

**HEC Montréal**

**Rapport d'un projet de recherche**

Nahla Slim

**Risque de défaut : évaluation du risque individuel dans le cadre des  
produits financiers bancaires destinés aux particuliers**

**Direction**

Mr. Georges Dionne

**Hiver 2018**

## Résumé

Ce rapport étudie le risque de défaut de produits financiers bancaires destinés aux particuliers. Nous utilisons une base de données provenant d'une banque nord-américaine qui contient des informations mensuelles sur tous les produits financiers détenus par les clients de la banque entre février 2012 et mars 2017. En premier lieu, nous construisons le groupe de variables pour estimer notre modèle. En deuxième lieu, nous présentons une vue générale sur le taux d'utilisation et la probabilité de défaut de deux produits financiers différents détenus par un même individu. Finalement, nous estimons la distribution de non-paiement pour prédire le risque de défaut en deux étapes : estimer la probabilité de défaut du prêt ainsi que les paramètres de la distribution conditionnelle des non-paiements.

L'objectif principal de ce rapport est de tester empiriquement l'effet de certains déterminants liés au prêt et au client qui sont faciles à obtenir le jour de la signature du contrat. Nous procédons en deux étapes. Nous proposons un modèle économétrique afin de déterminer les variables qui expliquent le défaut ainsi que la distribution des non-paiements. Cette dernière pourrait donner des estimations des coûts de gestion. Nous effectuons une analyse de mesures des erreurs de Type I et de Type II afin de voir la précision du modèle dans la prédiction du défaut. Dans des recherches futures, il sera intéressant de tester empiriquement la présence d'asymétrie d'information.

**Mots clés :** prêt à terme, ligne de crédit, risque de crédit, risque de défaut.

## I- INTRODUCTION

Le surendettement des particuliers est un phénomène très important en finance et a attiré l'attention des chercheurs ces dernières années. Plusieurs facteurs ont été évoqués dans la littérature et plusieurs travaux empiriques ont démontré que certains de ces facteurs peuvent expliquer ce surendettement. Parmi ces facteurs, citons par exemple les taux d'intérêt très bas et l'accessibilité des cartes et des lignes de crédit. Si un individu a le choix entre la consommation et l'épargne, alors il dispose d'une allocation en argent et doit savoir l'utiliser de sorte à couvrir ses dépenses (liées à sa consommation) et ses investissements (liés à son épargne).

La littérature propose que lorsque le consommateur est confronté à un certain choc négatif, il puisse ne pas être en mesure d'exploiter de nouvelles sources de crédit. Cependant, les prêteurs ne peuvent pas exiger le remboursement immédiat des soldes impayés. Pour cette raison, certains consommateurs peuvent profiter de l'argent avancé sur les lignes de crédit ou autres prêts pour construire une certaine assurance financière. Ils peuvent aussi choisir de ne pas payer le solde au complet et conserver des liquidités pour des dépenses futures qui pourraient dépasser leurs revenus. Les recherches tentent de découvrir les différentes raisons qui déterminent le comportement de tirage sur les lignes de crédit ainsi que le comportement des emprunteurs face au défaut. Selon Bergerès et al. (2015), il existe une relation entre les différentes dettes d'un consommateur et il semble évident qu'une décision prise pour l'un des produits financiers affecte l'autre. Dans leur étude, les auteurs proposent une analyse conjointe du taux d'utilisation d'une ligne de crédit et de la probabilité de défaut sur un prêt à terme. Ils montrent que l'utilisation de la ligne de crédit influence la probabilité de défaut de paiement sur le prêt à terme. La relation inverse (où la probabilité de défaut affecte le taux d'utilisation) est aussi vraie. D'autres études suggèrent qu'il y a une certaine corrélation entre les différents risques des produits financiers détenus par un même individu.

Dans ce rapport, nous présentons une extension de cette étude à l'aide d'une base mensuelle de données. Pour ce faire, nous estimons la probabilité de défaut pour les nouveaux prêts, ainsi que la distribution des non-paiements. L'objectif est d'estimer la

distribution de non-paiement pour prédire le risque de défaut. Nous suivrons le travail de Dionne et al. (1996). Notre base de données comporte 1900 individus pour la période allant de février 2012 à décembre 2016. L'échantillon prélevé contient un total de 1779 nouveaux prêts à terme, chaque prêt ayant un montant autorisé moyen de 17 006\$ chacun. Au total, la base de données contient 39 627 observations mensuelles sur une durée de 62 mois.

Ce rapport consiste à estimer la probabilité de défaut et estimer la distribution des non-paiements. Nous présentons le cadre statique comme première étape. L'objectif est de trouver un lien entre la probabilité de défaut du prêt à terme et les caractéristiques de l'emprunteur et de son prêt. Ces dernières proviennent de caractéristiques de crédit connues par la banque au moment de l'ouverture du prêt et de l'utilisation des différents autres produits durant la vie du prêt à terme. Notre échantillon nous permet d'identifier une probabilité de défaut de 2.97%. Nous vérifions ensuite s'il existe une relation entre ces caractéristiques et la distribution des non-paiements. Pour ce faire, nous estimons un modèle de Poisson. Nous définissons les non-paiements (0, 1, 2 et 3) comme étant les non-défauts et les (4 et +) représentent les défauts. Dans un premier lieu, nous essayons d'identifier les caractéristiques qui prédisent le mieux la probabilité de défaut (il faut rappeler qu'un prêt est considéré en défaut, si le nombre de jours en retard de paiement dépasse 90 jours). Dans un deuxième lieu, nous voulons voir si ces variables sont toujours valides pour estimer la distribution des non-paiements. En troisième lieu, nous voulons voir la performance du modèle à travers la ROC curve et la matrice de confusion qui nous permet d'analyser l'erreur de Type I et celle de Type II en nous basant sur le seuil de fréquence moyenne des défauts observés. Ainsi, nous estimons les pertes des « bons » et des « mauvais » prêts, dans le but de bien calculer les coûts de gestion qui y sont relatifs.

Notre base de données comporte seulement les nouveaux prêts ayant en même temps une ligne de crédit. Nous voulons aussi étudier la relation entre le taux d'utilisation des lignes de crédit et la probabilité de non-paiement. Pour ce faire, nous effectuons une étude d'événements pour estimer la variation du taux d'utilisation des lignes de crédit sur une fenêtre de 12 mois autour de la réalisation de l'un des deux événements : manquer ou effectuer un paiement. Les résultats montrent qu'il y a une certaine baisse d'environ 40%

du taux moyen avant le premier évènement, et une augmentation d'environ 20% avant le deuxième. Cela suggère la présence d'une certaine relation entre le taux d'utilisation des lignes de crédit et les événements réalisés sur les nouveaux prêts à terme.

Des recherches futures pourraient analyser la présence d'asymétrie d'information entre la banque et ses clients en s'inspirant du modèle de Dionne et al. (1997, 2001) et Dionne et al.(2009.b).

La suite du rapport se divise comme suit : la section II présente une brève revue de littérature; la section III présente une description de notre base de données; la section IV décrit les principales variables de notre étude; la section V présente les statistiques descriptives de l'analyse; la section VI donne une vue générale sur les cycles de transition. La section VII décrit la méthodologie et les résultats. Dans la section VIII, nous présentons un test de robustesse. La section IX propose des recherches futures. Finalement, la section X présente la conclusion de notre travail.

## **II-REVUE DE LITTÉRATURE**

La relation entre l'utilisation des lignes de crédit et la probabilité de défaut a attiré l'attention des chercheurs ainsi que celles des praticiens. Beaucoup d'études ont été réalisées dans ce domaine pour mettre l'accent sur le rôle crucial que joue l'utilisation des lignes de crédit dans la détermination du risque de défaut. Parmi ces études, on trouve les travaux séminaux d'Agarwal et al. (2006a) et Jiménez et al. (2009a) qui montrent qu'il y a une relation positive entre l'utilisation des lignes de crédit et la probabilité de défaut.

Premièrement, Jiménez et al. (2009a) examinent les lignes de crédit destinées aux entreprises en Espagne pour une période de 20 ans entre 1986 et 2005. Leur analyse montre que les emprunteurs défaillants ont des taux d'utilisation des lignes de crédit significativement plus élevés. Ils observent que cette relation peut être étendue jusqu'à 5 ans avant leur défaut réel.

Dans une autre étude, Jiménez et al. (2009b) étendent leur premier papier en examinant les déterminants de l'utilisation des lignes de crédit. Leur objectif est d'extraire empiriquement les principaux facteurs influençant la décision des entreprises

d'utiliser leurs lignes de crédit. Leurs résultats soulignent l'importance du risque de défaut des entreprises pour déterminer l'utilisation des lignes de crédit. En particulier, ils constatent que les entreprises en défaut de paiement ont un taux d'utilisation élevé. Effectivement, le taux d'utilisation des lignes de crédit est très différent entre les entreprises qui font défaut et celles qui ne le font pas, même plusieurs années avant l'année du défaut.

Une autre analyse effectuée par Agarwal et al. (2006a) utilise une base de données contenant 34 384 lignes de crédit sur la période de janvier 1998 à mai 2001. Les auteurs testent deux hypothèses à propos de la relation entre le risque de crédit de l'emprunteur et son utilisation de la ligne de crédit. La première hypothèse est qu'en début de période, le taux d'utilisation de la ligne sera plus faible pour les emprunteurs qui anticipent la détérioration de leur qualité de crédit dans le futur. La deuxième hypothèse est que l'utilisation des lignes de crédit dans une période ultérieure est corrélée avec les changements dans la qualité de crédit de l'emprunteur. Leurs résultats montrent que les emprunteurs ayant des perspectives de détérioration de leur qualité de crédit dans le futur optent pour ouvrir des lignes de crédit afin de préserver leur flexibilité financière. En outre, les auteurs trouvent que les emprunteurs ayant un faible score de crédit au début utilisent un faible pourcentage de leur ligne de crédit comparé aux emprunteurs avec un score de crédit plus élevé. Leurs résultats suggèrent aussi que les emprunteurs avec des qualités de score de crédit inférieurs reconnaissent les avantages de maintenir la flexibilité financière en conservant des lignes de crédit inutilisées. En revanche, les emprunteurs ayant de faibles attentes en termes de besoin de crédit futur supplémentaire utilisent une proportion plus élevée de leurs lignes de crédit en début de période.

Ces constatations ont poussé les chercheurs à soulever la crainte que ce produit financier, la ligne de crédit, puisse également avoir des effets sur d'autres produits financiers que l'emprunteur possède. Dans ce cadre, Agarwal et al. (2006b) examinent cette corrélation dans le cadre de prêts hypothécaires et de marges de crédit hypothécaire.

Tout d'abord, leur analyse montre l'existence d'une différence significative de la propension de prépaiement et de la probabilité de défaut entre les prêts hypothécaires et les marges de crédit hypothécaire. Ils constatent que les ménages ayant un prêt

hypothécaire sont relativement plus sensibles aux changements des taux d'intérêt, tandis que ceux ayant des marges de crédit hypothécaire sont plus sensibles à l'appréciation de leur valeur hypothécaire. Cependant, ils n'ont pas analysé le lien simultané qui existe entre les deux produits, car les emprunteurs ont soit un prêt hypothécaire soit une marge de crédit hypothécaire et non pas les deux en même temps.

Des études récentes ont également souligné l'importance de la liquidité sur le défaut. Elul et al. (2010) considèrent que les emprunteurs accordent une priorité stratégique au paiement mensuel des multiples dettes afin de préserver leur liquidité. Leur hypothèse principale est que les emprunteurs qui ne remboursent pas leurs dettes hypothécaires conservent leur accès au crédit en demeurant à jour sur leur compte de carte de crédit ou de Ligne de Crédit sur Valeur Domiciliaire (LDCVD). Cela leur permet de maintenir leur consommation essentielle des autres biens et services.

Deux autres études soulignent également le rôle de la liquidité dans la détermination de la probabilité de défaut. Cohen-Cole et Morse (2010) examinent le choix des consommateurs de ne pas respecter leur obligation sur les cartes de crédit ou sur l'hypothèque. Leur étude comporte 2,2 millions de ménages de 2006 à 2007. Ils concluent que les consommateurs préfèrent garder leur compte de carte de crédit à jour que de garder leur dette hypothécaire, préservant ainsi la liquidité à un coût élevé. D'autre part, Jagtiani et Lang (2011) examinent les « décisions de faire défaut », prises par les emprunteurs, entre l'hypothèque de premier et de deuxième rang (c'est-à-dire prêt hypothécaire ou LDCVD) pour la période de 2004 à 2009. Ils constatent que les propriétaires sont plus susceptibles de rester à jour sur leur LDCVD afin de préserver leur liquidité.

