HEC MONTRÉAL

Gestion de la trésorerie pour une compagnie d'assurance

Par

Émilie Matton

Sciences de la gestion
Finance corporative

Projet supervisé en organisation : mandat d'intervention
Présenté comme exigence partielle pour l'obtention
de la maîtrise en science de la gestion (M.Sc.)

Décembre 2018
© Émilie Matton 2018

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage chez Intact Corporation Financière et qui m'ont aidée lors de la rédaction de ce projet supervisé.

J'adresse, en premier lieu, mes plus sincères remerciements à mon professeur référent, M. Georges Dionne, pour ses judicieux conseils qui ont contribué à alimenter ma réflexion et pour sa remarquable disponibilité dont il a fait preuve tout au long de son accompagnement dans la réalisation de mon projet supervisé. Je tiens également à remercier l'étudiant de doctorat de M. Georges Dionne, Helmi Jedidi, pour le temps qu'il a consacré à m'apporter les outils méthodologiques indispensables à la conduite de ce projet.

Je suis aussi redevable des cursus de grande qualité du programme en sciences de la gestion (M. Sc.) offerts à HEC Montréal et de ses intervenants professionnels responsables de ma formation, pour avoir assuré la partie théorique de celle-ci.

Je remercie spécialement M. Nicolas Delisle, directeur de la trésorerie de Intact Corporation Financière, de m'avoir offert l'inestimable opportunité d'effectuer un stage au sein de son équipe au cours de l'été 2018. En sa qualité de directeur de stage, j'ai pu bénéficier de son expertise et apprécier son implication à mettre à ma disposition tous les moyens nécessaires à la réalisation de ma tâche tout en m'accordant sa confiance dans l'exécution de l'objectif. Enfin, je tiens à souligner la collaboration remarquable de toute l'équipe de trésorerie chez Intact Corporation Financière, qui aura su contribuer à mon expérience enrichie au sein d'une entreprise reconnue pour son intégrité.

Table des matières

1 Profil de l'organisation **8**

2 Trésorerie **10**
2.1 Description de la fonction de trésorerie 10
2.2 Gestion de la trésorerie 11
2.3 Enjeu adressé 11

3 Revue de littérature **12**
3.1 Modèles quant au niveau optimal du solde bancaire 13
3.2 Modèles d'estimation des flux monétaires prévisionnels 16
3.2.1 Approche Payment Pattern (PP) 17
3.2.2 Approche par distribution 18

4 Collecte de données **20**

5 Méthodologie **20**
5.1 Modèles de distribution d'un montant mensuel prévisionnel . . . 21
5.1.1 Modèle additif 21
5.1.2 Modèle multiplicatif 22
5.1.3 Modèle généralisé mixte 23
5.2 Estimation des paramètres 24
5.3 Problèmes d'estimation 24
5.3.1 Distorsion selon la durée du mois 24
5.3.2 Variation dans la combinaison des tendances hebdomadaires et mensuelles 25
5.3.3 Cohérence quotidienne-mensuelle 25
5.4 Résolution de la cohérence quotidienne-mensuelle 25
5.4.1 Renormalisation 25
5.4.2 Mois standardisé 26
5.5 Sélection de variables et de modèle 27
5.5.1 Critères de performance 28
5.5.2 Méthodes de rééchantillonnage 32
5.5.3 Recherche automatique du meilleur modèle 33
5.6 Synthèse 36

6 Résultats **36**
6.1 Description statistique de l'échantillon et visualisation des données 36
6.2 Régressions des différentes spécifications 43
6.2.1 Résultats du modèle additif 43
6.2.2 Résultats du modèle multiplicatif 46
6.2.3 Résultats du modèle généralisé mixte 50
6.3 Résultats des procédures de sélection 54
6.3.1 Stepwise avec Pr = 0.15 55
6.3.2 Stepwise avec Pr = 0.60 suivi d'un all-subset 58
6.4 Comparaison des modèles 71

7 Conclusion **72**

A Annexes **74**
A.1 Figures complémentaires 74
A.1.1 Pattern hebdomadaire par type de flux monétaires 74
A.1.2 Pattern journalier moyen par type de flux monétaires . . 78
A.1.3 Pattern journalier par mois - Payables totaux 82
A.2 Liste des variables 88
A.3 Tableaux supplémentaires 92
A.3.1 Statistiques descriptives supplémentaires 92

Références **95**

Liste des tableaux

Tableau I Statistiques descriptives de l'échantillon total (n=1304 par type), sur une période 2013-2017 38
Tableau II Statistiques descriptives de l'échantillon d'apprentissage (n=1044), sur une période 2013-2016 39
Tableau III Statistiques descriptives de l'échantillon de validation (n=260), pour l'année 2017 39
Tableau IV Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au lundi (n=260), sur une période 2013-2017 40
Tableau V Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au mardi (n=261), sur une période 2013-2017 40
Tableau VI Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au mercredi (n=261), sur une période 2013-2017 41
Tableau VII Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au jeudi (n=261), sur une période 2013-2017 41
Tableau VIII Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au vendredi (n=261), sur une période 2013-2017 41
Tableau IX Estimation des paramètres du modèle additif à 28 variables sur l'échantillon apprentissage (n=1044) 45
Tableau X Évaluation du modèle additif à 28 variables 46
Tableau XI Estimation des paramètres du modèle multiplicatif à 138 variables sur l'échantillon apprentissage (n=1044) 47
Tableau XII Évaluation du modèle multiplicatif à 138 variables . . . 50
Tableau XIII Estimation des paramètres du modèle généralisé mixte à 143 variables sur l'échantillon apprentissage (n=1044) 51
Tableau XIV Évaluation du modèle généralisé mixte à 143 variables . 54
Tableau XV Résumé de la procédure de sélection stepwise, Pr=0.15 . 56
Tableau XVI Estimation des paramètres du modèle à 27 variables retenues par la méthode stepwise (Pr=0.15) sur l'échantillon apprentissage (n=1044) 57
Tableau XVIIÉvaluation du modèle à 27 variables retenues par la méthode stepwise (Pr=0.15) 58
Tableau XVIIIRésumé de la procédure de sélection stepwise, Pr=0.60 . 58
Tableau XIX Résumé de la procédure de sélection all-subset, à partir de la pré-sélection stepwise Pr=0.60 60
Tableau XX Estimation des paramètres du modèle à 31 variables retenues par la méthode all-subset (AIC) sur l'échantillon apprentissage (n=1044) 69
Tableau XXI Évaluation du modèle à 31 variables retenues par la méthode all-subset (AIC) 70

Tableau XXIIIEstimation des paramètres du modèle à 12 variables retenues par la méthode all-subset (SBC) sur l'échantillon apprentissage (n=1044) 71
Tableau XXIIIÉvaluation du modèle à 12 variables retenues par la méthode all-subset (SBC) 71
Tableau A.1 Liste des variables dépendantes 88
Tableau A.2 Liste des variables indépendantes binaires 88
Tableau A.3 Statistiques descriptives de l'échantillon des premiers semestres (n=647), sur une période 2013-2017 92
Tableau A.4 Statistiques descriptives de l'échantillon des deuxièmes semestres (n=657), sur une période 2013-2017 92
Tableau A.5 Statistiques descriptives de l'échantillon des premiers trimestres (n=322), sur une période 2013-2017 93
Tableau A.6 Statistiques descriptives de l'échantillon des deuxièmes trimestres (n=325), sur une période 2013-2017 93
Tableau A.7 Statistiques descriptives de l'échantillon des troisièmes trimestres (n=329), sur une période 2013-2017 93
Tableau A.8 Statistiques descriptives de l'échantillon des quatrièmes trimestres (n=328), sur une période 2013-2017 94

Table des figures

Figure 1 Croissance des primes directes souscrites(%) (base 100 = 2008) 9
Figure 2 Ratio Combiné (%) 9
Figure 3 Rendement des capitaux 9
Figure 4 Part de marché par société 9
Figure 5 Comportement aléatoire des flux de trésorerie 14
Figure 6 Pattern hebdomadaire - Payables totaux 42
Figure 7 Pattern journalier moyen - Payables totaux 43
Figure 8 Résumé des approches de sélection de variables ou de modèles 55
Figure 9 Synthèse des évaluations de performance de chaque modèle 72
Figure A.1 Structure Corporative de IFC 74
Figure A.2 Pattern hebdomadaire - Primes totales 74
Figure A.3 Pattern hebdomadaire - Primes PAP 75
Figure A.4 Pattern hebdomadaire - Primes chèques 75
Figure A.5 Pattern hebdomadaire - Primes téléphone 76
Figure A.6 Pattern hebdomadaire - Primes cartes de crédit 76
Figure A.7 Pattern hebdomadaire - Réclamations 77
Figure A.8 Pattern hebdomadaire - Payables chèques 77
Figure A.9 Pattern hebdomadaire - Payables EFT 78
Figure A.10 Pattern journalier - Primes Totales 78
Figure A.11 Pattern journalier - Primes PAP 79
Figure A.12 Pattern journalier - Primes chèques 79
Figure A.13 Pattern journalier - Primes téléphone 80
Figure A.14 Pattern journalier - Primes par cartes de crédit 80
Figure A.15 Pattern journalier - Réclamations 81
Figure A.16 Pattern journalier - Payables chèques 81
Figure A.17 Pattern journalier - Payables EFT 82
Figure A.18 Pattern journalier (Janvier) - Payables totaux 82
Figure A.19 Pattern journalier (Février) - Payables totaux 83
Figure A.20 Pattern journalier (Mars) - Payables totaux 83
Figure A.21 Pattern journalier (Avril) - Payables totaux 84
Figure A.22 Pattern journalier (Mai) - Payables totaux 84
Figure A.23 Pattern journalier (Juin) - Payables totaux 85
Figure A.24 Pattern journalier (Juillet) - Payables totaux 85
Figure A.25 Pattern journalier (Août) - Payables totaux 86
Figure A.26 Pattern journalier (Septembre) - Payables totaux 86
Figure A.27 Pattern journalier (Octobre) - Payables totaux 87
Figure A.28 Pattern journalier (Novembre) - Payables totaux 87
Figure A.29 Pattern journalier (Décembre) - Payables totaux 88

Introduction

La prévision des flux de trésorerie est un élément clé de la gestion financière des entreprises. Les prévisions de trésorerie à court terme sont une condition préalable à la gestion de la liquidité et à la couverture des risques financiers, alors que les prévisions de trésorerie à plus long terme constituent la base des décisions d'investissement et de financement. Ceci-dit, l'optimisation du processus de prévision des flux monétaires représente un important potentiel d'amélioration de la performance financière pour les entreprises. Grâce à la mise en place de meilleures pratiques, les organisations obtiennent une meilleure visibilité de leur liquidité et sont en mesure de l'optimiser, permettant ainsi de gérer le risque de liquidité plus efficacement. De plus, la technologie spécialisée permet aux entreprises d'exploiter d'importants gains d'efficacité. Les tâches manuelles lors de la collecte des données et de la création de rapports sont réduites grâce à des processus hautement automatisés et les entreprises obtiennent une visibilité globale de leur trésorerie et de leurs objectifs en matière de prévisions. En général, le processus de prévision d'une entreprise peut être influencé par la taille et la structure organisationnelle ainsi que la complexité résultante dans l'organisation de la tâche et des connaissances spécialisées. Toutefois, la sophistication du processus doit à la fois permettre d'obtenir une meilleure précision et efficacité des prévisions tout en restant applicable et flexible dans un environnement hautement dynamique.

Le mandat du projet supervisé effectué chez Intact Corporation Financière a pour but d'élaborer un processus de prévision, en commençant avec l'intégration automatisée des données sous une banque de données structurée. Des analyses historiques de tendance, de répartition et de propriétés statistiques pourront s'appuyer sur cette banque de données pour faciliter des décisions financières ultérieurement et améliorer les prévisions. Dans un deuxième temps, la tâche consiste à structurer le problème de prévision quotidienne des flux de trésorerie afin d'implanter un modèle de prévision pour plusieurs types de flux monétaires en tenant compte de plusieurs particularités concernant la nature bancaire de règlement. Le projet répond donc à un double objectif : l'amélioration de la prévision des flux monétaires quotidiens en s'appuyant sur des méthodes plus statistiques et sophistiquées ainsi que l'automatisation de certaines étapes afin de réduire l'erreur humaine et d'accélérer le processus de prévision.

L'étude s'appuie sur une base de données collectées à l'interne de l'entreprise, couvrant une période de 2013 à 2017 sur une base quotidienne. Les méthodes de prévision retenues pour les fins du projet seront estimées et évaluées sur cet échantillon. Par la suite, les procédures de sélection de variables et de modèles pourront être comparées selon le pouvoir prédictif des différentes spécifications de modèles. Selon l'approche employée et le critère de performance sélectionné, le modèle qui apparaît comme étant le plus performant varie. Néanmoins, les résultats démontrent que la spécification additive n'est jamais retenue selon chaque critères.

Le rapport est subdivisé de la manière suivante. La revue de littérature présente d'abord un survol de plusieurs contributions proposant différentes alternatives de modélisation pour prévoir des flux de trésorerie. La section suivante décrit le processus de collecte de données. La section méthodologie poursuit en exposant les méthodes de prévision retenues pour la société ainsi que les techniques de sélection de variables et de modèles utilisées dans le projet. Par la suite, la section des résultats présente une description statistique de l'échantillon et enchaîne avec l'exposition des résultats obtenus quant à l'application des modèles sur nos données. Finalement, une synthèse du travail est présentée dans la conclusion ainsi que les limites et les recommandations pour de prochaines avenues de recherche.

1 Profil de l'organisation

Intact Corporation Financière (IFC) est le chef de file de l'assurance incendie, accidents et risques divers (IARD) et l'un des principaux fournisseurs d'assurance spécialisée avec près de 10 milliards de dollars de primes annuelles souscrites en Amérique du Nord. L'entreprise dessert à la fois des clients du secteur public et des clients institutionnels à partir de bureaux localisés au Canada et aux États-Unis.

En termes de primes directes souscrites ($CA), 85% proviennent de l'activité canadienne et 15% sont générées par l'activité américaine. Au Canada, les produits d'assurance sont offerts sous la marque Intact Assurance par l'entremise d'un vaste réseau de courtiers, y compris la filiale en propriété exclusive, BrokerLink, ainsi que directement aux consommateurs à travers belairdirect. Aux États-Unis, la filiale en propriété exclusive One Beacon offre des produits d'assurance spécialisés par l'intermédiaire d'agences indépendantes, de courtiers, de grossistes et d'agence de gestion générale. (Rapport Annuel, 2017)

Depuis 1983, Intact Corporation Financière s'engage dans des acquisitions notamment celle de One Beacon, sa plus récente aux États-Unis. Ceci-dit, la diversification géographique et l'expansion de certaines lignes d'affaires lui ont permis d'accéder à de nouveaux marchés propices à la croissance future. L'entreprise a su soutenir sa solidité financière à travers les années, surperformant ainsi l'industrie canadienne pour 10 années consécutives. Ces maintes acquisitions témoignent également d'une industrie en phase de consolidation dans un marché de grande taille fragmenté par plusieurs joueurs, dont IFC avec une taille de marché environ 18 fois celle de la moyenne des compagnies dans l'industrie canadienne.

FIGURE 1 – Croissance des primes directes souscrites(%) (base 100 = 2008)

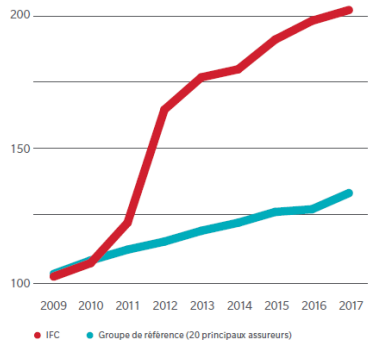


FIGURE 2 – Ratio Combiné (%)

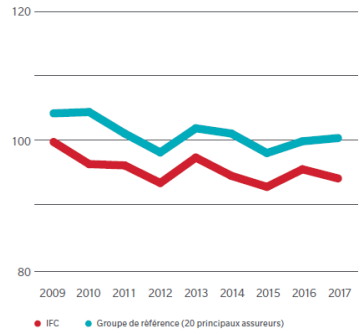


FIGURE 3 – Rendement des capitaux

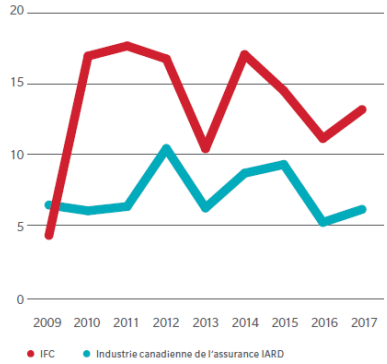
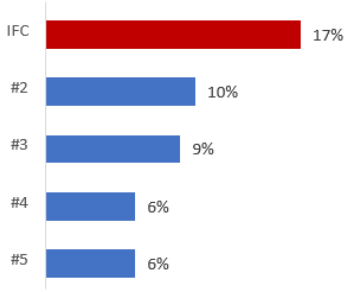


FIGURE 4 – Part de marché par société



Ces graphiques démontrent la détermination d'une des sociétés les plus respectées au Canada de surclasser l'industrie, année après année. En effet, d'après la figure 1, la croissance des primes de IFC a surperformé ses comparables depuis 2010 grâce à sa croissance organique et le succès de ses acquisitions. De plus, selon la figure 2, IFC a su se démarquer en affichant une profitabilité exemplaire par rapport à ses pairs. Le ratio combiné pour une compagnie d'assurance est formulé comme suit :

$$Ratio\ combiné = \frac{Frais\ de\ gestion + pertes\ encourrues}{primes\ encaissées}$$

Le ratio combiné est une mesure importante de profitabilité pour une compagnie d'assurance. Ce critère mesure l'efficacité de gestion des pertes et des frais de gestion par rapport aux primes récoltées et témoigne ainsi de la profitabilité

des activités de souscriptions de l'entreprise. Un ratio inférieur à 100% indique que les primes encaissées excèdent les dépenses à payer en terme de pertes et de frais de gestion, tel est le cas pour IFC depuis 2010. La surperformance de IFC par rapport à l'industrie peut être expliquée par leur expertise interne en indemnisation, par leur approche sophistiquée de tarification des primes et par leur discipline en souscription. Par ailleurs, la figure 3 supporte à son tour le surclassement de IFC au plan du rendement des capitaux (RCP), soit une mesure des bénéfices nets sur les capitaux propres moyens. Ce ratio mesure la capacité de l'entreprise à générer des profits à partir des investissements de ses actionnaires. La supériorité des résultats de souscription, les rendements d'investissement et la gestion du capital de IFC ont contribué à un meilleur RCP que celui de l'industrie. Finalement, la figure 4 compare la taille de marché de IFC évaluée à 17% par rapport aux 4 autres principaux assureurs de l'industrie canadienne.

Il est à spécifier que les activités d'assurance de base ne sont pas les seules à contribuer à cet objectif. L'équipe interne de gestion des placements, également reconnue comme l'une des meilleures au Canada par ses pairs, gère une vaste gamme de classes d'actifs en Amérique du Nord et participe au succès de l'entreprise. Elle collabore étroitement avec les équipes d'assurance, de gestion de trésorerie et des capitaux, de fiscalité et optimise la répartition des actifs sous gestion pour refléter la nature unique des activités d'assurance IARD de la société. Intact Corporation Financière jouit donc d'une équipe en matière de placement adaptée à ses besoins et qui s'est montrée des plus efficaces jusqu'à ce jour, affichant une surperformance du ratio combiné d'approximativement 160 points de base depuis 2010.

2 Trésorerie

2.1 Description de la fonction de trésorerie

La trésorerie est une fonction financière essentielle à la santé financière et au cœur du succès de chaque entreprise, qu'elle soit petite ou grande. Son rôle évolue continuellement à mesure que des nouvelles exigences corporatives se présentent, que des outils technologiques innovatifs sont offerts et que les demandes économiques se modifient. Cependant, certaines fonctions et processus clefs de la trésorerie restent fondamentaux à travers ces étapes évolutives. Ces piliers sont centraux à tout investissement, décision financière, objectifs à court et long terme, et généralement au maintien de l'entreprise dans les paramètres qu'elle s'est fixée.

La trésorerie implique la gestion des liquidités et des risques financiers dans une entreprise. Sa priorité est de s'assurer que l'entreprise dispose des fonds nécessaires pour rencontrer ses obligations quotidiennes, tout en supportant la

prise de décision sur l'élaboration de sa stratégie et de ses politiques financières à long terme. Plus particulièrement chez IFC, la fonction de trésorerie a pour responsabilités de :

- Optimiser les liquidités et la gestion du capital à travers l'organisation (capital optimal, coût des fonds, implications fiscales)
- Financer les acquisitions et les initiatives, préserver la flexibilité financière et innover
- Mitiger l'exposition au marché (activités de couverture)
- Superviser d'autres initiatives stratégiques/tactiques (ex : analyse des comparables et métriques de fusion et acquisition)
- Gérer les comptes bancaires et les relations avec les investisseurs des titres à revenus fixes
- Gérer les relations avec les agences de notation de crédit
- Superviser la gouvernance et les mesures de contrôles de la gestion de trésorerie

2.2 Gestion de la trésorerie

Le présent travail se concentrera uniquement sur la fonction de gestion de trésorerie qui peut paraître simple à première vue, mais requiert beaucoup de ressources et de connaissances sur la trésorerie. La capacité de transférer des actifs à plus long terme en actifs à court terme, tels que des liquidités, est cruciale pour la solidité financière de la firme. L'accumulation excessive de liquidités dans les comptes bancaires entraîne un coût d'opportunité, soit une perte de rendement. À l'inverse, si la société ne dispose pas suffisamment de liquidités, elle pourrait faire l'objet de pénalités si elle est dans l'incapacité de rencontrer ses passifs à court terme. De plus, pour une entreprise d'envergure comme Intact Corporation Financière, la fonction de gestion de trésorerie est responsable de la circulation des fonds dans les multiples systèmes bancaires et entre les juridictions. Cette tâche nécessite une somme considérable de connaissances afin d'être en mesure de circuler les montants exacts à travers les différentes unités tout en conservant des réserves de liquidités adéquates pour faire face aux imprévus. La structure corporative de la société est présentée en annexe (A.1) pour illustrer l'écosystème dans lequel la fonction de trésorerie doit coordonner et planifier les activités opérationnelles en matière de liquidité.

2.3 Enjeu adressé

À même la fonction de gestion de trésorerie décrite précédemment, l'une des tâches fondamentales est la prévision des flux monétaires. La capacité des

entreprises à prévoir avec précision les besoins futurs en liquidité est certainement importante pour maintes raisons. D'après les répondants au sondage 2017 Strategic Role Of Treasury, 64% ont déclaré que la gestion de la trésorerie et le processus de prévision des flux monétaires étaient des aspects prioritaires pour leur département de trésorerie au cours des trois prochaines années. Ceci-dit, le développement d'un processus de prévision permet entre autres de :

- Minimiser les liquidités excédentaires, la dette court à terme et les découverts bancaires
- Optimiser le financement à court terme et les décisions de prêts
- Fournir de la liquidité pour tous les besoins en devise
- Gérer les transactions clés
- Fournir une vue courante de la santé financière et de la position de l'entreprise
- Suivre étroitement le processus de génération de liquidité de l'entreprise
- Soutenir les décisions de déploiement de trésorerie telles que les réserves de liquidité, les dépenses en capital et les rachats d'actions
- Valider d'autres prévisions, telles que les prévisions de l'état des résultats et du bilan

La prévision des flux de trésorerie par comptes bancaires et la visibilité des liquidités à travers l'entreprise facilitent l'opération des comptes bancaires à des niveaux optimaux. Elle fournit également un cadre d'orientation tant au plan corporatif qu'au plan opérationnel et permet à une entreprise d'examiner ses capacités – à la fois actuelles et futures. Ces prévisions mettent en lumière la position financière de la firme et contribuent à déterminer comment diriger les investissements futurs et rembourser les dettes. Ainsi, des prévisions de liquidités fiables se traduisent par une utilisation plus efficace du capital, une réduction des coûts de financement et moins de risques.

