ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2018-ENST-0038

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Economie »

présentée et soutenue publiquement par

Yann BALGOBIN

le 9 juillet 2018

Contributions à l'économie de la vie privée et des données financières

Directeur de thèse : **Patrick WAELBROECK**
Co-encadrement de la thèse : **David BOUNIE**

Jury

M. Marvin SIRBU, Professeur, Carnegie Mellon University	Président
M. Fabrice LE GUEL, Maître de Conférences, Université Paris Sud	Rapporteur
M. Guillaume BLOT, Chief Digital Officer, Sopra Banking Software	Examinateur
Mme Lola Hernandez, Banque Centrale Européenne	Examinateur

TELECOM ParisTech
école de l'Institut Mines-Télécom - membre de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

T H È S E

THÈSE DE DOCTORAT DE
TÉLÉCOM PARISTECH

Spécialité

Économie

École doctorale Informatique, Télécommunications et électronique (Paris)

Présentée par

Yann BALGOBIN

Pour obtenir le grade de

DOCTEUR de TÉLÉCOM PARISTECH

Sujet de la thèse :

THREE ESSAYS ON THE ECONOMICS OF PRIVACY AND FINANCIAL INFORMATION

devant le jury composé de :

M. Patrick WAELBROECK	Directeur de thèse - Télécom ParisTech
M. David BOUNIE	Directeur de thèse - Télécom ParisTech
M. Marvin SIRBU	Rapporteur - Carnegie Mellon University
M. Fabrice LE GUEL	Rapporteur - Université Paris Sud
M. Guillaume BLOT	Examinateur - Sopra Banking Software
Mme. Lola HERNANDEZ	Examinateur - Banque Centrale Européenne

DECLARATION

I declare that the thesis has been composed by myself and that the work has not be submitted for any other degree or professional qualification. I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Chapters 3 and 4 have benefited from the co-authorship of Prof. David Bounie, Prof. Patrick Waelbroeck and Martin Quinn. Chapters 2 and 3 use data from the "Baromètre de la confiance des Français dans le numérique" surveys, conducted by the ACSEL (Association pour le commerce et les services en ligne) and the Caisse des dépôts et consignations (CDC).

REMERCIEMENTS

Je tiens à adresser mes premiers remerciements à Patrick Waelbroeck et David Bounie pour m'avoir encadré dans la rédaction de cette thèse pendant un peu plus de trois ans. Mes premiers pas dans la vie académique n'auraient certainement pas été les mêmes sans leur soutien sans failles ainsi que leur disponibilité. Je ne pourrais jamais les remercier assez pour ce qu'ils ont pu m'apprendre et m'apporter.

Je souhaite également remercier les personnes qui ont accepté d'être membres du jury de cette thèse : Marvin Sirbu, Lola Hernandez, Guillaume Blot et Fabrice Le Guel. Je n'oublie pas que ce dernier m'a encouragé à débuter une thèse sur l'économie numérique lors de ma deuxième année de Master.

La chaire "Valeurs et Politiques des Informations Personnelles" m'aura permis de mieux appréhender les enjeux liés aux données personnelles et à la vie privée, sujet souvent complexe mais toujours passionnant. Je remercie les membres de la chaire d'avoir accompagné mes débuts dans ce monde aux enjeux si cruciaux pendant le déroulement de cette thèse.

J'exprime ma profonde gratitude à mes collègues et amis Mattias Mano, Martin Quinn et Antoine Dubus pour les nombreuses discussions que nous avons pu avoir, leurs conseils avisés, sans oublier nos collaborations passées, présentes et, je l"espère bien, futures.

Cette thèse doit également beaucoup aux différents membres du département dont j'ai pu croiser le chemin. Ils m'auront permis de travailler dans une ambiance sympathique et chaleureuse, et je les en remercie.

Enfin, il m'est impossible de mentionner ici toutes les personnes qui, par leur présence et leur soutien inépuisable, ont participé au déroulement de cette thèse. Je souhaite que tous, famille, amis, collègues, sachent qu'ils ont grandement compté tout au long de ce travail. Pour cela, je leur en suis humblement reconnaissant.

ABSTRACT

This thesis investigates the economic consequences of consumers' control over the level of personal information they are willing to share with firms.

This subject is vital for the digital economy as many firms collect and use information about consumers to increase their revenues. Firms may face greater difficulty to generate profit from personal data. Firstly, because consumers are increasingly concerned about their privacy. Secondly, because more and more privacy-enhancing technologies (PETs) become available. We find in the thesis that the use of PETs could positively influence consumers' willingness to share personal information, enabling a data collection that takes privacy concerns into account.

We make similar conclusions in the case of financial information. Developing the use of non-bank payment instruments (and thus allowing consumers to hide some information to banks) could benefit e-commerce, leading consumers to buy more online. Finally, in a context where consumers are concerned with their privacy, banks may benefit from making screening less intrusive, as it would improve their lending strategy.

RÉSUMÉ

Cette thèse étudie les conséquences économiques du contrôle par les consommateurs de la quantité de de données personnelles qu'ils sont prêts à partager avec des firmes.

Ce sujet est d'une importance vitale pour l'économie numérique dans la mesure où de nombreuses firmes collectent et utilisent des informations sur les consommateurs afin d'augmenter leurs profits. Ces entreprises pourraient toutefois avoir de grandissantes difficultés à générer des profits à partir des données personnelles. Premièrement, de plus en plus de consommateurs sont inquiets au sujet du respect de leur vie privée en ligne. Deuxièmement, un nombre croissant d'outils permet de contrôler la collecte de données personnelles. Nous montrons dans la thèse que l'usage de tels outils ont un effet positif sur la disposition des consommateurs à partager leurs informations personnelles, permettant ainsi une collecte de données plus respectueuse de la vie privée.

Ces conclusions s'appliquent également au sujet des données financières. Encourager l'usage de moyens de paiement non-bancaires (et ainsi permettre aux consommateurs de cacher des informations aux banques) pourrait être bénéfique au commerce en ligne, en conduisant les consommateurs à faire plus d'achats. Enfin, dans un contexte où les consommateurs sont inquiets quant au respect de leur vie privée, les banques pourraient bénéficier du fait de rendre leurs pratiques de *screening* moins intrusives, dans la mesure où cela rendrait plus efficaces leurs politiques de crédit.

RÉSUMÉ ÉTENDU

Contributions à l'économie de la vie privée et des données financières

Chaque navigation sur Internet est source de traces numériques. L'exploitation de ces traces est au centre du succès de certaines des plus fameuses entreprises actuelles, à l'instar de Google ou Facebook. Grâce à leur large base d'utilisateurs[1] et leur capacité à collecter des informations, Google et Facebook sont en mesure de faire des données personnelles des sources de revenu publicitaire. En 2017, ces revenus étaient estimés à 109,65 milliards de dollars US[2], et environ 40,7 milliards de dollars US pour Facebook. [3] Fin octobre 2017, Alphabet (la maison-mère de Google) a atteint les 700 milliards de dollars US de capitalisation boursière [4], un montant supérieur au PIB estimé de la Suisse cette même année.[5]

Les firmes collectent et utilisent l'information à leur disposition depuis des années afin d'augmenter leurs revenus, par exemple en ayant une meilleure maîtrise des coûts ou bien en ayant une meilleure compréhension du marché. Ce phénomène a toutefois pris une importance sans précédent avec l'émergence d'Internet, la réduction des coûts de collecte de données et le développement de nouvelles techniques d'analyse statistique. Il est possible de nos jours de transformer les traces numériques des consommateurs en profit. Par exemple, les courtiers de données (*data brokers*) collectent des données personnelles afin de les revendre à d'autres firmes, aidant ainsi ces firmes à mieux connaître leurs consommateurs. Une étude de la *Federal Trade Commission* (FTC) américaine de 2014 estimait ainsi que ces courtiers de données avaient des informations sur "presque chaque foyer américain et chaque transaction marchande".[6]

La collecte et l'analyse de données personnelles peuvent bénéficier aux firmes de plusieurs façons. Utiliser des données personnelles peut aider les firmes à avoir des revenus plus élevés par le biais de la personnalisation de produits et de services (Acquisti and Varian, 2005). Les données personnelles peuvent être aussi à la source de réduction des coûts. Dans le cas de