Dans une autre étude qui appuie l'importance de la liquidité sur le défaut, Andersson et al. (2013) examinent comment les emprunteurs priorisent l'ordre de payer leur dette hypothécaire mensuelle et leur dette de carte de crédit sur une période de neuf ans (2001-2009). Ils analysent l'importance empirique de facteurs reflétant à la fois le comportement stratégique de défaut et la capacité de payer selon l'ordre de défaillance entre la dette hypothécaire et la dette de carte de crédit. Leurs données permettent de suivre le comportement de remboursement mensuel d'un consommateur au fil du temps et

leur permettent de mesurer les caractéristiques clés des consommateurs et de leurs produits de dette. Ils constatent que le comportement de défaut sur la dette hypothécaire précède le défaut sur la carte de crédit; en d'autres mots, les consommateurs risquent davantage de ne pas rembourser leur prêt hypothécaire que de rembourser leur carte de crédit.

Ces études confirment que les emprunteurs préfèrent rester à jour sur le produit qui leur permet de préserver leur liquidité, et de faire défaut sur le produit hypothécaire.

Un autre volet de la littérature étudie le contenu informationnel des comptes chèque et justifie la nécessité de surveiller l'utilisation du crédit renouvelable en même temps que la probabilité de défaut. Norden et Weber (2010) enquêtent sur la vérification des comptes et des informations qu'ils fournissent concernant l'utilisation des lignes de crédit tant pour les entreprises que pour les consommateurs. En utilisant des données sur les grandes entreprises, sur les petites entreprises et sur les particuliers pour la période de 2002 à 2006, ils constatent que l'activité du compte est un déterminant important pour surveiller les petites entreprises et les individus, mais pas pour les grandes entreprises. Ils étudient le lien entre l'activité du compte et la production de l'information sur la qualité de l'emprunteur. Ils proposent que l'analyse de l'activité du compte des emprunteurs est une source importante d'informations privées pour les banques. Les deux principaux résultats de leur étude montrent que l'activité du compte est très informative, et que les banques utilisent cette information pour gérer leurs prêts. Ils affirment que l'incorporation d'informations sur l'activité du compte améliore considérablement les modèles de prédiction de défaut, et qu'elle est particulièrement utile pour surveiller les petites entreprises et les particuliers.

Dans le même contexte, Mester et al. (2007) constatent également que les comptes chèque des entreprises peuvent fournir des informations utiles sur la qualité du crédit de l'emprunteur. Ils montrent que les banques intensifient l'activité de vérification des comptes de transaction lorsque les prêts sont perçus en phase de délinquance (manquer des paiements). Les comptes chèque de l'entreprise donnent de l'information à la banque sur leurs liquidités. Cette dernière peut se servir de cette information pour prédire la probabilité de défaut. Comme les transactions sont informatives, l'observation des



transactions devrait aider les intermédiaires financiers à surveiller les emprunteurs. Les auteurs analysent un ensemble unique de données qui comprend des informations mensuelles et annuelles sur les soldes des transactions et les garanties posées par les emprunteurs de petites entreprises à une banque canadienne. Ils constatent que (1) les changements mensuels dans les comptes débiteurs sont perceptibles dans les changements des soldes des transactions des comptes lorsque l'emprunteur a une relation bancaire exclusive avec le prêteur; (2) le total d'emprunts antérieurs dépassant la garantie est un prédicteur de détérioration du crédit et de dépréciations de prêts, et que le prêteur utilise cette information rapidement; et (3) le prêteur intensifie la surveillance quand les prêts se détériorent. En d'autres termes, les revues de prêt deviennent plus longues et plus fréquentes. Pris ensemble, ces résultats établissent un ensemble de liens montrant que les banques peuvent utiliser les transactions des comptes pour surveiller les emprunteurs. Cet article décrit les efforts d'une banque canadienne qui utilise l'information contenue dans les transactions afin de bien surveiller les activités des petites entreprises. Lorsque les emprunts dépassent la propre évaluation de la banque des comptes débiteurs et des actifs, il y a signal de détérioration du crédit. De plus, les mouvements dans la vérification des soldes de comptes sont étroitement liés aux mouvements de l'évaluation par la banque, suggérant fortement que le compte courant fournit une fenêtre transparente sur ces aspects de l'activité d'une entreprise.

Cela incite la recherche à s'intéresser de plus près aux informations provenant des comptes chèque pour prédire la probabilité de défaut sur d'autres produits financiers. Dans ce cadre, Hibbeln et al. (2016) examinent si les entrées et les sorties des fonds dans les comptes chèque prédisent le défaut sur la carte de crédit. Ils étudient la collecte d'information dans le crédit à la consommation, en utilisant des données de panel pour les comptes individuels de carte de crédit entre 2007 et 2014. Ils examinent à quel point les informations personnelles confidentielles provenant des comptes chèque et des comptes de cartes de crédit sont utiles pour évaluer le risque de défaut des consommateurs. Pour ces produits de crédit, il est essentiel que les prêteurs produisent continuellement des informations confidentielles, car, contrairement aux hypothèques, ces produits de crédit ne sont pas garantis et leur exposition au risque varie en fonction du temps. Ils obtiennent trois résultats principaux. Tout d'abord, la vérification des comptes et des cartes de crédit

contient des informations importantes sur le risque de défaut au-delà des cotes de crédit, des caractéristiques des emprunteurs et des caractéristiques de la relation banque-emprunteur. Ils montrent que les flux de trésorerie des deux comptes fournissent des informations complémentaires et dépendent du type de choc que le consommateur subit. Ils constatent aussi que les rentrées et les sorties de fonds diminuent avant le défaut du compte chèque, mais elles augmentent avant les défauts des comptes de cartes de crédit. Il est important de noter que les informations sur plusieurs produits sont précieuses, car l'activité d'un compte aide à prédire le risque de défaut de l'autre compte et vice versa. Deuxièmement, les mesures d'activité des comptes chèques sont plus utiles pour la prédiction du défaut que les mesures d'activité des comptes de cartes de crédit. Troisièmement, les erreurs de décision sont significativement plus faibles lorsque le prêteur prend en compte les informations croisées sur les produits.

Les banques peuvent rassembler et valider des informations confidentielles privées en interagissant de manière répétée avec le même emprunteur et/ou produit au fil du temps. De telles informations croisées sur les produits sont utiles si elles sont complémentaires les unes aux autres. En d'autres termes, les informations privées provenant de différentes sources sont similaires aux pièces d'un puzzle qui, uniquement lorsqu'il est entièrement assemblé, montre l'ensemble de l'image.

Deux autres études, plus proches de l'analyse de ce travail, soulignent également l'importance d'une étude conjointe de deux produits financiers. Strahan (1999) étudie le marché du crédit d'entreprises et soutient que les besoins de liquidité et la probabilité de défaut sont liés. Il trouve que le premier, le besoin de liquidité, est plus élevé lorsque le financement d'autres sources est plus difficile à obtenir, généralement en raison du risque de défaut élevé. L'auteur note qu'un accord entre une entreprise et sa banque devrait inclure à la fois un prêt à terme et une ligne de crédit, ce qui pourrait aider la banque à mieux comprendre la qualité du crédit de l'emprunteur. Les informations sur les deux instruments de crédit doivent être combinées pour que le comportement de prélèvement sur les lignes de crédit puisse être relié à l'évaluation des risques de défaut d'un prêt à terme, et vice versa.

Récemment, Bergerès et al. (2015) ont poursuivi empiriquement cette idée pour les produits destinés aux particuliers en montrant que l'information sur l'utilisation des lignes de crédit peut fournir des informations sur la qualité de crédit de leurs prêts à terme personnels. Ils analysent également l'autre direction de la dépendance entre les deux instruments. Les résultats du travail de Bergerès et al. (2015) soulignent la dépendance entre l'utilisation de la ligne de crédit et l'état de défaut sur le prêt à terme. Les auteurs confirment qu'un défaut de prêt à terme a un effet positif sur l'utilisation de la ligne de crédit et ils constatent que l'utilisation de la ligne de crédit réduit la probabilité de défaut sur le prêt à terme.

En nous basant sur ces résultats, notre objectif est d'étendre les recherches de Bergerès et al. (2015) et de proposer un modèle économétrique afin de déterminer les variables qui expliquent le défaut ainsi que la distribution des non-paiements.

Mais nous commençons par présenter l'environnement empirique. Nous suivons le travail de Dionne et al. (1996) pour estimer la probabilité de défaut et aussi estimer la distribution des non-paiements. C'est la meilleure façon d'estimer la distribution des non-paiements (en deux étapes) pour prédire le risque de défaut. Dionne et al. (1996) analysent la sélection des risques pour le crédit à la consommation pour 2446 clients d'une banque espagnole à une date donnée (mai 1989). Pour procéder, les auteurs montrent que la distribution des non-paiements n'est pas un simple processus de Poisson, mais contient une barrière à la valeur 4 des non-paiements (ce qui est expliqué par le fait que le prêt fait défaut au quatrième non-paiement). Les auteurs expliquent que cette valeur permet de séparer les non-défauts des défauts. Pour estimer la probabilité d'obtenir plus de trois non-paiements les auteurs présentent un modèle Logit, et pour estimer la distribution conditionnelle des non-paiements ils utilisent des distributions de comptage tronquées : modèle Binomial Négatif. Leurs résultats ont montré que plusieurs caractéristiques facilement observables des consommateurs peuvent affecter leurs probabilités de défaut et leur distribution de non-paiements des mensualités. Ils ont aussi présenté comment calculer les espérances des coûts des non-paiements et les primes de risque à demander sur les taux d'intérêt.

### **III-DESCRIPTION DES DONNÉES :**

Le but de ce travail est d'estimer la probabilité de défaut ainsi que la distribution des non-paiements dans le cadre des produits financiers bancaires destinés aux particuliers.

La base de données utilisée aux fins de notre recherche provient d'une banque nord-américaine. Elle contient de l'information sur 74 417 individus et s'étale de février 2010 jusqu'à mars 2017. Cette base contient cinq différents produits financiers : comptes chèques, prêt à terme, marge de crédit, hypothèque, prêt à terme sur carte de crédit. Au total, l'échantillon comporte 9 045 335 observations mensuelles de tous les produits. La qualité de cet échantillon permet d'obtenir une vaste sélection de variables concernant les consommateurs nord-américains.

Aux fins de notre étude, nous sommes intéressés par les prêts à terme ouverts au cours de la période allant de février 2010 à décembre 2016. Nous avons procédé à un nettoyage des données dans le but de garder uniquement 7 196 individus possédant de nouveaux prêts à terme (nous avons supprimé tous les prêts à terme ouverts avant février 2010). Chaque individu pourrait avoir de 1 à 6 prêts à terme au cours de cette période.

Pour chaque nouveau prêt, l'ensemble d'observations des autres produits fournit des informations détaillées obtenues au moment de la signature du contrat. Les principales variables révèlent les caractéristiques de l'emprunteur, du prêt en question et des autres produits qui sont en possession de l'emprunteur au moment de l'ouverture du prêt. Pour les caractéristiques liées aux emprunteurs, la base de données fournit des informations démographiques, la cote de risque et sa situation financière. Pour les nouveaux prêts à terme, la base de données signale la date d'ouverture, la date d'échéance de chaque prêt et aussi son âge. En ce qui concerne les caractéristiques des produits financiers, l'ensemble de données comprend diverses variables. Elles décrivent : le type de prêt; le but du prêt; le montant et la fréquence de remboursement; le montant autorisé; le solde; le taux d'intérêt et le taux écart; le type et le taux d'assurance; le type et les valeurs retenues de la garantie; ainsi que le nombre de jours et le montant total en retard. Les mêmes caractéristiques sont présentes pour les lignes de crédit.