Toutefois, cette tâche complexe reste difficile à appliquer et à maîtriser pour plusieurs trésoriers. En raison des multiples entités, comptes bancaires et flux de trésorerie, l'organisation des données est un défi en soi sans compter la complexité de plusieurs systèmes ERP (Enterprise Resource Planning) décentralisés, plateformes bancaires et fichiers Excel individuels. Bien que l'importance et la pertinence du défi de processus de prévision soit reconnue, peu de littérature empirique se penche sur le sujet.

3 Revue de littérature

Les décisions liées à la fonction de trésorerie ont suscité l'attention grandissante des académiciens et des praticiens à partir des années 1950. Rappelons

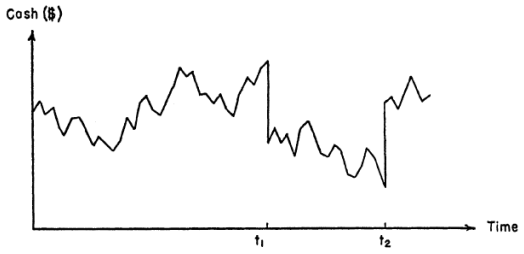
que l'un des objectifs de la fonction de trésorerie est de maintenir un niveau de liquidités le plus bas possible dans ses comptes bancaires tout en s'assurant que l'entreprise puisse opérer efficacement. Ceci-dit, si le solde bancaire de la firme est trop faible, les coûts de détention seront réduits, mais des situations de déficit obligeront l'entreprise à vendre des titres de son portefeuille de placement à des conditions de marché possiblement défavorables, augmentant ainsi les coûts de transaction. À l'opposé, si le solde bancaire est trop élevé, les coûts de transaction seront faibles mais l'entreprise supportera des coûts de détention élevés puisque les liquidités excédentaires ne produisent aucun rendement. Par conséquent, il existe un niveau de solde bancaire cible qui devrait être optimisé par la fonction de trésorerie en considérant les particularités de ses flux monétaires.

Plusieurs modèles de gestion de trésorerie ont été proposés pour contrôler les soldes bancaires depuis la publication de Baumol en 1952. Ces modèles se basent généralement sur les propriétés statistiques spécifiques aux balances de liquidités ou sur les prévisions de flux de trésorerie. Ils s'appuient principalement sur une hypothèse quant à la distribution des flux de trésorerie. Par ailleurs, la littérature empirique se concentre davantage sur la recherche d'un modèle permettant d'établir le niveau optimal des soldes bancaires auquel opérer plutôt que sur la recherche d'un modèle permettant la prévision des flux monétaires. Le premier type de modèle est important en guise d'indication sur le niveau optimal des liquidités à opérer certes, mais une gestion de qualité nécessite préalablement des prévisions fiables de flux monétaires pour juger adéquatement le solde bancaire requis. Les prochaines lignes du rapport décrivent les principaux modèles proposés dans la littérature qui ont guidé la méthodologie concernant la prévision des flux monétaires de l'entreprise. Par ailleurs, la bibliographie inclut des publications supplémentaires pertinentes à la recherche de modèles prévisionnels et qui ont également nourri le développement du sujet dans le présent rapport.

3.1 Modèles quant au niveau optimal du solde bancaire

Le modèle de Miller et Orr (1966) considère seulement un univers à deux actifs, soit la trésorerie et un investissement liquide à faible risque. Ce modèle suppose également que le flux de trésorerie varie aléatoirement, représenté à la figure 5 ci-bas, avec des décisions au moment T1 d'investissement dans des obligations et des décisions au moment T2 de ventes d'actifs.

FIGURE 5 – Comportement aléatoire des flux de trésorerie



Ceci-dit, le modèle tente de définir deux limites pour le niveau du solde bancaire : une limite supérieure (h) et inférieure (égale à 0 dans la publication originale) qui entraînent l'achat ou la vente d'actifs du portefeuille de placement afin de rééquilibrer le niveau des fonds au niveau optimal (z). Ainsi, le modèle optimise le niveau des fonds en considérant les coûts de transaction (γ) et les coûts d'opportunité (ν) Le modèle de Miller et Or se définit comme suit :

$$Z^* = (\frac{3\gamma}{4\nu}\sigma^2)^{\frac{1}{3}}$$

Ici, * indique la valeur estimée optimale et σ^2 représente la variance historique des soldes bancaires, soit les flux de trésorerie nets. Ainsi, le solde bancaire optimal obtenu par le modèle proposé est :

$$Solde\ bancaire\ optimal = 4\nu/3$$

Après les années 1980, les recherches se sont concentrées sur l'amélioration des techniques méthodologiques concernant le problème d'optimisation du niveau des fonds. Les approches subséquentes au modèle de Miller et Or se basent principalement sur des processus stochastiques.

Dans ce sens, Tapiero et Zuckerman (1980) présentent un modèle stochastique sous l'hypothèse que les flux de trésorerie varient aléatoirement suivant un processus Poisson, et Smith (1986) développe un modèle dynamique stochastique supposant un processus de diffusion des flux monétaires temporairement indépendants et suivant un mouvement brownien. Par la suite, Barbosa et Pimentel (2001) ont proposé une technique alternative, soit la programmation linéaire, qui permet d'estimer le flux monétaire sur une base de fréquence spécifique en supportant les entrées monétaires et le processus de paiement. Quelques années après, Premachandra (2004) propose une approche alternative au modèle de Miller et Orr (1966)), soit la technique de diffusion approximative. Ce modèle plus généralisé permet d'assouplir les hypothèses restrictives du modèle de Miller et Or concernant les coûts fixes de transaction et la distribution normale des flux monétaires afin de représenter plus justement la réalité, ce qui se traduit

par des résultats significativement plus précis que ceux du modèle de Miller et Orr. Volosov et al. (2005) incorporent l'aspect international dans leurs modèles afin de prendre en compte le risque de variation des taux de change. Les auteurs présentent un modèle de programmation stochastique en deux étapes, basé sur des arbres de scénarios, qui considère non seulement le niveau des fonds mais également l'exposition aux devises internationales. Ils obtiennent des résultats positifs quant à la détermination du niveau des fonds optimal ainsi qu'à la manière d'utiliser la couverture.

De ces travaux et d'autres, Salas-Molina et collab. (2016) relèvent que les hypothèses communes quant aux propriétés statistiques des séries temporelles de flux monétaires incluent :

— Normalité : les flux monétaires suivent une distribution Gaussienne avec des observations symétriquement centrées autour de la moyenne avec une variance constante.
— Absence de corrélation : les flux monétaires passés n'affectent pas la probabilité d'évènement des flux monétaires futurs
— Stationnarité : la distribution de probabilité des flux monétaires est constante dans le temps et, conséquemment, ses propriétés statistiques telles que la moyenne et la variance restent constantes.
— Linéarité : les flux monétaires sont proportionnels soit à une autre variable explicative ou à une combinaison de variables explicatives

Dans leur travail, les auteurs fournissent une analyse des propriétés statistiques de 54 ensembles de données sur les flux monétaires de petites et moyennes entreprises en Espagne, soit un échantillon représentatif des types d'entreprises les plus communes en Europe. Leurs résultats démontrent que les hypothèses de normalité, d'absence de corrélation et de stationnarité sont à peine présentes dans l'ensemble des données à l'étude. Toutefois, les résultats confirment l'influence saisonnière telle que suggérée dans les travaux de Miller et Stone (1985) et Stone et Miller (1987) décrite dans la prochaine section. Ceci-dit, Salas-Molina et collab. (2016) concluent en soutenant qu'un certain nombre d'étapes préliminaires sont nécessaires à la prévision des flux monétaires avant de choisir un modèle : (i) résumé statistique incluant la normalité, la corrélation et la stationnarité ; (ii) l'impact de la transformation des données tel que le traitement des données aberrantes et la transformation Box-Cox ; (iii) le test de non-linéarité afin de déterminer le type de model plus susceptible de livrer une meilleure performance.

Finalement, il est à mentionner que Gormley et Meade (2007) ont contribué à la recherche de modèles en présentant une politique dynamique du solde bancaire qui minimise les coûts de transactions lorsque les flux monétaires ne sont pas distribués uniformément et identiquement (i.i.d.) et sont temporellement

dépendants. Avec cette méthodologie, aucune hypothèse est nécessaire sur la distribution des flux monétaires, ce qui représente un grand avantage.

3.2 Modèles d'estimation des flux monétaires prévisionnels

Les approches concernant la prévision des flux monétaires peuvent être classifiées en deux catégories : une approche traditionnelle basée sur une agrégation d'estimations locales et une approche statistique basée sur un processus statistique qui tente d'extrapoler les relations causales observées dans les données.

L'approche traditionnelle consiste typiquement à répondre à des exigences de rapports, notamment en vue de fournir un montant prévisionnel pour le budget mensuel. Plusieurs départements de l'entreprise produisent des estimés d'entrées et sorties de fonds pour des intervalles de temps prédéterminés sur un horizon de planification. Par exemple, le département des taxes et des comptes à payer peuvent fournir des estimés de prévision sur une base mensuelle pour l'année prévisionnelle au meilleur de leur connaissance spécifique. Ces prévisions dispersées à travers l'entreprise sont agrégées pour former un budget de trésorerie pro forma qui servira par la suite au processus de planification de la trésorerie. L'approche par planification fait partie de cette catégorie. Plus précisément, elle consiste à prévoir les entrées et sorties de fonds selon les données primitives quant au processus de facturation-règlement-livraison, les montants prévus selon les contrats signés, le calendrier de production, etc. À titre d'exemple, si l'entreprise effectue ses achats selon des termes de crédit net trente jours et suit une politique de paiement qui la contraint à toujours payer au trentième jour, alors la tâche de prévision se limite seulement à prévoir des déboursements selon le terme de crédit et les factures contractées. Le développement d'un tel planning nécessite seulement un système d'information pour capturer et organiser les données de facturation. En ce sens, la planification pure peut être considérée comme une prévision basée sur un système d'information et qui ne requiert aucune estimation statistique.

L'approche statistique est complémentaire à l'approche traditionnelle et sert également d'outil d'évaluation pour détecter les anomalies de performance. L'une des procédures standards est de projeter une série temporelle selon une moyenne mobile qui reflète les tendances et les cycles sur une certaine période. Toutefois, cette technique purement extrapolative ne permet pas de lier le budget de caisse aux flux monétaires projetés. Les prochains paragraphes détaillent la méthode Payment Pattern (PP) et la méthode de distribution, deux approches statistiques généralement utilisées pour établir des prévisions de flux monétaires.

3.2.1 Approche Payment Pattern (PP)

En ce qui concerne l'approche PP, la collecte de flux monétaires est liée à la source de génération de ces flux et cette approche permet ainsi d'estimer la distribution du temps de délais entre la génération et la collection réelle. Par exemple, si l'entreprise collecte généralement 60% des ventes du mois en cours dans le même mois et les 40% restants au cours du prochain mois, alors la collection du mois courant est de 60% de la somme des ventes du même mois et 40% des ventes du mois précédent. Stone (1976) examine en détail les procédures couramment utilisées en pratique et souligne les lacunes de ces approches conventionnelles comparativement à l'approche PP. Sa publication s'appuie sur la prévision des ventes et des comptes recevables en indiquant que les mesures de contrôle populaires dépendent fortement de la tendance des ventes. Ainsi, plus les ventes présentent une cyclicité, plus les résultats obtenus se sont révélés insatisfaisants. L'utilisation de mesures saisonnières peut aider à faire face aux tendances cycliques, toutefois cette extension est adéquate seulement dans la mesure où les comportements passés se répètent. L'incapacité de dissocier l'effet des variations de ventes et l'effet de décalage des paiements combinée au fait que la variation des ventes est généralement beaucoup plus significative que les changements dans le comportement des paiements se traduit par un niveau d'insatisfaction élevé des utilisateurs. Les méthodes conventionnelles tentent de résumer le comportement du processus de paiement en un seul paramètre (ADO, ratio et proportion des soldes) pour estimer les prévisions.

Par exemple, le paramètre ADO (nombre moyen de jours en souffrance) est un ratio des comptes recevables divisé par une mesure du montant quotidien de ventes. Étant donnée une prévision annuelle des ventes à crédit et une valeur pro-forma du ratio ADO, une valeur des recevables dérivée de l'équation suivante est projetée :

$$TAR_t = ADO_t(ventes \;\; crédit \; annuelles)/365,$$

TAR_t représente la valeur totale des comptes à recevoir à la fin du mois t comme étant la somme des comptes recevables provenant des ventes à crédit au mois t, t-H et ainsi de suite :

$$TAR_t = F_0 CS_t + F_1 CS_{t-1} + ... + F_{H-1} CS_{t-H+1} = \sum_{i=0}^{H-1} F_i CS_{t-i}$$

Dans la même logique, la prévision basée sur un ratio fait référence à toute procédure de prévision qui suppose que les recevables à un certain moment dans le temps est proportionnel à une mesure des ventes. La méthode de prévision basée sur la proportion du solde suppose que les paiements reçus au cours d'un mois donné représentent une proportion constante du compte recevables

au début du mois. Ces paramètres résumés à un seul chiffre ne reflètent pas correctement la dépendance des recevables par rapport aux ventes. Ainsi, l'utilisation de ces techniques introduit une erreur de prévision structurelle à la fois dans la projection des recevables et dans la projection des flux monétaires.

La prise en compte des proportions de paiement et des fractions de solde variables dans le temps fournit la généralité conceptuelle nécessaire pour élaborer des prévisions en présence d'un changement dans le comportement des paiements. Stone (1976) propose donc une méthode avec régression qui s'appuie sur un historique de données afin de projeter des flux monétaires selon des proportions variables. Entre autres, l'un des modèles de régression présentés permet d'estimer conjointement l'effet du pattern de paiement de base, de la saisonnalité et de la déviation du taux d'intérêt sur la proportion de paiement :

$$\frac{CF_{ti}}{CS_t} = P_i + \sum_{m=1}^{12} S_{mi} D_{tm} + a_i(I_t - \overline{I}) + u_{ti}$$

La proportion de paiement est ainsi modélisée en réponse à différentes variables et permet également un ensemble différent de proportions à chaque mois selon la saisonnalité. CF_{ti} est le flux monétaire au mois $t+i$ découlant des ventes à crédit du mois t. S_{mi} représente la déviation saisonnière dans la proportion de paiement P_i pour le mois m dans l'année. D_{tm} est une variable binaire qui prend la valeur de un lorsque le temps t survient dans le mois m de l'année et zéro sinon. Le coefficient a_i mesure la sensibilité des paiements survenant i mois après la vente à la déviation du taux d'intérêt par rapport à son niveau normal et u_{ti} est le terme d'erreur résiduel.

3.2.2 Approche par distribution

Cette approche purement statistique permet de répartir un montant mensuel prévisionnel sur les jours du mois afin de refléter le pattern intra-mois des flux monétaires. Cette méthode tente de dériver des prévisions quotidiennes à partir d'une prévision mensuelle en s'appuyant sur une estimation statistique des cycles mensuels et hebdomadaires. Implicitement, cela implique que les flux monétaires projetés suivent le même pattern de distribution quotidienne à l'intérieur de chaque mois.

Stone et Wood (1977) élaborent cette approche et supposent que le timing et le montant de la plupart des paiements majeurs, tels que les paiements de dividendes et de dettes, sont connus avec considérablement plus de précision que les paiements mineurs. Ils argumentent également que les paiements majeurs inattendus sont généralement traités en une seule transaction financière, comme l'emprunt sur une marge de crédit. Ceci-dit, la tâche de prévision par l'approche de distribution se concentre principalement sur la projection quoti-

dienne des flux monétaires mineurs. Les auteurs affirment que l'approche par distribution est plus facile et moins coûteuse à implanter et à gérer que l'approche par planification, par contre sa validité requiert une stabilité statistique des patterns des flux monétaires à l'intérieur de chaque mois. D'un point de vue statistique, la méthode par distribution nécessite toutefois des hypothèses de stabilité moins strictes par rapport aux techniques extrapolatives (moyenne mobile). De plus, d'un point de vue administratif, cette méthode permet au gestionnaire de lié le budget de caisse aux prévisions de flux monétaires quotidiens. Cet aspect est non négligeable puisque les gestionnaires sont souvent évalués selon l'atteinte des objectifs fixés et le suivi du plan budgétaire.

En l'absence de cycles mensuels ou hebdomadaires, une technique simple pour évaluer le pattern intra-mensuel pour chaque composante de flux monétaires mineurs est de calculer la fraction moyenne du total mensuel pour un jour ouvrable donné :

$$a_t = \frac{1}{N} \sum_{n=1}^{N} f_{tn}$$

D'après cette équation, les valeurs de a_t résument les estimés des patterns intra-mensuel où a_t représente la fraction moyenne au jour t du flux monétaire sur le total mensuel pour les N mois. Toutefois, en présence de cycles mensuels et hebdomadaires, Stone et Wood (1977) suggèrent l'utilisation d'une régression par variables binaires afin d'estimer simultanément ces patterns. Le traitement des jours fériés dans la régression est également adressé par les auteurs et ceux-ci élaborent subséquemment des techniques supplémentaires pour traiter davantage de particularités, notamment sur la longueur variable du mois, dans leur publication de 1985.

Stone et Wood (1977) rapportent les résultats de leur technique par distribution sur une compagnie réelle et concluent que leur modèle performe relativement bien sur un intervalle de trois à quatre jours. Le modèle produit toutefois des erreurs quotidiennes importantes, soit une magnitude des erreurs quotidiennes moyennes entre 12% et 20% pour les différentes composantes de flux monétaires. Leur application suggère également que les erreurs quotidiennes sont négativement corrélées, c'est-à-dire qu'une sous-estimation du montant quotidien est suivie par une surestimation du montant quotidien le jour suivant. D'après les auteurs, la source primaire d'erreur provient du décalage d'encaissement ou de déboursement de l'ordre d'un ou de deux jours. Il est important de mentionner qu'aucune indication est fournie quant à la description de leur échantillon ayant servi à l'évaluation du modèle. Les coefficients estimés par la régression ne sont également pas présentés dans leur publication.

4 Collecte de données

La fonction de trésorerie au sein d'Intact Corporation Financière archive les données de flux monétaires quotidiens par entité sous des fichiers Excel à chaque année. Le solde bancaire de début, suivit des flux monétaires et du solde bancaire de fin sont inscrits dans ces fichiers pour chaque entité sous IFC. À même ce fichier, les données actuelles et prévisionnelles des flux de trésorerie sont disponibles à chaque jour ouvrable et ces données sont mises à jour par l'analyste en charge. L'écart de prévision peut ainsi être suivi par le gestionnaire et permet également d'ajuster les prévisions ultérieures.

Ceci-dit, la construction de l'échantillon s'appuie sur ces fichiers Excel annuels et considère seulement les comptes bancaires d'une de leurs institutions financières, soit ceux de la Banque Royale, en devise canadienne. Une base de données industrialisée sous un seul fichier a alors été construite à partir des fichiers Excels de 2013 à 2018 pour structurer plus efficacement les données. Nous avons toutefois choisi d'arrêter notre période d'étude en décembre 2017 afin de constituer un échantillon avec des années complètes de données. Tel que mentionné, les fichiers initiaux archivaient les valeurs de flux monétaires par entité. Cependant, les flux monétaires ont été aggrégés selon leur type, sans égard à l'entité concernée, pour contruire l'échantillon sur lequel se bâtira le modèle de prévision. Au final, 1304 observations actuelles quotidiennes ont été retenues aux fins de l'étude pour 9 types de flux monétaires, totalisant 11 736 observations en dollars canadiens.

5 Méthodologie

Le processus de prévision en place chez Intact Corporation Financière se base principalement sur des moyennes mobiles pour chaque type de flux de trésorerie en considérant leur nature de règlement et selon la tendance de croissance de chaque entité d'assurance sous IFC (voir figure A.1). L'analyste en charge réflète les prévisions quotidiennes selon ces moyennes historiques et ajuste celles-ci de manière subjective en fonction de ses connaissances spécifiques. Ainsi, un fichier Excel est monté pour chaque mois où les différents flux monétaires quotidiens sont documentés pour chaque entité. Ceci-dit, l'objectif du mandat consistait à proposer un processus de prévision plus sophistiqué tout en produisant des résultats cohérents avec les prévisions budgétaires mensuelles. Cette particularité pratique a guidé notre choix de méthodologie vers l'approche de distribution décrite par Miller et Stone (1985).

Tel que mentionné brièvement dans la revue de littérature, cette méthode permet d'allouer un montant total d'une période de temps à des sous-périodes individuelles dans cette période. Par exemple, l'estimé d'un montant total mensuel obtenu selon un budget prévisionnel serait distribué sur les jours du mois pour

obtenir des prévisions quotidiennes. Selon le contexte de prévision, le montant mensuel peut être alloué sur les jours calendriers ou les jours ouvrables du mois. L'allocation quotidienne est effectuée à partir d'une régression par variables binaires afin d'estimer la tendance typique des flux de trésorerie au cours des jours de la semaine et des jours du mois. Les coefficients obtenus permettent par la suite de distribuer le montant budgétaire prévisionnel de la firme sur les jours du mois.

Soit l'index n représentant le jour du mois allant de 1 à N, où N désigne le nombre de jours dans un mois en particulier. Dans un contexte de jours ouvrables, N peut prendre les valeurs 20,21,22 ou 23. En termes de jours calendriers, les valeurs possibles de N sont 28,29,30 et 31. Supposons également l'index w représentant le jour de semaine où w prend la valeur de 1 lorsque le jour de la semaine est un lundi, 2 pour mardi et ainsi de suite pour le reste de la semaine. La valeur maximale que peut prendre w dans un contexte de jours ouvrables est 5 et 7 pour un contexte de jours calendriers. La prévision quotidienne, au jour n du mois lorsque celui-ci s'avère le jour w de la semaine, peut être obtenue par l'équation suivante :

$$FORECAST_{nw} = p_{nw} * TCF_m \tag{1}$$

p_{nw} : proportion du total mensuel
TCF_m :total mensuel

Le point central est donc de formuler un modèle pour p_{nw} en termes de paramètres estimables pouvant être liés aux flux de trésorerie. L'estimation de ces paramètres requiert une base de données historique pour les flux de trésorerie de l'entreprise. De plus, la proportion p_{nw} doit pouvoir capter la cyclicité dans les processus de commande-livraison-règlement généralement présente dans la plupart des firmes. Une telle résolution de tendances requiert un modèle d'interaction entre l'effet du jour de semaine et l'effet du jour du mois.

Nous utiliserons un contexte de jours ouvrables pour distribuer le montant mensuel budgétaire et comparerons les trois différentes spécifications proposées par Miller et Stone (1985), soient le modèle additif, le modèle multiplicatif et le modèle généralisé mixte.

5.1 Modèles de distribution d'un montant mensuel prévisionnel

5.1.1 Modèle additif

Soit la variable a_n représentant la proportion du flux de trésorerie qui se produirait au jour ouvrable n en l'absence de l'effet du jour de semaine et la variable b_w pour mesurer l'effet du jour de la semaine. La proportion pour le

jour du mois n et le jour de la semaine w serait spécifié de la façon suivante :

$$p_{nw} = a_n + b_w \tag{2}$$

L'ensemble des a_n caractérise la tendance mensuelle alors que l'ensemble des b_w (avec w égale à 1 pour lundi et 5 pour vendredi) caractérise la tendance hebdomadaire.

5.1.2 Modèle multiplicatif

Selon la spécification additive, l'effet du jour du mois et l'effet du jour de la semaine seraient structurellement indépendants. En d'autres mots, ceci implique que l'effet hebdomadaire ne varie pas avec l'effet mensuel et vice-versa.