[1] La part de marché de Google parmi les moteurs de recherche était de 91,74% en janvier 2018 (Statcounter - Search Engine Market Share Worldwide) tandis que quatre des cinq services de réseau social et de messagerie les plus utilisés en 2017 appartiennent à Facebook (F.Richter, Statista - "Facebook Inc. Dominates the Social Media Landscape"

[2] Statista - Annual revenue of Google from 2002 to 2017 (in billion U.S. dollars)

[3] Statista - Facebook's revenue and net income from 2007 to 2017 (in million U.S. dollars)

[4] J.C.Owens, Marketwatch "Alphabet heads for $700 billion market cap after earnings"

[5] IMF - World Economic Outlook A Shifting Global Economic Landscape

[6] Federal Trade Commission, 2014, "Data-brokers: A Call for Transparency and Accountability''.

l'industrie de la publicité en ligne par exemple, les données personnelles améliorent la correspondance entre les publicités et les besoins des consommateurs (Blattberg and Deighton, 1991). Les informations personnelles peuvent également donner aux firmes un avantage compétitif, en rendant par exemple plus facile la mise en place de la discrimination par les prix (Varian, 1985). En ligne, les vendeurs peuvent ainsi pratique un prix différent selon la localisation (Mikians et al., 2012). Les firmes peuvent également obtenir un autre avantage compétitif par le biais des données personnelles en personnalisant les services, ce qui augmentent les coûts de transaction du fait des effets de réseaux (Ball et al., 2006). L'offre de services personnalisés est rendue plus simple par l'analyse de données personnelles, les firmes pouvant prédire les préférences des consommateurs de façon plus précise (Linden et al., 2003). Ce type d'analyse peut de même aider les firmes à optimiser leurs processus d'innovation.

Afin de collecter des informations personnelles à propos des Internautes, les sites Internet et les firmes utilisent des outils come les *web trackers*, qui leur permettent de collecter des données de navigation. Les firmes ont ainsi développé des techniques dans le but de suivre les consommateurs en ligne, mais les contributions volontaires en données personnelles gardent tout de même leur importance. Par exemple, le succès des réseaux sociaux repose sur une participation active des consommateurs, principalement sous la forme d'échange de contenus entre utilisateurs. De façon générale, toutes les industries qui profitent des contenus crées par les utilisateurs (ou *user generated contents*) ne peuvent fonctionner efficacement sans une disposition des consommateurs à contribuer. Ce constat vaut également pour les plateformes come eBay, où l'évaluation des vendeurs et des acheteurs joue un rôle crucial.[7]

La disposition des consommateurs à partager des informations personnelles est limitée par leurs inquiétudes au sujet de la vie privée. Quatre types d'inquiétudes ont été identifiées par Smith et al. (1996): le volume de données collectées, l'usage secondaire non autorisé (qu'il s'agisse de la firme collectant les données ou bien une tierce partie), les accès non désirés (c'est-à-dire des données accessibles tandis qu'elles ne devraient pas l'être), et enfin les erreurs (dans la collecte des données, mais aussi la difficulté à les corriger). Ces inquiétudes concernant la vie privée sont devenus de plus en plus courantes ces dernières années. En 2013, les révélations d'Edward Snowden ont montré comment plusieurs Etats avaient mis en place des systèmes de surveillance globale.[8] Les fuites de données et le piratage de bases de données ont également été

[7]Sur ce type de plateformes, les utilisateurs sont invités à évaleuer leurs pairs afin de créer un système de réputation

[8]G.Greenwald, The Guardian - "NSA collecting phone records of millions of Verizon customers daily"

un sujet d'inquiétude grandissant. Ainsi, Equifax a découvert en juillet 2017 que des données concernant 145 500 000 citoyens américains avaient été piratées.[9]

La multiplication de tels exemples a endommagé la confiance des consommateurs au sujet du respect de leur vie privée de façon profonde. Dans un sondage américain de 2014 du *Pew Research Center*, 91% des personnes interrogées déclaraient qu'elles étaients inquiètes concernant leur perte de contrôle sur la façon dont leurs données sont collectées et utilisées par les firmes. [10] De façon similaire, 85% des personnes interrogés dans un sondage français de 2017 disaient s'inquiéter à propos de la protection de leurs données personnelles.[11]

Les consommateurs sont cependant de plus en plus actifs au sujet du respect de leur vie privée. Dans une étude de 2013 du *Pew Research Center*, 86% des personnes interrogées indiquaient qu'elles avaient pris des mesures afin d'effacer certaines de leurs informations disponibles en ligne. 55% déclaraient qu'elles avaient adopté des outils ou des stratégies afin d'éviter d'être suivies en ligne par des personnes, des firmes ou des gouvernements. [12] Dans un sondage de 2015 auprès d'un échantillon représentatif des Internautes français, 84% des personnes interrogées disaient avoir pris des mesures afin d'améliorer leur vie privée en ligne.[13] Dans le but d'avoir davantage confiance dans le respect de leur vie privée en ligne, les consommateurs ont donc adopté des outils et des stratégies de protection de la vie privée (*privacy-enhancing tools ou PETs*), comme le fait de supprimer les *cookies* de façon régulière ou d'installer certaines extensions pour navigateurs Internet.[14] Une autre solution pour les consommateurs est la mise en place de stratégies d'obfuscation, que l'on peut définir comme "l'addition délibéré d'informations ambiguës, fausses ou propices à la confusion dans le but d'interférer avec la surveillance et la collecte de données".(Brunton and Nissenbaum, 2015). Les stratégies d'obfuscation et l'usage de PETs peuvent alors être une façon pour les consommateurs de réduire l'asymétrie d'information qui caractérise souvent la collecte et l'usage de données personnelles par les firmes, la nature et l'objet de ces pratiques n'étant pas toujours transparent ou compréhensible.

Les firmes font donc face à un dilemme : s'il est avantageux à court terme d'utiliser les

[9] L.Mathews, Forbes - "Equifax Data Breach Impacts 143 Million American"

[10] Pew Research Center - "Public Perception of Privacy and Security in the post-Snowden era"

[11] CSA Research - "La protection des données personnelles"

[12] Pew Research Center - "Anonymity, Privacy, and Security Online"

[13] ACSEL and CDC - "Baromètre 2015 ACSEL-CDC de la Confiance des Français dans le numérique"

[14] Ces extensions comprennent des outils permettant l'anonymat des communications comme les *virtual private networks (VPN)* ou celles conçues pour bloquer le suivi en ligne (*tracking*) comme Ghostery ou DoNotTrack.

données personnelles afin d'en tirer du profit, cette pratique pourrait à long terme éloigner les consommateurs qui perdent de plus en plus confiance concernant le respect de leur vie privée. Les inquiétudes au sujet du respect de la vie privée peuvent avoir des effets économiques négatifs de plusieurs façons. Dans le cas de la publicité en ligne par exemple, Tucker (2014) montre comment la personnalisation des annonces peut être contreproductive si le ciblage conduit les consommateurs à s'inquiéter au sujet de leur vie privée. Ce type d'inquiétudes peut aussi réduire les dépenses en ligne (Akhter, 2012). Un autre exemple concerne les choix en matière de moyens de paiement : la possibilité de rester anonyme a ainsi un effet positif sur son adoption (von Kalckreuth et al., 2014), en partie du fait du risque de vol d'identité qui devient alors nul (Kahn and Linares-Zegarra, 2015).

Le but de cette thèse est d'étudier les différentes façons dont le contrôle croissant des consommateurs sur le partage de leurs données personnelles peut affecter la profitabilité des firmes. Ce phénomène résulte-t-il en une plus grande difficulté pour les firmes à tirer profit des données personnelles ? Ce phénomène reflète-t-il une perte de confiance si importante que les consommateurs préféreraient dès lors renoncer à des activités comme le commerce en ligne ? Existe-t-il un moyen de réconcilier la recherche de profits basée sur les données personnelles et les inquiétudes de vie privée des consommateurs ?