Une étape importante dans notre procédure de gestion des données est d'assortir pour chaque enregistrement de nouveaux prêts à terme des données historiques extraites des quatre autres produits financiers. Étant donné que nous souhaitons étudier la probabilité de défaut des nouveaux prêts à terme, la procédure de nettoyage des données est basée sur une combinaison: garder uniquement les nouveaux prêts à terme ayant au moins un autre produit financier au moment de l'ouverture comme première étape, ensuite associer tout produit parmi les quatre autres liés respectivement à chaque nouveau prêt. Cette association est basée sur deux importants identifiants. Le premier représente l'identifiant unique pour chaque individu (cette variable est donnée par la banque). Le deuxième représente l'identifiant unique pour chaque produit et est construit par nous-mêmes (individus, date d'ouverture et type de produit). Les données du panel résultantes et suivies mensuellement sont importantes en termes de taille. Après la procédure d'association, nous effectuons ensuite un nettoyage de nombreuses données, tel que la suppression des doublons, le remplissage des observations manquantes <sup>1</sup>, la suppression des prêts ayant un montant autorisé inférieur à 100\$, la suppression des prêts ayant moins de trois observations, la suppression des prêts n'ayant aucun produit financier autre que le prêt à terme le jour de l'ouverture, la suppression des prêts destinés aux étudiants, etc. À la suite de ce nettoyage, nous avons réorganisé notre base de données afin de créer un ensemble complet de prêts que nous pouvions suivre dans le temps. Par ailleurs, la variable *cote de risque* est une variable importante dans notre analyse, cependant elle est manquante pour toute la période entre 2010 et janvier 2012. Alors, nous avons décidé de garder uniquement les nouveaux prêts pour lesquels nous avons cette variable, ceux ouverts à partir de février 2012. Notre travail porte sur l'analyse entre des produits différents, alors nous avons gardé uniquement les nouveaux prêts associés à au moins une ligne de crédit.

À la fin, l'échantillon comporte 1 900 individus pour la période de février 2012 à décembre 2016. Tout en sachant que chaque individu pourrait avoir de 1 à 5 nouveaux

---

<sup>1</sup>Étant donné qu'un individu pourrait avoir plusieurs produits en même temps, nous avons essayé de chercher l'information qui pouvait exister pour les autres produits, mais pas pour le nouveau prêt à terme. Pour chaque produit la banque enregistre les caractéristiques de l'emprunteur. Alors, nous pouvons récupérer l'information sur sa cote de crédit, son âge, son revenu, son ancienneté avec la banque, etc, qui existe pour un produit autre que son nouveau prêt à terme. La récupération est faite par individu et par date d'observation.

prêts à terme, nous obtenons un total de 1 779 nouveaux prêts à terme, ayant un montant autorisé moyen de 17 006\$ chacun et que nous allons suivre jusqu'en mars 2017. Ainsi, la base de données comporte 39 627 observations mensuelles pour 62 mois.

L'objectif principal est de lier la probabilité de défaut au niveau du prêt à terme ainsi que la distribution conditionnelle de non-paiements avec des informations fondamentales sur l'emprunteur, qui proviennent de son comportement et de son utilisation de lignes de crédit. Nous avons dans cette base 53 nouveaux prêts à terme considérés en défaut de paiement par la banque, soit un taux de 2.97% et 407 nouveaux prêts à terme en phase de délinquance (manque des paiements) selon la définition de la banque, soit 22.87%. Cet échantillon permet d'établir avec confiance des analyses empiriques avec les différentes variables retenues.

#### **IV-VARIABLES :**

Notre variable dépendante est la probabilité de défaut. Elle est représentée par une variable indicatrice, égale à 1 si le nombre de jours en retard dépassent les 90 et 0 sinon. Notre deuxième variable dépendante est le non-paiement. C'est une variable de comptage modélisée par un processus de Poisson tel que (0, 1, 2, et 3) représentent le groupe de non-défaut, et (4 et +) représentent le groupe de défaut.

Le taux d'utilisation est une variable qui représente la valeur utilisée sur chaque produit financier pour chaque individu et à chaque observation. Nous expliquons le calcul fait dans le tableau 1.

Pour contrôler les caractéristiques des emprunteurs de l'échantillon, nous allons nous référer aux mêmes variables de contrôle déjà développées par la littérature.

La cote de risque est l'une des variables de contrôle des plus importantes dans notre analyse. Dionne (2017) explique l'importance de la notation de crédit comme un outil utilisé par les banques pour gérer le risque de défaut. Cette variable présente un effet d'incitation dans le sens où une personne ayant un mauvais score sera plus incitée à fournir de l'effort pour garder sa ligne de crédit renouvelable et pour pouvoir obtenir d'autres crédits si elle le souhaite. Elle permet aussi à la banque de mieux évaluer son

portefeuille de risque. Pour ces raisons, nous pensons que cette variable est importante dans notre étude. La cote de crédit utilisée dans notre analyse est une cote interne. Dans notre régression, nous prévoyons que le prêteur accordera une cote élevée aux emprunteurs à risque, c'est-à-dire les personnes ayant de mauvais antécédents de paiement et/ou des niveaux élevés d'endettement. D'autre part, nous nous attendons à ce que les prêteurs aient tendance à accorder une faible cote de crédit à des personnes ayant un faible risque de défaut. Ainsi, une relation positive est attendue entre la cote de risque et la probabilité de défaut.

Nous considérons également le taux d'intérêt comme une variable clé de risque de crédit dans notre régression. Comme les taux d'intérêt plus élevés reflètent un risque de défaut plus élevé, nous nous attendons à ce que cette variable soit positivement corrélée avec la probabilité de défaut sur le prêt à terme.

Bergerès et al. (2015) utilisaient la variable *collatéral* sur le prêt comme instrument dans leur équation du taux d'utilisation des lignes de crédit. Les auteurs argumentent que la présence d'une garantie n'affecte pas les besoins de liquidité de l'emprunteur, mais pouvait être un signal de son risque, car un individu qui présente une garantie ne veut pas la perdre. Nous nous attendons à une relation négative avec la probabilité de défaut sur le prêt à terme.

Dionne et al. (1996) présentent une variable *durée* comme une mesure d'exposition au risque. Ils argumentent que c'est une variable importante à considérer pour analyser la probabilité de non-paiement. Elle représente la durée de vie du contrat dans notre échantillon. Intuitivement, plus la durée est élevée plus la probabilité de défaut devrait être grande (Dionne et al. (1996)). Nous nous attendons à une relation positive avec la probabilité de défaut.

Nous considérons également le montant de remboursement du prêt comme une caractéristique de risques supplémentaires pour les prêts. D'une part, plus le solde du prêt est élevé, plus les versements mensuels du principal et des intérêts sont importants. Ainsi, plus la probabilité pour l'emprunteur de manquer un ou deux paiements, et même de faire défaut, est élevée. Par conséquent, une relation positive est attendue, un montant de

remboursement de prêt plus élevé est probablement associé à un risque de défaut plus élevé.

Si nous considérons le montant de remboursement du prêt, nous ne pouvons pas ne pas considérer le nombre de mois restants à payer, l'échéance sur le prêt est importante aussi, et plus l'échéance est longue, plus la probabilité de manquer des paiements et, par la suite le défaut, est grande. Ainsi une relation positive est attendue.

Dans la littérature, l'ancienneté du client avec la banque est utilisée comme un proxy à l'étroitesse de la relation que la banque entretient avec son client. Bergères et al. (2015) décrivent cette variable en termes de nombre de mois. Un emprunteur ayant une relation avec la banque depuis 10 ans n'a pas le même risque qu'un emprunteur qui vient d'ouvrir son compte. Une relation négative est attendue.

Une classification par tranches d'âge de l'emprunteur est souhaitable pour capter l'effet sur le comportement d'utilisation et la probabilité de défaut.

Finalement, le sexe est représentée par une variable binaire afin de tenir compte de la relation entre le sexe de l'emprunteur et sa probabilité de défaut.

Nous présentons les deux variables dépendantes, les variables de contrôle et leurs descriptions dans le tableau 1.

## **V- STATISTIQUES DESCRIPTIVES :**

Dans cette section, nous présentons des statistiques descriptives pour les variables utilisées dans l'analyse économétrique. Le panel A du tableau 2 présente des caractéristiques liées au prêt. Nous avons 1779 nouveaux prêts à terme avec un taux d'intérêt moyen de 7.96%. Le paiement mensuel moyen est de 167 \$. Le montant est calculé sur une base mensuelle pour chaque nouveau prêt à terme même si la fréquence de remboursement n'est pas toujours mensuelle et diffère d'un prêt à un autre.



**Tableau 1 : liste des variables utilisées dans l'analyse économétrique**

Variable	Définition	Description
Statut du défaut	Plus que 90 jours de retard sur le prêt à terme de l'emprunteur	=1 si le prêt à terme est en défaut, 0 sinon
Nb. Non-paiements	De 0 à 8 qui reflète le nombre de non-paiements mensuels pour chaque prêt. 0, 1, 2, 3 : groupe de non-défaut 4, 5, 6, 7, 8 : groupe des défauts	=0 : aucun non-paiement =1 : 1 non-paiement manqué =2 : 2 non-paiements manqués =3 : 3 non-paiements manqués =4 : 4 non-paiements manqués =5 : 5 non-paiements manqués =6 : 6 non-paiements manqués =7 : 7 non-paiements manqués =8 : 8 non-paiements manqués
Taux d'utilisation	La valeur de l'utilisation sur le produit financier pour chaque individu à chaque observation	=Montant autorisé <sub>it</sub> - solde <sub>it</sub> / montant autorisé <sub>it</sub>
Cote de risque	Groupe de catégorie de 1 à 3 qui reflète le niveau de risque de l'emprunteur	=Groupe1 : faible risque (groupe de référence) =Groupe 2 : risque moyen =Groupe 3 : risque élevé
Taux d'intérêt	Taux attribué par la banque pour chaque emprunteur	=Pourcentage
Collatéral	La garantie mise par l'emprunteur sur le prêt à terme	=1 s'il y a un collatéral, 0 sinon
Durée	La durée du prêt : Une mesure d'exposition au risque	=Le nombre de mois entre le jour d'ouverture du prêt et le dernier mois d'observation
Montant de remboursement	Valeur mensuelle du montant à payer pour acquitter le prêt à terme	=Montant monétaire
Échéance de remboursement	Le nombre de mois restant pour payer le prêt	=La différence entre la date d'ouverture et la date d'échéance en termes de mois (la durée mensuelle restante sur le prêt)
Ancienneté	L'ancienneté de l'emprunteur avec la banque. Un indicateur sur la relation entre la banque et l'emprunteur	=Le nombre de mois depuis l'ouverture du compte chèque.
Sexe	Sexe de l'emprunteur	=1 si l'emprunteur est un homme, 0 si c'est une femme.
Âge	Âge de l'emprunteur (catégorie de 1 à 3) Age1 : entre 16 et 29 ans Age2 : entre 30 et 49 ans Age3 : plus que 50 ans	=1 si ça coïncide avec l'une des catégories, 0 sinon Age1 est la catégorie de référence
Nombre de prêts	Nombre de prêts additionnels actives pour chaque nouveau prêt à terme	=le nombre de prêts lié à chaque solde à chaque date d'observation
Nombre d'hypothèques	Nombre d'hypothèques actives pour chaque nouveau prêt	=le nombre d'hypothèques à chacune d'observation
Nombre de prêts sur carte	Nombre de prêts sur carte active pour chaque nouveau prêt	=le nombre de prêts sur carte à chaque date d'observation
Nombre de lignes de crédit	Nombre de lignes de crédit actives pour chaque nouveau prêt	=le nombre de prêts sur ligne de crédit à chaque date d'observation

Note : Ce tableau présente les variables, leurs définitions, et leurs descriptions utilisées dans l'analyse économétrique. Les deux premières variables sont utilisées comme variable dépendante dans l'analyse, les autres représentent les variables de contrôle de l'analyse.

L'échéance de remboursement moyenne est de 72 mois. Environ 18% de ces nouveaux prêts présentent une garantie sur le compte. Nous connaissons la date initiale du contrat et, par conséquent, le nombre de périodes contractuelles pour chaque prêt à terme représenté par la variable durée (allant de 3 à 62 mois). Le nombre de jours en retard de paiement qui permet de différencier le statut de défaut ou non depuis le début du contrat est connu. Cependant, étant donné que la date d'ouverture du prêt n'est pas la même pour tous les contrats, nous devons contrôler l'exposition au risque de défaut qui diffère d'un prêt à un autre en utilisant l'âge observée du contrat. La variable *durée* est utilisée à cette fin. Dionne et al (1996) expliquent que cette variable n'a pas été discutée dans le modèle théorique parce qu'ils supposaient implicitement que tous les individus avaient la même durée contractuelle. Cependant, dans le cas où, comme dans notre échantillon, *durée* ne mesure pas la durée du contrat, mais l'âge du contrat lors de l'échantillonnage, elle peut être considérée comme une variable exogène comme toute variable d'exposition au risque (Dionne et al. (1995), et Dionne et al. (1996)). Notre échantillon possède une durée mensuelle moyenne d'environ 23 mois. La cote de crédit sur le compte provient d'une mesure interne de la banque. Cette mesure varie de 1 à 8 avec des cotes inférieures indiquant une plus grande qualité de crédit. Nous avons regroupé ces cotes en 3 catégories selon un taux de défaut similaire entre les cotes ainsi qu'un nombre de prêts total pour avoir des groupes homogènes. Ainsi on trouve le groupe 1 (les cotes 1, 2, 3, et 4) représentant les emprunteurs avec un faible risque, le groupe 2 (les cotes 5 et 6) représentant les emprunteurs avec un risque moyen et le groupe 3 (les cotes 7 et 8) représentant les emprunteurs avec risque élevé. En moyenne, nous avons un échantillon avec une cote de risque de 1,64.