Miller et Stone (1985) argumentent toutefois qu'on observe des cycles hebdomadaires en présence de variations dans le timing du jour de la semaine pour le processus de livraison, le traitement bancaire, le traitement informatique interne et le règlement bancaire. L'amplitude du bruit provenant de l'effet du jour de la semaine devrait être relativement moins important en période de creux du cycle mensuel et inversement en période de sommet du cycle mensuel.

Ainsi, l'une des possibilités pour modéliser l'interaction entre l'effet hebdomadaire et l'effet mensuel serait d'employer une mise à l'échelle multiplicative. On peut distinguer deux méthodes pour la mise en échelle : l'effet du jour de la semaine modélisé comme un facteur de perturbation pour la mise à l'échelle joint à un cycle mensuel de base ou l'effet du jour du mois modélisé comme un facteur de perturbation pour la mise à l'échelle joint à un cycle hebdomadaire de base.

La première alternative modélise la proportion pour le jour du mois n et le jour de la semaine w de la façon suivante :

$$p_{nw} = a_n^*(1 + b_w^*) = a_n^* + a_n^* b_w^* \tag{3}$$

Les variables a_n^* et b_w^* sont des analogues multiplicatifs à a_n et b_w. Ainsi, a_n^* est l'estimé multiplicatif de la proportion du total mensuel qui se produirait au jour ouvrable n en l'absence d'effet hebdomadaire. Dans cette spécification, b_w^* est le facteur d'ajustement pour l'effet du jour de la semaine. Lorsque b_w^* est supérieur(inférieur) à zéro, il s'agit d'un ajustement à la hausse(baisse) et lorsqu'il est égale à zéro, il n'y a aucun ajustement pour l'effet du jour de la semaine. En pratique, un cycle mensuel de base couplé au facteur de perturbation pour l'effet du jour de la semaine est plus souvent observé.

Néanmoins, la proportion pour le jour du mois n et le jour de la semaine w selon la deuxième alternative est spécifiée de la façon suivante :

$$p_{nw} = b_w^*(1 + a_n^*) = b_w^* + b_w^* a_n^* \tag{4}$$

Ici, b_w^* est l'estimé multiplicatif de la proportion du total mensuel qui se produirait au jour de la semaine w en l'absence d'effet mensuel. Dans cette spécification, a_n^* est le facteur d'ajustement pour l'effet du jour du mois. Lorsque a_n^* est supérieur(inférieur) à zéro, il s'agit d'un ajustement à la hausse(baisse) et lorsqu'il est égale à zéro, il n'y a aucun ajustement pour l'effet du jour du mois.

5.1.3 Modèle généralisé mixte

La possibilité d'une formulation additive (niveau indépendant) et multiplicative (mise à l'échelle) des perturbations cycliques laisse entrevoir une troisième forme de modèle qui permettrait d'agencer ces deux formulations. En d'autres mots, il est possible que le traitement pour capter la saisonnalité dans le jour du mois et le jour de la semaine devrait coupler les composantes additive et multiplicative. Pour un cycle mensuel de base dominant, la spécification serait de la forme suivante :

$$p_{nw} = a_n^*(1 + b_w^*) + b_w^{**} = a_n^* + a_n^* b_w^* + b_w^{**} \tag{5}$$

Pour un cycle hebdomadaire de base dominant, la spécification serait de la forme suivante :

$$p_{nw} = b_w^*(1 + a_n^*) + a_n^{**} = b_w^* + b_w^* a_n^* + a_n^{**} \tag{6}$$

La variable b_w^{**} est un facteur additif de perturbation hebdomadaire tandis que la variable a_n^{**} représente un facteur additif de perturbation mensuelle. Une formulation générale pour prendre en compte les composantes additive et multiplicative modéliserait la proportion pour le jour du mois n et le jour de la semaine w selon l'équation suivante :

$$p_{nw} = a_n^{'} + i_{nw} + b_w^{'} \tag{7}$$

Dans ce dernier cas, $a_n^{'}$ est un paramètre général pour le jour ouvrable du mois, $b_w^{'}$ est un paramètre général pour le jour de la semaine et i_{nw} est un paramètre général d'interaction.

5.2 Estimation des paramètres

Afin de pouvoir calculer la proportion selon le modèle au choix, il faut préalablement estimer les paramètres pour les effets mensuel et hebdomadaire. Miller et Stone (1985) proposent d'estimer ces paramètres à l'aide d'une régression linéaire multiple avec des variables indicatrices binaires :

$$\frac{CF_t}{TCF_m} = \sum_{n=1}^{N} A_n m_{tn} + \sum_{w=1}^{W} B_w d_{tw} + \sum_{n=1}^{N}\sum_{w=1}^{W} I_{nw} m_{tn} d_{tw} + \epsilon_{tnw} \qquad (8)$$

Dans ce modèle de régression, t indexe les observations quotidiennes de l'échantillon, CF_t représente le flux monétaire observé au jour indiqué par l'indice t, A_n est le coefficient qui mesure l'effet du jour ouvrable n, N est le maximum du nombre de jours ouvrables pour n'importe quel mois (i.e. 23), B_w est le coefficient qui mesure l'effet du jour de la semaine w, m_{tn} est une variable indicatrice binaire qui vaut 1 lorsque l'observation t se produit au jour ouvrable n, d_{tw} est une variable indicatrice binaire qui vaut 1 lorsque l'observation t se produit au jour de la semaine w et ϵ_{tnw} est le terme d'erreur. I_{nw} est le coefficient d'interaction et il existe un terme d'interaction pour chaque paire de n et w. Ainsi, cette spécification contient NxW coefficients supplémentaires par rapport à la spécification simplement additive. La généralité du modèle permet d'estimer une variété de modèles, cependant cette généralité engendre l'estimation de beaucoup plus de paramètres. Davantage de données sont donc nécessaires pour obtenir une estimation fiable. Avec trois ou quatre années de données quotidiennes, il est possible que certaines combinaisons, par exemple lorsque le 15^e jour ouvrable est un jeudi, ne soient pas observées dans l'échantillon. Néanmoins, un modèle général est avantageux pour trois principales raisons :

1. Développement de théorie : La théorie développée pour ce modèle est applicable pour tous les cas particuliers incluant les spécifications additives et multiplicatives.
2. Architecture des systèmes et programmation : Ce modèle général est flexible pour incorporer plusieurs cas particuliers et simplifie alors le design des systèmes et la programmation puisqu'un seul modèle peut être traité.
3. Validation des spécifications : le modèle général peut être utilisé sur une base de comparaison pour tester les différentes spécifications à savoir si les paramètres omis sont non significativement différents de zéro.

5.3 Problèmes d'estimation

5.3.1 Distorsion selon la durée du mois

Les mois n'ont pas tous le même nombre de jours soit en termes de jours ouvrables ou jours calendriers. Ainsi, l'unité de temps pour un mois n'est pas

exactement régulière, elle varie plutôt par 4 jours (20 à 23 jours ouvrables ou 28 à 31 jours calendriers). Selon l'approche de distribution, le total mensuel serait alloué sur plus de jours pour un mois de plus longue durée et engendrerait un décalage dans les proportions observés à certains jours.

5.3.2 Variation dans la combinaison des tendances hebdomadaires et mensuelles

L'élément central de l'approche proposée par les auteurs s'appuie sur l'estimation jointe de l'effet du jour de la semaine et du jour du mois en vue de prévoir des flux monétaires quotidiens. Ces deux effets se produisent simultanément et il faut donc les traiter simultanément. Toutefois, un mois n'est jamais un multiple intégral de nombre de semaines, à l'exception du mois de février lorsque l'année n'est pas bissextile. En effet, tous les mois sont composés de 4 semaines et une fraction d'une autre semaine. Cette fraction supplémentaire dépend de la durée du mois et donc varie à chaque mois.

5.3.3 Cohérence quotidienne-mensuelle

Il est nécessaire que la somme des prévisions quotidiennes pour chaque mois équivaille au montant total mensuel. De la même façon, la somme des proportions p_{nw} doit être égale à 1 pour satisfaire l'exigence de cohérence. Dans l'optique des jours ouvrables, il existe 20 paires possibles entre la durée du mois et la combinaison hebdomadaire-mensuel. Soit quatre durées possibles (20,21,22 et 23) multiplié par 5 combinaisons hebdomadaire-mensuel possibles (le premier jour ouvrable du mois peut soit être un lundi, mardi, mercredi, jeudi ou vendredi).

5.4 Résolution de la cohérence quotidienne-mensuelle

Il existe plusieurs alternatives pour assurer que la somme des prévisions quotidiennes équivaille au total mensuel prévu. Nous traiterons deux approches, soit la renormalisation et l'utilisation d'un mois standardisé, pour faire face au problème de la cohérence quotidienne-mensuelle.

5.4.1 Renormalisation

La technique de renormalisation utilisée par Stone et Wood (1977) permet d'assurer la cohérence quotidienne-mensuelle pour chaque mois. L'idée est d'estimer dans un premier temps les paramètres de la régression du modèle généralisé mixte sans contraintes pour le problème de cohérence. Les prévisions quotidiennes, obtenues par les coefficients estimés, sont additionnées pour obtenir un montant total prévisionnel, TCF*. En général, ce montant TCF* n'équivaut

pas l'entrée de flux monétaire mensuel utilisé pour générer les prévisions. Leur approche consistait ensuite à mettre à l'échelle les montant quotidiens prévisionnels par le facteur $\text{TCF}_M/\text{TCF}^*$. Cette mise à l'échelle permet d'assurer que la cohérence quotidienne-mensuelle soit respectée et ainsi que le flux monétaire mensuel prédit soit le même que celui utilisé pour générer la prévision.

La technique de renormalisation est équivalente à additionner les proportions prédites et ensuite mettre à l'échelle chacune de ces proportions par rapport à leur total. En supposant que $\hat{p}_{nw}$ est l'estimé de la proportion non contrainte et que $\overline{p}_{nw}$ est l'estimé renormalisé, alors la formule pour $\overline{p}_{nw}$ est la suivante :

$$\overline{p}_{nw} = \hat{p}_{nw} / \sum_{t=1}^{N'} \sum_{n=1}^{N} \sum_{w=1}^{W} \hat{p}_{nw} m_{t,n} d_{t,w} \tag{9}$$

Cette technique est facile à implanter en pratique, toutefois l'erreur de prévision est distribuée uniformément sur tout le mois. En réalité, Miller et Stone (1985) stipulent que ces erreurs ont tendance à être plutôt concentrées au début et à la fin du mois en raison du cycle hebdomadaire et de la longueur des différents mois.

5.4.2 Mois standardisé

L'utilisation d'un mois standardisé est une manière alternative pour faire face au problème de cohérence quoditienne-mensuelle ainsi que la distorsion selon la durée du mois. Un mois standard est défini comme étant les 20 premiers jours ouvrables du mois, sans considération des jours fériés. Il existe cinq combinaisons possibles des jours de la semaine et du mois standardisé, soit une pour chaque possibilité que le mois débute selon un jour de semaine (lundi, mardi,...,vendredi). Dépendamment de la spécification additive, multiplicative ou généralisée mixte, cette technique implique 1 à 5 contraintes.

En premier lieu, le flux monétaire total du mois TCF_M est remplacé par son équivalent standardisé :

$$STCF_M = \sum_{t=1}^{20} CF_t \tag{10}$$

Lors de l'utilisation d'un mois standard, la régression (8) doit être appliquée sur la variable dépendante $CF_t/STCF_M$ au lieu de CF_t/TCF_M. L'estimation des paramètres de cette équation est contrainte pour que la somme des proportions quotidiennes du mois égale à 1. Pour la spécification généralisée mixte (7),

les 5 contraintes sont formulées comme suit :

$$\sum_{n=1}^{N} a_n^{'} + \sum_{C_j} i_{nw} + 4 \sum_{w=1}^{5} b_w^{'} = 1, \quad j = 1,...,5 \tag{11}$$

Ici, C_1 représente la combinaison (n,w) qui se produit lorsque le premier jour ouvrable du mois est un lundi, C_2 lorsque le premier jour ouvrable du mois est un mardi, et ainsi de suite jusqu'à C_5. Ainsi, la spécification généralisée mixte contient 5 différentes contraintes selon le premier jour ouvrable du mois.

5.5 Sélection de variables et de modèle

Tel que mentionné précédemment, le modèle généralisé mixte requiert généralement plus de données afin d'obtenir des estimations fiables puisque les termes d'interaction augmentent considérablement le nombre de paramètres dans le modèle. Par le fait même, plusieurs coefficients captant l'effet du jour de la semaine, du jour ouvrable du mois et de l'interaction entre ceux-ci seront statistiquement non significatifs. Des techniques de sélection de variables seront utilisées dans le but de réduire le nombre de variables et de déterminer quels coefficients devraient être retenus ou exclus du modèle. Par la suite, la performance des modèles retenus par différentes méthodes de sélection de variables ainsi que les modèles des différentes spécifications proposées par Miller et Stone (1985) sera ultimement comparée dans la section des résultats.

Rappelons que le but de l'étude concerne le développement d'un modèle pour faire des prévisions sans qu'il soit important de tester formellement les effets des paramètres individuellement pour répondre à une question de recherche scientifique. Dans ce contexte, il importe seulement que le modèle soit performant pour prévoir les valeurs futures de la variable dépendante, soit $\frac{CF_t}{TCF_m}$ la proportion quotidienne du flux monétaire au jour t. Le choix de variables et de modèle peut alors être basé sur des critères de performance prévisionnelle future du modèle. À partir du modèle généralisé mixte, des variables seront enlevées ou ajoutées au besoin afin d'améliorer le pouvoir de prédiction. Par ailleurs, les mêmes observations ne peuvent être utilisées pour évaluer la performance d'un modèle et pour estimer les paramètres du modèle. Ce principe fondamental permet d'éviter la surestimation de la performance. En d'autres mots, si les observations qui ont servi à ajuster le modèle sont également utilisées pour évaluer sa performance, l'estimation de l'erreur de prédiction du modèle sera biaisée à la baisse. Le modèle présentera un pouvoir prédictif meilleur qu'en réalité.

Afin de contourner ce problème principal, il existe deux grandes approches pour faire de la sélection de variables ou de modèle :

— Pénaliser le critère de performance pour tenir compte de la complexité du modèle (AIC, SBC, R^2 ajusté)
— Évaluer la performance en utilisant des méthodes de rééchantillonnage

Les outils et les méthodes présentés dans cette section permettront d'obtenir une évaluation objective quant à la performance prévisionnelle des modèles.

5.5.1 Critères de performance

La performance prévisionnelle d'un modèle peut être quantifiée selon plusieurs critères, notamment le GMSE, le TMSE, le R^2, le R^2 ajusté, le AIC et le SBC. Ces critères seront comparés et s'appliquent généralement dans à peu près n'importe quel contexte de modélisation (régression linéaire, forêt aléatoire, réseau de neurones). Il est important de souligner que l'évaluation de ces critères doit être appliquée exactement aux mêmes observations pour chacun des modèles pour pouvoir les comparer sur la même base. Afin d'illustrer leurs applications dans la présente étude, les critères seront décrits selon une approche par régression.

Ainsi, supposons que l'on dispose de n observations indépendantes sur (Y, X_1,..., X_p) et que l'on a estimé un modèle $\hat{f}(X_1,\ldots,X_p)$ avec ces données pour prévoir une variable continue Y.

$$\hat{f}(X_1,...,X_p) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + ... + \hat{\beta}_p X_p$$

Dans notre contexte, il s'agit de la régression du modèle généralisé mixte avec 143 variables binaires :

$$\frac{\hat{CF_t}}{TCF_m} = \sum_{n=1}^{23} \hat{A_n} m_{tn} + \sum_{w=1}^{5} \hat{B_w} d_{tw} + \sum_{n=1}^{23} \sum_{w=1}^{5} \hat{I_{nw}} m_{tn} d_{tw} \tag{12}$$

GMSE (General Mean Squared Error)

L'erreur quadratique moyenne de généralisation (GMSE) mesure l'erreur que fait le modèle en moyenne pour l'ensemble de la population.

$$GMSE = [(Y_i - \hat{f}(X_{1i},...,X_{pi}))^2]$$

Plus cette quantité est petite, meilleur est le pouvoir prédictif du modèle. Par contre, étant donné qu'il est impossible d'obtenir les observations de toute la population, cette mesure ne peut être calculée. Il existe toutefois d'autres mesures qui tentent de l'estimer ou d'autres fonctions qui permettront de classifier les modèles dans le même ordre que le GMSE.

TMSE (Training Mean Squared Error)

L'erreur quadratique moyenne de l'échantillon d'apprentissage (TMSE) est l'équivalent du GMSE calculé sur l'échantillon.

$$TMSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{f}(X_{1i},...,X_{pi}))^2$$

D'après le principe fondamental, cette quantité n'est pas un bon estimateur du GMSE puisque les mêmes observations ont été utilisées pour ajuster le modèle et pour évaluer sa performance. Le TMSE aura tendance à diminuer lorsque le nombre de paramètres dans le modèle augmente. Autrement dit, le TMSE tend à surestimer la performance du modèle en sous-estimant le GMSE.

R^2 (coefficient de détermination)

Le R^2, une mesure très connue de la régression linéaire, est souvent utilisé comme une mesure de qualité du modèle. Il quantifie la proportion de la variance de la variable dépendante Y expliquée par le modèle. Ainsi, plus la valeur du R^2 est élevée, plus le modèle est exhaustif et permet d'expliquer la variation de la variable dépendante. À l'inverse, un R^2 faible indique que le modèle explique peu la variation de la variable dépendante. Le R^2 se défini comme suit :

$$R^2 = \frac{SST - SSE}{SST} = 1 - SSE/SST$$

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y_i})^2 = n * TMSE$$

$$SST = \sum_{i=1}^{n}(y_i - \overline{y})^2 \; avec \; \overline{y} = moyenne \; des \; y_i \; de \; l'échantillon$$

Toutefois, il est démontré qu'un modèle avec plus de termes aura toujours un meilleur R^2 simplement parce qu'il a plus de variables. L'ajout d'une variable quelconque au modèle implique la diminution de la valeur de SSE. L'explication est la suivante : en régression linéaire, les estimations sont obtenues par la méthode des moindres carrés ordinaires qui consiste à minimiser la distance entre les valeurs prédites et les valeurs réelles, soit le SSE. En ajoutant une variable au modèle, le SSE va nécessairement diminuer puisque, dans le pire scénario, la nouvelle variable n'explique aucunement la variable Y et donc son coefficient serait de zéro. D'après ces équations, le SST ne varie pas en fonction du modèle et l'ajout de paramètres implique la diminution du SSE et par conséquent, l'augmentation du R^2. De plus, si le modèle contient trop de prédicteurs ou de termes d'ordre élevé, il modélisera en réalité le bruit aléatoire dans les données. Un modèle sur-ajusté produira des valeurs de R^2 très élevées, mais trompeuses étant

donné une mauvaise capacité de prédire des nouvelles observations. Ceci-dit, un modèle inutilement complexe dont la relation entre les variables explicatives et la variable dépendante est peu défendable et explicable n'est pas nécessairement souhaitable. Le R^2 ne peut donc pas être utilisé pour comparer la performance prévisionnelle de divers modèles avec un nombre de variables différent.

R^2 ajusté

Cette mesure alternative vient corriger le R^2 pour l'ajout de variables (p) dans le modèle et permet ainsi de procéder à la sélection de variables et de modèle. Il s'interprète de la même manière que le R^2, soit une valeur élevée du R^2 ajusté implique un meilleur pouvoir prévisionnel du modèle. Le R^2 ajusté est défini comme suit :

$$R^2_{ajusté} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}$$

D'après cette équation, l'ajout d'une variable diminue le SSE mais aussi la quantité (n-p). Ceci-dit, le R^2 ajusté peut soit augmenter, soit diminuer avec l'ajout de variables et constitue donc une base valide pour choisir entre des modèles avec un nombre de termes différents. Toutefois, ce critère s'applique seulement aux modèles de régression linéaire.

AIC (Akaike Information Criterion)et SBC (Schwarz Bayesian Criterion)

Ces critères découlent de la méthode du maximum de vraisemblance lors de l'estimation des paramètres du modèle. En supposant la normalité des termes d'erreurs du modèle, les estimateurs des paramètres obtenus par la méthode des moindres carrés en régression linéaire sont équivalents à ceux provenant de la méthode du maximum de vraisemblance ce qui permet d'obtenir les mesures AIC et SBC. Plus la valeur du AIC ou du SBC est petite, meilleur est le pouvoir prédictif du modèle.

Supposons un modèle ajusté avec p paramètres (incluant l'ordonnée à l'origine). Les critères AIC et SBC se définissent alors comme suit :

$$AIC = -2 * ln(L) + 2p$$

$$SBC = -2 * ln(L) + pln(n)$$

En contexte de régression linéaire et en supposant la normalité des termes d'erreurs, on obtient les formes pratiques suivantes :

$$AIC = n * ln(SSE/n) + 2p$$

$$SBC = n * ln(SSE/n) + pln(p)$$

Ces critères permettent de modifier la quantité du TMSE pour tenir compte de la complexité du modèle et se prémunir contre le sur-ajustement. Tout comme le R^2 ajusté, la valeur du AIC et du SBC peut soit augmenter, soit diminuer avec l'ajout de variables. D'après les équations ci-haut, L est la vraisemblance maximisée et p représente le nombre de paramètres dans le modèle. Ainsi, avec le AIC, la déviance du modèle $(-2 * ln(L))$ est pénalisée par 2 fois le nombre de termes. Le critère AIC représente donc un compromis entre la vraisemblance du modèle (qui diminue avec le nombre de paramètres) et la parcimonie du modèle. Ce critère va diminuer seulement si la baisse du SSE est suffisante pour compenser l'augmentation du terme $2p$ à $2(p+1)$. Le SBC pénalise le modèle par $(ln(n))$ fois le nombre de termes, ce qui est plus sévère que pour le AIC. Conséquemment, le modèle retenu par le SBC contiendra soit le même nombre, soit moins de paramètres que le modèle retenu par le AIC. Par ailleurs, alors que le R^2 et le R^2 ajusté sont disponibles que pour la régression linéaire, les critères AIC et SBC ont une plus grande portée puisqu'ils sont obtenus avec d'autres types de modèles telle que la régression logistique. Il est toutefois impossible de statuer d'avance lequel de ces deux critères sera préférable pour un contexte donné.

Les modèles de prévision pourront être comparés selon ces trois derniers critères et le nombre de paramètres du modèle retenu dépendra du critère de sélection :

AICc

Lorsque le nombre de paramètres est relativement élevé par rapport à la taille de l'échantillon, il existe une version pour corriger le biais des petits échantillon (correction de biais de second ordre), soit l'AICc. Celui-ci devrait être employé pour évaluer la performance du modèle lorsque n/p est plus petit que 40.

$$AIC_c = AIC + \frac{2 * (p+1) * (p+2)}{n-p-2} \quad si\ n/p < 40$$

Hurvich et Tsai (1989) démontrent que l'équation peut également être formulée comme suit :

$$AIC_c = n * ln(\frac{SSE}{n}) + \frac{n * (n+p)}{n-p-2}$$

Étant donné que AICc converge vers AIC lorsque la taille de l'échantillon augmente, le critère d'AICc devrait être utilisé en pratique si certains modèles

comparés contiennent beaucoup de termes par rapport au nombre d'observations.

5.5.2 Méthodes de rééchantillonnage

Pour contrer le problème de sur-ajustement, nous pouvons tenter d'estimer le GMSE directement par des méthodes de rééchantillonnage au lieu de pénaliser le TMSE à l'aide de critères comme le SBC. La technique de division de l'échantillon et la technique de validation-croisée permettent d'évaluer la performance du modèle sur des observations différentes que celles qui ont servi à l'ajuster.