Une première conséquence possible de la perte de confiance des consommateurs est une difficulté grandissante des firmes à générer des profits à partir des données personnelles. La publicité en ligne est un exemple d'industrie qui pourrait souffrir de façon importante de ce phénomène. Si certaines firmes de l'économie numérique tirent leurs profits de la vente de services, la majorité d'entre elles ont conçu leurs modèles d'affaires autour de la vente d'espaces pour la publicité ciblée. La valeur de ces espaces augmente avec l'audience et la capacité des firmes à transformer l'usage de leurs services en données personnelles, qui sont ensuite utilisées pour améliorer le ciblage publicitaire. C'est ce type de pratiques qui a permis aux firmes de proposer des services gratuits aux utilisateurs : Google, Facebook, mais également les médias en ligne par exemple. Bounie et al. (2018) montrent comment les inquiétudes des consommateurs concernant leur vie privée peuvent endommager l'efficacité des algorithmes publicitaires, conduisant à une hausse du désagrément publicitaire et donc des pratiques de blocage de publicités. De plus en plus de consommateurs installent en effet des logiciels de blocage de publicités sur leurs navigateurs Internet, et dès lors toutes ces firmes ont de plus en plus de mal à transformer leur audience en profit, et donc à proposer des services gratuits. De façon plus générale, toutes

les firmes qui base leurs modèles d'affaires autour de la collection et de l'usage des données personnelles sont menacées par l'usage grandissant des *privacy-enhancing tools* et des stratégies de protection de la vie privée par les consommateurs. Cela concerne l'industrie de la publicité en ligne, mais également les réseaux sociaux, les sites marchands, tout comme les firmes qui utilisent des données afin de mieux évaluer les risques comme les banques ou les assurances.

Une autre conséquence à l'usage grandissant des PETs par les consommateurs est toutefois possible. En dehors des potentiels effets négatifs[15] il pourrait également s'agir d'un moyen de restaurer la confiance entre les firmes et les consommateurs. Cela indiquerait qu'il y a une valeur économique de la vie privée, pour les consommateurs mais également pour les firmes. En analysant comment les consommateurs sont à la recherche de moyens d'avoir plus de contrôle sur le respect de leur vie privée en ligne, cette thèse étudie la viabilité économique de cette option.

Cette thèse contribue à l'économie de la vie privée en analysant les conséquences économiques d'une gestion active du partage des données personnelles par les consommateurs. Avec le développement d'Internet et de l'analyse des données des consommateurs, de nombreux chercheurs se sont intéressés à la théorie économique de la vie privée, en étudiant les différents dilemmes pour les consommateurs comme pour les firmes. Ceci a conduit à de nombreux articles au sujet de la discrimination par les prix, le rôle des tierces parties, ou encore l'usage de données personnelles dans les techniques de *marketing* par exemple. (Acquisti et al. 2015). Des travaux empiriques (y compris en économie comportementale) se sont attaqués à des sujets similaires, mais ont également analysé les différents déterminants des préférences de vie privée. Un sujet d'intérêt a par exemple été l'évaluation économique de la vie privée par les consommateurs. Une illustration en est le fait que certains consommateurs sont prêts à payer des sites marchands plus cher s'ils sont davantage respectueux de leur vie privée. (Tsai et al. 2011). Jusqu'à présent, peu de travaux ont été dédiés sur la façon dont la profitabilité des firmes évolue en fonction de la disposition des consommateurs à partager leurs données personnelles.[16] Cette question est l'objet principal de cette thèse.

La thèse est constituée de trois chapitres. Les chapitres 2 et 3 sont basés sur des données

[15] Les consommateurs pourraient par exemple utiliser les PETs afin de piéger les algorithmes développés par les firmes. Cet usage pourrait aussi rendre les données moins disponibles, maiségalement les rendre moins fiables

[16] Il convient tout de même de mentionner Villas-Boas (2004) et Taylor (2004), qui montrent comment des consommateurs peuvent reporter leurs achats afin d'éviter d'être identifés et donc souffrir de la discrimination par les prix. Reporter leurs achats de façon stratégique permet aux consommateurs d'accéder à des prix uniformes, et inférieurs.

de sondages français.[17] Ces sondages contiennent des questions à propos de divers aspects d'Internet : vie privée, commerce en ligne, administration en ligne, etc. Le chapitre 4 présente un modèle théorique. La thèse montre qu'il est dans l'intérêt économique des firmes d'offrir aux consommateurs des manières de mieux contrôler leur vie privée et l'usage de leurs données personnelles. Donner plus de contrôle aux consommateurs est une façon pour les firmes de résoudre le dilemme auquel elles font face, entre tirer leur profit des données personnelles et la perte de confiance des consommateurs. En effet, cette thèse montre que des consommateurs ayant plus de contrôle sur leur vie privée en ligne demeure disposé à partager leurs informations personnelles.

Le chapitre 2 de cette thèse étudie de façon empirique à l'effet global de l'usage des PETs par les consommateurs sur la quantité de données personnelles disponibles pour les firmes. L'objectif principal est de savoir s'il y a un risque pour les firmes que l'usage de PETs et de stratégies d'obfuscation résulte en une plus grande difficulté pour les firmes de générer du profit à partir des données personnelles. Reflétant la nature diverse des informations personnelles, cette question concerne de nombreux acteurs de l'économie numérique : réseaux sociaux, plateformes d'*open source*, entreprises publicitaires, etc. En utilisant simultanément des données d'Internautres français représentatifs sur l'adoption de PETs et la disposition de consommateurs à partager des données personnelles, le but est d'avoir comment les deux influencent l'un l'autre. Etudier cette question permet d'évaluer si les capacités de collecte de données des firmes sont menacées par le contrôle grandissant des consommateurs sur leur vie privée en ligne. Ce chapitre montre toutefois que l'usage de PETs a un effet positif sur la disposition des consommateurs à partager des données personnelles.

Ce résultat suggère que la perte de confiance des consommateurs issue de leurs inquiétudes au sujet de leur vie privée représente une opportunité économique pour les firmes. Elles pourraient en effet bénéficier du fait de développer des produits et des services plus respectueux de la vie privée, dans la mesure où cela ne conduirait pas nécessairement à une diminution des données disponibles. Chellappa and Sin (2005) montrent de façon similaire que la confiance des consommateurs est un moyen d'obtenir davantage d'informations des consommateurs. De plus, comme suggéré par Tsai et al. (2011), le respect de la vie privée pourrait être utilisé par les firmes comme un argument de vente. Le chapitre 2 montre comment améliorer la vie privée en ligne pourrait augmenter la participation des consommateurs en terme de données personnelles,

[17] ACSEL and CDC "Baromètre 2015 ACSEL-CDC de la Confiance des Français dans le numérique"

indiquant que cela pourrait être un moyen de résoudre le dilemme auquel elles font face entre usage des données personnelles et perte de confiance.

Une autre contribution du chapitre 2 est que ses résultats vont dans le même sens que l'approche régulatoire adoptée par l'Union Européenne à propos des données personnelles et de la vie privée. [18] Ce "réglement général sur la protection des données" (RGPD) est en application depuis le 25 mai 2018. Une des caractéristiques principales du RGPD est que les firmes doivent avoir le consentement des consommateurs avant de collecter et utiliser leurs données personnelles. Nous montrons dans le chapitre 2 que si le RGPD est respecté, il pourrait être plus facile pour les firmes de collecter des informations personnelles.

La publicité en ligne, la vente en ligne et le marché des données personnelles sont des industries qui ont été le plus étudiées par l'économie de la vie privée. Même si les banques ont été parmi les premières a tiré profit de l'usage d'information à propos de leurs clients, peu de recherche académique a été dédiée aux liens entre les données financières (données de paiement, historique de crédit, etc.) et la vie privée. Les chapitres 3 et 4 se concentrent principalement sur comment un contrôle croissant des consommateurs au sujet de leur vie privée change les enjeux économiques autour des données financières. Traiter ce sujet permet de faire le lien entre plusieurs domaines de la littérature économique disjoints jusqu'à présent : vie privée, choix de moyens de paiement, achats en ligne et intermédiation financière.