Le panel B du tableau 2 présente des informations sur les autres produits financiers actifs le jour d'ouverture du nouveau prêt à terme. Nous avons 538 nouveaux prêts à terme ayant au moins un deuxième prêt à terme actif le jour d'ouverture. Chaque nouveau prêt pourrait avoir entre 1 et 4 autres prêts à terme avec un taux d'utilisation total moyen de 37.31% et un taux d'intérêt moyen de 3.71%. Nous avons aussi 333 nouveaux prêts à terme ayant entre 1 et 10 hypothèques avec un taux d'utilisation total de 16.11% et un taux d'intérêt moyen de 3.19%. En ce qui concerne les prêts à terme sur carte de crédit, il y a 481 nouveaux prêts à terme ayant entre 1 et 32 produits de ce type avec un taux

d'utilisation total moyen de 33.51% et un taux d'intérêt moyen de 10.01%. Et finalement, nous avons 1 779 nouveaux prêts ayant une ligne de crédit le jour d'ouverture avec un taux d'utilisation total moyen de 89.04% et un taux d'intérêt moyen de 4.27%. Ce dernier produit est le seul produit renouvelable dans notre base.

Le panel C du tableau 2 présente des informations qui concernent spécifiquement l'emprunteur. Notre échantillon est composé d'environ 53% d'hommes ayant un âge moyen de 37 ans. La variable ancienneté pourrait être un indicateur sur la relation entre l'emprunteur et sa banque. En moyenne, le client avait ouvert son premier compte depuis environ 148 mois soit un peu plus que 12 ans.

**Tableau 2 : Statistiques descriptives (1779 prêts à terme ouverts entre février 2012 et décembre 2016)**

Variable	N	Moyenne	P50	Ecart-type	Min	Max
<b>Panel A : Caractéristique du prêt</b>						
Cote de risque	1779	1.64	1	0.76	1	3
Taux d'intérêt (%)	1779	7.96	6.18	4.13	0	16.25
Collatéral	1779	0.18	0	0.39	0	1
Durée (mois)	1779	22.27	18	14.76	3	62
Montant de remboursement (\$)	1779	167	107.1	182.3	0	2472
Échéance de remboursement (mois)	1779	71.31	60	61.28	1	344
<b>Panel B : Autres produits financiers</b>						
Taux d'utilisation prêt (%)	538	37.31	30.90	32.75	0	100
Nombre de prêts par nouveau prêt	538	1.54	1	0.65	1	4
Taux d'intérêt sur le prêt (%)	538	3.71	3.50	2.05	0	13.75
Taux d'utilisation d'hypothèque (%)	333	16.11	6.54	23.13	0	100
Nombre d'hypothèques par nouveau prêt	333	1.95	2	1.28	1	10
Taux d'intérêt sur hypothèque (%)	333	3.19	2.99	0.71	1.34	6.95
Taux d'utilisation prêt sur carte (%)	481	33.51	27.91	27.14	0	99.85
Nombre de prêts sur carte par nouveau prêt	481	2.60	2	2.89	1	32
Taux d'intérêt sur prêt sur carte (%)	481	10.01	13.25	6	0	22.32
Taux d'utilisation des lignes de crédit (%)	1779	89.04	100	27.55	0	100
Nombre de lignes de crédit par nouveau prêt	1779	1	1	-	1	1
Taux d'intérêt sur ligne de crédit (%)	1779	4.27	4.50	0.41	2.70	4.50
<b>Panel C : caractéristique démographique</b>						
Ancienneté (mois)	1779	147.90	139	104.3	0	734
Sexe	1779	0.53	1	0.49	0	1
Âge (mois)	1779	36.73	33	13.80	17	86

Note : les statistiques descriptives sont calculées pour les nouveaux prêts à terme ayant au moins un produit financier le jour d'ouverture. La table présente des informations liées aux comptes, des informations liées aux emprunteurs, et les taux d'utilisation des différents produits financiers détenus par les mêmes emprunteurs. Pour chaque type de produit nous présentons le taux d'utilisation sur le produit, le nombre de produits détenus par un nouveau prêt, et le taux d'intérêt du produit financier

### **a-Comparaison des moyennes :**

Dans cette section, nous voulons voir la différence des moyennes pour toutes les variables de contrôle entre les trois groupes de délinquance de notre échantillon. Le tableau 3 présente le nombre de nouveaux prêts à terme et la moyenne pour chaque variable, dépendamment si le prêt est à jour, en délinquance ou en défaut. Nous testons si la différence entre les moyennes des groupes est statistiquement significative ou non. La première colonne représente les variables de contrôle utilisées dans cette analyse. Les colonnes 2 et 3 correspondent au nombre de nouveaux prêts à terme ainsi que la moyenne des variables pour le groupe des prêts à jour (les prêts ayant 0 jour de retard). Ce groupe contient uniquement les prêts qui n'ont jamais été dans les autres cycles, autrement dit, le cycle maximum atteint par ce prêt est le premier cycle (C). Les colonnes 4 et 5 décrivent le nombre et les moyennes pour le groupe des délinquants. Ce groupe contient tous les prêts qui sont passés aux cycles de délinquance (L1, L2, ou L3) sans jamais atteindre le défaut, le maximum atteint pour chaque prêt est soit L1, L2, ou L3. Les colonnes 6 et 7 correspondent aux nombres et moyennes respectivement, si un prêt à un certain moment de son contrat avait atteint le défaut, il serait considéré dans ce groupe uniquement et il ne serait pas compté dans les deux autres groupes. Ainsi notre échantillon comporte 1319 prêts à jour (0 jour de retard), 407 prêts délinquants (L1 (entre 1 et 29 jours de retard), L2 (entre 30 et 59 jours de retard), et L3 (entre 60 et 89 jours de retard)) et 53 prêts considérés en défaut (plus que 90 jours de retard). Les trois dernières colonnes représentent les trois tests de comparaison, dont l'hypothèse nulle est « DIFF=0 » dont « DIFF » est la différence entre chaque deux groupes (moyenne groupe1-moyenne groupe2=0, moyenne groupe1-moyenne groupe3=0, moyenne groupe2-moyenne groupe3=0).

Dans notre échantillon, plus la variable cote de crédit est élevée, plus l'emprunteur sera considéré comme risqué par la banque. Ceci est reflété dans les trois groupes; les moyennes sont différentes et les différences sont statistiquement significatives. En ce qui concerne l'âge et l'ancienneté de l'emprunteur avec l'institution, les moyennes montrent que les emprunteurs en défaut sont plus âgés, mais non significative, et ont une relation d'affaires significativement plus courte avec la banque. Le test de comparaison n'est pas

significatif pour la variable âge de l'emprunteur. Le test de comparaison des moyennes entre le groupe des prêts à jour et le groupe des prêts délinquants est significatif pour toutes les variables de contrôle de l'analyse sauf sexe, ancienneté et terme de remboursement. Les variables collatéral et âge ne sont pas statistiquement différentes entre les deux groupes des prêts délinquants et des prêts en défaut. Les personnes faisant face à un défaut sur le prêt à terme ont une moyenne plus grande de la mesure d'exposition au risque représentée par la variable durée. La différence entre le groupe à jour et le groupe des délinquants ainsi que celle entre le groupe à jour et le groupe défaut est significative, mais elle ne l'est pas entre le groupe délinquant et le groupe défaut. Les emprunteurs ayant le taux d'intérêt le plus élevé sont plus susceptibles d'appartenir au groupe de défaut. Les emprunteurs ayant présenté un collatéral le jour d'ouverture de leurs prêts sont plus probables de rester à jour. Les prêts défailants présentent un montant de remboursement moyen d'environ 174\$.

**b- L'effet marginal des variables de contrôle sur la probabilité de défaut du prêt à terme :**

Nous voulons voir l'effet des variables de contrôle liées aux caractéristiques du prêt et de l'emprunteur sur la probabilité de défaut. Nous commençons par tester empiriquement quatre équations *Probit*. Chacune comporte certaines variables de contrôles. Les équations seront décrites comme suit :

$$\Pr (\text{Defaut}_i=1|\mathbf{X}_i) = f (\beta_i \mathbf{X}_i + \varepsilon_i)$$

$f(.)$  représente la fonction de distribution normale dans le cas du modèle *Probit*.

Le vecteur  $\mathbf{X}_i$  contient les variables cote de crédit, taux d'intérêt, collatéral et durée. Pour la deuxième équation, nous ajoutons les variables échéance de remboursement, ancienneté du client, le sexe et les catégories d'âge. En ce qui concerne la troisième équation, nous ajoutons la variable montant de remboursement, et nous enlevons la variable échéance de remboursement. Ainsi, le vecteur  $\mathbf{X}_i$  comporte la cote de crédit, le taux d'intérêt, le collatéral, la durée, le montant de remboursement, l'ancienneté du client, le sexe et les différentes catégories d'âge. Finalement, nous ajoutons l'échéance de remboursement pour la quatrième équation et le vecteur  $\mathbf{X}_i$  contient toutes les variables de

contrôle liées aux caractéristiques de l'emprunteur et du prêt à terme. Les résultats sont présentés dans le tableau 4.

Nous présentons les effets marginaux moyens obtenus à partir d'un modèle *Probit*, afin de capturer les effets conditionnels. La fréquence de défaut de l'échantillon est de 2.97%.

Les variables taux d'intérêt et cote de risque sont significatives dans l'équation de probabilité du défaut et sont quantitativement les facteurs les plus importants affectant la probabilité de défaut pour les quatre régressions. Le résultat suggère qu'un taux d'intérêt plus élevé et une appartenance au troisième groupe de cote de risque ont un impact statistiquement significatif sur l'augmentation de la probabilité de défaut sur un nouveau prêt à terme. Ces variables sont importantes économiquement, une personne ayant une cote appartenant au groupe 3 (de cote de risque) est susceptible d'être en défaut de 4.9% de plus qu'une personne ayant une cote appartenant au premier groupe selon la première régression et d'environ 5.4% de plus selon la quatrième régression, tout en contrôlant le reste des variables. La cote de crédit est donc un signal potentiel de probabilité de défaut d'un nouveau prêt à terme à la consommation. La cote de crédit est également très significative et positive. Cela montre que la banque a réussi à classer les clients en fonction du risque qu'ils présentent. Un consommateur plus à risque a une plus grande probabilité de défaut. L'effet marginal du taux d'intérêt sur la probabilité de défaut du nouveau prêt à terme est d'environ 0.2%. Alternativement, une augmentation de 1% du taux d'intérêt entraîne une augmentation de 0.2 point de pourcentage dans la probabilité de défaut (tout en sachant que la fréquence de défaut de l'échantillon est de 2.97%).

**Tableau 3 : Table de comparaison des moyennes ((date d'ouverture >2012m2) 1779 prêts à terme)**

Variable	Groupe non délinquant (1)		Groupe délinquant (2)		Groupe Défaut (3)		Test de comparaison		
	N	Moyenne	N	Moyenne	N	Moyenne	DIFF : (1) - (2)	DIFF : (1) -(3)	DIFF : (2) -(3)
Age	1319	37.02 (14.06)	407	35.69 (13.14)	53	37.04 (11.97)	1.32*	-0.57	-1.88
Sexe	1319	0.51 (0.49)	407	0.56 (0.49)	53	0.66 (0.47)	-0.04	-0.14**	-0.10
Ancienneté	1319	169.85 (109.18)	407	166.46 (104)	53	134.85 (87.32)	3.38	35**	31.62**
MNT remb	1319	177.44 (194.23)	407	132.28 (133.74)	53	174.10 (167.77)	45.17***	3.34	-41.82**
Terme remb	1319	71.98 (62.55)	407	67.33 (55.99)	53	85.07 (66.52)	4.65	-13.08	-17.73**
Taux intérêt	1319	7.58 (3.99)	407	8.92 (4.31)	53	9.90 (4.20)	-1.34***	-2.32***	-0.98
Cote risque	1319	1.48 (0.67)	407	2.10 (0.80)	53	2.24 (0.82)	-0.62***	-0.78***	-0.15
Collatéral	1319	0.20 (0.40)	407	0.14 (0.35)	53	0.19 (0.39)	0.05**	0.07	-0.04
Durée	1319	21.26 (14.65)	407	25.06 (14.66)	53	25.98 (14.15)	-3.80***	-4.71**	-0.92

Ce tableau présente le nombre et la moyenne pour chaque variable explicative par statut des emprunteurs. Non-délinquant représente le groupe des prêts à jour, délinquant représente les prêts ayant entre 1 et 89 jours de retard, et défaut représentent les prêts ayant >90 jours de retard. L'erreur standard est entre parenthèses. Le seuil du test de comparaison sur la différence des moyennes \* 10%; \*\* 5% et \*\*\* 1%.

