-09     0.32   0.746    -1.63e-08    2.27e-08
ratiocouve~e |  -.0208274   .0111669    -1.87   0.062    -.0427142    .0010594
engagement~l |  -8.98e-09   1.69e-08    -0.53   0.594    -4.20e-08    2.41e-08
fondderoul~t |   .5342362   .0492813    10.84   0.000     .4376467    .6308257
changement~o |   .1660444    .069345     2.39   0.017     .0301307    .3019581
 manufacture |   .3765389   .0600012     6.28   0.000     .2589386    .4941392
agricultur~n |   .1944355   .0678146     2.87   0.004     .0615213    .3273497
venteaudtail |   .0819527   .0677848     1.21   0.227    -.0509031    .2148086
       _cons |  -1.213146   .1414747    -8.58   0.000    -1.490431   -.9358605
------------------------------------------------------------------------------
```

Interprétation des coefficients – effets marginaux :

```
Marginal effects after probit
      y  = Pr(dfaut) (predict)
         =  .10425242
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    anne*|  -.0095269      .00806   -1.18   0.237   -.02533  .006276   .588023
taill~es |  -1.71e-10      .00000   -1.63   0.103  -3.8e-10  3.5e-11   2.4e+07
      de |  -.0037066      .00335   -1.11   0.268  -.010266  .002852    1.8947
dbanqu~t |  -.0172598      .01708   -1.01   0.312  -.050731  .016212   .809997
   eatot |  -.0735349      .03975   -1.85   0.064  -.151444  .004374   .415452
      fr |  -.0006198       .0012   -0.52   0.604  -.002962  .001723   2.68065
 actdtot |   .0023152      .00228    1.02   0.310  -.002155  .006786   1.33785
croiss~s |   .0029786      .00128    2.33   0.020   .000477   .00548    .85337
     roa |  -.0065288      .00441   -1.48   0.139  -.015172  .002114   .401418
fondsd~s |   5.82e-10      .00000    0.32   0.746  -2.9e-09  4.1e-09   1.3e+06
ratioc~e |  -.0037676      .00202   -1.86   0.062  -.007728  .000193   1.45677
engage~l |  -1.62e-09      .00000   -0.53   0.594  -7.6e-09  4.4e-09   1.1e+06
fondde~t*|   .1142104        .012    9.51   0.000   .090684  .137737   .241509
change~o*|   .0324725      .01455    2.23   0.026   .003954  .060991   .119442
manufa~e*|   .0746129      .01289    5.79   0.000   .049348  .099878   .306973
agricu~n*|   .0377671      .01408    2.68   0.007   .010173  .065361   .205086
ventea~l*|   .0152446      .01296    1.18   0.239  -.010156  .040645    .22822
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

RÉGRESSION (4) : Modèle probit en ne conservant que les variables significatives (robust)

```
Probit regression                                 Number of obs   =       6095
                                                  Wald chi2(8)    =     190.15
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -2058.0567                 Pseudo R2       =     0.0475

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tailleave~es |  -8.24e-10   4.22e-10    -1.95   0.051    -1.65e-09    4.22e-12
       eatot |  -.1661851   .1405244    -1.18   0.237    -.4416078    .1092377
croissance~s |   .0164984   .0069756     2.37   0.018     .0028265    .0301703
ratiocouve~e |  -.0223819   .0115524    -1.94   0.053    -.0450243    .0002604
fondderoul~t |   .5469215   .0482259    11.34   0.000     .4524004    .6414425
changement~o |   .1594745   .0683966     2.33   0.020     .0254196    .2935295
 manufacture |   .3398972   .0489606     6.94   0.000     .2439363    .4358582
agricultur~n |   .1519925   .0585529     2.60   0.009     .0372308    .2667541
       _cons |  -1.430393   .0663816   -21.55   0.000    -1.560499   -1.300288
------------------------------------------------------------------------------
```

RÉGRESSION (5) : Modèle logit en ne conservant que les variables significatives (robust)

```
Logistic regression                               Number of obs   =       6095
                                                  Wald chi2(8)    =     196.35
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -2058.3124                 Pseudo R2       =     0.0474

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tailleave~es |  -1.73e-09   8.66e-10    -2.00   0.046    -3.43e-09   -3.45e-11
       eatot |   -.305151   .2640778    -1.16   0.248     -.822734     .212432
croissance~s |   .0284639   .0105488     2.70   0.007     .0077886    .0491391
ratiocouve~e |  -.0365312   .0215274    -1.70   0.090    -.0787241    .0056617
fondderoul~t |    1.02165   .0890835    11.47   0.000     .8470491     1.19625
changement~o |   .3217182   .1319549     2.44   0.015     .0630914     .580345
 manufacture |   .6439768   .0926287     6.95   0.000     .4624279    .8255258
agricultur~n |   .2914086   .1132248     2.57   0.010     .0694921    .5133251
       _cons |  -2.503646   .1284261   -19.49   0.000    -2.755357   -2.251935
------------------------------------------------------------------------------
Note: 0 failures and 1 success completely determined.
```