Division de l'échantillon

L'idée est de séparer l'échantillon en deux parties, soit un échantillon d'apprentissage de taille n_1 et un échantillon de validation de taille n_2. Les observations contenues dans l'échantillon d'apprentissage serviront à ajuster le modèle alors que les observations contenues dans l'échantillon de validation serviront à évaluer la performance prévisionnelle du modèle. Ainsi, puisque les observations utilisées pour l'évaluation de la performance n'ont pas servi à estimer les paramètres, nous pouvons utiliser directement le critère de performance choisi sans pénaliser celui-ci pour la complexité du modèle. Par exemple, l'erreur quadratique moyenne de l'échantillon de validation serait une estimation valable du GMSE du modèle. Chaque modèle sera évalué de la sorte et celui avec le meilleur pouvoir prévisionnel sera retenu.

L'approche par division de l'échantillon a une portée encore plus générale que les critères AIC et SBC découlant de la méthode du maximum de vraisemblance et peut s'appliquer à d'autres types de modèles telles que les forêts aléatoires ou les réseaux de neurones. Ceci-dit, la division de l'échantillon a pour avantage de permettre la comparaison entre des modèles de différentes familles mais nécessite une grande taille d'échantillon initialement. En effet, le fait de diviser l'échantillon en deux parties implique que l'échantillon d'apprentissage doit contenir suffisamment d'observations pour estimer les paramètres et l'échantillon de validation nécessite également suffisamment d'observations pour évaluer de manière fiable la performance.

Validation-croisée

Lorsque la taille de l'échantillon n'est pas suffisante pour procéder à la technique de division de l'échantillon, la validation croisée est une option alternative. L'idée est similaire, soit l'estimation des paramètres sur une partie de l'échantillon et l'évaluation du modèle sur une autre partie de l'échantillon, mais

l'échantillon est divisé en K parties plutôt que deux. En pratique, il est recommandé de diviser l'échantillon soit en 5 ou 10 groupes pour de meilleurs résultats.

Les étapes suivantes décrivent le processus de validation croisée à K groupes :

1. Diviser l'échantillon en K parties P1, P2,..., PK contenant approximativement le même nombre d'observations
2. Pour j=1 à K,
   i Réserver la partie j
   ii Estimer les paramètres du modèle en combinant les observations des K-1 autres parties
   iii Calculer la mesure de performance de ce modèle sur la partie Pj
3. Faire la moyenne des K estimations de performance pour obtenir une mesure de l'estimation finale

Ce processus sera appliqué sur 5 groupes. Ainsi, chaque modèle est estimé sur 80% des données et le 20% résiduel sert à la prévision. Puisque ces étapes sont effectuées en boucle sur les 5 sous-échantillons, chaque observation est prédite une seule fois pour chaque modèle.

5.5.3 Recherche automatique du meilleur modèle

Lorsqu'il y a plusieurs variables en jeu, le nombre de modèles potentiels augmente considérablement. À titre d'information, pour un nombre p de variables, le nombre de modèles potentiels se chiffre à 2^p. En effet, chaque variable peut soit être incluse ou exclue du modèle. Ceci-dit, pour l'inspection du meilleur modèle jusqu'à 30 variables, il existe 1 073 741 824 modèles potentiels. Vous pouvez donc vous imaginer le bassin de modèles possibles avec la spécification généralisée mixte qui comporte 143 variables. Il s'avère alors fastidieux de comparer chacune des performances prévisionnelles des modèles et de sélectionner un modèle selon un certain critère sans algorithme.

Dans le contexte du projet, les spécifications additive, multiplicative et mixte comportent un nombre considérable de variables. Des algorithmes efficaces seront utilisés afin de sélectionner le meilleur modèle sans devoir examiner tous les modèles possibles. Dépendamment de la puissance de calcul, certaines procédures sont applicables à un nombre élevé de variables. Deux méthodes seront employées pour sélectionner le meilleur modèle, soit par une recherche limitée de type « stepwise » ou par une recherche générale de type « al-subset ». Les modèles retenus selon ces différentes approches seront par la suite comparés à la section des résultats.

Méthode « stepwise »

Cette méthode consiste à ajouter et retirer une variable sur la base de sa valeur p selon un critère d'entrée et un critère de sortie prédéterminés. Plus précisément, cette procédure permet de tester l'ajout de chaque variable individuellement et si cette variable additionnelle est assez significative, elle sera conservée dans le modèle.

Les étapes suivantes décrivent le processus de la méthode « stepwise » :

1. Modèle de départ : $Y = \beta_0 + \epsilon$
2. Pour i = 1 à p (où p représente le nombre de variables, soit 143 dans notre cas)
   - i Ajuster à chaque tour le modèle de départ avec l'ajout d'une variable

     $Y = \beta_0 + \beta_1 X_1 + \epsilon$

   - ii Tester la significativité de la nouvelle variable et calculer sa valeur p

     $H_0 : \beta_1 = 0$

     $H_1 : \beta_1 \neq 0$

   - iii Si pmin (la plus petite des valeurs p parmi toutes les variables) est inférieure au critère d'entrée, passer à l'étape 3. Sinon, passer à l'étape 4.
3. Le modèle résultant en supposant que pmin correspond à la variable Xj est alors :

   $Y = \beta_0 + \beta_1 X_j + \epsilon$

4. Pour i = 1 à k (où k représente le nombre de variables incluses dans le modèle jusqu'à présent)
   - iv Tester le retrait de chaque variable individuellement et retirer la moins significative si sa valeur p excède le critère de sortie.
   - v Si pmax (la plus grande des valeurs p parmi l'ensemble des variables du modèle) est supérieure au critère de sortie et qu'elle corresponde à la variable Xj, retirer cette variable du modèle.
   - vi Répéter les étapes iv à vi avec le modèle réduit jusqu'à temps qu'aucune autre variable ne peut être retirée.
5. Ajouter une nouvelle variable au modèle et répéter ces étapes jusqu'à temps qu'aucune variable ne satisfasse le critère d'entrée ou de sortie.

En d'autres mots, le fait d'ajouter des variables au fur et à mesure peut rendre certaines variables initialement dans le modèle non significatives. Suite à l'ajout d'une nouvelle variable, l'algorithme inspecte alors les variables du modèle résultant et retire les variables non significatives selon le critère de sortie déterminé. Une nouvelle variable sera alors incluse au modèle et les mêmes étapes seront répétées tant et aussi longtemps que les variables du modèle ne satisfont pas le critère de sortie et qu'aucune variable ne peut être rajoutée selon le critère d'entrée. Ainsi, avec la méthode « stepwise », une variable peut être ajoutée et retirée ultérieurement dans le processus.

Le choix du critère d'entrée et de sortie est arbitraire. Plus le critère d'entrée est élevé, plus le processus automatique laissera entrer des variables dans le modèle et conséquemment, le modèle final contiendra plus de termes. Dans un même sens, plus le critère de sortie est élevé, moins l'algorithme sera contraignant.

Méthode « all-subset»

La sélection de modèle peut se faire par une recherche plus globale, soit la méthode « all-subset ». Cette dernière permet de sélectionner le meilleur modèle pour tous les nombres de variables fixées. En considérant le modèle généralisé mixte à 143 variables, la procédure cible le meilleur modèle parmi toutes les possibilités de combinaison des variables pour un modèle à 2 variables, à 3 variables, et ainsi de suite jusqu'à 143 variables. Ceci-dit, la méthode « all-subset » propose un sous-ensemble de 143 modèles retenus selon leur R2, soit un pour chaque nombre de variables fixées. Par exemple, pour le modèle à 2 variables, supposons que le meilleur modèle parmi toutes les combinaisons possibles à 2 variables soit celui où la variable binaire du 15e jour ouvrable du mois ainsi que la variable indicatrice du jour de semaine équivalent au lundi :

$$\frac{CF_t}{TCF_m} = A_{15} m_{t15} + I_{15,1} m_{t15} d_{t1}$$

Pour le modèle à 3 variables, supposons que la procédure sélectionne le modèle avec les variables indicatrices du 8^e jour ouvrable du mois, de l'interaction entre le 1er jour ouvrable du mois et le vendredi puis du dernier jour ouvrable du mois pour prédire la proportion du total mensuel :

$$\frac{CF_t}{TCF_m} = A_8 m_{t1} + B_5 d_{t5} + I_{1,5} m_{t1} d_{t5}$$

Ainsi, la procédure « all-subset » sélectionne un sous-ensemble de modèles retenus selon leur R^2 et le meilleur modèle globalement parmi ce sous-ensemble peut alors être déterminé en sélectionnant celui qui minimise le AIC ou le SBC. Ces étapes seront illustrées à la section des résultats et il sera alors plus facile

pour le lecteur de comprendre la procédure. Il est à noter toutefois que cette procédure prend un temps colossal lorsqu'un grand nombre de variables est en jeu. Étant donné que le modèle généralisé mixte comprend un total de 143 variables, une méthode « stepwise » peut être employée initialement pour réduire le nombre de variables avant d'effectuer une recherche globale « all-subset ». Plusieurs techniques peuvent être combinées pour arriver à un modèle final. Les modèles finaux retenus par ces différentes combinaisons de techniques et selon le choix du critère de performance peuvent finalement être comparés pour sélectionner le meilleur modèle final.

5.6 Synthèse

Pour récapituler, la régression de la spécification généralisée mixte (8) est d'abord estimée puisque ses coefficients permettent également d'obtenir les proportions pour les spécifications additives et multiplicatives. Le principe fondamental stipule toutefois que l'estimation des paramètres ne peut se faire sur les mêmes observations que celles utilisées pour évaluer la performance du modèle. Ainsi une technique de division de l'échantillon et de validation croisée seront utilisées et comparées pour évaluer la performance des 3 différentes spécifications. Afin d'obtenir des prévisions qui respectent la cohérence quotidienne-mensuelle, les techniques de renormalisation et de standardisation du mois seront employées. Par la suite, les techniques de sélection de variables et de modèle seront appliquées sur le modèle généralisé mixte à 143 variables. Les modèles retenus par ces différentes techniques et selon le choix du critère de performance seront également exposés et pourront être comparés pour sélectionner le meilleur modèle final.

6 Résultats

6.1 Description statistique de l'échantillon et visualisation des données

L'échantillon comporte 9 types de flux monétaires pour lesquels des modèles de prévision distincts doivent être élaborés. Les entrées et sorties de fonds majeures sont généralement connues par IFC, tels que les montants en dividendes versés, en réassurance, en émission d'actions, etc. Ainsi, les primes, les réclamations et les comptes payables sont les principaux types de flux monétaires qui doivent être modélisés selon l'approche par distribution. Ces catégories se distinguent également selon leur mode de règlement, soit par prélèvement préautorisé (PAP), par chèques, par téléphone, par carte de crédit ou par transfert électronique (EFT). Ceci-dit, la base de données regroupe 1304 observations par type de flux monétaires, sur une période 2013 à 2017, ce qui totalise 11 736 observations de flux monétaires en devise canadienne.

Une prime d'assurance est la somme d'argent qu'un particulier ou une entreprise doit payer pour une police d'assurance. Elle constitue un revenu pour la compagnie d'assurance et représente à la fois un passif puisque l'assureur doit fournir une couverture pour les réclamations faites contre cette police. Les polices d'assurance peuvent se distinguer entre autres selon la ligne d'affaire et la catégorisation du risque. Par exemple, IFC opère au Canada dans 3 principaux segments : l'assurance des biens des particuliers, l'assurance automobile des particuliers et l'assurance des entreprises. Toutefois, la trésorerie chez IFC s'intéresse aux soldes des comptes bancaires. Les primes sont donc ségrégées selon l'entité légale et les différents comptes bancaires plutôt que la ligne d'affaire. De plus, puisque les polices sont des contrats prédéterminés, le montant et le timing de ce type de flux monétaire peuvent être plus facilement estimés. Les réclamations quant à elles posent un plus grand défi pour la prévision. À titre d'indication, une réclamation d'assurance est une demande officielle adressée à une compagnie d'assurance pour obtenir une couverture ou une indemnisation en cas de sinistre couvert ou d'événement lié à une police. La compagnie d'assurance valide la demande et, une fois approuvée, émet le paiement à l'assuré. Cependant, il est important de distinguer les pertes payées des pertes encourues. Une réclamation est encourue lorsque une perte assurée se produit et, dépendemment de plusieurs facteurs, la réclamation est payée par l'assureur dans un délai variable. Ceci-dit, le département de trésorerie chez IFC prévoient seulement les réclamations payées et les chèques de réclamation en circulation. Tout comme les primes, la prévision du niveau des soldes bancaires pour la gestion de trésorerie implique que les réclamations sont prévues et différenciées selon l'entité plutôt que la ligne d'affaire. Les payables sont des dépenses généralement administratives tels que les paiements aux fournisseurs, le loyer, etc.

En moyenne sur l'échantillon, 31 031 295 dollars de primes totales sont versées quotidiennement, 14 045 942 dollars de réclamations et 2 802 600 dollars de comptes payables sont payés quotidiennement.

Le développement d'un modèle de prévision basé sur l'approche par distribution requiert que chaque type de flux de trésorerie soit exprimé en proportion. Plus spécifiquement, les valeurs en dollars canadiens ont été transformées pour obtenir la proportion du flux monétaire quotidien par rapport au total mensuel. Ainsi, pour chaque type de flux monétaire, la variable dépendante s'exprime sous forme de proportions, CF_t/TCF_m. Les variables indépendantes sont des variables binaires indiquant le jour de la semaine, le jour ouvrable du mois et le terme d'interaction entre ces deux effets. La liste des variables ainsi que leur description est disponible en annexes. Le tableau I suivant résume les statistiques descriptives de l'échantillon total à l'étude, pour chaque type de flux de trésorerie sous forme de proportions.

TABLE I – Statistiques descriptives de l'échantillon total (n=1304 par type), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04601	0.0394	0	0.2812	0.02956	0.0233	0.0847
Primes PAP	0.04601	0.0382	0	0.2475	0.02750	0.0218	0.0952
Primes chèques	0.04601	0.0166	0	0.1567	0.04471	0.0320	0.0632
Primes téléphone	0.04601	0.0276	0	0.3032	0.04083	0.0267	0.0683
Primes cartes de crédit	0.04601	0.0157	0	0.1070	0.04234	0.0355	0.0675
Réclamations	0.04601	0.0221	0	0.1874	0.04195	0.0310	0.0653
Payables chèques	0.04601	0.0630	-0.082	0.4815	0.02426	0.0049	0.1107
Payables EFT	0.04601	0.0704	0	0.8670	0.00372	0.000	0.1374
Payables totaux	0.04601	0.0511	-0.023	0.4591	0.0300	0.0026	0.1139

D'après ces résultats, 4.601% du total mensuel est distribué quotidiennement, en moyenne, indépendamment du type de flux de trésorerie. Par contre, les types de flux monétaires semblent se distinguer d'après les autres statistiques descriptives. L'écart type des proportions quotidiennes des primes totales se chiffre à 0.0394 alors que celui des proportions quotidiennes des payables totaux s'élève à 0.0511. Ce premier aperçu des propriétés de distribution laisse croire que chaque type de flux monétaire doit être modélisé séparément afin de mieux capter chaque comportement et, par conséquent, prédire plus précisemment les flux monétaires.

Par ailleurs, puisque nous estimerons les modèles de prévision dans un premier temps sur l'échantillon d'apprentissage et que nous évaluerons dans un deuxième temps leur performance prévisionnelle sur l'échantillon de validation, les tables II et III présentent les statistiques descriptives de ces deux sous-échantillons.

TABLE II – Statistiques descriptives de l'échantillon d'apprentissage (n=1044), sur une période 2013-2016

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04598	0.0403	0	0.2812	0.0295	0.0232	0.0838
Primes PAP	0.04598	0.0381	0	0.2475	0.0275	0.0218	0.0953
Primes chèques	0.04598	0.0168	0	0.1567	0.0445	0.0318	0.0634
Primes téléphone	0.04598	0.0270	0	0.3032	0.0413	0.0276	0.0673
Primes cartes de crédit	0.04598	0.0160	0	0.1070	0.0421	0.0355	0.0687
Réclamations	0.04598	0.0196	0	0.1601	0.0431	0.0314	0.0656
Payables chèques	0.04598	0.0632	-0.0816	0.4815	0.0243	0.0056	0.1102
Payables EFT	0.04598	0.0687	-0.0002	0.8670	0.0037	0.0000	0.1386
Payables totaux	0.04598	0.0489	-0.0226	0.2933	0.0271	0.0029	0.1160

TABLE III – Statistiques descriptives de l'échantillon de validation (n=260), pour l'année 2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04615	0.0358	0	0.2260	0.0301	0.0243	0.0858
Primes PAP	0.04615	0.0387	0	0.2087	0.0272	0.0215	0.0940
Primes chèques	0.04615	0.0158	0	0.1336	0.0459	0.0326	0.0624
Primes téléphone	0.04615	0.0301	0	0.2678	0.0391	0.0251	0.0735
Primes cartes de crédit	0.04615	0.0144	0	0.0903	0.0443	0.0353	0.0640
Réclamations	0.04615	0.0302	0	0.1874	0.0388	0.0290	0.0618
Payables chèques	0.04615	0.0623	-0.039	0.3806	0.0238	0.0032	0.1170
Payables EFT	0.04615	0.0769	0	0.4931	0.0045	0.0000	0.1304
Payables totaux	0.04615	0.0594	0	0.4591	0.0285	0.0017	0.1113

Les propriétés statistiques de l'échantillon où les paramètres des régressions seront ajustés ne semblent pas différer considérablement de l'échantillon où les modèles de prévision seront évalués. Par exemple, pour les payables totaux, les proportions quotidiennes de flux monétaires se chiffrent en moyenne à 4.598% pour l'échantillon d'apprentissage comparativement à 4.615% pour l'échantillon de validation. Les autres propriétés statistiques sont sensiblement les mêmes entre ces deux échantillons pour chaque variable dépendante.

Il est également intéressant d'extraire les propriétés statistiques pour les données de flux monétaires regroupées sous le même jour de semaine afin de déceler si certains jours de semaine sont plus susceptibles de subir des encais-

sements/décaissements plus importants. Les tableaux IV à VIII présentent les statistiques descriptives pour les sous-échantillons représentant chaque jour de la semaine (lundi, mardi, ...,vendredi) selon chaque type de flux monétaires.

TABLE IV – Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au lundi (n=260), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.07167	0.0420	0	0.2674	0.0639	0.0491	0.1168
Primes PAP	0.09064	0.0444	0	0.2475	0.0826	0.0649	0.1579
Primes chèques	0.04620	0.0222	0	0.1439	0.0470	0.0126	0.0701
Primes téléphone	0.05341	0.0302	0	0.2626	0.0501	0.0280	0.0770
Primes cartes de crédit	0.06068	0.0216	0	0.1070	0.0651	0.0338	0.0769
Réclamations	0.04676	0.0293	0	0.1685	0.0431	0.0161	0.0673
Payables chèques	0.03370	0.0541	-0.082	0.4010	0.0184	0.0005	0.0685
Payables EFT	0.02429	0.0641	0	0.4887	0.0000	0.000	0.1000
Payables totaux	0.02675	0.0487	-0.023	0.4591	0.0080	0.0009	0.0991

TABLE V – Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au mardi (n=261), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04189	0.0333	0	0.2260	0.0301	0.0247	0.0791
Primes PAP	0.04016	0.0358	0	0.2203	0.0265	0.0218	0.0978
Primes chèques	0.04651	0.0164	0	0.1567	0.0445	0.0330	0.0625
Primes téléphone	0.05298	0.0210	0	0.2183	0.0495	0.0365	0.0740
Primes cartes de crédit	0.04874	0.0136	0	0.0967	0.0468	0.0387	0.0655
Réclamations	0.04568	0.0164	0	0.1299	0.0424	0.0325	0.0660
Payables chèques	0.06269	0.0768	0	0.4121	0.0345	0.0055	0.1393
Payables EFT	0.07289	0.0668	0	0.4931	0.0607	0.0000	0.1477
Payables totaux	0.07076	0.0502	0	0.4263	0.0642	0.0193	0.1278

TABLE VI – Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au mercredi (n=261), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.03456	0.0273	0	0.2059	0.0266	0.0221	0.0567
Primes PAP	0.03275	0.0229	0	0.1388	0.0250	0.0207	0.0526
Primes chèques	0.04506	0.0148	0	0.1059	0.0425	0.0306	0.0638
Primes téléphone	0.03919	0.0178	0	0.1381	0.0345	0.0256	0.0578
Primes cartes de crédit	0.04115	0.0093	0	0.0754	0.0400	0.0355	0.0514
Réclamations	0.04779	0.0249	0	0.1874	0.0394	0.0307	0.0765
Payables chèques	0.04268	0.0575	-0.0379	0.3615	0.0214	0.0054	0.1000
Payables EFT	0.02236	0.0701	0	0.8670	0.0000	0.000	0.0700
Payables totaux	0.02823	0.0420	-0.0164	0.3543	0.0117	0.0020	0.0807

TABLE VII – Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au jeudi (n=261), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.03713	0.0338	0	0.2812	0.0274	0.0232	0.0578
Primes PAP	0.03312	0.0216	0	0.1468	0.0265	0.0220	0.0495
Primes chèques	0.04493	0.0123	0	0.0893	0.0435	0.0325	0.0596
Primes téléphone	0.04286	0.0358	0	0.3032	0.0339	0.0253	0.0637
Primes cartes de crédit	0.03975	0.0077	0	0.0677	0.0392	0.0350	0.0463
Réclamations	0.04312	0.0169	0	0.1323	0.0394	0.0302	0.0573
Payables chèques	0.04046	0.0531	-0.0018	0.4815	0.0255	0.0063	0.0950
Payables EFT	0.01534	0.0402	0	0.2580	0.0000	0.0000	0.0586
Payables totaux	0.02268	0.0344	0	0.2806	0.0099	0.0028	0.0580

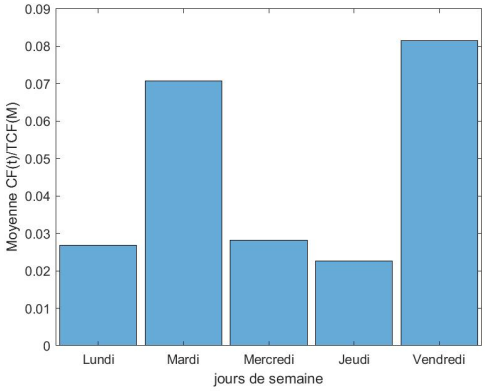
TABLE VIII – Statistiques descriptives de l'échantillon comprenant seulement les jours de semaine équivalent au vendredi (n=261), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04491	0.0464	0	0.2635	0.0282	0.0236	0.0923
Primes PAP	0.03357	0.0239	0	0.1459	0.0269	0.0221	0.0548
Primes chèques	0.04736	0.0156	0	0.1021	0.0471	0.0332	0.0631
Primes téléphone	0.04165	0.0263	0	0.2488	0.0361	0.0269	0.0624
Primes cartes de crédit	0.03980	0.0105	0	0.0775	0.0400	0.0350	0.0472
Réclamations	0.04671	0.0200	0	0.1646	0.0445	0.0333	0.0625
Payables chèques	0.05049	0.0669	-0.0119	0.3757	0.0280	0.0055	0.1283
Payables EFT	0.09510	0.0680	-0.0002	0.4146	0.0890	0.0000	0.1882
Payables totaux	0.08157	0.0465	0	0.2933	0.0758	0.0313	0.1393

À première vue, ces résultats démontrent les différences statistiques des proportions quotidiennes selon le jour de la semaine. En effet, pour les payables totaux, la proportion quotidienne moyenne de flux monétaires s'élève à 7.076% et 8.157% le mardi et le vendredi respectivement, comparativement à une proportion quotidienne moyenne de 2.5% approximativement pour les autres jours de la semaine. Ceci-dit, les proportions de flux monétaires quotidiennes semblent, en moyenne, être considérablement plus élevées lorsque le jour de la semaine est un mardi ou un vendredi pour ce type de flux de trésorerie. Les modèles de prévision devraient donc pouvoir capter un décaissement plus important pour les payables totaux lorsqu'il s'agit d'un mardi ou d'un vendredi. Toutefois, ces données statistiques ne discocient pas l'effet du jour ouvrable soulevé par Miller et Stone (1985).