Les banques et les vendeurs en ligne utilisent des données financières pour plusieurs raisons, de la détection de la fraude à la proposition d'offres personnalisés (des produits financiers par exemple). Dans les chapitres 3 et 4, la thèse étudie la notion de *financial privacy*, c'est-à-dire les inquiétudes de vie privée au sujet des données financières. Les consommateurs peuvent parfois avoir envie de cacher certaines informations financières, par exemple pour éviter d'être surveillés par les banques ou les assurances. Une autre raison peut être la volonté de cacher certains achats à leurs proches dans le cas d'un compte commun. Ainsi, 55% des personnes interrogées par KPMG dans un sondage de 2016 déclaraient avoir renoncé à certains achats en ligne pour des raisons de vie privée. [19]. De façon similaire, un sondage auprès d'un échantillon de 2000 Internautes français représentatifs indique que 35% des personnes interrogées ont renoncé à l'achat ou changé de moyen de paiement au moins une fois pour ces mêmes raisons.[20] Les con-

[18]Le texte complet peut être retrouvé à cette adresse : http://data.europa.eu/eli/reg/2016/679/oj.

[19]Crossing the line - Staying on the right side of consumer privacy report.

[20]Chaire Valeurs et Politiques des Information Personnelles - Données personnelles et confiance : quelles stratégies pour les citoyens-consommateurs en 2017 ?

sommateurs ont parfois une préférence pour certains moyens de paiement anonymes comme l'argent liquide. Aux Etats-Unis en 2015, l'argent liquide était toujours le moyen de paiement le plus fréquemment utilisé, comme l'indique une note de la *Federal Reserve*.[21]

Dans le chapitre 3, la thèse étudie si les inquiétudes de vie privée concernant les données financières peuvent avoir des effets économiques positifs. Plus spécifiquement, il s'agit d'évaluer comment l'usage de différents moyens de paiement peut être un moyen pour les consommateurs d'améliorer le respect de leur *financial privacy*. Les consommateurs peuvent échapper à la surveillance des banques en utilisant des moyens de paiement non-bancaires comme PayPal ou Bitcoin. L'offre de moyens de paiement se diversifie de plus en plus, certains étant proposés par des opérateurs comme Vodafone ou des acteurs du numérique comme Apple. Cela permet aux consommateurs de répartir leurs achats entre différents moyens de paiement, limitant ainsi le volume de donnée collecté par chaque acteur. Avoir de multiples moyens de paiement à disposition pourrait encourager les consommateurs à acheter en ligne davantage.

La littérature existante a montré que les inquiétudes de vie privée influençaient de façon négative le commerce en ligne (cf. Akhter (2012) par exemple). La grande majorité des achats en ligne sont réglés avec des cartes bancaires, ce qui permet aux sites de vente comme aux banques de garder des traces de ces transactions. Certains consommateurs pourraient dès lors être réticents à utiliser leur carte bancaire en ligne. Le chapitre 2 de cette thèse montre que l'usage de PETs conduit à une plus grande disposition à partager des données personnelles. En appliquant ce résultat au contexte du commerce en ligne, on peut s'attendre à ce que l'usage de moyens de paiement non-bancaires conduit à plus d'achats. En utilisant les mêmes données de sondage que la chapitre 2, la thèse montre de façon empirique que les consommateurs qui utilisent des moyens de paiement non-bancaires ont une fréquence d'achat plus élevée. De plus, en utilisant des techniques bayésiennes MCMC et des variables instrumentales, la thèse montre que l'usage de moyens de paiement non-bancaires s'explique en partie par des inquiétudes au sujet de la *financial privacy*.

Comme dans le chapitre 2, les résultats du chapitre 3 indiquent qu'il peut y avoir des effets positifs pour les firmes du fait que les consommateurs aient plus de contrôle sur leur vie privée. En effet, le chapitre 3 suggère que développer une offre de moyens de paiement respectueux de la *financial privacy* pourrait bénéficier au commerce en ligne. Ceci pourrait passer par une

[21] The note "The State of Cash - Preliminary Findings from the 2015 Diary of Consumer Payment Choice" by Wendy Matheny, Shaun O'Brien, and Claire Wang can be retrieved here.

plus grande diversité de moyens de paiement. A l'inverse, ne pas prendre en considération les inquiétudes des consommateurs pourrait conduire à une baisse des achats en ligne. Cela pourrait également conduire à une baisse de données disponibles pour les banques.

Ce dernier point est l'objet du chapitre 4, qui explore comment la *financial privacy* peut affecter la profitabilité des banques en tant qu'intermédiaires financiers. En effet, dans un contexte dominé par des inquiétudes grandissantes au sujet des données personnelles et par l'adoption croissante de moyens de paiement non-bancaires, une collection massive de données financières par les banques pourrait pousser les consommateurs à utiliser davantage les moyens de paiement non-bancaires, et donc conduire à une baisse de disponibilité des données financières.

Ce risque est abordé dans le chapitre 4 par le biais d'un modèle théorique dans lequel une banque offreuse de crédit veut être en mesure de distinguer les consommateurs selon leur solvabilité. Dans ce but, la banque utilise des données issues de l'usage de la carte bancaire par les consommateurs. Certains consommateurs peuvent alors être tenté de dissimuler une partie de leurs achats en utilisant des moyens de paiement non-bancaires afin de cacher leur faible solvabilité. D'autres consommateurs peuvent ne pas user la carte bancaire en raison d'inquiétudes au sujet de l'usage de leurs données par la banque.

Le modèle démontre comment les inquiétudes à propos de la vie privée financière représentent un coût pour les banques. En premier lieu, l'usage de la carte bancaire diminue, ce qui correspond à une baisse de profits sur le marché des moyens de paiement. De plus, la banque perd en efficacité dans l'évaluation de la solvabilité des consommateurs. En effet, la banque accorde un crédit à des consommateurs non solvables et ne l'accorde pas à certains consommateurs pourtant solvables. Il pourrait donc être plus intelligent pour la banque de proposer des moyens de paiement moins intrusifs. Cela permettrait notamment d'éviter de confondre des consommateurs solvables avec des consommateurs non solvables. Encore une fois, ce chapitre montre qu'il peut y avoir des incitations économiques pour les firmes à respecter la vie privée des consommateurs.

Les trois chapitres de cette thèse démontrent tous, à leur façon, qu'améliorer la vie privée en ligne aurait divers effets économiques positifs. Si améliorer la vie privée pourrait être interpréter comme un coût par les firmes, il pourrait en réalité davantage s'agir d'un investissement. En effet, cela pourrait permettre aux firmes d'avoir accès à plus de données personnelles, et plus fiables. De la même façon, le commerce en ligne pourrait bénéficier du développement de

moyens de paiement non-bancaires.

Chapitre 1 - *Privacy-enhancing technologies* (PETs) et disposition à partager des données personnelles

Deux aspects majeurs de l'usage des données personnelles dans l'économie numérique sont en jeu actuellement. Le premier aspect est le manque de confiance croissant des Internautes en ce qui concerne la collecte et l'usage de leurs données personnelles. Le second aspect est l'usage grandissant de PETs et de stratégies d'obfuscation qui rendent les données moins fiables et plus difficiles à collecter. A court terme, les firmes bénéficient certes de la collecte de données personnelles, mais elles perdent la confiance des consommateurs à long terme.

Le but de ce chapitre est d'étudier l'effet de l'usage de PETs sur la disposition des consommateurs à partager leurs données personnelles. En d'autres termes, l'objectif est d'évaluer si l'utilisation de PETs a un effet positif ou négatif sur le volume de données auquel les firmes ont accès. Si l'on peut s'attendre à ce que l'utilisation de PETs correspondent à une baisse de la disposition à partager, il se pourrait également que cet usage donne suffisamment de contrôle aux consommateurs en ce qui concerne leur vie privée sans mettre de côté le partage d'informations.

La contribution de ce chapitre est un résultat économétrique qui peut paraître contre-intuitif : les Internautes qui ont adopté des solutions pour mieux contrôler l'usage de leurs données personnelles sont davantage disposés à les partager. Développer l'utilisation de PETs serait un moyen de réconcilier les consommateurs avec l'usage de leurs données personnelles. Ce résultat est robuste et s'applique pour quatre types d'acteurs : les services gouvernementaux, les banques, les sites de vente en ligne et les réseaux sociaux. L'implication générale de ce chapitre est qu'encourager le développement de solutions respectueuses de la vie privée est un moyen de restaurer la confiance entre les Internautes et les acteurs du numérique. Il existe une valeur économique de la vie privée, et améliorer la vie privée en ligne est une opportunité économique pour les firmes.