**Table 4 : Probabilité de défaut pour prêt à terme (Probit marginal effects)**

**(Nouveaux prêts à terme ouverts entre février 2012 et décembre 2016 et ayant au moins un autre produit financier)**

	(1)	(2)	(3)	(4)
<b>Cote risque</b>				
1	Ref	Ref	Ref	Ref
2	0.016 (1.55)	0.020 (1.95)*	0.018 (1.72)*	0.020 (1.92)*
3	0.049 (4.42)**	0.055 (4.82)**	0.053 (4.68)***	0.054 (4.76)***
Taux intérêt	0.002 (2.14)*	0.002 (1.72)*	0.001 (1.29)	0.002 (1.72)*
<b>Collatéral</b>				
0	Ref	Ref	Ref	Ref
1	0.010 (0.95)	-0.012 (1.08)	0.00002 (0.00)	-0.012 (1.09)*
Durée	0.0004 (1.87)***	0.0002 (1.06)	0.0005 (1.93)*	0.0002 (1.05)
Montant de remboursement			0.017 (1.02)	0.004 (0.24)
Échéance de remboursement		0.0002 (2.87)**		0.0002 (2.77)**
Ancienneté		-0.0001 (3.02)**	-0.0001 (3.29)***	-0.0001 (3.01)**
Sexe		0.01 (1.27)	0.01 (1.21)	0.009 (1.24)
Age1		Ref	Ref	Ref
Age2		0.022 (2.16)**	0.023 (2.27)**	0.022 (2.14)**
Age3		0.018 (1.75)*	0.020 (1.87)*	0.018 (1.71)*
N	1779	1779	1779	1779
Pseudo R2	0.1105	0.1568	0.1424	0.1569
Année. E.F	OUI	OUI	OUI	OUI
Mois. E.F	OUI	OUI	OUI	OUI

Note : Ce tableau présente l'effet marginal des variables explicatives estimées par une régression Probit P ( $\text{Défaut}_{it}|X_{it}$ ) utilisant de nouveaux prêts à terme ouverts entre février 2012 et décembre 2016. Le z statistique est présenté entre parenthèses et l'erreur standard est corrigée pour l'hétéroscédasticité au sein du compte. Toutes les variables sont contrôlées pour le jour d'ouverture du prêt à terme. Pour la cote de risque, 1 représente le groupe ayant la meilleure cote et 3 représente le groupe ayant la plus mauvaise cote. Le collatéral vaut 1 s'il y a un collatéral sur le prêt et 0 sinon, le montant de remboursement est sur une base mensuelle, l'échéance est représentée par le nombre de mois restant sur le prêt, le sexe est égal à 1 pour un homme et 0 pour une femme. La variable âge est représentée par le groupe age1 (entre 16 et 29 ans), age2 (entre 30 et 44 ans), age3 (plus que 45ans). Le seuil statistique est\*\*\* 1%, \*\* 5%, et \* 10%.



La durée est l'un des déterminants les plus importants de la probabilité de défaut. Cette variable est significative pour le modèle (1) et pour le modèle (3), mais elle ne l'est pas pour les modèles (2) et (4) où nous ajoutons la variable échéance de remboursement qui pourrait être complémentaire. La différence c'est que les emprunteurs peuvent négocier la date d'échéance de leurs prêts et même la renégocier au court du contrat, mais la durée est une mesure d'exposition au risque des emprunteurs. Un prêt ouvert pour trois mois a une probabilité de défaut plus petite, qu'un prêt pour une durée de 12 mois. Il est à noter que nous avons certains prêts dont la date d'échéance est la dernière date d'observation (91 prêts). Autrement dit, une durée plus élevée conduit à une augmentation de probabilité de défaut sur le nouveau prêt à terme. L'effet marginal de cette variable indique que, pour les prêts ayant un mois de plus, la probabilité de défaut augmente de 0.04 point de pourcentage. L'interprétation économique est qu'un prêt avec une exposition au risque plus grande a une probabilité de défaut plus élevée. Ces résultats confirment notre intuition que plus le prêt a une durée élevée plus la probabilité de défaut est grande.

Le fait d'avoir un collatéral ou non est particulièrement intéressant dans cette équation. Les résultats montrent clairement que les emprunteurs avec un collatéral ont une probabilité de défaut plus faible sur leur prêt à terme. Le signe négatif de cette variable implique que les personnes ayant un collatéral ont une probabilité de défaut plus faible d'environ 1.2% qu'un prêt sans collatéral, tout en contrôlant pour les autres variables. Cette variable est significative pour le modèle (4), et elle ne l'est pas pour les modèles (1), (2) et (3). Nous notons cependant que cette variable n'est pas très significative (seuil de 10%) dans le modèle (4) seulement.

Nous pensons que plus le montant de remboursement du prêt est élevé plus la probabilité de défaut sera élevée. Une augmentation de 1\$ entraîne une augmentation de 0.4 point de pourcentage dans la probabilité de défaut. Cette variable n'est pas significative dans le modèle (3) et (4).

L'échéance de remboursement est aussi une variable importante dans notre analyse. Elle représente le nombre de mois au cours desquels les prêts à terme seront payés. Intuitivement, plus l'échéance est longue, plus le prêt a une probabilité de défaut élevée.

Nos modèles (2) et (4) le confirment; un mois de plus en échéance entraîne une augmentation de la probabilité de défaut d'environ 0.02 point de pourcentage.

Nous avons aussi la variable ancienneté qui est fortement significative dans notre modèle. La variable ancienneté est un *proxy* de la relation entre l'institution et l'emprunteur. Un emprunteur avec une ancienneté élevée implique qu'il est fidèle à l'institution et qu'elle peut lui faire confiance étant donné qu'ils font affaire ensemble depuis une certaine période. Ainsi, nous devons avoir un signe négatif pour cette variable. Effectivement, nos résultats suggèrent qu'un emprunteur plus ancien d'un mois a une probabilité de défaut moins élevée de 0.01 point de pourcentage. Ce qui confirme notre description pour cette variable vu qu'elle est fortement significative.

La variable sexe n'est pas significative dans notre analyse.

Quant à l'âge de l'emprunteur, les résultats montrent que, par rapport à la première catégorie omise (les plus jeunes âgés de 16 à 29 ans), ce sont les emprunteurs âgés de 30 à 44 ans qui ont une probabilité de défaut plus élevée. En outre, l'effet sur la probabilité de défaut pour la catégorie 2 (30 à 44 ans) est plus probable d'environ 2.2 point de pourcentage que celui de la catégorie 1 (16 à 29 ans), ce qui souligne le fait que ces emprunteurs sont plus risqués. Ceux de la catégorie 3 ont une probabilité de défaut supérieure d'environ 1.8 point de pourcentage que les emprunteurs dans la catégorie 1. Ainsi les individus âgés de 29 et 44 ans sont les plus risqués, cela est possible vu qu'ils ont plus de charges.

## **VI- VUE GÉNÉRALE SUR LE TAUX D'UTILISATION ET LES CYCLES DE TRANSITION :**

D'Astous et Shore (2017) examinent l'effet d'une augmentation des paiements minimums requis pour demeurer à jour sur les cartes de crédit pour une banque nord-américaine. Ils montrent que les emprunteurs les plus touchés répondent aux changements de politique en augmentant leurs paiements. Ils fournissent aussi des résultats sur les comportements prospectifs, les emprunteurs les plus susceptibles d'être touchés par les paiements minimums à l'avenir réagissent en réduisant leurs dépenses et leurs soldes renouvelables. Les auteurs soutiennent l'idée que ce phénomène pourrait ainsi expliquer la croissance de l'endettement des ménages et leurs attitudes

envers le crédit. Leurs résultats montrent que les ménages les plus touchés sont ceux qui s'exposent à une situation financière vulnérable. Ce sont les emprunteurs délinquants qui font les paiements minimums et qui sont les plus susceptibles à ne pas parvenir à confronter un changement dans leurs paiements.

Il sera intéressant, alors pour nous, de comprendre aussi l'effet de l'utilisation d'un produit renouvelable (ligne de crédit) sur la probabilité de défaut d'un nouveau prêt à terme détenu par un individu

**a- Taux d'utilisation et statut de défaut :**

Le panel A du tableau 5 présente la proportion des individus ayant 0% ou 100% de l'utilisation de leurs différents produits financiers (autres que le nouveau prêt), ainsi que la proportion des observations entre les deux extrêmes. Nous avons environ 5.95% des prêts à terme ayant un autre prêt à terme, mais qui n'a pas été utilisé et seulement 0.52% des prêts ayant des lignes de crédit non utilisées le jour d'ouverture, mais environ 82% ont une utilisation de 100%.

Le panel B du tableau 5 montre les individus ayant des valeurs extrêmes de l'utilisation de chaque type de produits en proportion du nombre total d'emprunteurs dans chaque statut de prêt (défaut et non-défaut). Il y a 39 prêts qui sont en défaut et qui ont un taux d'utilisation de 100% sur leur ligne de crédit le jour de la signature du contrat. Ce qui justifie la dépendance entre le taux d'utilisation et le défaut sur le prêt à terme (Bergerès et al. (2015)).

**Tableau 5 : taux d'utilisation des différents produits financiers**

	Prêt		hypot		Carte		ligne	
<b>Panel A : taux d'utilisation</b>	N	%	N	%	N	%	N	%
0%	32	5.95	11	3.30	42	8.72	8	0.52
0%<utilisation<100%	452	84.01	310	93.09	439	91.28	261	14.15
100%	54	10.04	12	3.61	-	-	1253	82.33
total	538	100	333	100	481	100	1522	100

<b>Panel B : utilisation par statut du prêt</b>								
Statut	0%	100%	0%	100%	0%	100%	0%	100%
Défaut	2	2	-	-	3	-	-	39
Non-défaut	30	52	11	12	39	-	8	1214
total	32	54	11	12	42	-	8	1253

Note : Le panel A présente le nombre de nouveaux prêts à terme ayant un taux d'utilisation 0%, entre 0% et 100%, et 100%, et leurs proportions pour chaque type de produit financier. Le panel B présente le nombre et les proportions de nouveaux prêts à terme par statut (défaut et non-défaut) ayant un taux d'utilisation 0% ou 100% pour chaque type de produit financier.

### **b- Les cycles de délinquance :**

La variable de cycles de délinquance indique le nombre jour en retard du prêt à terme. Les prêts peuvent être à jour, en retard ou en défaut. Les prêts de 90 jours de retard ou plus sont considérés comme étant en défaut de paiement. Il est à noter qu'un prêt pourrait rester un mois délinquant pour toujours, si l'emprunteur a toujours effectué le paiement minimum sans payer le montant en souffrance sur le prêt. Ainsi, le prêt pourrait transiter d'un cycle à un autre chaque mois selon le paiement de l'emprunteur.

Pour la définition des cycles de délinquance nous nous référons au travail de d'Astous et Shore (2017). Les auteurs présentent trois états de transition  $\Omega = \{C, L, W\}$ , où (C) définit comme à jour, (L) comme délinquant et (W) comme défaut. Dans ce qui suit nous utilisons leur modèle et nous présentons le groupe de délinquant en trois sous-groupes. Ainsi à chaque mois, un prêt peut transiter entre cinq états  $\Omega = \{C, L1, L2, L3, D\}$  où D est utilisé pour défaut en fonction d'un retard de paiement. Les états sont

définis comme (C) si le prêt est à jour, ce qui équivaut à 0 jour de retard, (L1) représente le premier cycle de délinquance, autrement dit, le nombre de jours de retard sur le prêt est entre 1 et 29 jours. (L2) est le deuxième cycle de délinquance ce qui équivaut à un nombre de jours de retard entre 30 et 59 jours. (L3) est le dernier cycle de délinquance représenté par un nombre de jours de retard entre 60 et 89 jours. Une fois qu'un prêt indique un nombre de jours en retard supérieur à 90, il est considéré en défaut et est représenté par le cycle (D).