De manière plus visuelle, les graphiques suivants illustrent l'effet moyen des jours de la semaine et des jours ouvrables pour les payables totaux.

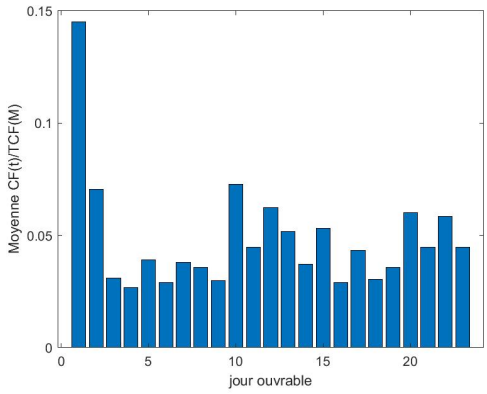
FIGURE 6 – Pattern hebdomadaire - Payables totaux



L'axe vertical représente la moyenne des proportions quotidiennes sur l'échantillon total pour chaque jour de la semaine en contexte de jours ouvrables. Ce graphique supporte également que les flux monétaires prévisionnels lors du mardi et du vendredi devraient être plus importants par rapport aux autres jours de la semaine. La réprésentation graphique du pattern hebdomadaire sert simplement d'indication générale de la structure de règlement des payables totaux.

D'après Miller et Stone (1985), le jour ouvrable du mois devrait également être considéré dans l'élaboration du modèle de prévision. Le graphique suivant illustre alors le pattern journalier moyen des payables totaux afin d'avoir une première intuition concernant la distribution des flux monétaires sur les jours ouvrables d'un mois.

FIGURE 7 – Pattern journalier moyen - Payables totaux



L'axe vertical représente la moyenne des proportions quotidiennes sur l'échantillon total pour chaque jour ouvrable du mois, allant de 1 à 23. En d'autres mots, le graphique représente la moyenne des proportions quotidiennes pour le 1^{er} jour ouvrable de chaque mois, la moyenne des proportions quotidiennes pour le 2^{e} jour ouvrable de chaque mois, et ainsi de suite sur l'échantillon total. Nous pouvons alors remarquer qu'en moyenne, le premier jour ouvrable du mois subit des décaissements plus importants comparativement aux autres jours ouvrables du mois. Encore une fois, ces moyennes par jour ouvrable ne discocient pas l'effet du jour de la semaine et servent de premières intuitions quant aux prévisions.

Dans le but d'alléger la présentation du présent rapport, l'estimation des modèles,les procédures de sélection de variables et de modèles ainsi que les résultats seront exposés et analysés uniquement pour la variable dépendante CF_t/TCF_m des payables totaux. Toutefois, les résultats concernant les autres variables dépendantes sont disponibles en annexes.

6.2 Régressions des différentes spécifications

Cette section présente les résultats des régressions des trois spécifications proposées par Miller et Stone (1985) afin de prédire les proportions quotidiennes des payables totaux.

6.2.1 Résultats du modèle additif

Rappelons que le modèle additif modélise la proportion quotidienne comme étant l'addition de l'effet du jour de la semaine et de l'effet du jour ouvrable du mois indépendemment. Ainsi, les variables binaires $x1$ à $x5$, réprésentant chacune un jour de la semaine, capturent les tendances hebdomadaires et les

variables binaires $x6$ à $x28$, représentant chacune le jour ouvrable du mois, capturent les tendances mensuelles. Une liste des variables indépendantes est disponible en annexes pour une description plus détaillée. La régression linéaire du modèle additif se définit comme suit :

$$\frac{CF_t}{TCF_m} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_{28} x_{28} + \epsilon \tag{13}$$

Le tableau suivant présente les résultats de la régression du modèle additif sur l'échantillon d'apprentissage qui regroupe 1044 observations sur les proportions quotidiennes des payables totaux.

TABLE IX – Estimation des paramètres du modèle additif à 28 variables sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
x1	-0.061280***	0.004424	-13.85	<.0001
x2	-0.012270***	0.004436	-2.77	0.0058
x3	-0.060965***	0.004441	-13.73	<.0001
x4	-0.061689***	0.004405	-14.00	<.0001
x5	0.090846***	0.010465	8.68	<.0001
x6	0.068512***	0.012704	5.39	<.0001
x7	0.039504***	0.012336	3.20	0.0014
x8	-0.003543	0.012103	-0.29	0.7698
x9	-0.012365	0.012134	-1.02	0.3084
x10	-0.015849	0.012124	-1.31	0.1914
x11	-0.018378	0.012082	-1.52	0.1285
x12	-0.013965	0.012076	-1.16	0.2478
x13	-0.016136	0.012103	-1.33	0.1828
x14	-0.014969	0.012134	-1.23	0.2176
x15	0.010988	0.012124	0.91	0.3650
x16	0.007087	0.012082	0.59	0.5576
x17	0.007474	0.012076	0.62	0.5361
x18	0.005744	0.012103	0.47	0.6352
x19	0.00602	0.012134	0.54	0.5921
x20	-0.004101	0.012124	-0.34	0.7353
x21	-0.008616	0.012082	-0.71	0.4759
x22	-0.007147	0.012076	-0.59	0.5541
x23	-0.020752*	0.012103	-1.71	0.0867
x24	-0.011308	0.012134	-0.93	0.3516
x25	-0.005283	0.012124	-0.44	0.6631
x26	0.005190	0.012276	0.42	0.6725
x27	0.013724	0.013188	1.04	0.2983
x28	0.004264	0.016482	0.26	0.7959
R^2 ajusté	0.3425			
Valeur F	21.13			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

Les variables du jour de la semaine, $x1$ à $x5$ sont toutes significatives à 1% ainsi que les variables $x6$ et $x7$ qui représentent le premier et deuxième jour ouvrable du mois respectivement. La variable $x23$ est statistiquement significative au niveau de 10% seulement. On peut donc dire que pour ces 5 variables, le coefficient est statistiquement différent de zéro. Il est également important de mentionner que certaines valeurs des coefficients sont négatives puisque les payables totaux comportent des entrées ainsi que des sorties de flux monétaires.

Toutefois, nous nous intéressons pas à expliquer l'effet de ces variables mais plutôt à développer un modèle pour faire des prévisions.

Après avoir estimé ces paramètres, nous avons évalué la performance prévisionnelle du modèle additif selon différentes approches. À partir du modèle ajusté, nous avons d'abord évaluer la performance sur l'échantillon de validation, contenant 260 observations, en estimant directement le GMSE par des techniques de rééchantillonage. La validation-croisée a été effectuée sur l'échantillon total, soit les 1304 observations. Le modèle a donc été estimé sur 5 segments différents de l'échantillon et l'évaluation de sa performance a été obtenue en calculant la moyenne des 5 GMSE sur les segments différents. La validation-croisée est particulièrement intéressante lorsque l'échantillon comporte peu d'observations par rapport au nombre d'effets à estimer, tel que pour le modèle généralisé mixte estimé un peu plus loin. Les différents critères de performance, qui pénalisent pour le nombre de paramètres, sont également des outils alternatifs pour évaluer la performance du modèle. Le tableau suivant présente une compilation des évaluations par les approches alternatives.

TABLE X – Évaluation du modèle additif à 28 variables

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0027554	0.0022386
Pénalisation des critères de performance		
R^2 ajusté	0.3147	
AIC	-8220.440	
SBC	-8185.452	
AICc	-6913.166	

6.2.2 Résultats du modèle multiplicatif

Le modèle multiplicatif retenu est celui modélisant un cycle mensuel avec une perturbation du jour de la semaine, soit l'équation 3 définie précédemment. Pour cette spécification, les effets du jour ouvrable du mois, couplés d'un ajustement pour l'effet du jour de la semaine représentent les variables indépendantes du modèle. La régression linéaire du modèle multiplicatif se définit comme suit :

$$\frac{CF_t}{TCF_m} = \beta_1 x_6 + \beta_2 x_7 + \beta_3 x_8 + ... + \beta_{143} x_{143} + \epsilon \tag{14}$$

Le tableau suivant présente les résultats de la régression du modèle multiplicatif à 138 variables.

TABLE XI – Estimation des paramètres du modèle multiplicatif à 138 variables sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
x6	0.076433***	0.018747	4.08	<.0001
x7	0.043696**	0.020830	2.10	0.0362
x8	-0.014786	0.020830	-0.71	0.4780
x9	0.008444	0.018747	0.45	0.6525
x10	0.008869	0.019667	0.45	0.6521
x11	0.013340	0.015364	0.87	0.3855
x12	0.005987	0.016402	0.37	0.7152
x13	0.042289**	0.020830	2.03	0.0426
x14	0.027364	0.018747	1.46	0.1447
x15	0.040538**	0.019667	2.06	0.0396
x16	0.055613***	0.015364	3.62	0.0003
x17	0.048424***	0.016402	2.95	0.0032
x18	0.063931***	0.020830	3.07	0.0022
x19	0.027669	0.018747	1.48	0.1403
x20	0.043468**	0.019667	2.21	0.0273
x21	0.026136*	0.015364	1.70	0.0892
x22	0.025864	0.016402	1.58	0.1152
x23	0.006094	0.020830	0.29	0.7699
x24	0.022018	0.018747	1.17	0.2405
x25	-0.008768	0.019667	-0.45	0.6558
x26	0.041461	0.016011	2.59	0.0098
x27	0.038552***	0.020830	1.85	0.0645
x28	0.041677*	0.024470	1.70	0.0889
x29	-0.025900	0.023021	-1.13	0.2609
x30	0.006636	0.022241	0.30	0.7655
x31	-0.035204	0.024023	-1.47	0.1431
x32	-0.019133	0.024023	-0.80	0.4260
x33	0.059817	0.010205	5.86	<.0001
x34	-0.017530	0.021487	-0.82	0.4148
x35	0.023097	0.024747	0.93	0.3509
x36	-0.046967*	0.024023	-1.96	0.0509
x37	-0.043131*	0.025681	-1.68	0.0934
x38	.	.	.	.
x39	0.028139	0.022241	1.27	0.2061
x40	0.021487	0.021487	1.12	0.2616
x41	-0.028445	0.024747	-1.15	0.2503
x42	-0.038900	0.024023	-1.62	0.1057
x43	.	.	.	.
x44	-0.029636	0.024023	-1.23	0.2176
x45	-0.012130	0.020303	-0.60	0.5503
x46	-0.044916**	0.019474	-2.31	0.0213
x47	-0.053360**	0.023021	-2.32	0.0207
x48	.	.	.	.
x49	-0.048858**	0.023021	-2.12	0.0341
x50	-0.020185	0.024747	-0.82	0.4149

x51	-0.054052**	0.021155	-2.56	0.0108
x52	-0.050482**	0.020361	-2.48	0.0133
x53	.	.	.	.
x54	-0.059608***	0.020361	-2.93	0.0035
x55	-0.023544	0.019474	-1.21	0.2270
x56	-0.054056**	0.021487	-2.52	0.0120
x57	-0.061354***	0.017228	-3.56	0.0004
x58	.	.	.	.
x59	-0.044257**	0.017228	-2.57	0.0104
x60	0.002843	0.021155	0.13	0.8931
x61	-0.049105**	0.020303	-2.42	0.0158
x62	-0.049069**	0.022241	-2.21	0.0276
x63	.	.	.	.
x64	-0.094691***	0.022241	-4.26	<.0001
x65	-0.046477**	0.021487	-2.16	0.0308
x66	-0.088674***	0.024747	-3.58	0.0004
x67	-0.087559***	0.024023	-3.64	0.0003
x68	.	.	.	.
x69	-0.080911***	0.024023	-3.37	0.0008
x70	-0.015040	0.020303	-0.74	0.4590
x71	-0.078965***	0.019474	-4.05	<.0001
x72	-0.079052***	0.023021	-3.43	0.0006
x73	.	.	.	.
x74	-0.074792***	0.023021	-3.25	0.0012
x75	0.010153	0.024747	0.41	0.6817
x76	-0.062770***	0.021155	-2.97	0.0031
x77	-0.057274***	0.020361	-2.81	0.0050
x78	.	.	.	.
x79	-0.087793***	0.020361	-4.31	<.0001
x80	-0.004579	0.019474	-0.24	0.8142
x81	-0.095818***	0.021487	-4.46	<.0001
x82	-0.103913***	0.017228	-6.03	<.0001
x83	.	.	.	.
x84	-0.075061***	0.017228	-4.36	<.0001
x85	-0.024769	0.021155	-1.17	0.2420
x86	-0.075917***	0.020303	-3.74	0.0002
x87	-0.072079***	0.022241	-3.24	0.0012
x88	.	.	.	.
x89	-0.104136***	0.022241	-4.68	<.0001
x90	-0.050991**	0.021487	-2.37	0.0178
x91	-0.053577**	0.024747	-2.16	0.0306
x92	-0.094218***	0.024023	-3.92	<.0001
x93	.	.	.	.
x94	-0.073913***	0.024023	-3.08	0.0022
x95	0.020276	0.020303	1.00	0.3182
x96	-0.047950**	0.019474	-2.46	0.0140
x97	-0.066913***	0.023021	-2.91	0.0037
x98	.	.	.	.
x99	-0.070305***	0.023021	-3.05	0.0023
x100	-0.041245*	0.024747	-1.67	0.0959

x101	-0.084227***	0.021155	-3.98	<.0001
x102	-0.079604***	0.020361	-3.91	<.0001
x103	.	.	.	.
x104	-0.076992***	0.020361	-3.78	0.0002
x105	-0.019036	0.019474	-0.98	0.3286
x106	-0.078611***	0.021487	-3.66	0.0003
x107	-0.054083***	0.017228	-3.14	0.0017
x108	.	.	.	.
x109	-0.072090***	0.017228	-4.18	<.0001
x110	-0.020889	0.021155	-0.99	0.3237
x111	-0.038483**	0.020303	-1.90	0.0583
x112	-0.070433***	0.022241	-3.17	0.0016
x113	.	.	.	.
x114	-0.056593**	0.022241	-2.54	0.0111
x115	-0.017895	0.021487	-0.83	0.4052
x116	-0.055685**	0.024747	-2.25	0.0247
x117	-0.037689	0.024023	-1.57	0.1170
x118	.	.	.	.
x119	-0.075038***	0.024023	-3.12	0.0018
x120	-0.018114	0.020303	-0.89	0.3725
x121	-0.057412***	0.019474	-2.95	0.0033
x122	-0.063240***	0.023021	-2.75	0.0061
x123	.	.	.	.
x124	-0.023197	0.023021	-1.01	0.3139
x125	0.050219**	0.024747	2.03	0.0427
x126	-0.025493	0.021155	-1.21	0.2285
x127	-0.024927	0.020361	-1.22	0.2212
x128	.	.	.	.
x129	-0.089617***	0.020853	-4.30	<.0001
x130	-0.034574*	0.019988	-1.73	0.0840
x131	-0.079622***	0.021954	-3.63	0.0003
x132	-0.032425*	0.019289	-1.68	0.0931
x133	.	.	.	.
x134	-0.075880**	0.031453	-2.41	0.0160
x135	0.032642	0.025681	1.27	0.2040
x136	-0.054818**	0.024023	-2.28	0.0227
x137	-0.077526***	0.025681	-3.02	0.0026
x138	.	.	.	.
x139	.	.	.	.
x140	.	.	.	.
x141	-0.084505**	0.033973	-2.49	0.0130
x142	-0.062886**	0.029839	-2.11	0.0353
x143	.	.	.	.
R^2 ajusté	0.3427			
Valeur F	5.81			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

D'après ces résultats, certains jours ouvrables du mois ainsi que certains paramètres d'ajustement pour tenir compte du jour de la semaine sont déterminants pour expliquer les proportions quotidiennes des payables totaux. En effet, plusieurs coefficients sont statistiquement différents de zéro. Par contre, plusieurs effets n'ont pu être estimés en raison d'un nombre d'observations trop faible. Ceci-dit, certaines combinaisons de jours ouvrables et de jours de semaine ne sont pas observées et ne peuvent donc pas être estimées fiablement. D'après le R^2 ajusté, le modèle multiplicatif expliquerait 34.27% de la variation des proportions des payables totaux comparativement à 34.25% pour le modèle additif. En se fiant à ce critère, le modèle multiplicatif ne semble guère plus performant que le modèle additif. Toutefois, il se peut que ce résultat soit causé par le manque d'observations de plusieurs termes d'intéraction. Ce modèle pourrait être estimé à nouveau lorsque plus d'observations seront accumulées.

TABLE XII – Évaluation du modèle multiplicatif à 138 variables

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0029914	0.0024315
Pénalisation des critères de performance		
R^2 ajusté	0.3427	
AIC	-5345.993	
SBC	-5827.601	
AICc	-5317.243	

6.2.3 Résultats du modèle généralisé mixte

Le "coffre à outils" du prévisionniste pour la résolution saisonnière a été augmenté en fournissant un ensemble de spécifications enrichi. Il est toutefois souligné qu'une spécification n'est pas nécessairement meilleure qu'une autre. Étant donné que les trois types de modèles nécessitent les mêmes observations, nous avons estimé chacune des spécifications pour ensuite établir une comparaison des prévisions. La meilleure performance indiquera quelle spécification est la plus appropriée. Ainsi, le modèle généralisé mixte, soit la spécification la plus générale, est estimé afin de le comparer aux modèles additif et multiplicatf. Le modèle mixte regroupe les variables indiquant le jour de la semaine, le jour ouvrable du mois et tous les termes d'intéraction, ce qui totalise 143 variables indépendantes.La régression linéaire du modèle généralisé mixte se définit alors comme suit :

$$\frac{CF_t}{TCF_m} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_{143} x_{143} + \epsilon \tag{15}$$

Le tableau suivant présente les résultats de la régression du modèle généralisé mixte à 143 variables.

TABLE XIII – Estimation des paramètres du modèle généralisé mixte à 143 variables sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
x1	-0.061280***	0.004424	-3.87	0.0001
x2	-0.012270	0.004436	-1.17	0.2442
x3	-0.060965**	0.004441	-2.51	0.0123
x4	-0.061689***	0.004405	-3.04	0.0024
x5	0.090846***	0.010465	6.60	<.0001
x6	0.026182	0.022833	1.15	0.2518
x7	-0.006554	0.024545	-0.27	0.7895
x8	-0.065036***	0.024545	-2.65	0.0082
x9	-0.041806*	0.022833	-1.83	0.0674
x10	-0.041382*	0.023582	-1.75	0.0796
x11	-0.036911*	0.020195	-1.83	0.0679
x12	-0.044264**	0.020982	-2.11	0.0352
x13	-0.007962	0.024545	-0.32	0.7457
x14	-0.022887	0.022833	-1.00	0.3164
x15	-0.009713	0.023582	-0.41	0.6805
x16	0.005363	0.020195	0.27	0.7906
x17	-0.001827	0.020982	-0.09	0.9306
x18	0.013681	0.024545	0.56	0.5774
x19	-0.022581	0.022833	-0.99	0.3229
x20	-0.006782	0.023582	-0.29	0.7737
x21	-0.024114	0.020195	-1.19	0.2327
x22	-0.024387	0.020982	-1.16	0.2454
x23	-0.044157	0.024545	-1.80	0.0723
x24	-0.028233	0.022833	-1.24	0.2166
x25	-0.059018	0.023582	-2.50	0.0125
x26	-0.008789	0.020683	-0.42	0.6710
x27	-0.011698	0.024545	-0.48	0.6338
x28	-0.008573	0.027653	-0.31	0.7566
x29	0.069195**	0.033524	2.06	0.0393
x30	0.042111	0.037596	1.12	0.2630
x31	0.049301	0.041269	1.19	0.2325
x32	0.073455*	0.038660	1.90	0.0577
x33	.	.	.	.
x34	0.077565**	0.032506	2.39	0.0172
x35	0.058572	0.039107	1.50	0.1345
x36	0.037538	0.041269	0.91	0.3633
x37	0.049458	0.039695	1.25	0.2131
x38	.	.	.	.
x39	0.123234***	0.033001	3.73	0.0002
x40	0.059609	0.037162	1.60	0.1090
x41	0.056060	0.041269	1.34	0.1790
x42	0.053689	0.038660	1.39	0.1652
x43	.	.	.	.

x44	0.065459**	0.034208	1.91	0.0560
x45	0.023345	0.036502	0.64	0.5226
x46	0.039589	0.038839	1.02	0.3083
x47	0.039228	0.038056	1.03	0.3029
x48	.	.	.	.
x49	0.046237	0.033524	1.39	0.1682
x50	0.015290	0.039107	0.39	0.6959
x51	0.030453	0.039695	0.77	0.4432
x52	0.042106	0.036533	1.15	0.2494
x53	.	.	.	.
x54	0.035486	0.031785	1.12	0.2645
x55	0.011931	0.036055	0.33	0.7408
x56	0.030448	0.039870	0.76	0.4452
x57	0.031235	0.034912	0.89	0.3712
x58	.	.	.	.
x59	0.050838*	0.029907	1.70	0.0895
x60	0.038318	0.036975	1.04	0.3003
x61	0.035400	0.039255	0.90	0.3674
x62	0.043519	0.037596	1.16	0.2374
x63	.	.	.	.
x64	0.000404	0.033001	0.01	0.9902
x65	-0.011002	0.037162	-0.30	0.7673
x66	-0.004169	0.041688	-0.10	0.9204
x67	0.005030	0.038660	0.13	0.8965
x68	.	.	.	.
x69	0.014184	0.034208	0.41	0.6785
x70	0.020435	0.036502	0.56	0.5757
x71	0.005540	0.038839	0.14	0.8866
x72	0.013537	0.038056	0.36	0.7221
x73	.	.	.	.
x74	0.020303	0.033524	0.61	0.5449
x75	0.045628	0.039107	1.17	0.2436
x76	0.021735	0.039695	0.55	0.5841
x77	0.035314	0.036533	0.97	0.3340
x78	.	.	.	.
x79	0.007302	0.031785	0.23	0.8184
x80	0.030896	0.036055	0.86	0.3917
x81	-0.011313	0.039870	-0.28	0.7767
x82	-0.011325	0.034912	-0.32	0.7457
x83	.	.	.	.
x84	0.020033	0.029907	0.67	0.5031
x85	0.010706	0.039255	0.29	0.7722
x86	0.008588	0.039255	0.22	0.8269
x87	0.020510	0.037596	0.55	0.5855
x88	.	.	.	.
x89	-0.009041	0.033001	-0.27	0.7842
x90	-0.015516	0.037162	-0.42	0.6764
x91	0.030928	0.041688	0.74	0.4583
x92	-0.001629	0.038660	-0.04	0.9664
x93	.	.	.	.

x94	0.021182	0.034208	0.62	0.5359
x95	0.055751	0.036502	1.53	0.1270
x96	0.036555	0.038839	0.94	0.3468
x97	0.025676	0.038056	0.67	0.5000
x98	.	.	.	.
x99	0.024790	0.033524	0.74	0.4598
x100	-0.005770	0.039107	-0.15	0.8827
x101	0.000278	0.039695	0.01	0.9944
x102	0.012985	0.036533	0.36	0.7223
x103	.	.	.	.
x104	0.018103	0.031785	0.57	0.5691
x105	0.016439	0.036055	0.46	0.6485
x106	0.005894	0.039870	0.15	0.8825
x107	0.038506	0.034912	1.10	0.2703
x108	.	.	.	.
x109	0.023005	0.029907	0.77	0.4420
x110	0.014586	0.036975	0.39	0.6933
x111	0.046022	0.039255	1.17	0.2413
x112	0.022156	0.037596	0.59	0.5558
x113	.	.	.	.
x114	0.038502	0.033001	1.17	0.2436
x115	0.017580	0.037162	0.47	0.6363
x116	0.028820	0.041688	0.69	0.4895
x117	0.054899	0.038660	1.42	0.1559
x118	.	.	.	.
x119	0.020057	0.034208	0.59	0.5578
x120	0.017361	0.036502	0.48	0.6344
x121	0.027093	0.038839	0.70	0.4856
x122	0.029348	0.038056	0.77	0.4408
x123	.	.	.	.
x124	0.071898	0.033524	2.14	0.0322
x125	0.085694	0.039107	2.19	0.0287
x126	0.059011	0.039695	1.49	0.1375
x127	0.067662	0.036533	1.85	0.0643
x128	.	.	.	.
x129	0.005477	0.032098	0.17	0.8645
x130	0.000901	0.036330	0.02	0.9802
x131	0.004883	0.040119	0.12	0.9032
x132	0.060160	0.035956	1.67	0.0946
x133	.	.	.	.
x134	0.019215	0.039695	0.48	0.6285
x135	0.068117	0.039695	1.72	0.0865
x136	0.029687	0.041269	0.72	0.4721
x137	0.015062	0.039695	0.38	0.7044
x138	.	.	.	.
x139	.	.	.	.
x140	.	.	.	.
x141	.	.	.	.
x142	0.029703	0.042459	0.70	0.4844

x143	.	.	.	.
R^2 ajusté	0.3534			
Valeur F	5.91			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

Sans suprise, plusieurs termes n'ont pu être estimés de manière fiable puisque cette spécification requiert davantage de paramètres à estimer que le modèle multiplicatif, et ce sur les mêmes observations. Ce modèle explique 35.34% des variations des proportions pour les payables totaux, soit une légère augmentation par rapport aux R^2 ajusté des deux autres spécifications. Le tableau suivant compile les différents critères de performance du modèle généralisé mixte.