Source des données

Ce chapitre est basé sur une série de sondages qui ontété menés tous les 18 mois de 2009 à 2015 par la Caisse des Dépôts et l'ACSEL (Association pour le commerce et les services en ligne). Les échantillons de chaque sondage sont représentatifs des Internautes français.

Le principal objectif de ces sondages est de mesurer le niveau de confiance des Internautes dans différents services en ligne (banque, administration, etc.). Ces sondages sont divisés en plusieurs parties, chacune ayant un objet spécifique : commerce en ligne, vie privée, réseaux sociaux, sécurité, etc. Le sondage de 2015 contient ainsi des questions relatives à l'utilisation de PETs.

Disposition à partager des données personnelles

Les sondages interrogent aussi les participants à propos de leur disposition à partager des données personnelle avec le tableau suivant :

Avec quels acteurs êtes-vous prêts à partager ces informations personnelles ?

	Local territories	State services	Banks	Operators	Online retailers	Internet actors	Social networks	I do not want to share this information
Name								
Adress								
Phone number								
ID card number								
Bank card or bank account number								
Health information								
Information about tastes or preferences								

La disposition à partager des données personnelles est interprétée comme étant la somme des informations (parmi le nom, l'adresse, le numéro de téléphone, le numéro de carte d'identité, le numéro de compte et/ou de carte bancaire, les informations de santé, les goûts et les préférences) qu'une personne est prête à partager avec n'importe quel des acteurs proposés (collectivités locales, services gouvernementaux, banques, opérateurs, sites de vente en ligne, acteurs Internet comme Google ou Microsoft, réseaux sociaux).

Une part importante des personnes interrogées en 2015 avait déjà utilisé un PET ou mis en place une stratégie de protection de la vie privée : 64% avaient déjà effacé leur historique de navigation, 71% effaçaient leurs *cookies* de façon régulière et/ou utilisaient des extensions pour navigateur dédiés à la protection de la vie privée, et enfin 44% utilisaient des bloqueurs de publicité.

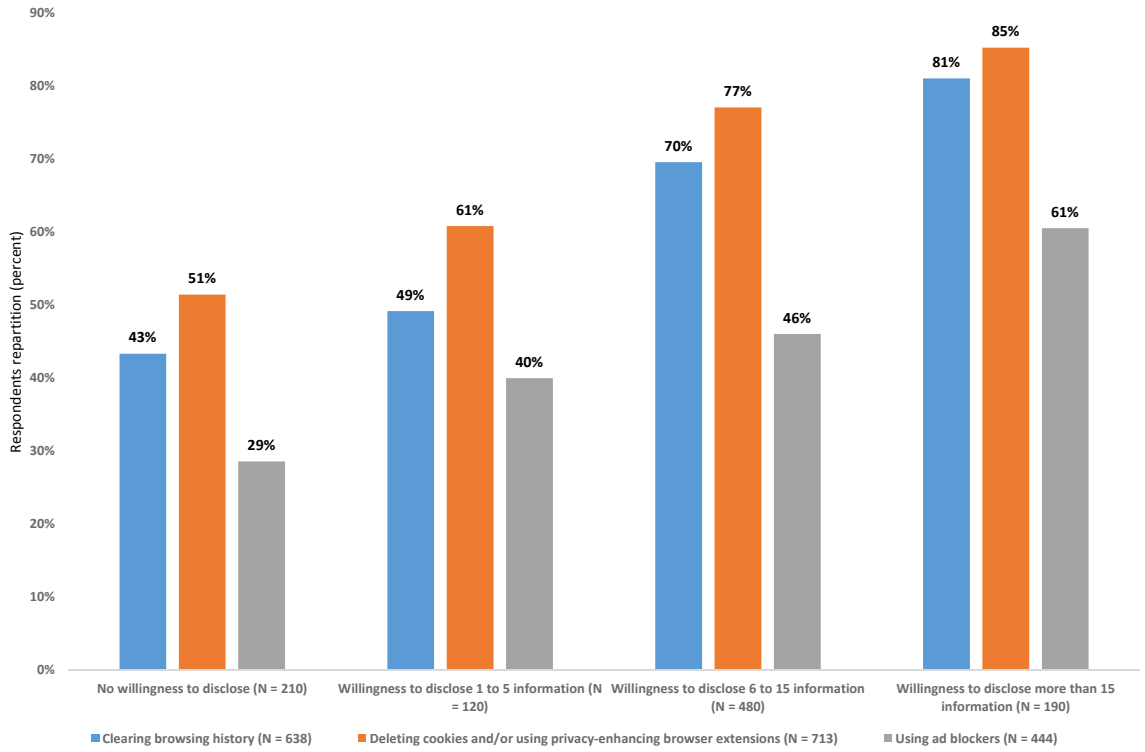


Figure 1 – Utilisation de PETs par niveau de disposition à partager des informations personnelles

La proportion d'individus utilisant des PETs ou des stratégies de protection de la vie privée augmente clairement avec le niveau de disposition à partager des informations personnelles : 43% des personnes qui ne souhaitent rien partager ont déjà effacé leur historique de navigation tandis qu'ils sont 81% parmi ceux disposés à partager plus de 15 informations par exemple. Cette observation est confirmée par les résultats économétriques.

Il n'y a donc pas de corrélation négative entre l'utilisation de PETs ou de stratégies de protection de la vie privée en ligne et la disposition à partager des données personnelles. En réalité, il semble que les Internautes qui ont pris des mesures pour avoir plus de contrôle sur leur vie privée en ligne sont disposés à partager davantage d'informations personnelles.

On spécifie l'équation suivante :

$$\begin{aligned}(\text{Disposition partager des donnes personnelles})_i = \alpha &+ \Sigma_{j=1}^{3}\beta_j(\text{Utilisation de PETs})_i \\ &+ \Sigma_{j=4}^{7}\beta_j(\text{Risques de } \textit{privacy})_i \\ &+ \Sigma_{j=8}^{9}\beta_j(\text{Activités en ligne})_i \\ &+ \Sigma_{j=10}^{16}\beta_j(\text{Variables sociodémographiques})_i\end{aligned}$$

Tout d'abord, trois variables binaires sont utilisées pour indiquer si un individu a adopté une approche stratégique à propos de sa vie privée en ligne ("Utilisation de PETs") : avoir effacé son historique de navigation, avoir effacé ses *cookies* et/ou installé des extensions de *privacy* sur leur navigateur Internet, et enfin avoir installé des bloqueurs de publicité.

Ensuite, quatre variables binaires sont utilisées pour prendre en compte les risques liés à la vie privée en ligne : le risque que des proches aient accès à des informations personnelles par le biais des réseaux sociaux, le risque que des services gouvernementaux gardent des informations personnelles indéfiniment, le risque de surveillance gouvernementale et enfin le risque d'être géolocalisé.

Deux variables sont utilisées pour prendre en compte le niveau d'activité en ligne : la fréquence de connexion à Internet et la fréquence d'achats en ligne. Enfin, des variables sociodémographiques sont inclues : niveau d'éducation, genre, CSP, âge.

Un probit ordonné est utilisé, ainsi qu'une régression des moindre carrés ordinaires dans un but de comparaison. Les estimations montrent bien que, toutes choses égales par ailleurs, utiliser des PETs a un effet positif sur la disposition à partager des données personnelles. Ainsi, les Internautes qui ont pris des mesures afin d'avoir plus de contrôle sur le respect de leur vie privée en ligne sont disposés à partager davantage d'informations personnelles.

Moyens de paiement, *financial privacy* et achats en ligne

La protection des données personnelles est devenue un sujet de grand intérêt pour les consommateurs (Acquisti et al., 2015). Les données financières ne font pas exception. Comme décrit par Lacker (2002) : "les transactions financières d'un consommateur sont une importance source de données personnelles. Chaque achat par carte, chaque retrait, chaque paiement de loyer, chaque dépôt laisse une trace électronique chez une banque." Avec le développement

technologique, un nombre croissant de firmes est en mesure de surveiller les achats des consommateurs, leurs montants, leurs localisations, etc.