RÉGRESSION (6) : Test de stabilité – Modèle probit en éliminant les premiers et derniers percentiles (robust)

```
Probit regression                                 Number of obs   =       5973
                                                  Wald chi2(17)   =     207.00
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -2015.9899                 Pseudo R2       =     0.0515

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anne |  -.0550161    .044479    -1.24   0.216    -.1421933    .0321611
tailleave~es |  -2.12e-10   8.16e-10    -0.26   0.795    -1.81e-09    1.39e-09
          de |  -.0237224   .0210874    -1.12   0.261    -.0650529    .0176082
 dbanquedtot |  -.1033763     .09553    -1.08   0.279    -.2906117    .0838591
       eatot |   -.419871   .2323879    -1.81   0.071    -.8753429    .0356009
          fr |  -.0024153   .0058292    -0.41   0.679    -.0138402    .0090097
     actdtot |   .0104176   .0128832     0.81   0.419     -.014833    .0356681
croissance~s |   .0165478   .0069416     2.38   0.017     .0029424    .0301531
         roa |  -.0282823   .0245878    -1.15   0.250    -.0764736    .0199089
fondsdispo~s |  -1.64e-08   1.16e-08    -1.41   0.158    -3.92e-08    6.37e-09
ratiocouve~e |  -.0185612   .0106503    -1.74   0.081    -.0394353     .002313
engagement~l |  -7.38e-09   1.70e-08    -0.44   0.663    -4.06e-08    2.59e-08
fondderoul~t |   .5443599   .0496469    10.96   0.000     .4470537    .6416661
changement~o |   .1682161   .0698962     2.41   0.016      .031222    .3052101
 manufacture |   .3900785    .060783     6.42   0.000     .2709459     .509211
agricultur~n |   .2146833   .0686546     3.13   0.002     .0801227    .3492438
venteaudtail |   .0892753   .0688605     1.30   0.195    -.0456889    .2242394
       _cons |  -1.209676   .1502793    -8.05   0.000    -1.504218   -.9151345
------------------------------------------------------------------------------
```

RÉGRESSION (7) : Test de stabilité – Modèle probit en éliminant les cinq premiers et derniers percentiles (5%) (robust)

```
Probit regression                                 Number of obs   =       5485
                                                  Wald chi2(17)   =     202.85
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood =  -1834.654                 Pseudo R2       =     0.0540

------------------------------------------------------------------------------
             |               Robust
       dfaut |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anne |   -.080971   .0466015    -1.74   0.082    -.1723082    .0103661
tailleave~es |  -2.11e-09   1.89e-09    -1.12   0.265    -5.82e-09    1.60e-09
          de |  -.0180239   .0181931    -0.99   0.322    -.0536817     .017634
 dbanquedtot |  -.0896563   .0999227    -0.90   0.370    -.2855012    .1061885
       eatot |  -.4576379   .2258553    -2.03   0.043    -.9003062   -.0149696
          fr |  -.0017668   .0052018    -0.34   0.734    -.0119621    .0084285
     actdtot |   .0148176   .0136723     1.08   0.278    -.0119797    .0416149
croissance~s |   .0181003   .0078646     2.30   0.021      .002686    .0335146
         roa |  -.0518357   .0244725    -2.12   0.034     -.099801   -.0038703
fondsdispo~s |   2.12e-08   1.38e-08     1.53   0.126    -5.98e-09    4.83e-08
ratiocouve~e |  -.0151888    .010621    -1.43   0.153    -.0360056    .0056279
engagement~l |  -3.98e-09   1.81e-08    -0.22   0.826    -3.95e-08    3.16e-08
fondderoul~t |   .5613513   .0519543    10.80   0.000     .4595228    .6631798
changement~o |   .1783079   .0738126     2.42   0.016     .0336379    .3229778
 manufacture |   .4243719   .0645356     6.58   0.000     .2978845    .5508594
agricultur~n |    .232274   .0731675     3.17   0.002     .0888684    .3756796
venteaudtail |   .1470151   .0730009     2.01   0.044     .0039359    .2900943
       _cons |  -1.271286   .1447422    -8.78   0.000    -1.554976    -.987597
------------------------------------------------------------------------------
```