La matrice de transition conditionnelle pour notre échantillon est représentée dans le tableau 6. L'emprunteur peut effectuer un paiement sans acquitter le solde complètement, pourvu qu'il puisse empêcher le passage en défaut. 97.4% des prêts à jour restent à jour dans le prochain mois. Environ 2.4% des prêts passent au premier cycle de délinquance le mois suivant. La proportion des prêts qui sont en deuxième cycle de délinquance est si faible qu'elle est nulle lorsqu'elle est arrondie à la deuxième décimale. Les prêts qui sont en troisième cycle de délinquance ainsi que le cycle de défaut est nul le mois suivant. Environ 61% des prêts qui sont considérés en retard dans le premier cycle de délinquance reviennent à jour le mois suivant, 26.1 % restent délinquants au premier cycle, 11.7% passent au cycle (L2), 0.8% se retrouvent aux (L3), mais la proportion des prêts qui sont en défaut de paiement est nulle le mois suivant. Environ 32.7% des prêts considérés en retard entre 30 et 59 jours reviennent à 0 jour de retard le mois suivant, 14.4% reviennent au premier cycle de délinquance, 12.9% restent au même cycle, 38.9% passent au cycle (L3) le mois suivant, et environ 0.9% des prêts sont considérés en défaut de paiement le prochain mois. Environ 12.6% des prêts qui sont dans le cycle (L3) y restent le mois suivant, 3.8% reviennent au cycle (L2), 26.2% paient tout le solde qui leur est dû, 53.4% passent au statut de défaut. Plus de la moitié des prêts qui se retrouvent en défaut de paiement y restent le mois suivant et 25% des prêts acquittent leurs soldes et reviennent à jour le mois suivant, de très faibles proportions passent aux différents statuts de délinquances.

**Tableau 6 : Matrice de transition**

	$C_t$	$L1_t$	$L2_t$	$L3_t$	$D_t$
$C_{t+1}$	0.974	0.613	0.327	0.262	0.251
$L1_{t+1}$	0.024	0.261	0.144	0.038	0.063
$L2_{t+1}$	0.0009	0.117	0.129	0.038	0
$L3_{t+1}$	0	0.008	0.389	0.126	0.021
$D_{t+1}$	0	0	0.009	0.534	0.664

Note : Ce tableau montre les probabilités de transition entre les différents états  $w=\{C, L1, L2, L3, D\}$ , pour les nouveaux prêts à terme ayant au moins un produit financier. Les états sont définis comme (C) si le prêt à terme est à jour (L1) si le prêt à terme note un retard entre 1 et 29 jours, (L2) si le prêt à terme note un retard entre 30 et 59 jours, (L3) si le prêt à terme note un retard entre 60 et 89 jours, et (D) si le prêt à terme note un retard  $\geq 90$  jours et ainsi considéré en défaut.

### **c-Le taux d'utilisation avant et après la réalisation d'un événement :**

D'Astous et Shore (2017) comparent les transitions entre les différents cycles de délinquance avant et après l'augmentation du paiement minimum sur les cartes de crédit. En se référant à leurs comparaisons, nous analysons les taux d'utilisation moyens des lignes de crédit (vu que c'est le seul produit renouvelable dans notre base) avant et après la réalisation de l'un des événements suivants: manquer un paiement ou effectuer un paiement.

Nous voulons voir où la variation du taux d'utilisation sur une fenêtre de 12 mois autour de la réalisation de l'événement représenté par « 0 ». La ligne horizontale représente le mois au cours duquel il y a eu un événement. La ligne verticale représente le taux d'utilisation moyen entre 0 et 100%. Le panel (a) montre le changement des taux d'utilisation moyens sur les lignes de crédit six mois avant et six mois après l'événement de manquer un paiement sur le nouveau prêt à terme à un moment de la durée du contrat représenté par 0. Nous pouvons bien voir une baisse du taux d'utilisation moyen en nous rapprochant de l'événement pour atteindre une valeur d'environ 50%. Le panel (b) représente le changement des taux d'utilisation moyens avant et après l'événement d'effectuer un paiement. Nous voyons bien que le taux augmente juste avant le jour d'événement (effectuer un paiement '0') d'environ 20% et une fois que le paiement a été fait, nous remarquons une baisse du taux. Nous interprétons cela comme une suggestion

que les emprunteurs peuvent utiliser la liquidité fournie par leur ligne de crédit pour réduire la probabilité de défaut sur leur prêt à terme et que la détresse financière, mesurée par le fait de manquer un paiement sur le prêt à terme, conduit à une utilisation plus élevée des lignes de crédit.

La figure 1 montre qu'il y a une relation entre le taux d'utilisation des lignes de crédit et le fait d'effectuer ou de manquer un paiement sur le prêt à terme. Cela prouve que les changements des taux d'utilisation des lignes de crédit sont un signal sur les probabilités de défaut des nouveaux prêts à terme pour les prochaines périodes de paiements. Ainsi, les banques devraient gérer soigneusement les interactions entre les instruments de crédit d'un emprunteur pour détecter un tel comportement et prendre des mesures préventives pour atténuer les pertes potentielles.

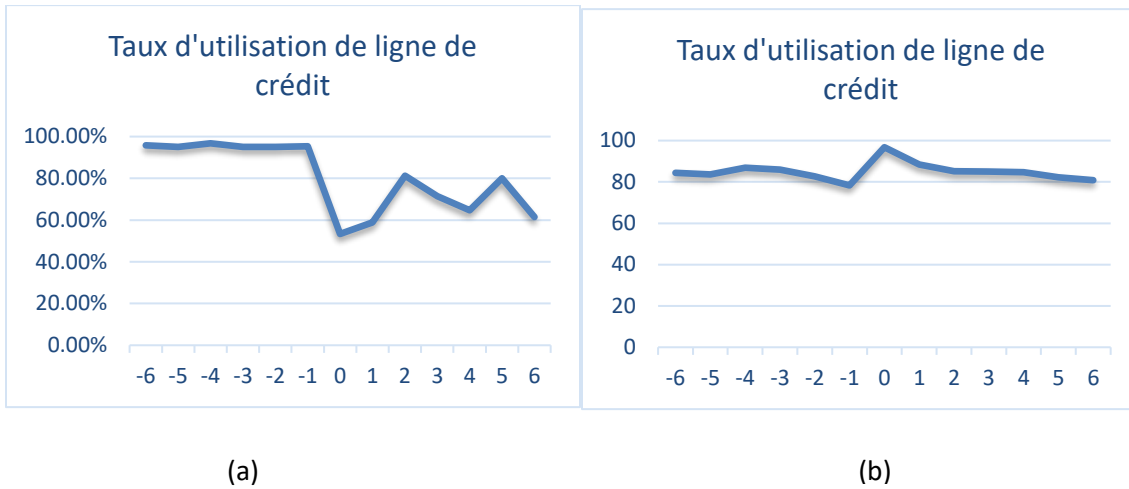
Ces premiers résultats nous motivent à voir cette relation de plus près, et pourraient permettre d'étudier la présence d'asymétrie d'information entre la banque et ses clients dans des recherches futures.

## **VII-MÉTHODOLOGIE ET RÉSULTATS :**

L'objectif de cette section est de présenter un modèle économétrique pour estimer les déterminants à la fois de la probabilité de défaut et le coût espéré conditionnel en se référant à l'article de Dionne et al (1996).

Étant donné la connaissance de certaines caractéristiques des clients à la date de signature de contrat, nous pouvons prédire les probabilités de défaut à la date de signature en estimant les paramètres d'un modèle Probit qui admet deux événements possibles, défaut ou non-défaut. Nous associons ici le défaut à un retard supérieur à 90 jours.

**Figure 1 : Les taux d'utilisation de ligne de crédit pour chaque événement**



Note : Cette figure trace le changement des taux d'utilisation pour les lignes de crédit 6 mois avant et 6 mois après l'événement de manquer ou faire un paiement sur le prêt à terme correspondant, représenté par  $t=0$ . Panel (a) représente le taux d'utilisation avant et après l'événement de manquer un paiement, panel (b) représente le taux d'utilisation avant et après l'événement de faire un paiement.

Avec ces mêmes caractéristiques nous pouvons estimer aussi les distributions conditionnelles des non-paiements en utilisant des distributions de comptage (modèle Poisson). Voici un tableau qui montre la fréquence du nombre de non-paiements :

**Tableau 7 : Fréquence de non-paiement**

Nombre	Fréquence	Pourcentage	Cumulatif
0	1319	74.27	74.27
1	331	18.61	92.75
2	49	2.75	95.50
3	27	1.52	97.02
4	26	1.46	98.48
5	11	0.62	99.10
6	5	0.28	99.38
7	3	0.17	99.55
8+	8	0.45	100.00
Total	1779	100	

La description des principales variables utilisées est donnée au tableau 1. Les signes des effets prédits de ces variables sur la probabilité de défaut sont également présentés à la section « variables ».

Nous voulons isoler les caractéristiques significatives influençant la probabilité des emprunteurs qui pourraient être considérés comme en défaut avant la fin de leurs



contrats. Nous souhaitons prendre en compte le comportement des clients en examinant le nombre attendu de non-paiements et, par conséquent, à leurs coûts conditionnels attendus.

Les modèles présentés ci-dessous présentent le fait qu'il pourrait y avoir deux processus pour les prêts : un pour les prêts qui ne sont pas considérés comme étant en défaut, et un autre pour les prêts qui ont manqué suffisamment de paiement pour être considérés en défaut. Dionne et al (1996) les considèrent comme étant deux modèles différents, en expliquant que la possibilité de défaut postérieure mérite d'être considérée.

Les modèles de Poisson sont fréquemment utilisés pour estimer des résultats entiers non négatifs ayant des restrictions d'égalité de la moyenne et de la variance (conditionnelle à un ensemble de variables explicatives) du modèle à estimer. Le modèle de Poisson établit que  $Y_i$ , le nombre de non-paiements du prêt  $i$ , la distribution du Poisson est :

$$\text{Prob}(Y_i = j) = \frac{e^{-\lambda_i} \lambda_i^j}{j!} \quad \begin{matrix} j = 0, 1, 2, \dots \\ \lambda_i > 0 \end{matrix} ,$$

où  $\lambda$  est le paramètre de moyenne et de variance qui peut différer pour chaque prêt,  $j$  est le nombre de non-paiements.

Le tableau 8 présente les principaux résultats économétriques de l'estimation de la probabilité de défaut et des paramètres de la distribution des non-paiements de ceux qui n'ont pas fait défaut (Colonnes 2 et 3) et la même information est présentée dans les colonnes 4 et 5 du même tableau, mais pour ceux ayant fait défaut. Ce qui est intéressant de mentionner que les variables significatives ne sont pas les mêmes pour expliquer la probabilité de défaut et les distributions conditionnelles des non-paiements. Pour chaque groupe, la première colonne représente (Probit) les résultats de l'estimation de la probabilité de défaut et la seconde donne ceux de l'estimation des non-paiements (0, 1, 2, 3) pour les non-défauts et (4, 5, 6, 7, 8) pour les défauts. La distribution de Poisson n'est pas rejetée pour les non-paiements. Les tests effectués sont non significatifs pour confirmer que le modèle Poisson est inapproprié. « Pearson statistic » et « Deviation statistic » montrent que nous ne rejetons pas l'hypothèse nulle : Le modèle Poisson est approprié.

#### **a-Pour le groupe de non-défaut :**

La variable durée (mesure d'exposition au risque) est positive et non significative pour expliquer la probabilité de défaut, alors qu'elle l'est pour expliquer les non-paiements. Ce qui implique qu'une durée plus élevée n'a pas d'effet sur la probabilité de défaut du prêt, mais pourrait influencer positivement le nombre des non-paiements. Un prêt ayant une cote de risque de catégorie 2 a une probabilité de défaut plus importante comparée à un prêt avec une cote de risque de catégorie 1. La cote de risque est moins significative pour l'estimation de la distribution des non-paiements. La variable montant de remboursement est positive et non significative pour expliquer la probabilité de défaut, mais elle l'est pour expliquer les non-paiements. La variable de catégorie 3 de la cote de risque est positive et significative dans les deux régressions et le signe indique que ce groupe de prêt est plus risqué que le groupe ayant une cote de risque de catégorie 1. La variable échéance de remboursement est positive et significative pour expliquer la probabilité de défaut, mais elle ne l'est pas pour expliquer les non-paiements. Avoir une échéance plus longue peut affecter positivement la probabilité de défaut du prêt, mais en aucun cas elle n'aura d'effet sur les non-paiements. La variable sexe n'influence ni la probabilité de défaut ni la distribution des non-paiements à un degré de confiance de 95%. La variable ancienneté a un signe négatif, indiquant que le prêt le plus ancien est moins risqué, même si le degré de signification est plus faible pour estimer le nombre de non-paiements par rapport à estimer celui de la probabilité de défaut. La variable collatéral a un signe négatif ce qui explique qu'un prêt avec collatéral est moins risqué que celui qui n'a pas de collatéral, mais cette variable n'a pas d'effet sur la distribution de non-paiement.