Les banques peuvent par exemple utiliser les données de paiement pour vendre des produits financiers comme les crédits (Mester et al., 2007), détecter la fraude ou créer des programmes de fidélité (Ching and Hayashi, 2010). Les réseaux de cartes bancaires peuvent quant à elles répartir les consommateurs en segments et vendre ces données à des *data brokers*, des publicitaires, etc.

Pour toutes ces raisons, les consommateurs peuvent être réticents à utiliser leur carte bancaire lors d'achats en ligne. Ne pas utiliser la carte bancaire peut permettre en effet d'éviter d'être surveillé ou sollicité par sa banque. Cette réticence peut être renforcée lorsque les consommateurs désirent acheter des biens sensibles comme des médicaments, des produits illicites ou pornographiques, etc.

Afin de préserver la confidentialité de leurs données financières, certains consommateurs peuvent donc décider d'utiliser des moyens de paiement non-bancaires, c'est- à-dire non directement liés à leur compte bancaire. Ainsi, une transaction effectuée avec PayPal masque au moins une partie de l'information à la banque, même si cet achat est lié à une carte bancaire. De façon similaire, l'utilisation de monnaies virtuelles comme le Bitcoin est un moyen d'effectuer des achats en ligne sans surveillance bancaire.

Ce chapitre s'intéresse à la question de savoir si l'utilisation de moyens de paiement non-bancaires a un effet positif sur les achats en ligne. L'intuition est la suivante : utiliser la carte bancaire peut être considéré comme un coût par certains consommateurs, soucieux du respect de leur vie privée financière (*financial privacy*). En l'absence de moyens de paiement non-bancaires, ces consommateurs pourraient donc limiter leurs achats en ligne. Ce raisonnement s'appuie sur la recherche existante : les inquiétudes liées au respect de la vie privée en ligne ont un effet négatif sur les achats en ligne (Akhter, 2012). En revanche, en présence de moyens de paiement non-bancaires, les consommateurs sensibles à la *financial privacy* ont une solution pour acheter en ligne tout en évitant de partager des informations avec leur banque ou leurs proches (dans le cadre d'un compte commun).

Ce chapitre se base sur les données d'un sondage français de 2015, qui interroge un échantillon de 1000 Internautes français représentatifs. Ce sondage apporte simultanément des informations sur les achats en ligne et les choix de moyens de paiement. Deux méthodes économétriques

sont utilisées : une régression en deux étapes, ainsi qu'un modèle bayésien MCMC (*Markov Chain Monte Carlo*) pour prendre en compte des éventuels problèmes d'endogénéité.

Le commerce en ligne français est un des plus développés d'Europe. En 2014, 34,7 millions de consommateurs (soit 79% des Internautes en France) ont dépensé environ 57 milliards d'euros auprès de 164 000 vendeurs en ligne (FEVAD, 2015). Dans le sondage utilisé dans ce chapitre, 81% des personnes interrogées avaient réalisé au moins un achat en ligne au cours des 12 derniers mois.

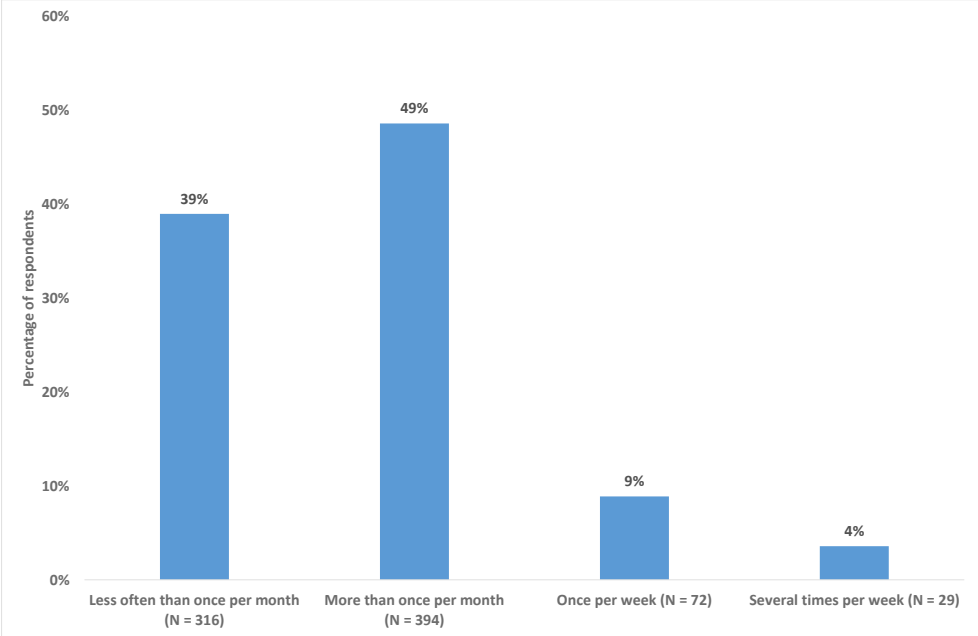


Figure 2 – Fréquence d'achat

Les consommateurs peuvent utiliser différents moyens de paiement afin de réaliser leurs achats en ligne : cartes bancaires, PayPal, chèques, etc. (Figure 3.2) Les cartes fournies par les banques et PayPal sont les moyens de paiement les plus utilités : 95% pour les cartes et 36 pour PayPal.

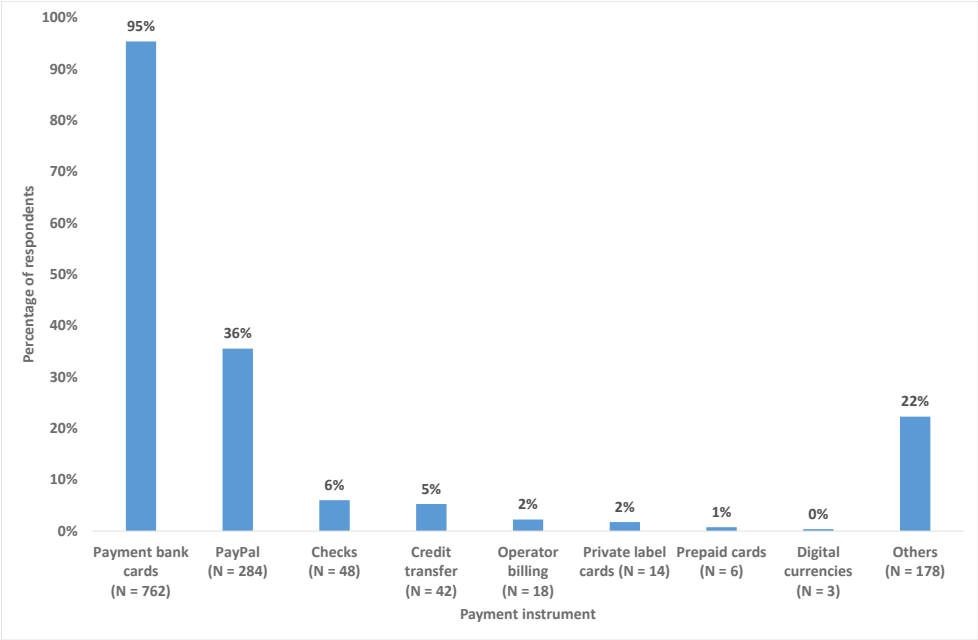


Figure 3 – Utilisation de moyens de paiement

Parmi les consommateurs, 53% utilisent un seul moyen de paiement (la carte à 94%), 29% en utilisent deux et 17% en utilisent trois ou plus. Il est important de noter que contrairement aux cartes bancaires, tous les vendeurs en ligne n'acceptent pas PayPal. Utiliser PayPal correspond donc bien à un choix du consommateur, au détriment de la carte bancaire.

En résumé, les consommateurs ont le choix d'utiliser les moyens de paiement fournis par leur banque afin de régler leurs achats en ligne. Ils peuvent également utiliser des moyens de paiement disjoints de leur compte bancaire : les moyens de paiement dits "non-bancaires". Il est supposé dans ce chapitre que l'utilisation de moyens de paiement non-bancaires est au moins en partie motivée par des raisons de vie privée (*financial privacy*). La facilité d'usage et la sécurité des transactions sont deux exemples d'autres motivations possibles. Ces autres facteurs sont pris en compte dans la modélisation économétrique.