TABLE XIV – Évaluation du modèle généralisé mixte à 143 variables

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0029721	0.0033437
Pénalisation des critères de performance		
R^2 ajusté	0.3534	
AIC	-5360,490	
SBC	-5827,250	
AICc	-5330,129	

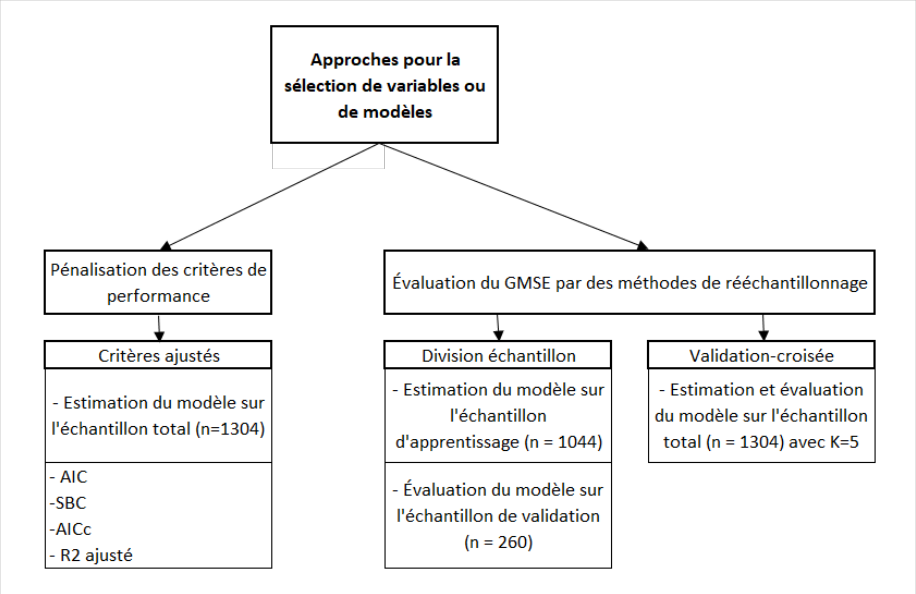
6.3 Résultats des procédures de sélection

Tel que mentionné à plusieurs reprises, le modèle généralisé mixte proposé par Miller et Stone (1985) ayant pour but de capturer la saisonnalité des effets hebdomadaires et mensuels simultanément, comporte un nombre considérable de coefficients dont plusieurs d'entre eux s'avèrent non significatifs. Les techniques de sélection de variables seront utilisées afin de déterminer quels paramètres devraient être retenus ou enlevés du modèle généralisé mixte. Cet écrémage de variables permettra de simplifier la spécification du modèle généralisé mixte pour s'en tenir aux coefficients les plus significatifs.

La méthode *stepwise* sera employée en faisant varier la valeur du critère d'entrée et de sortie, puis les différents modèles résultants contenant nombre de variables réduit seront comparés en terme de différents critères de performance. Plus particulièrement, les modèles retenus par les procédures *stepwise* à différents niveaux de valeurs p seront évalués selon des critères de performance pénalisés pour le nombre de paramètres inclus dans la régression, et selon l'évaluation directe du GMSE par la division de l'échantillon et par la validation-croisée. De plus, la performance en terme de GMSE sera évaluée en utilisant la technique

de rééchantillonage par validation-croisée. Par la suite, la méthode *stepwise* sera conjuguée à la méthode *all-subset* pour obtenir les meilleures spécifications de modèle pour chaque nombre de variables fixées. La figure suivante résume les différentes approches utilisées pour faire de la sélection de variables ou de modèles :

FIGURE 8 – Résumé des approches de sélection de variables ou de modèles



Ainsi, dépendamment du critère de performance choisi et de la technique d'évaluation du modèle employée, les modèles retenus peuvent varier. Les paragraphes suivants décrivent en détail les procédures de sélection utilisées ainsi que leurs résultats.

6.3.1 Stepwise avec Pr = 0.15

En premier lieu, la méthode *stepwise* est employée pour sélectionner un sous-ensemble de variables en fixant le critère d'entrée et de sortie à une valeur Pr de 0.15. Afin de respecter le principe fondamental concernant le surajustement, une technique de division de l'échantillon est utilisée pour ségréger l'estimation du modèle et l'évaluation de sa performance. Ainsi, la sélection de variables est effectuée sur l'échantillon d'apprentissage alors que l'évaluation du modèle résultant est effectuée sur l'échantillon de validation. Le résumé de la procédure de sélection avec la méthode *stepwise* est présenté dans le tableau ci-dessous.

TABLE XV – Résumé de la procédure de sélection stepwise, Pr=0.15

Étape	**variable entrée**	**variable supprimée**	**nb. var.**	**R^2 du modèle**	**Valeur F**	**Pr > F**
1	x5		1	0.1305	156.40	<.0001
2	x2		2	0.2467	160.58	<.0001
3	x6		3	0.3021	82.50	<.0001
4	x7		4	0.3274	39.20	<.0001
5	x39		5	0.3366	14.40	0.0002
6	x135		6	0.3433	10.53	0.0012
7	x132		7	0.3493	9.50	0.0021
8	x95		8	0.3548	8.80	0.0031
9	x15		9	0.3599	8.24	0.0042
10	x91		10	0.3652	8.65	0.0033
11	x80		11	0.3701	8.02	0.0047
12	x83		12	0.3747	7.62	0.0059
13	x17		13	0.3791	7.29	0.0070
14	x93		14	0.3826	5.75	0.0166
15	x125		15	0.3851	4.31	0.0381
16	x43		16	0.3898	4.17	0.0414
17	x128		17	0.3921	3.67	0.0510
18	x33		18	0.3962	3.82	0.0662
19	x38		19	0.3942	3.64	0.0567
20	x111		20	0.3962	3.38	0.0662
21	x96		21	0.3983	3.53	0.0605
22	x136		22	0.3999	2.80	0.0948
23	x63		23	0.4015	2.67	0.1026
24	x75		24	0.4030	2.49	0.1148
25	x23		25	0.4043	2.20	0.1380
26	x11		26	0.4057	2.42	0.1200

Ce tableau résume l'historique des étapes de la procédure *stepwise*. À l'étape 1, la variable x_5 est entrée puisque la valeur Pr de celle-ci est inférieure au critère d'entrée de 0.15. Un peu plus loin, à l'étape 6, la variable x_135 est entrée puisque sa valeur p, évaluée à 0.0012, est inférieure au critère d'entrée. Au total, 26 étapes ont été effectuées et 27 variables (incluant la constante) ont été retenues parmi les 143 variables initiales du modèle généralisé mixte. Le R^2 pour l'ajustement du modèle à 27 variables sur l'échantillon d'apprentissage est de 0.4057.

Le modèle obtenu par la procédure stepwise avec un critère d'entrée et de sortie fixé à 0.15 est défini comme suit :

$$\frac{CF_t}{TCF_m} = \beta_0 + \beta_1 x_5 + \beta_2 x_2 + \beta_3 x_6 + ... + \beta_{26} x_{11} + \epsilon \tag{16}$$

Les estimations des 27 paramètres retenus sont présentées dans le tableau

suivant. Toutes les variables sont significatives au niveau 0.15.

TABLE XVI – Estimation des paramètres du modèle à 27 variables retenues par la méthode stepwise (Pr=0.15) sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
Intercept	0.019403***	0.002029	9.56	<.0001
x5	0.069239***	0.003992	17.34	<.0001
x2	0.044498***	0.003771	11.80	<.0001
x6	0.086838***	0.008446	10.28	<.0001
x7	0.055019***	0.007370	7.47	<.0001
x39	0.053767***	0.012531	4.29	<.0001
x135	0.067110***	0.017816	3.77	0.0002
x132	0.049446***	0.014422	3.43	0.0006
x95	0.043860***	0.012826	3.42	0.0007
x15	0.016776**	0.006865	2.44	0.0147
x91	0.050768***	0.016317	3.11	0.0019
x80	0.046950***	0.015523	3.02	0.0026
x83	0.026788**	0.011647	2.30	0.0216
x17	0.016307**	0.006420	2.54	0.0112
x93	0.035106**	0.017864	1.97	0.0497
x125	0.037367**	0.017816	2.10	0.0362
x43	-0.043611**	0.017864	-2.44	0.0148
x128	-0.037593**	0.016597	-2.27	0.0237
x33	-0.039231**	0.017680	-2.22	0.0267
x38	-0.040148**	0.019287	-2.08	0.0376
x111	0.027794*	0.015280	1.82	0.0692
x96	0.020133*	0.011245	1.79	0.0737
x136	0.024148	0.015280	1.58	0.1143
x63	-0.022839*	0.012893	-1.77	0.0768
x75	0.029830	0.019055	1.57	0.1178
x23	-0.010196	0.006426	-1.59	0.1129
x11	-0.010013	0.006435	-1.56	0.1200
R^2 ajusté	0.3905			
Valeur F	26.70			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

La performance du modèle à 27 variables peut être évaluée selon les critères de performance pénalisés pour le nombre de termes ou en estimant directement le GMSE selon les 2 techniques de rééchantillonage. Le tableau suivant présente la performance du modèle pour chaque approche d'évaluation. L'évaluation de la performance se fait donc sur l'échantillon de validation en ce qui a trait à la

technique de division de l'échantillon. En ce qui concerne la validation-croisée, le modèle à 27 variables est ajusté et évalué sur une rotation de 5 différents segments de l'échantillon total d'après la procédure décrite à la section Méthodologie. Les valeurs des critères de performance pour la technique de validation-croisée correspondent alors aux moyennes des critères des 5 segments distincts (K=5).

TABLE XVII – Évaluation du modèle à 27 variables retenues par la méthode stepwise (Pr=0.15)

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0029416	0.0021943
Pénalisation des critères de performance		
R^2 ajusté	0.3905	
AIC	-5505.368	
SBC	-6417.696	
AICc	-5503.768	

Ce tableau récapitulatif pour l'évaluation de la performance sera compilé pour chaque modèle alternatif afin de faciliter leur comparaison.

6.3.2 Stepwise avec Pr = 0.60 suivi d'un all-subset

En deuxième lieu, la méthode *stepwise* est employée en fixant cette fois le critère d'entrée et de sortie à une valeur Pr de 0.60. L'augmentation da la valeur du critère a pour but de retenir un plus grand sous-ensemble de variables parmi les 143 paramètres initiaux. Cette étape préalable permet par la suite d'employer la méthode *all-subset* pour faire une recherche globale du meilleur modèle pour chaque nombre fixe de variables. Le modèle retenu sera déterminé en fonction du AIC ou du SBC des modèles ciblés par la procédure *all-subset*.

TABLE XVIII – Résumé de la procédure de sélection stepwise, Pr=0.60

Étape	**variable entrée**	**variable supprimée**	**nb. var.**	**R^2 du modèle**	**Valeur F**	**Pr > F**
1	x5		1	0.1305	156.40	<.0001
2	x2		2	0.2467	160.58	<.0001
3	x6		3	0.3021	82.50	<.0001
4	x7		4	0.3274	39.20	<.0001
5	x39		5	0.3366	14.40	0.0002
6	x135		6	0.3433	10.53	0.0012
7	x132		7	0.3493	9.50	0.0021
8	x95		8	0.3548	8.80	0.0031

9	x15		9	0.3599	8.24	0.0042
10	x91		10	0.3652	8.65	0.0033
11	x80		11	0.3701	8.02	0.0047
12	x83		12	0.3747	7.62	0.0059
13	x17		13	0.3791	7.29	0.0070
14	x93		14	0.3826	5.75	0.0166
15	x125		15	0.3851	4.31	0.0381
16	x43		16	0.3898	4.17	0.0414
17	x128		17	0.3921	3.67	0.0510
18	x33		18	0.3962	3.82	0.0662
19	x38		19	0.3942	3.64	0.0567
20	x111		20	0.3962	3.38	0.0662
21	x96		21	0.3983	3.53	0.0605
22	x136		22	0.3999	2.80	0.0948
23	x63		23	0.4015	2.67	0.1026
24	x75		24	0.4030	2.49	0.1148
25	x23		25	0.4043	2.20	0.1380
26	x11		26	0.4057	2.42	0.1200
27	x48		27	0.4068	1.92	0.1657
28	x117		28	0.4079	1.94	0.1636
29	x34		29	0.4091	1.92	0.1658
30	x53		30	0.4101	1.85	0.1742
31	x35		31	0.4109	1.40	0.2368
32	x44		32	0.4117	1.22	0.2699
33	x107		33	0.4123	1.08	0.2990
34	x25		34	0.4130	1.16	0.2815
35	x142		35	0.4136	1.16	0.2826
36	x20		36	0.4143	1.10	0.2953
37	x18		37	0.4150	1.21	0.2719
38	x133		38	0.4155	0.97	0.3255
39	x58		39	0.4160	0.72	0.3948
40	x77		40	0.4164	0.75	0.3873
41	x70		41	0.4168	0.70	0.4014
42	x14		42	0.4175	1.14	0.2850
43	x118		43	0.4179	0.67	0.4123
44	x64		44	0.4182	0.61	0.4352
45	x50		45	0.4185	0.57	0.4503
46	x42		46	0.4189	0.56	0.4532
47	x21		47	0.4192	0.53	0.4673
48	x123		48	0.4195	0.61	0.4361
49	x105		49	0.4199	0.65	0.4219
50		x133	50	0.4198	0.25	0.6177
51	x113		51	0.4202	0.78	0.3765
52	x76		52	0.4206	0.72	0.3953
53		x15	53	0.4205	0.18	0.6689
54	x98		54	0.4209	0.59	0.4422
55	x99		55	0.4212	0.50	0.4788

56		x20	56	0.4211	0.12	0.7247
57	x119		57	0.4214	0.49	0.4827
58	x13		58	0.4217	0.50	0.4779
59		x64	59	0.4216	0.22	0.6380
60	x32		60	0.4218	0.49	0.4833
61	x45		61	0.4221	0.44	0.5091
62	x82		62	0.4224	0.44	0.5059
63	x109		63	0.4226	0.35	0.5558
64	x89		64	0.4228	0.35	0.5524
65	x88		65	0.4229	0.31	0.5765
66	x29		66	0.4231	0.29	0.5875
67	x55		67	0.4233	0.29	0.5916

Ce tableau résume l'historique des étapes de la procédure *stepwise* avec un critère de Pr=0.60. À l'étape 1, la variable x_5 est entrée puisque la valeur p de celle-ci est inférieure au critère d'entrée de 0.60. Dans la même logique, la variable x_133 est entrée à l'étape 38. Toutefois, celle-ci est subséquemment sortie à l'étape 50 puisque sa valeur p, évaluée à 0.6177, est supérieure au critère de sortie. Au total, 67 étapes ont été effectuées et 60 variables (incluant la constante) ont été retenues parmi les 143 variables initiales du modèle généralisé mixte. Le R^2 pour l'ajustement du modèle à 60 variables sur l'échantillon d'apprentissage est de 0.4233.

Par la suite, la procédure *all-subset* est utilisée sur ce sous-ensemble de 60 variables retenues par la procédure *stepwise*. La tableau suivant synthétise les modèles retenus par la méthode *all-subset* pour chaque nombre fixé de variables

TABLE XIX – Résumé de la procédure de sélection all-subset, à partir de la pré-sélection stepwise Pr=0.60

Nombre dans le modèle	**R carré**	**AIC**	**SBC**	**Variables du modèle**
1	0.1305	-6204.1424	-6194.24075	x5
2	0.2467	-6351.9109	-6337.05841	x5 x2
3	0.3021	-6429.6071	-6409.80387	x5 x2 x6
4	0.3274	-6466.2735	-6441.51944	x5 x2 x6 x7
5	0.3366	-6478.6541	-6448.94923	x5 x2 x6 x7 x39
6	0.3433	-6487.2038	-6452.54812	x5 x2 x6 x7 x39 x135
7	0.3493	-6494.7345	-6455.12801	x5 x2 x6 x7 x39 x135 x132
8	0.3548	-6501.5781	-6457.02079	x5 x2 x6 x7 x39 x135 x132 x95
9	0.3598	-6507.7959	-6458.28776	x5 x2 x6 x7 x39 x135 x132 x95 x91

10	0.3646	-6513.5563	-6459.09734	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80
11	0.3690	-6518.8080	-6459.39823	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83
12	0.3728	-6523.1710	-6458.81044	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17
13	0.3767	-6527.7015	-6458.39005	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x75
14	0.3800	-6531.2581	-6456.99592	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x75
15	0.3827	-6533.7875	-6454.57446	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x75
16	0.3827	-6536.2227	-6452.05889	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x75
17	0.3877	-6538.2090	-6449.09430	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x75 x77
18	0.3900	-6540.1591	-6446.09359	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x75 x77
19	0.3923	-6542.1951	-6443.17884	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x75 x77
20	0.3946	-6544.0118	-6440.04464	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x75 x77
21	0.3964	-6545.2844	-6436.36648	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x75 x77
22	0.3984	-6546.6326	-6432.76385	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x75 x77
23	0.4001	-6547.6478	-6428.82821	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x63 x75 x77

24	0.4017	-6548.3337	-6424.56338	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x77
25	0.4033	-6549.1456	-6420.42437	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x23 x11 x63 x75 x77
26	0.4047	-6549.5910	-6415.91899	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x77
27	0.4059	-6549.7422	-6411.11939	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x77
28	0.4071	-6549.8777	-6406.30405	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x77
29	0.4083	-6549.9038	-6401.37934	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x53 x77
30	0.4094	-6549.9585	-6396.48329	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x53 x77 x76
31	0.4105	-6549.9420	-6391.51591	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x77 x76
32	0.4114	-6549.3862	-6386.00936	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x35 x77 x76

33	0.4121	-6548.6401	-6380.31243	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x35 x44 x77 x76
34	0.4121	-6548.1444	-6374.86590	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x35 x44 x77 x14 x76
35	0.4136	-6547.3901	-6369.16075	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x35 x44 x77 x14 x76 x13
36	0.4142	-6546.4968	-6363.31669	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x23 x11 x63 x75 x48 x34 x53 x35 x44 x77 x70 x14 x76 x13
37	0.4149	-6545.6865	-6357.55549	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x76
38	0.4156	-6544.8866	-6351.80481	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x76 x99
39	0.4161	-6543.9200	-6345.88737	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x76

40	0.4168	-6543.0346	-6340.05116	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x76 x99
41	0.4173	-6541.9667	-6334.03251	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x76 x99
42	0.4177	-6540.7376	-6327.85259	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x123 x76 x99
43	0.4182	-6539.5724	-6321.73658	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x21 x123 x113 x76 x99
44	0.4188	-6538.6050	-6315.81830	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x21 x123 x105 x113 x76 x99

45	0.4192	-6537.3707	-6309.63321	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x21 x123 x105 x113 x76 x99
46	0.4196	-6536.1696	-6303.48129	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x21 x123 x105 x113 x76 x98 x99 x13
47	0.4200	-6534.9226	-6297.28351	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x21 x123 x105 x113 x76 x98 x99 x13
48	0.4205	-6533.6645	-6291.07458	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x42 x21 x123 x105 x113 x76 x98 x99 x13
49	0.4209	-6532.4062	-6284.86549	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x42 x21 x123 x105 x113 x76 x98 x99 x13 x89

50	0.4213	-6531.1025	-6278.61095	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x13
51	0.4216	-6529.6464	-6272.20401	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13
52	0.4218	-6528.1644	-6272.77120	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32
53	0.4221	-6526.6243	-6259.28026	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45

54	0.4224	-6525.0917	-6252.79687	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82
55	0.4226	-6523.4585	-6246.21291	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82 x109
56	0.4228	-6521.8322	-6239.63578	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82 x109 x89
57	0.4228	-6520.1626	-6233.01534	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82 x109 x89 x88

58	0.4231	-6518.4747	-6226.37665	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82 x109 x89 x88 x29
59	0.4233	-6516.7802	-6219.73134	x5 x2 x6 x7 x39 x135 x132 x95 x91 x80 x83 x17 x93 x125 x43 x128 x33 x38 x111 x96 x136 x63 x75 x23 x11 x48 x117 x34 x53 x35 x44 x107 x25 x142 x18 x58 x77 x70 x14 x118 x50 x42 x21 x123 x105 x113 x76 x98 x99 x119 x13 x32 x45 x82 x109 x89 x88 x29 x55

Ainsi, le meilleur modèle (selon le R^2) parmi tous les modèles potentiels du sous-ensemble de variables avec une seule variable est celui avec la variable x_5. Le meilleur modèle (selon le R^2) parmi tous les modèles avec deux variables est celui avec x_5 et x_2. Il est important de préciser que, pour un nombre de variables fixé, le meilleur modèle selon le R^2, le AIC ou le SBC est le même puisque les modèles sont comparés avec le même nombre de variables sur le même échantillon. Ceci-dit, afin de déterminer le meilleur modèle parmi ces différentes spécifications, il suffit de déterminer celui qui minimise le AIC ou le SBC. Dans notre cas, le meilleur modèle en fonction du AIC est celui à 31 variables (incluant la constante) . En effet, d'après le tableau ci-haut, le modèle avec 31 variables est celui qui minimise le AIC globalement (AIC = -6549.9585). Le modèle obtenu est défini comme suit :

$$\frac{CF_t}{TCF_m} = \beta_0 + \beta_1 x_5 + \beta_2 x_2 + \beta_3 x_6 + ... + \beta_{30} x_{76} + \epsilon \tag{17}$$

Les estimations des 31 paramètres retenus, sur l'échantillon d'apprentissage, sont présentées dans le tableau suivant.