#### **b-Pour le groupe de défaut :**

Certains changements pour le groupe de défaut. L'âge et le taux d'intérêt ont un effet sur la probabilité de défaut, mais une fois en défaut ces variables n'ont plus d'effet sur le nombre de non-paiements. Le collatéral et la durée n'ont aucun effet sur le risque des prêts. La variable du montant de remboursement affecte positivement et significativement l'estimation des non-paiements. L'ancienneté est significative

seulement pour l'estimation des probabilités de défaut, mais elle ne l'est pas pour l'estimation de non-paiements. La cote de risque perd aussi son effet sur le nombre de non-paiements pour les prêts en défaut. L'échéance de remboursement affecte positivement et significativement la probabilité de défaut, mais elle est négative et moins significative pour l'estimation de la distribution des non-paiements (l'échéance n'a plus d'effet sur les non-paiements une fois que le prêt est en défaut).

### **c-La matrice de confusion**

La capacité d'un modèle à prédire les défauts et les non-défauts est donnée par la matrice de confusion. Les résultats de la classification des défauts et non défauts déterminés au moyen d'un seuil préétabli (ici le seuil est la fréquence moyenne des défauts observés), sont comparés aux défauts et non défauts réalisés. Fondamentalement, il s'agit d'un tableau qui fournit le type de classification et le nombre associé à chaque type. Ces types sont définis en termes de bonnes classifications et de mauvaises classifications. Les classifications sont dites bonnes lorsqu'un défaut prédit correspond à un défaut actuel (sensibilité du modèle) ou un non-défaut prédit correspond à un non-défaut actuel (spécificité du modèle). Les classifications de type mauvais sont celles qui classent les défauts prédits comme vrais alors qu'ils correspondent à des non-défauts actuels. Cette mauvaise classification représente une erreur de type II. Ce type d'erreur peut pousser la banque à revoir son niveau de capital économique pour se prémunir d'éventuels chocs. Ainsi, cela peut se traduire par un manque à gagner et par ailleurs peut affecter la qualité de crédit du client. Inversement, une mauvaise classification qui classe les non-défauts prédits comme vrais alors qu'ils correspondent à des défauts actuels représente une erreur de type I. Les conséquences d'une telle classification sont encore plus importantes puisqu'elles se matérialisent en termes de capital et d'intérêt. Ces classifications sont très sensibles au degré de conservatisme du modèle, soit le seuil que l'on choisit. Ainsi, plus ce seuil est restrictif, moins nous avons d'erreurs de type I mais plus d'erreurs de type II et inversement. La prochaine section nous permettra de donner la mesure de cette sensibilité à travers la ROC Curve. De façon similaire, nous pouvons présenter une définition des mauvaises classifications comme suit : Si nous supposons une hypothèse nulle définie comme telle :  $H_0 : Y_i = 1$  (*défaut*). Les mauvaises

classifications seront alors traduites ainsi : - L'erreur de type I correspond au rejet de  $H_0$  quand  $H_0$  est vraie - L'erreur de type II correspond à ne pas rejeter  $H_0$  quand  $H_0$  est fausse. Le tableau 9 présente une matrice de confusion générale.

Le tableau 9 analyse les erreurs de classification de Type 1 et de Type 2 du modèle d'estimation des probabilités de défaut en utilisant la fréquence moyenne des défauts observés (2.97%). Il est à noter que d'autres seuils peuvent être utilisés, ce choix revient aux gestionnaires de risque. Les 53 mauvais risques (11+42) sont ceux ayant déjà fait défaut et les 1726 bons risques (1252+585) sont ceux qui n'ont pas fait défaut.

Le modèle prédit 1252 bons risques, qui sont effectivement des bons risques. Il prédit également 42 mauvais risques, qui sont effectivement des mauvais risques. Mais il prédit 485 prêts comme étant des bons risques, alors qu'ils se révèlent être des mauvais risques (erreur Type 2). Cette erreur peut être onéreuse, car elle implique que la banque demandera des taux d'intérêt moins élevés à des mauvais risques. De plus, le modèle prédit 11 prêts comme étant des mauvais risques alors qu'ils sont des bons risques, ceci pourrait être très coûteux, car ceux-ci peuvent quitter la banque. Cette relation d'arbitrage entre les erreurs de Type 1 et Type 2 peut être gérée plus en détail, en considérant les différents bénéfices et coûts qui y sont associés.

**Tableau 9 : Matrice de confusion avec un seuil de 2.97%**

	Défaut actuel	Nn-défaut actuel
Défaut prédit	Vrai positif 42	Erreur de type II 485
Nn-défaut prédit	Erreur de type I 11	Vrai négatif 1252

**Tableau 8 : Probabilité de défaut (Probit) et distribution conditionnelle de non-paiement (Poisson) pour les groupes de défaut et non-défaut**

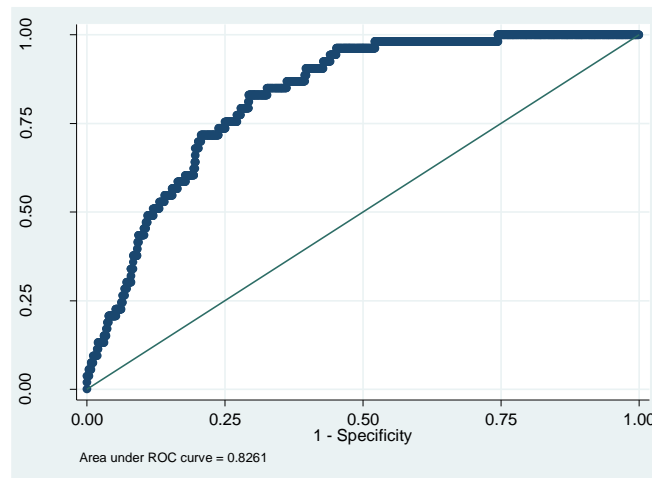
Variable	Non-défaut		Défaut	
	Probit	Poisson	Probit	Poisson
Constante		-2.557 (-12.71)***		1.585 (9.28)***
Cote de risque				
1	Ref		Ref	
2	0.020 (1.92)*	0.877 (6.66)***	0.020 (1.92)*	-0.0737 (-0.66)
3	0.054 (4.76)***	1.563 (12.30)***	0.054 (4.76)***	-0.1671 (-1.60)*
Taux d'intérêt	0.002 (1.72)*	0.029 (2.35)**	0.002 (1.72)*	0.0023 (0.24)
Collatérale				
0	Ref		Ref	
1	-0.012 (1.09)*	-0.063 (-0.45)	-0.012 (1.09)*	0.170 (1.53)
Durée	0.0002 (1.05)	0.014 (4.26)***	0.0002 (1.05)	0.001 (0.46)
Montant de remboursement	0.004 (0.24)	-0.665 (-1.75)*	0.004 (0.24)	0.538 (2.09)**
Échéance de remboursement	0.0002 (2.77)**	0.002 (1.64)	0.0002 (2.77)**	-0.0009 (-1.61)*
Ancienneté	-0.0001 (3.01)**	-0.0007 (-1.58)*	-0.0001 (3.01)**	-0.0002 (-0.39)
Sexe	0.009 (1.24)	0.129 (1.41)	0.009 (1.24)	0.050 (0.64)
Age1	Ref		Ref	
Age2	0.022 (2.14)**	0.160 (1.36)	0.022 (2.14)**	-0.063 (0.76)
Age3	0.018 (1.71)*	0.135 (1.00)	0.018 (1.71)*	-0.075 (-0.69)
N	1779	1726	1779	53
N. Paramètres	11	12	11	12
Pseudo R2	0.1569	0.1094	0.1569	0.0281

Note : Ce tableau présente l'effet marginal des variables explicatives estimées par une régression Probit P (Défaut<sub>it</sub>|X<sub>it</sub>) utilisant de nouveaux prêts à terme ouverts entre février 2012 et décembre 2016, ainsi que la distribution conditionnelle des non-paiements estimée par un modèle de Poisson. Les deux premières colonnes représentent le groupe de non-défaut et les deux dernières celle du défaut. Le z statistique est présenté entre parenthèses et l'erreur standard est corrigée pour l'hétéroscédasticité au sein du compte. Toutes les variables sont contrôlées pour le jour d'ouverture du prêt à terme. Pour la cote de risque, 1 représente le groupe ayant la meilleure cote et 3 représente le groupe ayant la plus mauvaise cote. Le collatéral vaut 1 s'il y a un collatéral sur le prêt et 0 sinon. Le montant de remboursement est sur une base mensuelle, l'échéance est représentée par le nombre de mois restant sur le prêt. Le sexe est égal à 1 pour un homme et 0 pour une femme. La variable âge est représentée par le groupe age1 (entre 16 et 29 ans), age2 (entre 30 et 44 ans), age3 (plus que 45ans). Le seuil statistique est\*\*\* 1%, \*\* 5%, et \* 10%

#### d- La ROC curve :

La ROC curve (Receiver Operating Characteristic curve) correspond à une mesure de la performance d'un classificateur binaire comme c'est le cas avec nos modèles Probit. C'est une courbe qui permet de présenter une synthèse sur l'information résultant de différentes matrices de confusions associées à des seuils de conservatisme correspondants. La mesure de spécificité du modèle, soit le pourcentage d'erreur de type II, est représenté par son axe des abscisses. La mesure de sensibilité est reportée sur son axe des ordonnées, soit le pourcentage de défauts correctement classés comme des défauts par le modèle. La courbe est alors obtenue en faisant varier le seuil de conservatisme du modèle qui permet de générer différentes combinaisons de classification de défaut et d'erreurs de type II. Ces différentes combinaisons permettent ainsi de déterminer dans quelle mesure le modèle est précis pour classer les défauts et les non-défauts. L'aire sous la ROC curve du nom de mesure AUC (area under curve) est une mesure de la performance du modèle dans la prédiction du défaut actuel. Un modèle parfait aura une mesure AUC de 1. Ainsi, plus la courbure de la ROC curve est proche du coin gauche du graphique vers le haut et la mesure AUC est proche de 1, plus le modèle est précis.

**Graphique 2 : ROC curve**



La ROC curve tient essentiellement son avantage de sa capacité à tenir compte des erreurs de type II ainsi que de la surface sous la courbe qui représente une excellente

mesure de performance d'un modèle. La valeur de l'AUC de notre modèle est de 0.8261. Elle est proche de 1, ce qui nous permet de confirmer que le modèle est assez précis pour prédire la probabilité de défaut.

### **VIII-TEST DE ROBUSTESSE**

Nous présentons des résultats économétriques additionnels obtenus de la même base de données. Cette analyse estime un modèle de Logit ordonné des non-paiements des nouveaux prêts à terme le jour d'ouverture en fonction des variables  $X_i$  (caractéristiques le jour d'ouverture du prêt) déjà expliquées dans les sections précédentes. Ainsi l'étude porte sur 1779 nouveaux prêts divisés en deux groupes : le groupe non-défaut  $y_i = (0, 1, 2 \text{ ou } 3)$  qui contient 1726 prêts, et le groupe défaut  $y_i = (4, 5, 6, 7 \text{ ou } 8)$  comportant 53 prêts.

Nous remarquons qu'une cote de risque plus élevée (une cote moins bonne), un taux d'intérêt plus élevé, la durée (le nombre de mois à compter du début du contrat jusqu'à la date de l'échantillonnage), augmentent le nombre de non-paiements pour le groupe des non-défauts. Ces variables sont très significatives pour les non-paiements de 0 à 3, mais ils n'ont plus d'effet une fois que le prêt est en défaut. Ces résultats confirment ce que nous avons obtenu. À l'inverse, le montant de remboursement affecte négativement le nombre de paiements pour les deux groupes même s'il est moins significatif pour le groupe de défaut. Ainsi que l'ancienneté du client avec sa banque affecte négativement le nombre de non-paiements. Elle est significative seulement pour le groupe de non-défaut. Les variables, collatéral, terme de remboursement et âge, n'ont aucun effet sur le nombre de non-paiements. Pas de constante dans le modèle, car l'effet est absorbé par les points seuils (permet de calculer la probabilité que  $y_i=j$ ).