Modélisation économétrique

Il existe en effet un potentiel problème d'endogénéité, si l'utilisation de moyens de paiement non-bancaires est corrélée avec des variables non-observées qui influencent également sur la fréquence d'achats en ligne. Ce problème potentiel est pris en compte par un système de deux

équations, qui modélise l'utilisation d'un moyen de paiement non-bancaire par un consommateur i :

$$w_i = \begin{cases} 0 & \text{si } w_i^* \leq 0, \\ 1 & \text{si } w_i^* > 0, \end{cases} \tag{1}$$

avec

$$w_i^* = X_i\beta + \varepsilon_i, \tag{2}$$

où w_i^* est la variable latente et X_i la matrice des variables de contrôle. L'équation principale est la suivante :

$$\begin{aligned} \mathrm{y}_i = \alpha + \delta w_i & \\ + \gamma_1 (\text{Facilité d' usage})_i & \\ + \Sigma_{j=2}^{3} \gamma_j (\text{Risques})_i & \\ + \Sigma_{j=4}^{7} \gamma_j (\text{Activités en line})_i & \\ + \Sigma_{j=8}^{14} \gamma_j (\text{Variables individuelles})_i & \\ + \eta_i & \end{aligned} \tag{3}$$

où V_i est une matrice de variables explicatives influençant la fréquence d'achats y.

Cette modélisation prend donc en compte la facilité d'usage qui peut être une des caractéristiques influençant sur le choix du moyen de paiement par le biais d'une variable binaire indiquant si les coordonnées de la carte ont été stockées sur au moins un site marchand ou non.

Deux types de risques sont introduits dans le modèle : le risque d'être piraté sur des sites de ventes, et le risque que des informations bancaires soient consultées par des tierces parties sans autorisation. Ces variables ont été utilisées par Akhter (2012) et Tsai et al. (2011). Ces auteurs démontrent que ces risques ont une influence négative sur les achats en ligne.

Trois variables liées à l'activité en ligne sont également introduites : une variable binaire indiquant si un individu se connecte à Internet de façon quotidienne ou non, une autre indiquant le nombre de mots de passe, et enfin une dernière variable catégorielle pour les montants dépensés en ligne. Des caractéristiques sociodémographiques sont également prises en compte : le fait d'avoir des enfants, la catégorie socioprofessionnelle, l'âge et enfin le niveau d'éducation.

La première équation, celle ayant pour but de prédire l'usage de moyens de paiement non-bancaires, est définie de la façon suivante :

$$\begin{aligned} w_i^* = \alpha &+ \beta_1(\text{Facilité d' usage})_i \\ &+ \Sigma_{j=2}^{3}\beta_j(\text{Risques})_i \\ &+ \Sigma_{j=4}^{7}\beta_j(\text{Activités en ligne})_i \\ &+ \Sigma_{j=8}^{14}\beta_j(\text{Variables individuelles})_i \\ &+ \Sigma_{j=15}^{17}\beta_j(\text{Instruments})_i \\ &+ \varepsilon_i \end{aligned} \tag{4}$$

En plus des variables décrites précédemment, trois instruments sont utilisés pour rendre la modélisation robuste malgré le risque d'endogénéité entre l'utilisation de moyens de paiement non-bancaires et la fréquence d'achat en ligne. Ces instruments doivent être corrélés avec la variable potentiellement endogène (l'utilisation de moyens de paiement non-bancaires) mais exogènes avec la variable dépendante (la fréquence d'achat en ligne).

Le premier instrument est le nombre de moyens de paiement utilisés par individu. En effet, les consommateurs n'ont pas toujours la possibilité d'utiliser un moyen de paiement non-bancaire.

Le second instrument est une variable binaire indiquant si la personne interrogée est disposée à partager ou non des informations personnelles avec les banques. Cet instrument a pour but de mettre en valeur un potentiel effet de la *financial privacy* sur le choix du moyen de paiement. Le résultat attendu ici est que les individus ne souhaitant pas partager des données personnelles avec les banques se tournent davantage vers les moyens de paiement non-bancaires.

Le troisième instrument est lié au contrôle des consommateurs sur le respect de leur vie privée en ligne : une variable binaire indiquant si un individu utilise des bloqueurs de publicité et/ou des extensions de *privacy* pour navigateur Internet (Ghostery ou HTTPSEverywhere par exemple). Il est ici attendu *a priori* que les personnes utilisant ce type d'outils ont plus tendance à recourir à des moyens de paiement non-bancaires.

Les résultats empiriques de chapitre confirment l'intuition selon laquelle l'utilisation de moyens de paiement non-bancaires peut être bénéfique pour le commerce en ligne. Si la ten-

dance actuelle est à une personnalisation croissante des moyens de paiement et des services financiers, il semble donc que prendre en compte les différents niveaux de sensibilité au respect de la vie privée est crucial.

Financial privacy, *screening* et moyens de paiement

Ce chapitre a pour but de savoir si la profitabilité des banques en tant qu'intermédiaires financiers est menacée par les inquiétudes de certains de leurs consommateurs au sujet du respect de leur vie privée financière (*financial privacy*). Sur le marché des moyens de paiement, la *financial privacy* peut conduire des consommateurs à privilégier des moyens de paiement non-bancaires afin de préserver la confidentialité de leurs transactions. En ce qui concerne l'attribution de crédits, les banques se basent sur les données de paiement afin d'évaluer la solvabilité de leurs clients. Des inquiétudes au sujet de la *financial privacy* pourraient diminuer le volume de données de paiement disponibles pour les banques, et donc rendre leur processus de *screening* moins efficace.

Le chapitre s'appuie sur un modèle théorique où une banque utilise des données de paiement pour prédire la solvabilité des consommateurs. Plus un consommateur utilise le moyen de paiement proposé par la banque, plus celle-ci évalue la solvabilité de ce consommateur de façon précise. Peu de recherche académique a été dédiée au lien entre moyens de paiement et attribution de crédits.

Les consommateurs choisissent leur moyen de paiement : pour un panier de biens donnée, ils déterminent la proportion de ce panier qui va être payée avec le moyen de paiement bancaire. Le reste est payé avec des moyens de paiement non-bancaires : argent liquide, Bitcoin, PayPal, etc. Le choix du moyen de paiement peut être influencé par des inquiétudes au sujet de la *financial privacy*, par le coût du moyen de paiement, et la volonté des consommateurs d'accéder au crédit. En fixant correctement le coût d'utilisation de son moyen de paiement, la banque peut arriver à faire des décisions d'attribution de crédits optimales. La banque doit toutefois prendre en considération l'existence de moyens de paiement concurrents.

Les inquiétudes au sujet de la *financial privacy* pourraient cependant intensifier la compétition sur le marché des moyens de paiement. Ce chapitre montre qu'il est impératif pour la banque de prendre ces inquiétudes en considération. Une manière de le faire est de diminuer le coût d'utilisation de ses moyens de paiement : facilité d'usage, tarifs moins élevés, etc. Cela ne

pourrait toutefois pas suffire, dans la mesure où il existe des façons de plus en plus diversifiées d'accéder au crédit. Il pourrait donc être plus judicieux pour les banques de moins s'appuyer sur l'analyse des données de paiement, et de proposer des moyens de paiement plus respectueux de la *financial privacy*.

Le modèle utilisé dans ce chapitre s'inspire de Hauswald and Marquez (2003), où les auteurs étudient comment les progrès dans les technologies de l'information influencent la compétition entre les intermédiaires financiers. Ce modèle est utilisé comme base dans ce chapitre car il fournit un moyen d'observer les liens entre le traitement de l'information et les décisions d'attribution de crédit. Hauswald and Marquez (2003) ne décrivent cependant pas la nature de l'information utilisée par les intermédiaires financiers. Dans le modèle présenté dans ce chapitre, l'information est considérée comme étant des données de paiement. En effet, l'analyse de ce type de données peut en effet fournir des informations cruciales en ce qui concerne l'attribution de crédits. Un autre élément du modèle de ce chapitre qui diffère de celui de Hauswald and Marquez (2003) est que la quantité de données à disposition des banques dépend directement des choix de moyens de paiement des consommateurs.