TABLE XX – Estimation des paramètres du modèle à 31 variables retenues par la méthode all-subset (AIC) sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
Intercept	0.019428***	0.002021	9.61	<.0001
x5	0.072776***	0.004191	17.37	<.0001
x2	0.044980***	0.003781	11.89	<.0001
x6	0.086670***	0.008434	10.28	<.0001
x7	0.054900***	0.007359	7.46	<.0001
x39	0.053742***	0.012515	4.29	<.0001
x135	0.066603***	0.017797	3.74	0.0002
x132	0.049421***	0.014404	3.43	0.0006
x95	0.043353***	0.012813	3.38	0.0007
x91	0.050743***	0.016297	3.11	0.0019
x80	0.046442***	0.015507	2.99	0.0028
x83	0.023226**	0.011694	1.99	0.0473
x17	0.015327**	0.006415	2.39	0.0171
x93	0.031544*	0.017883	1.76	0.0781
x125	0.036859**	0.017797	2.07	0.0386
x43	-0.047173***	0.017883	-2.64	0.0085
x128	-0.041156**	0.016620	-2.48	0.0134
x33	-0.042625**	0.017703	-2.41	0.0162
x38	-0.043591**	0.019305	-2.26	0.0242
x111	0.027769*	0.015261	1.82	0.0691
x96	0.020108*	0.011230	1.79	0.0737
x136	0.024123	0.015261	1.58	0.1143
x63	-0.026401**	0.012933	-2.04	0.0415
x75	0.046099***	0.017797	2.59	0.0097
x23	-0.014735**	0.007005	-2.10	0.0357
x11	-0.011223*	0.006435	-1.74	0.0814
x48	-0.023943	0.015606	-1.53	0.1253
x117	0.023528	0.016655	1.41	0.1581
x53	-0.023519	0.016620	-1.42	0.1573
x77	0.023652**	0.011230	2.11	0.0354
x76	0.018157	0.012515	1.45	0.1472
R^2 ajusté	0.3919			
Valeur F	23.41			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

TABLE XXI – Évaluation du modèle à 31 variables retenues par la méthode all-subset (AIC)

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0029599	0.0021965
Pénalisation des critères de performance		
R^2 ajusté	0.3919	
AIC	-5503.959	
SBC	-6396.483	
AICc	-5501.870	

Le meilleur modèle en fonction du SBC est celui à 12 variables (incluant la constante). Le critère SBC est donc plus contraignant que le critère AIC et la spécification à 12 variables minimise le SBC globalement (SBC = -6459.39823). Le modèle obtenu est défini comme suit :

$$\frac{CF_t}{TCF_m} = \beta_0 + \beta_1 x_5 + \beta_2 x_2 + \beta_3 x_6 + ... + \beta_{11} x_{83} + \epsilon \tag{18}$$

Les estimations des 12 paramètres retenus, sur l'échantillon d'apprentissage, sont présentées dans le tableau suivant.

TABLE XXII – Estimation des paramètres du modèle à 12 variables retenues par la méthode all-subset (SBC) sur l'échantillon apprentissage (n=1044)

Paramètres	**Estimation**	**Erreur Type**	**t Value**	**Pr > |t|**
Intercept	0.022177***	0.001840	12.05	<.0001
x5	0.061742***	0.003623	17.04	<.0001
x2	0.044551***	0.003703	12.03	<.0001
x6	0.076799***	0.007552	10.17	<.0001
x7	0.047571***	0.006921	6.87	<.0001
x39	0.050993***	0.012782	3.99	<.0001
x135	0.064283***	0.018184	3.54	0.0004
x132	0.046673***	0.014721	3.17	0.0016
x95	0.041033***	0.013064	3.14	0.0017
x91	0.047994***	0.016663	2.88	0.0041
x80	0.044123***	0.015832	2.79	0.0054
x83	0.031511***	0.011749	2.68	0.0074
R^2 ajusté	0.3622			
Valeur F	54.86			

*** : significatif à 1%
** : significatif à 5%
* : significatif à 10%

TABLE XXIII – Évaluation du modèle à 12 variables retenues par la méthode all-subset (SBC)

Critères	**Division échantillon**	**Validation-croisée K=5**
GMSE	0.0029798	0.0022259
Pénalisation des critères de performance		
R^2 ajusté	0.3622	
AIC	-5472.808	
SBC	-6459.398	
AICc	-5472.455	

6.4 Comparaison des modèles

Dans le but de faciliter la comparaison des différents modèles soulevés dans le présent travail, nous avons compiler un tableau récapitulatif des critères de performance selon les différentes approches pour chaque spécification. Le tableau suivant nous servira donc de guide pour déterminer quel modèle apparaît comme étant le plus performant.

FIGURE 9 – Synthèse des évaluations de performance de chaque modèle

	Additive	Multiplicative	Généralisée Mixte			
					Stepwise classique entrée=sortie=0.60 qui réduit à 60 variables suivi de all-subset	
	Aucune sélection	Aucune sélection	Aucune sélection	Stepwise classique entrée=sortie=0.15	AIC	SBC
Nombre de variables	28	138	143	27	31	11
Estimation du GMSE par division de l'échantillon	0,0027554	0,0029914*	0,0029721	0,0029416	0,0029599	0,0029798
Estimation du GMSE par validation	0,0022386	0,0024315	0,0033437*	0,0021943	0,0021965	0,0022259
AIC	-5425.34676	-5345.99347	-5360.49019	-5505,368*	-5503,959	-5472,808
SBC	-6332,724	-5827.60058	-5827,2449	-6417,696	-6396,483	-6459,398*
AICc	-5423,6308	-5317.24347	-5330,1291	-5503,768*	-5501,87	-5472,455
R2 ajusté	0,3425	0,3427	0,3534	0,3905	0,3919*	0,3622

À la lumière de ces résultats, nous pouvons constater que les différentes approches ne mènent pas toutes au même modèle final. En se basant sur l'estimation du GMSE directement avec une approche de division de l'échantillon, le meilleur modèle apparait comme étant le modèle multiplicatif à 138 variables puisque le GMSE est le plus élevé. Par contre, en se fiant sur l'AIC, c'est plutôt le modèle à 27 variables qui devrait être retenu. Ce modèle est également retenu par le critère du AICc, étant donné qu'il se démarque comme étant la plus petite valeur parmi les autres modèles. Ainsi, le choix du modèle final dépend sur quel critère le trésorier appuiera sa décision. Toutefois, nous pouvons constater que la spécification additive n'est retenue pour aucun des critères.

7 Conclusion

Le mandat du projet avait pour principal objectif de fournir un cadre théorique et structuré pour la prévision des flux monétaires du département de la trésorerie. Cette tâche complexe nécéssite à la fois de fournir des estimés plus précis et d'améliorer l'efficacité du processus tout en restant assez flexible et applicable dans un environnement dynamique. L'approche par distribution a donc été présentée ainsi que des spécifications alternatives pour résoudre la saisonalité des flux monétaires. Ce rapport a également traité un certain nombre de problèmes d'estimation pertinents à l'utilisation réussie de ces spécifications, à savoir la distorsion mensuelle et la cohérence journalière / mensuelle selon la renormalisation ou la régression contrainte. L'article de Miller et Stone (1985)

a ainsi été appliqué à un contexte d'entreprise en ajoutant des techniques de sélection pour proposer différentes spécifications avec un nomhre de variable réduit.

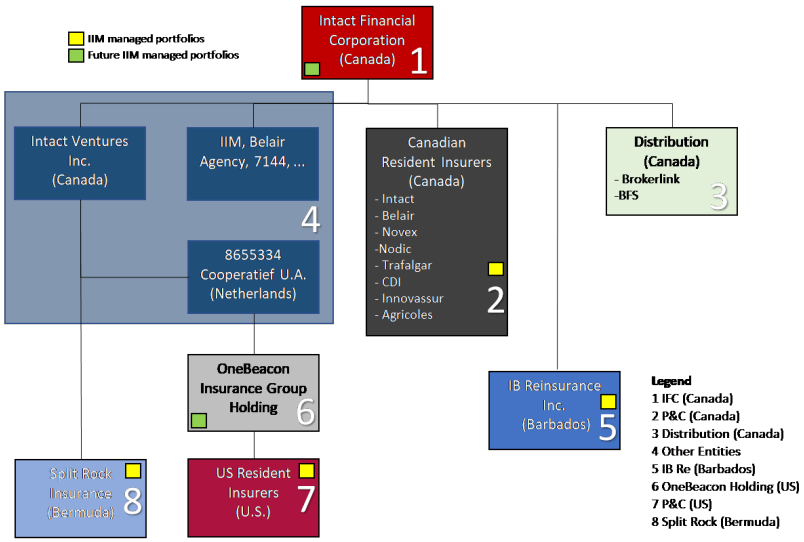
Il est toutefois important de mentionner qu'en utilisant l'approche statistique par distribution, la précision des prévisions dépend non seulement du modèle employé mais également des prévisions budgétaires utilisées allouer le montant mensuel sur les jours ouvrables. Les erreurs de prévisions provenant du budget affectent ainsi la fiabilité des prévisions quotidiennes et s'accumulent. De plus, les patterns de distribution des flux monétaires ne sont pas nécessairement constants et peuvent être appelés à changer dans le temps. Si ces changements sont répartis sur plusieurs mois, alors l'utilisation de paramètres mesurés de manière adaptative peut atténuer la détérioration des prévisions. Toutefois, un changement brusque (par exemple, lorsqu'une entreprise fait une acquisition ou qu'elle modifie ses politiques opérationnelles tels que ses conditions de crédit, son système bancaire ou ses pratiques de décaissement et de recouvrement) provoque généralement un délai avant que suffisamment de données soient disponibles pour estimer à nouveau les paramètres qui captent les cycles mensuels et hebdomadaires. Dans ces circonstances, les prévisions devront être basées sur des estimés discrétionnaires des patterns des flux de trésorerie, selon le jugement de l'analyste en question. Éventuellement, lorsque suffisamment de données deviennent disponibles, le jugement pourra être raffiné et l'estimé pourra être remplacé par une estimation statistique.

D'après l'étude de Stone et Wood (1977), la méthode de prévision basée sur une approche avec régression est probablement plus applicable pour les entreprises de taille moyenne, soit celles dont les ventes annuelles se chiffrent entre 50 et 500 milions de dollars. Par contre, les grandes entreprises qui ne sont pas dotées d'un système de prévision quotidienne sophistiqué peuvent bénéficier de cette méthodologie comme tremplin vers leur système ultime.

A Annexes

A.1 Figures complémentaires

FIGURE A.1 – Structure Corporative de IFC



A.1.1 Pattern hebdomadaire par type de flux monétaires

FIGURE A.2 – Pattern hebdomadaire - Primes totales

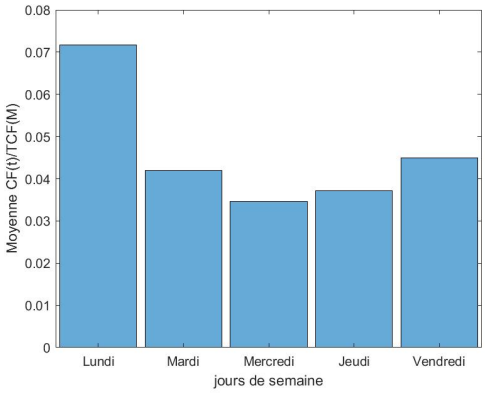


FIGURE A.3 – Pattern hebdomadaire - Primes PAP

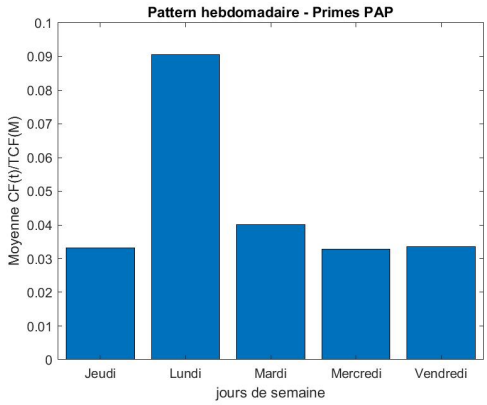


FIGURE A.4 – Pattern hebdomadaire - Primes chèques

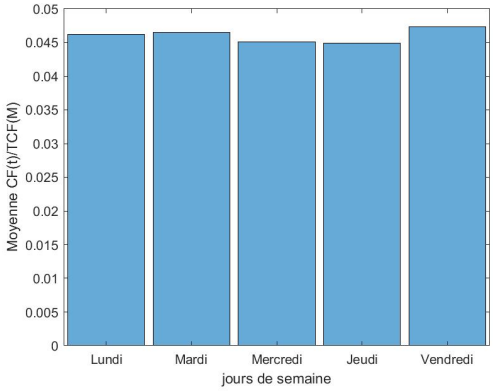


FIGURE A.5 – Pattern hebdomadaire - Primes téléphone

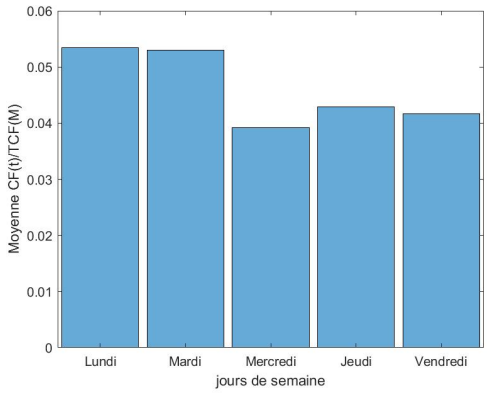


FIGURE A.6 – Pattern hebdomadaire - Primes cartes de crédit

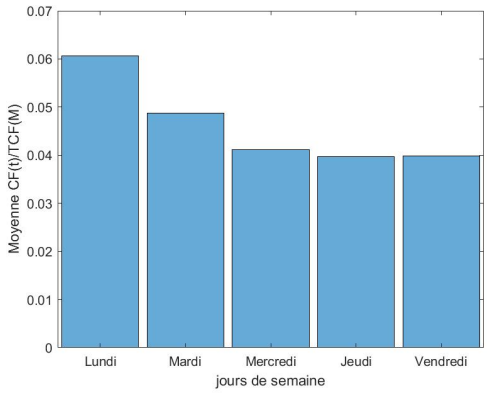


FIGURE A.7 – Pattern hebdomadaire - Réclamations

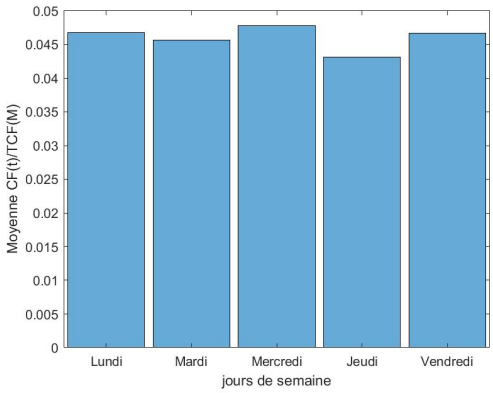


FIGURE A.8 – Pattern hebdomadaire - Payables chèques

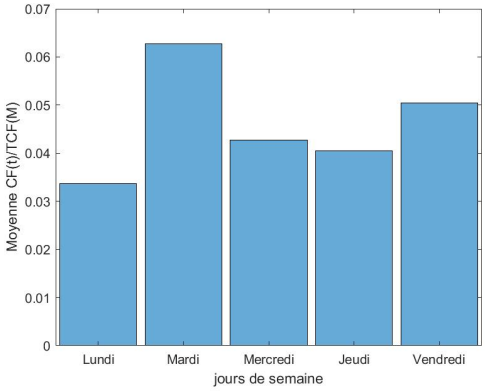
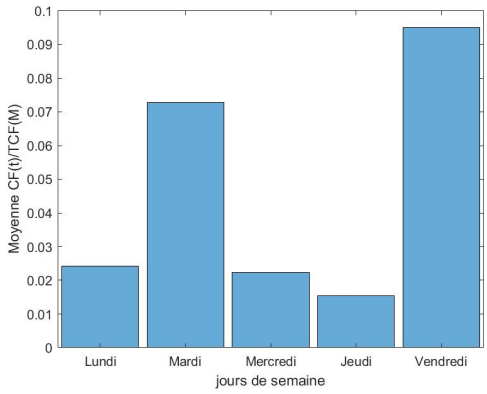


FIGURE A.9 – Pattern hebdomadaire - Payables EFT



A.1.2 Pattern journalier moyen par type de flux monétaires

FIGURE A.10 – Pattern journalier - Primes Totales

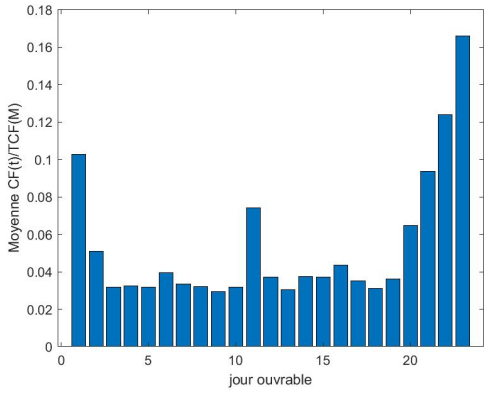


FIGURE A.11 – Pattern journalier - Primes PAP

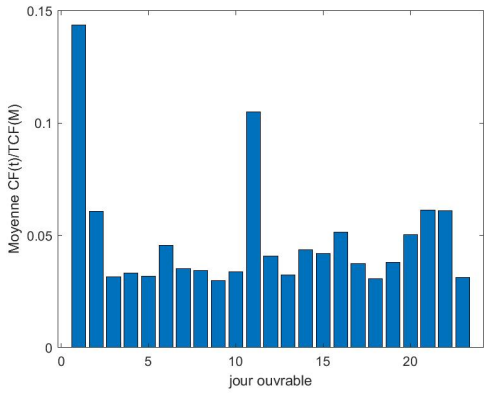


FIGURE A.12 – Pattern journalier - Primes chèques

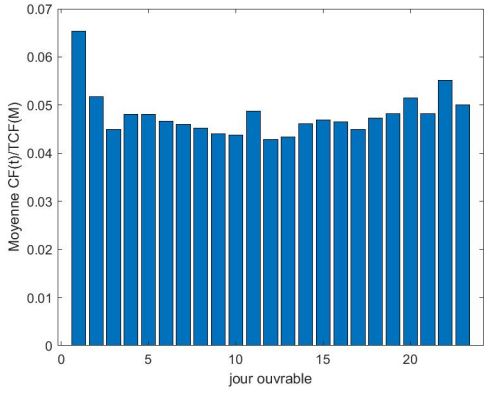


FIGURE A.13 – Pattern journalier - Primes téléphone

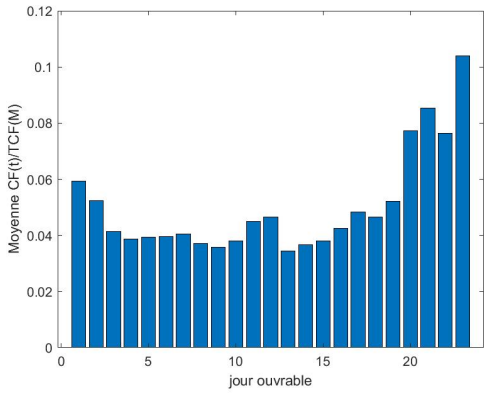


FIGURE A.14 – Pattern journalier - Primes par cartes de crédit

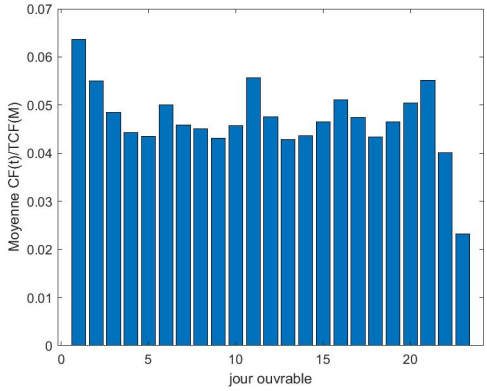


FIGURE A.15 – Pattern journalier - Réclamations

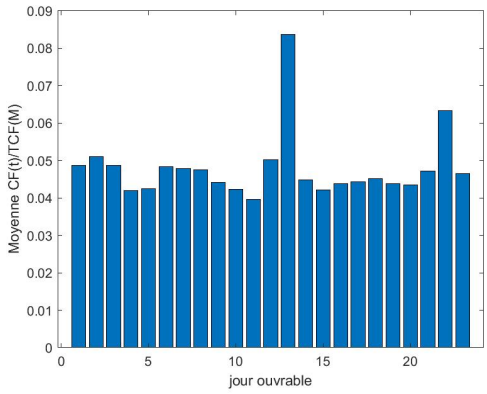


FIGURE A.16 – Pattern journalier - Payables chèques

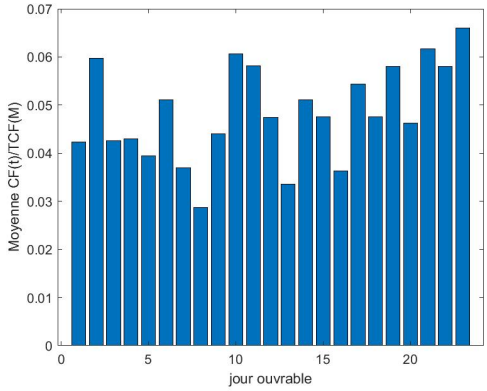
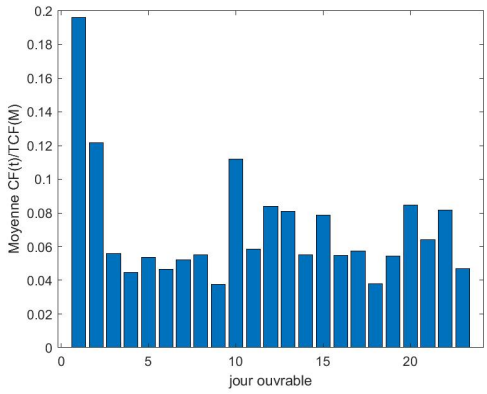