Cette analyse confirme les résultats déjà obtenus sur l'effet de la plupart des variables. Nous avons montré que plusieurs caractéristiques facilement observables des consommateurs peuvent affecter leur probabilité de défaut et leur distribution de non-paiement des mensualités.

## **IX-RECHERCHE FUTURE :**

Dans des recherches futures, nous pouvons analyser la présence d'asymétrie d'information entre la banque et ses clients. Pour ce faire, nous pouvons modéliser conjointement le taux d'utilisation des marges de crédit et la probabilité de défaut des prêts à terme. Le taux d'utilisation doit être analysé de façon endogène à la probabilité de défaut sur le prêt à terme doit donc être instrumentée. Le comportement de tirage sur la marge de crédit et la probabilité de défaut nous permettra de tester la présence d'une asymétrie d'information résiduelle conditionnelle entre la banque et son client. Le test nous donne la possibilité de vérifier si les emprunteurs profitent de l'information privilégiée à propos des caractéristiques de risque qu'ils possèdent au moment de l'ouverture du prêt. Le cas échéant, les résultats suggèreraient donc que l'asymétrie d'information a une influence sur la probabilité de défaut du nouveau prêt à terme. De tels résultats impliqueraient que l'institution financière n'a pas une gestion optimale de son risque de crédit et qu'elle devrait donc explicitement analyser les différents produits financiers détenus par un même individu de façon conjointe.

Afin de tester empiriquement la présence de l'asymétrie d'information entre les différents produits détenus par un même individu, nous pouvons analyser son choix d'utilisation de ses différents produits. En d'autres termes, nous pouvons vérifier s'il choisit d'utiliser plus ou moins ses différents produits, quand des probabilités de défaut sont plus élevées sur son nouveau prêt à terme. Le modèle proposé est inspiré des travaux de Chioporri et Salanié (2000) et Dionne et al. (1997-2001) dans le marché de l'assurance automobile.



**Tableau10 : test de robustesse avec un modèle de Logit ordonné pour la distribution des non-paiements**

Variable	Robustesse	Logit ordonné	
		Défaut	Non-défaut
<b>Cote de risque</b>			
1		Ref	Ref
2		-0.569 (-0.61)	0.951 (6.34)***
3		-1.075 (-1.13)	2.04 (12.27)***
Taux d'intérêt		0.0005 (0.01)	0.033 (1.89)**
<b>Collatérale</b>			
0		Ref	Ref
1		1.366 (1.28)	-0.097 (-0.53)
Durée		0.006 (0.25)	0.023 (5.53)***
Montant de remboursement		4.693 (1.55)*	-1.166 (-2.24)**
Échéance de remboursement		-0.001 (-1.75)*	0.001 (0.77)
Ancienneté		-0.0028 (-0.54)	-0.0007 (-1.22)
Sexe		0.554 (0.82)	0.156 (1.27)
Age1		Ref	Ref
Age2		-0.146 (-0.19)	0.164 (0.99)
Age3		0.035 (0.03)	0.102 (0.55)
<b>Seuil</b>			
Seuil 0			-
Seuil 1			2.658(0.245)
Seuil 2			4.765(0.262)
Seuil 3			5.862(0.314)
Seuil 4			-
Seuil 5		-0.513(1.791)	
Seuil 6		0.554(1.811)	
Seuil 7		1.188(1.796)	
Seuil 8		1.691(1.797)	
N		53	1726
N. Paramètres		11	11
Pseudo R2		0.0992	0.1069

Note : Ce tableau présente l'estimation de la distribution des non-paiements des nouveaux prêts à terme ouverts entre février 2012 et décembre 2016, par un modèle de Logit ordonné. La variable dépendante est le nombre de non-paiements (0, 1, 2, 3, 4, 5, 6, 7, et 8). Le z statistique est présenté entre parenthèses, pour les points seuils c'est l'erreur standard qui est entre parenthèses. Toutes les variables sont contrôlées pour le jour d'ouverture du prêt à terme. Pour la cote de risque, 1 représente le groupe ayant la meilleure cote et 3 représente le groupe ayant la plus mauvaise cote. Le collatéral vaut 1 s'il y a un collatéral sur le prêt et 0 sinon. Le montant de remboursement est sur une base mensuelle. L'échéance est représentée par le nombre de mois restants sur le prêt. Le sexe est égal à 1 pour un homme et 0 pour une femme. La variable âge est représentée par le groupe age1 (entre 16 et 29 ans), age2 (entre 30 et 44 ans), age3 (plus que 45ans). Le seuil statistique est\*\*\* 1%, \*\* 5%, et \* 10%.

Il sera important de séparer les comportements dynamiques et les caractéristiques du client pour comprendre si l'asymétrie d'information provient d'une sélection adverse ou d'un risque moral. Chioppori et Salenié (2013) mentionnent que le risque moral et la sélection adverse ont des implications empiriques semblables, mais avec une causalité inversée. En nous basant sur leurs idées, nous pouvons dire qu'en cas de sélection adverse, les emprunteurs sont caractérisés par différents niveaux de risque ex ante, qui se traduisent par des risques ex post différents (taux de défaillance), et éventuellement, conscients de ces différences de risque, ces emprunteurs utilisent leurs produits financiers et surtout les lignes de crédit, car c'est le seul produit renouvelable de manière différente (le taux d'utilisation des lignes de crédit varie d'un emprunteur à un autre). Dans un contexte de risque moral, les emprunteurs sont responsables de leurs différentes notations de crédit; ils sont donc confrontés à différentes incitations et adoptent des comportements plus ou moins prudents, ce qui aboutit finalement à des probabilités hétérogènes de défaut. Dans les deux cas, en tenant compte des observables, le taux d'utilisation des produits financiers sera corrélé à la probabilité de défaut du nouveau prêt à terme.

Nous pouvons étendre aussi nos recherches en appliquant la méthode de l'apprentissage automatique (Machine Learning). Nous pouvons construire un système pour la détection des produits défaillants le jour de la signature du contrat basé, exclusivement, sur la capacité prédictive hors échantillon du modèle.

## **X- CONCLUSION :**

Nous avons présenté une analyse du risque de défaut de prêts à la consommation d'une institution financière. Notre échantillon comporte un total de 1779 nouveaux prêts à terme (excluant les prêts destinés aux étudiants). Nous avons dans cette base 53 nouveaux prêts à terme considérés en défaut de paiement par la banque, soit un taux de 2.97% et 407 nouveaux prêts à terme en phase de délinquance (manque des paiements) selon la définition de la banque, soit 22.87%. Premièrement, nous avons commencé par lier la probabilité de défaut du prêt à terme avec les caractéristiques de l'emprunteur. Ces dernières proviennent de caractéristiques de crédit connues par la banque au moment de l'ouverture du prêt et de l'utilisation des lignes de crédit durant la vie du prêt à terme.

Nous constatons que les emprunteurs notant un défaut sur leur nouveau prêt à terme ont une cote de risque et un taux d'intérêt plus important comparé à ceux qui n'ont pas noté un défaut. Avec une durée plus grande, les nouveaux prêts à terme sont plus susceptibles d'être en défaut. Deuxièmement, nous avons présenté une vue générale sur les transitions des prêts entre les différents cycles de délinquance. Les résultats montrent que plus que la moitié des prêts qui se retrouvent en défaut y restent le mois suivant et 25% des prêts acquittent leurs soldes et reviennent à jour. Ainsi qu'une étude d'évènements pour voir la variation du taux d'utilisation de lignes de crédit sur une fenêtre de 12 mois autour de la réalisation de l'un des deux évènements : manquer ou effectuer un paiement. Les résultats montrent qu'il y a une baisse du taux moyen avant le premier évènement, et une augmentation avant le deuxième. Ce qui suggère la présence d'une certaine relation entre le taux d'utilisation de lignes de crédit et les évènements réalisés sur les nouveaux prêts à terme. Troisièmement, nous avons estimé la probabilité de défaut ainsi que la distribution conditionnelle des non-paiements. Nous avons observé que plusieurs caractéristiques facilement observables des consommateurs peuvent affecter leur probabilité de défaut et leur distribution de non-paiement des mensualités. Aussi, ce ne sont pas nécessairement les mêmes variables qui expliquent les deux estimations. Nous avons constaté que certaines variables expliquent la probabilité de défaut, mais une fois que le prêt est en défaut ces mêmes variables n'ont plus d'effet sur le nombre de non-paiements (4 ou +). Nous avons mesuré aussi les erreurs de Type I et de Type II pour estimer les pertes de « bons » et « mauvais » prêts. La ROC curve nous a permis de confirmer la performance de notre modèle pour la prédiction de la probabilité de défaut. Avec un AUC de 0.8261, le modèle est assez précis.

Les résultats montrent que ce ne sont pas nécessairement les mêmes variables qui expliquent les deux distributions.

#### **BIBLIOGRAPHIE :**

Agarwal, S., Ambrose, B. W., & Liu, C. (2006). Credit lines and credit utilization. *Journal of Money, Credit and Banking*, 1-22.

Agarwal, S., Ambrose, B. W., Chomsisengphet, S., & Liu, C. (2006). An empirical analysis of home equity loan and line performance. *Journal of Financial Intermediation*, 15(4), 444-469.

Andersson, Fredrik, et al. "The changing pecking order of consumer defaults." *Journal of Money, Credit and Banking* 45.2-3 (2013): 251-275.

Bergerès, A. S., d'Astous, P., & Dionne, G. (2015). Is there any dependence between consumer credit line utilization and default probability on a term loan? Evidence from bank-customer data. *Journal of Empirical Finance*, 33, 276-286.

Cohen-Cole, Ethan, and Jonathan Morse. (2010) "Your House or Your Credit Card, Which Would You Choose? Personal Delinquency Tradeoffs and Precautionary Liquidity Motives." Working Paper No. QAU09-05, Quantitative Analysis Unit, Federal Reserve Bank of Boston.

d'Astous, P., & Shore, S. H. (2017). Liquidity constraints and credit card delinquency: Evidence from raising minimum payments. *Journal of Financial and Quantitative Analysis*, 52(4), 1705-1730.

Dionne, Georges, *Gestion du risque : Théorie et application*, Economica, 2017

Dionne, Georges, Manuel Artís, and Montserrat Guillén. "Count data models for a credit scoring system." *Journal of Empirical Finance* 3.3 (1996): 303-325.

Dionne, G., Gouriéroux, C., & Vanasse, C. (2001). Testing for evidence of adverse selection in the automobile insurance market: A comment. *Journal of Political Economy*, 109(2), 444-453.

Dionne, G., St-Amour, P., & Vencatachellum, D. (2009). Asymmetric information and adverse selection in mauritian slave auctions. *The Review of Economic Studies*, 76(4), 1269-1295.

Dionne, G., La Haye, M., & Bergères, A. S. (2010). Does Asymmetric Information Affect the Premium in Mergers and Acquisitions?. *Cahier de recherche/Working Paper*, 10, 15.

Elul, R., and Souleles, N. S. and Chomsisengphet, S., and Glennon, D., and Hunt, R.M., What 'Triggers' Mortgage Default? (April 1, 2010). FRB of Philadelphia Working Paper No. 10-13.

Hibbeln, Martin Thomas and Norden, Lars and Usselman, Piet and Gürtler, Marc, *Economies of Scope in Consumer Credit* (August 9, 2017). Paris December 2015 Finance Meeting EUROFIDAI - AFFI. Available at SSRN: <https://ssrn.com/abstract=2536930> or <http://dx.doi.org/10.2139/ssrn.2536930>

Jagtiani, Julapa, and William W. Lang. "Strategic defaults on first and second lien mortgages during the financial crisis." *The Journal of Fixed Income* 20.4 (2011): 7-23.

Jiménez, G., Lopez, J. A., & Saurina, J. (2009). Empirical analysis of corporate credit lines. *The Review of Financial Studies*, 22(12), 5069-5098.

Jimenez, G., Lopez, J. A., & Saurina, J. (2009). Calibrating exposure at default for corporate credit lines. *Journal of Risk Management in Financial Institutions*, 2(2), 121-129.

Mester, L. J., Nakamura, L. I., & Renault, M. (2006). Transactions accounts and loan monitoring. *The Review of Financial Studies*, 20(3), 529-556.

Norden, L., & Weber, M. (2010). Credit line usage, checking account activity, and default risk of bank borrowers. *The Review of Financial Studies*, 23(10), 3665-3699.

Strahan, Philip E. "Borrower risk and the price and nonprice terms of bank loans." (1999).