FIGURE A.17 – Pattern journalier - Payables EFT



A.1.3 Pattern journalier par mois - Payables totaux

FIGURE A.18 – Pattern journalier (Janvier) - Payables totaux

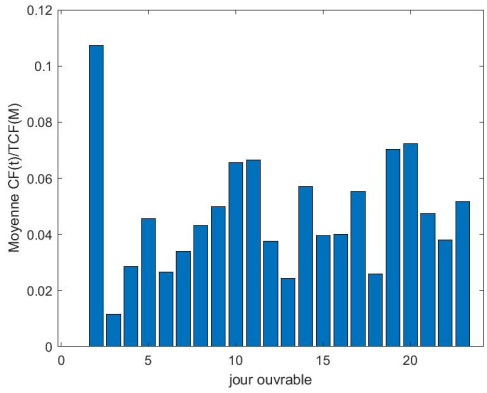


FIGURE A.19 – Pattern journalier (Février) - Payables totaux

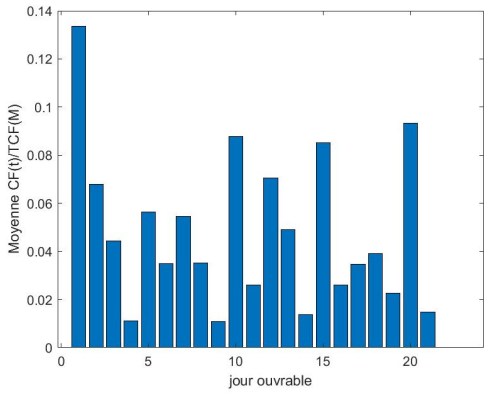


FIGURE A.20 – Pattern journalier (Mars) - Payables totaux

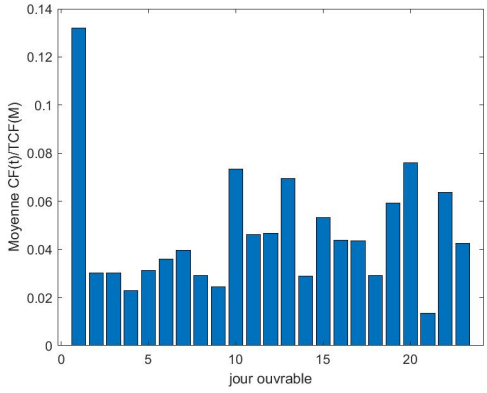


FIGURE A.21 – Pattern journalier (Avril) - Payables totaux

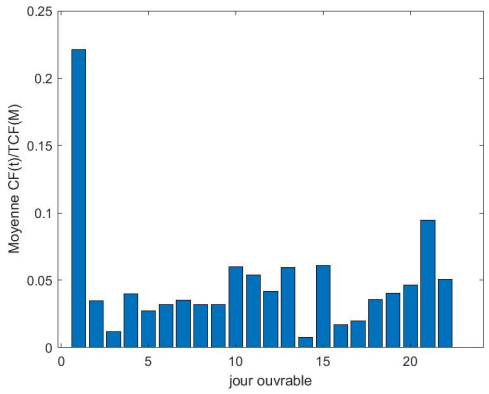


FIGURE A.22 – Pattern journalier (Mai) - Payables totaux

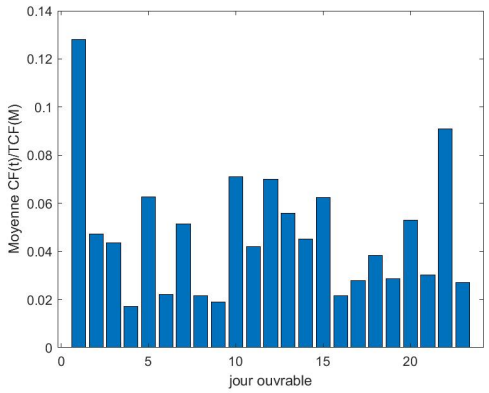


FIGURE A.23 – Pattern journalier (Juin) - Payables totaux

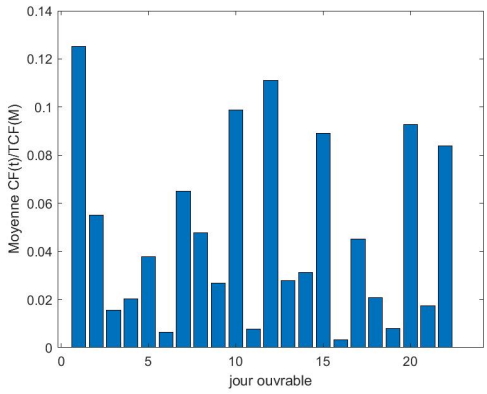


FIGURE A.24 – Pattern journalier (Juillet) - Payables totaux

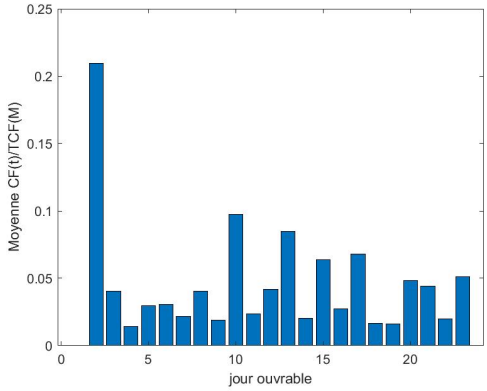


FIGURE A.25 – Pattern journalier (Août) - Payables totaux

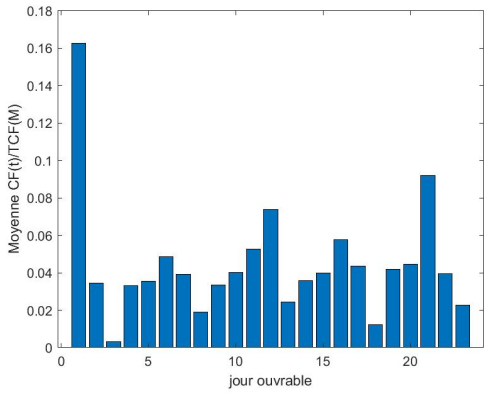


FIGURE A.26 – Pattern journalier (Septembre) - Payables totaux

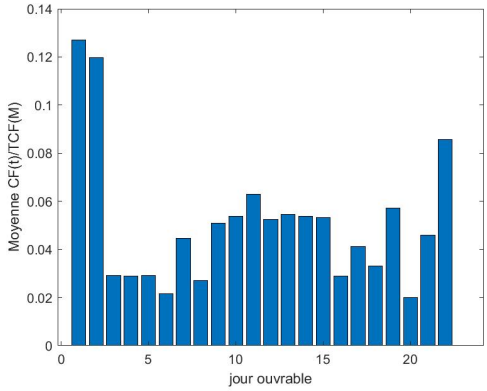


FIGURE A.27 – Pattern journalier (Octobre) - Payables totaux

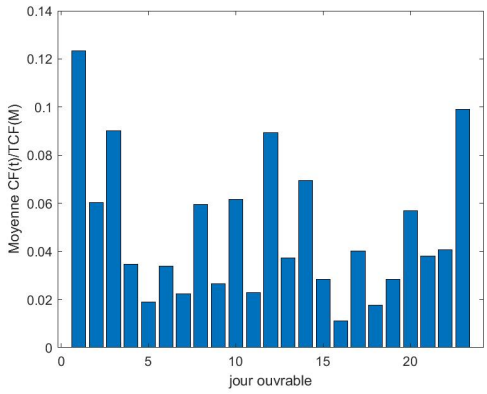


FIGURE A.28 – Pattern journalier (Novembre) - Payables totaux

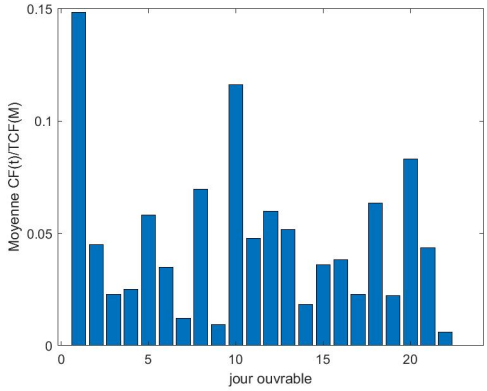
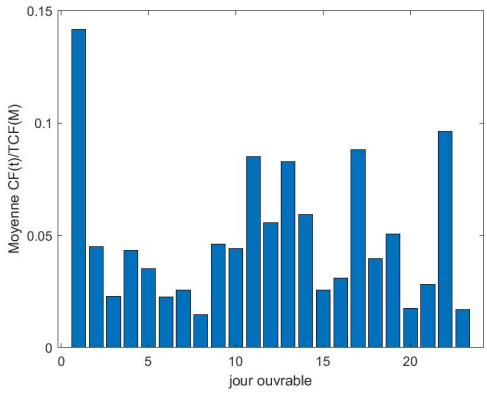


FIGURE A.29 – Pattern journalier (Décembre) - Payables totaux



A.2 Liste des variables

TABLE A.1 – Liste des variables dépendantes

Variable CF_t/TCF_m	**Description**
Primes totales	Proportion quotidienne des primes totales versées
Primes PAP	Proportion quotidienne des primes versées par prélèvement pré-autorisé
Primes chèques	Proportion quotidienne des primes versées par chèque
Primes téléphone	Proportion quotidienne des primes versées par téléphone/banking
Primes carte de crédit	Proportion quotidienne des primes versées par carte de crédit
Réclamations	Proportion quotidienne des réclamations totales payées
Payables chèques	Proportion quotidienne des comptes payables payés en chèques
Payables EFT	Proportion quotidienne des comptes payables payés par transfert électronique
Payables totaux	Proportion quotidienne des comptes payables totaux payés

TABLE A.2 – Liste des variables indépendantes binaires

Variable	**Description**	**Variable**	**Description**
x1	jour de semaine = lundi	x72	jour ouvrable=9 et jour semaine=jeudi
x2	jour de semaine = mardi	x73	jour ouvrable=9 et jour semaine=vendredi

x3	jour de semaine = mercredi	x74	jour ouvrable=10 et jour semaine=lundi
x4	jour de semaine = jeudi	x75	jour ouvrable=10 et jour semaine=mardi
x5	jour de semaine = vendredi	x76	jour ouvrable=10 et jour semaine=mercredi
x6	jour ouvrable 1 du mois	x77	jour ouvrable=10 et jour semaine=jeudi
x7	jour ouvrable 2 du mois	x78	jour ouvrable=10 et jour semaine=vendredi
x8	jour ouvrable 3 du mois	x79	jour ouvrable=11 et jour semaine=lundi
x9	jour ouvrable 4 du mois	x80	jour ouvrable=11 et jour semaine=mardi
x10	jour ouvrable 5 du mois	x81	jour ouvrable=11 et jour semaine=mercredi
x11	jour ouvrable 6 du mois	x82	jour ouvrable=11 et jour semaine=jeudi
x12	jour ouvrable 7 du mois	x83	jour ouvrable=11 et jour semaine=vendredi
x13	jour ouvrable 8 du mois	x84	jour ouvrable=12 et jour semaine=lundi
x14	jour ouvrable 9 du mois	x85	jour ouvrable=12 et jour semaine=mardi
x15	jour ouvrable 10 du mois	x86	jour ouvrable=12 et jour semaine=mercredi
x16	jour ouvrable 11 du mois	x87	jour ouvrable=12 et jour semaine=jeudi
x17	jour ouvrable 12 du mois	x88	jour ouvrable=12 et jour semaine=vendredi
x18	jour ouvrable 13 du mois	x89	jour ouvrable=13 et jour semaine=lundi
x19	jour ouvrable 14 du mois	x90	jour ouvrable=13 et jour semaine=mardi
x20	jour ouvrable 15 du mois	x91	jour ouvrable=13 et jour semaine=mercredi
x21	jour ouvrable 16 du mois	x92	jour ouvrable=13 et jour semaine=jeudi
x22	jour ouvrable 17 du mois	x93	jour ouvrable=13 et jour semaine=vendredi
x23	jour ouvrable 18 du mois	x94	jour ouvrable=14 et jour semaine=lundi
x24	jour ouvrable 19 du mois	x95	jour ouvrable=14 et jour semaine=mardi

x25	jour ouvrable 20 du mois	x96	jour ouvrable=14 et jour semaine=mercredi
x26	jour ouvrable 21 du mois	x97	jour ouvrable=14 et jour semaine=jeudi
x27	jour ouvrable 22 du mois	x98	jour ouvrable=14 et jour semaine=vendredi
x28	jour ouvrable 23 du mois	x99	jour ouvrable=15 et jour semaine=lundi
x29	jour ouvrable=1 et jour semaine=lundi	x100	jour ouvrable=15 et jour semaine=mardi
x30	jour ouvrable=1 et jour semaine=mardi	x101	jour ouvrable=15 et jour semaine=mercredi
x31	jour ouvrable=1 et jour semaine=mercredi	x102	jour ouvrable=15 et jour semaine=jeudi
x32	jour ouvrable=1 et jour semaine=jeudi	x103	jour ouvrable=15 et jour semaine=vendredi
x33	jour ouvrable=1 et jour semaine=vendredi	x104	jour ouvrable=16 et jour semaine=lundi
x34	jour ouvrable=2 et jour semaine=lundi	x105	jour ouvrable=16 et jour semaine=mardi
x35	jour ouvrable=2 et jour semaine=mardi	x106	jour ouvrable=16 et jour semaine=mercredi
x36	jour ouvrable=2 et jour semaine=mercredi	x107	jour ouvrable=16 et jour semaine=jeudi
x37	jour ouvrable=2 et jour semaine=jeudi	x108	jour ouvrable=16 et jour semaine=vendredi
x38	jour ouvrable=2 et jour semaine=vendredi	x109	jour ouvrable=17 et jour semaine=lundi
x39	jour ouvrable=3 et jour semaine=lundi	x110	jour ouvrable=17 et jour semaine=mardi
x40	jour ouvrable=3 et jour semaine=mardi	x111	jour ouvrable=17 et jour semaine=mercredi
x41	jour ouvrable=3 et jour semaine=mercredi	x112	jour ouvrable=17 et jour semaine=jeudi
x42	jour ouvrable=3 et jour semaine=jeudi	x113	jour ouvrable=17 et jour semaine=vendredi
x43	jour ouvrable=3 et jour semaine=vendredi	x114	jour ouvrable=18 et jour semaine=lundi
x44	jour ouvrable=4 et jour semaine=lundi	x115	jour ouvrable=18 et jour semaine=mardi
x45	jour ouvrable=4 et jour semaine=mardi	x116	jour ouvrable=18 et jour semaine=mercredi
x46	jour ouvrable=4 et jour semaine=mercredi	x117	jour ouvrable=18 et jour semaine=jeudi

x47	jour ouvrable=4 et jour semaine=jeudi	x118	jour ouvrable=18 et jour semaine=vendredi
x48	jour ouvrable=4 et jour semaine=vendredi	x119	jour ouvrable=19 et jour semaine=lundi
x49	jour ouvrable=5 et jour semaine=lundi	x120	jour ouvrable=19 et jour semaine=mardi
x50	jour ouvrable=5 et jour semaine=mardi	x121	jour ouvrable=19 et jour semaine=mercredi
x51	jour ouvrable=5 et jour semaine=mercredi	x122	jour ouvrable=19 et jour semaine=jeudi
x52	jour ouvrable=5 et jour semaine=jeudi	x123	jour ouvrable=19 et jour semaine=vendredi
x53	jour ouvrable=5 et jour semaine=vendredi	x124	jour ouvrable=20 et jour semaine=lundi
x54	jour ouvrable=6 et jour semaine=lundi	x125	jour ouvrable=20 et jour semaine=mardi
x55	jour ouvrable=6 et jour semaine=mardi	x126	jour ouvrable=20 et jour semaine=mercredi
x56	jour ouvrable=6 et jour semaine=mercredi	x127	jour ouvrable=20 et jour semaine=jeudi
x57	jour ouvrable=6 et jour semaine=jeudi	x128	jour ouvrable=20 et jour semaine=vendredi
x58	jour ouvrable=6 et jour semaine=vendredi	x129	jour ouvrable=21 et jour semaine=lundi
x59	jour ouvrable=7 et jour semaine=lundi	x130	jour ouvrable=21 et jour semaine=mardi
x60	jour ouvrable=7 et jour semaine=mardi	x131	jour ouvrable=21 et jour semaine=mercredi
x61	jour ouvrable=7 et jour semaine=mercredi	x132	jour ouvrable=21 et jour semaine=jeudi
x62	jour ouvrable=7 et jour semaine=jeudi	x133	jour ouvrable=21 et jour semaine=vendredi
x63	jour ouvrable=7 et jour semaine=vendredi	x134	jour ouvrable=22 et jour semaine=lundi
x64	jour ouvrable=8 et jour semaine=lundi	x135	jour ouvrable=22 et jour semaine=mardi
x65	jour ouvrable=8 et jour semaine=mardi	x136	jour ouvrable=22 et jour semaine=mercredi
x66	jour ouvrable=8 et jour semaine=mercredi	x137	jour ouvrable=22 et jour semaine=jeudi
x67	jour ouvrable=8 et jour semaine=jeudi	x138	jour ouvrable=22 et jour semaine=vendredi
x68	jour ouvrable=8 et jour semaine=vendredi	x139	jour ouvrable=23 et jour semaine=lundi

x69	jour ouvrable=9 et jour semaine=lundi	x140	jour ouvrable=23 et jour semaine=mardi
x70	jour ouvrable=9 et jour semaine=mardi	x141	jour ouvrable=23 et jour semaine=mercredi
x71	jour ouvrable=9 et jour semaine=mercredi	x142	jour ouvrable=23 et jour semaine=jeudi
		x143	jour ouvrable=23 et jour semaine=vendredi

A.3 Tableaux supplémentaires

A.3.1 Statistiques descriptives supplémentaires

TABLE A.3 – Statistiques descriptives de l'échantillon des premiers semestres (n=647), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04367	0.0395	0	0.2812	0.0300	0.0233	0.0850
Primes PAP	0.04367	0.0385	0	0.2475	0.0275	0.0219	0.0948
Primes chèques	0.04367	0.0158	0	0.1567	0.0445	0.0322	0.0639
Primes téléphone	0.04367	0.0290	0	0.3032	0.0405	0.0265	0.0683
Primes cartes de crédit	0.04367	0.0146	0	0.1070	0.0424	0.0358	0.0676
Réclamations	0.04367	0.0208	0	0.1874	0.0424	0.0313	0.0649
Payables chèques	0.04367	0.0586	-0.0816	0.4121	0.0276	0.0065	0.1050
Payables EFT	0.04367	0.0645	0	0.4887	0.0060	0.0000	0.1378
Payables totaux	0.04367	0.0508	-0.0226	0.4591	0.0255	0.0031	0.1162

TABLE A.4 – Statistiques descriptives de l'échantillon des deuxièmes semestres (n=657), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04566	0.0394	0	0.2674	0.0292	0.0233	0.0846
Primes PAP	0.04566	0.0380	0	0.2203	0.0275	0.0214	0.0952
Primes chèques	0.04566	0.0173	0	0.1439	0.0448	0.0310	0.0624
Primes téléphone	0.04566	0.0262	0	0.2678	0.0415	0.0267	0.0690
Primes cartes de crédit	0.04566	0.0167	0	0.0981	0.0423	0.0351	0.0675
Réclamations	0.04566	0.0233	0	0.1685	0.0413	0.0307	0.0654
Payables chèques	0.04566	0.0671	-0.0119	0.4815	0.0205	0.0036	0.1178
Payables EFT	0.04566	0.0757	-0.0002	0.8670	0.0028	0.0000	0.1370
Payables totaux	0.04566	0.0515	0	0.4263	0.0291	0.0021	0.1124

TABLE A.5 – Statistiques descriptives de l'échantillon des premiers trimestres (n=322), sur une période 2013-2017

Variable dépendante	Moyenne	Écart-type	Min	Max	Médiane	10 perc.	90 perc.
Primes totales	0.04658	0.0411	0	0.2812	0.0297	0.0233	0.0874
Primes PAP	0.04658	0.0391	0	0.2475	0.0276	0.0220	0.0952
Primes chèques	0.04658	0.0165	0	0.1567	0.0445	0.0325	0.0641
Primes téléphone	0.04658	0.0360	0	0.3032	0.0372	0.0250	0.0694
Primes cartes de crédit	0.04658	0.0152	0	0.1070	0.0417	0.0352	0.0687
Réclamations	0.04658	0.0208	0	0.1646	0.0431	0.0305	0.0646
Payables chèques	0.04658	0.0600	-0.0379	0.4121	0.0264	0.0062	0.1066
Payables EFT	0.04658	0.0644	0	0.4146	0.0029	0.0000	0.1381
Payables totaux	0.04658	0.0476	-0.0164	0.2933	0.0279	0.0035	0.1118

TABLE A.6 – Statistiques descriptives de l'échantillon des deuxièmes trimestres (n=325), sur une période 2013-2017

Variable dépendante	Moyenne	Écart-type	Min	Max	Médiane	10 perc.	90 perc.
Primes totales	0.04615	0.0378	0	0.2537	0.0302	0.0236	0.0847
Primes PAP	0.04615	0.0380	0	0.2057	0.0275	0.0218	0.0933
Primes chèques	0.04615	0.0152	0	0.1065	0.0447	0.0320	0.0636
Primes téléphone	0.04615	0.0199	0	0.1855	0.0427	0.0298	0.0669
Primes cartes de crédit	0.04615	0.0139	0	0.0957	0.0432	0.0365	0.0657
Réclamations	0.04615	0.0209	0	0.1874	0.0418	0.0314	0.0657
Payables chèques	0.04615	0.0573	-0.0816	0.3881	0.0285	0.0077	0.1027
Payables EFT	0.04615	0.0648	0.0000	0.4887	0.0088	0.0000	0.1372
Payables totaux	0.04615	0.0537	-0.0226	0.4591	0.0215	0.0029	0.1182

TABLE A.7 – Statistiques descriptives de l'échantillon des troisièmes trimestres (n=329), sur une période 2013-2017

Variable dépendante	Moyenne	Écart-type	Min	Max	Médiane	10 perc.	90 perc.
Primes totales	0.04559	0.0411	0	0.2674	0.0290	0.0237	0.0811
Primes PAP	0.04559	0.0375	0	0.2203	0.0274	0.0220	0.0939
Primes chèques	0.04559	0.0150	0	0.1275	0.0448	0.0325	0.0615
Primes téléphone	0.04559	0.0254	0	0.2678	0.0411	0.0284	0.0671
Primes cartes de crédit	0.04559	0.0145	0	0.0926	0.0422	0.0358	0.0651
Réclamations	0.04559	0.0223	0	0.1685	0.0409	0.0307	0.0653
Payables chèques	0.04559	0.0664	-0.0119	0.4815	0.0200	0.0035	0.1192
Payables EFT	0.04559	0.0863	0	0.8670	0.0003	0.0000	0.1435
Payables totaux	0.04559	0.0534	0	0.4263	0.0251	0.0022	0.1133

TABLE A.8 – Statistiques descriptives de l'échantillon des quatrièmes trimestres (n=328), sur une période 2013-2017

Variable dépendante	**Moyenne**	**Écart-type**	**Min**	**Max**	**Médiane**	**10 perc.**	**90 perc.**
Primes totales	0.04573	0.0378	0	0.2304	0.0298	0.0224	0.0908
Primes PAP	0.04573	0.0386	0	0.2087	0.0276	0.0210	0.0958
Primes chèques	0.04573	0.0194	0	0.1439	0.0449	0.0295	0.0630
Primes téléphone	0.04573	0.0269	0	0.2183	0.0417	0.0254	0.0699
Primes cartes de crédit	0.04573	0.0186	0	0.0981	0.0425	0.0344	0.0695
Réclamations	0.04573	0.0243	0	0.1628	0.0420	0.0305	0.0659
Payables chèques	0.04573	0.0679	-0.0039	0.4584	0.0216	0.0039	0.1174
Payables EFT	0.04573	0.0635	-0.0002	0.4388	0.0083	0.0000	0.1337
Payables totaux	0.04573	0.0497	0	0.3543	0.0324	0.0019	0.1117

Références

da Costa Moraes, M. B., M. S. Nagano et V. A. Sobreiro. 2015, «Stochastic cash flow management models : A literature review since the 1980s», dans *Decision models in engineering and management*, Springer, p. 11–28.

Gormley, F. M. et N. Meade. 2007, «The utility of cash flow forecasts in the management of corporate cash balances», *European journal of operational research*, vol. 182, n° 2, p. 923–935.

Hurvich, C. M. et C.-L. Tsai. 1989, «Regression and time series model selection in small samples», *Biometrika*, vol. 76, n° 2, p. 297–307.

Maier, S. F., D. W. Robinson et J. H. Vander Weide. 1981, «A short-term disbursement forecasting model», *Financial Management*, p. 9–20.

Miller, M. H. et D. Orr. 1966, «A model of the demand for money by firms», *The Quarterly journal of economics*, vol. 80, n° 3, p. 413–435.

Miller, T. W. et B. K. Stone. 1985, «Daily cash forecasting and seasonal resolution : Alternative models and techniques for using the distribution approach», *Journal of Financial and Quantitative Analysis*, vol. 20, n° 3, p. 335–351.

Penttinen, M. J. 1991, «Myopic and stationary solutions for stochastic cash balance problems», *European journal of operational research*, vol. 52, n° 2, p. 155–166.

Salas-Molina, F., J. A. Rodríguez-Aguilar, J. Serrà, M. Guillen et F. J. Martin. 2016, «Empirical analysis of daily cash flow time series and its implications for forecasting», *arXiv preprint arXiv :1611.04941*.

Srinivasan, V. et Y. H. Kim. 1986, «Deterministic cash flow management : state of the art and research directions», *Omega*, vol. 14, n° 2, p. 145–166.

Stone, B. K. 1976, «The payments-pattern approach to the forecasting and control of accounts receivable», *Financial Management*, p. 65–82.

Stone, B. K. et R. A. Wood. 1977, «Daily cash forecasting : a simple method for implementing the distribution approach», *Financial Management*, p. 40–50.