Le modèle décrit donc les relations entre une banque et des consommateurs demandeurs de crédit. Les consommateurs demandent un prêt auprès de la banque et dans le même temps consomment un panier de biens. Les consommateurs choisissent une fréquence d'usage du moyen de paiement bancaire, sachant qu'une plus grande utilisation améliore le processus de *screening* de la banque.

Le but de la banque est de distinguer entre les consommateurs solvables et les non-solvables. Il y a une proportion q de consommateurs solvables et une proportion $1-q$ de consommateurs non-solvables. L'efficacité du *screening* de la banque est de $\frac{1+\alpha}{2}$, où α est la proportion du panier de biens payé avec le moyen de paiement de la banque.

Les consommateurs arbitrent entre les moyens de paiement à leur disposition en fonction du coût d'utilisation. Pour une proportion ω des consommateurs s'ajoute toutefois un coût spécifique lié à l'utilisation du moyen de paiement bancaire, pour des raisons de *financial privacy*.

Prendre en compte ce coût lié à la *financial privacy* est impératif pour la banque si elle veut effectuer des décisions d'attribution de crédit optimales. Une manière de la faire est de proposer un moyen de paiement plus respectueux de la *financial privacy*.

Conclusion

Dans le chapitre 2 de cette thèse (***Privacy-enhancing technologies* (PETs) et disposition à partager des données personnelles**), il a été montré que toutes choses égales par ailleurs, les Internautes qui ont pris des mesures pour mieux contrôler leur vie privée en ligne sont prêts à partager plus d'informations personnelles que la moyenne.

En ligne, la collecte de données personnelles est la contrepartie de la plupart des services lorsque ceux-ci sont gratuits. De telles pratiques sont à l'origine d'une perte de confiance des consommateurs, qui se traduit par de plus en plus d'inquiétude au sujet du respect de la vie privée. Cette perte de confiance pourrait engendrer une baisse d'information disponible pour les firmes.

Les résultats du chapitre 2 suggèrent une solution : les consommateurs sont prêts à partager des informations personnelles s'ils ont à disposition des outils qui leur permettent de contrôler le partage de données. Ne pas être en mesure de contrôler le niveau d'identification ou le *tracking* par les firmes est en effet un coût pour les consommateurs. Comme le montre le chapitre 2, ce coût peut se traduire par une baisse de la disposition à partager des données personnelles. Proposer des outils comme les PETs ou bien des services plus respectueux de la vie privée peut être un moyen de réconcilier les préférences des consommateurs et les besoins de firmes en matière d'utilisation des données personnelles.

En faisant cela, les firmes pourraient encourager le partage d'informations personnelles, et même avoir accès à des données plus fiables. Mais cela n'est possible qu'à condition que les consommateurs ont le sentiment d'avoir assez de contrôle sur la collecte et l'utilisation de leurs données.

En ce qui concerne les informations de nature financière, un moyen pour les consommateurs d'avoir davantage de contrôle est d'utiliser plusieurs moyens de paiement : cartes bancaires, PayPal, monnaies virtuelles comme le Bitcoin, etc. Les consommateurs peuvent vouloir masquer une partie de leurs transactions pour différentes raisons : achat de biens sensibles (médicaments, paris, etc.) ou bien parce qu'ils veulent échapper à la surveillance des banques et autres intermédiaires financiers.

Le chapitre 3 de cette thèse (**Moyens de paiement, *financial privacy* et achats en ligne**, co-écrit avec D.Bounie, M.Quinn et P.Waelbroeck) montre que l'utilisation de moyens de paiement

non-bancaires a un effet positif sur les achats en ligne. Ce chapitre montre également comment des inquiétudes au sujet de la *financial privacy* rendent plus probable l'utilisation de moyens de paiement non-bancaires.

Comme dans le chapitre 2, ces résultats suggèrent qu'il existe des incitations économiques à développer des moyens de paiement non-bancaires. En effet, si les consommateurs ont de multiples moyens de paiement à disposition, non-bancaires en particulier, cela encouragerait leur recours au commerce en ligne. Ne pas prendre en compte l'envie de confidentialité financière de certains consommateurs pourrait ralentir le développement de ce commerce en ligne.

Un usage croissant de moyens de paiement non-bancaires se traduirait toutefois par une baisse de données disponibles pour les banques. Dans le chapitre 4 (***Financial privacy*, *screening* et moyens de paiement**, co-écrit with D.Bounie, M.Quinn and P.Waelbroeck), la thèse étudie comment cette perte d'information menace la profitabilité des banques.

Lorsque certains consommateurs sont inquiets au sujet du respect de leur vie privée financière, leur utilisation de moyens de paiement qui permettent d'échapper à la surveillance des banques et autres institutions financières augmente (principale les moyens de paiement non-bancaires et l'argent liquide). Dans le chapitre 4, la thèse montre que ce phénomène peut avoir des effets négatifs sur la profitabilité des banques.

Une des raisons expliquant cette baisse des profits est un usage inférieur des moyens de paiement bancaires, la compétition s'intensifiant sur le marché des moyens de paiement. Une autre raison est que les banques sont moins efficaces dans leur attribution de crédit, ayant moins de d'information à disposition.

Le principal risque pour les banques est de ne pas accorder de crédits à des consommateurs pourtant solvables en raison de leur faible utilisation de moyens de paiement bancaires. Cette faible utilisation pourrait être un signe de non-solvabilité, mais aussi le reflet d'inquiétudes au sujet du respect de la vie privée financière.

Le résultat du chapitre 2 peut fournir une solution pour les banques. Si les banques proposent aux consommateurs des moyens de paiement plus respectueux de leur vie privée financière, cela pourrait relancer l'utilisation de moyens de paiement bancaires. Cela rendrait les banques plus compétitives sur le marché des moyens de paiement. De plus, cela pourrait fournir suffisamment d'information aux banques pour améliorer leur processus d'attribution de crédit, tout en répondant à la demande de *financial privacy*.

Comme montré dans le chapitre 3 de cette thèse, une telle offre de moyens de paiement par les banques pourrait également avoir des effets positifs sur le commerce en ligne.

Plus généralement, il semblerait donc que proposer des outils, qu'il s'agisse de PETs ou de moyens de paiement, qui prennent en compte les préférences de respect de la vie privée (financière) des consommateurs a du sens du point de vue économique. Ne pas considérer ces préférences alimenterait la perte de confiance des consommateurs, et à plus long terme une baisse de données disponibles pour les firmes.

En revanche, si les firmes adaptent leurs pratiques de collecte de données, elles pourraient être en mesure de résoudre cette crise de la confiance, tout en gardant leur capacité à obtenir des données personnelles. Si cette obtention est le résultat du consentement du consommateur, cela pourrait se traduire par des données plus fiables, et en plus grande quantité.

Même si prendre une telle direction serait bénéfique pour l'ensemble des acteurs, les autorités de régulation ont tout de même un rôle important à jouer. En effet, il est crucial que les consommateurs aient le sentiment d'avoir un réel contrôle sur la collecte de leurs données personnelles. Cela veut dire être en maîtrise de quelles données sont collectées, quand, par qui et dans quel but.

Le Réglement Général de Protection des Données (RGPD) de l'Union Européenne a été conçu pour répondre à ces questions. Aux Etats-Unis, un ensemble de régulations sectorielles encadre ce qu'il est possible de faire avec les données personnelles. Une application stricte de ces régulations est cruciale. Sans une surveillance des autorités de régulation, les firmes pourraient piéger les consommateurs, les enfermant dans une illusion de contrôle. Elles pourraient alors continuer à collecter des données sans le consentement explicite des consommateurs.

BIBLIOGRAPHY

[1] **Acquisti, Alessandro and Hal R. Varian**, "Conditioning prices on purchase history," *Marketing Science*, 2005, *24* (3), 1–15.

[2] **____ , Curtis R. Taylor, and Liad Wagman**, "The Economics of Privacy," *Journal of Economic Literature*, July 2015.

[3] **Akhter, Syed H.**, "Who spends more online? The influence of time, usage variety, and privacy concern on online spending," *Journal of Retailing and Consumer Services*, October 2012, *19*, 109–115.

[4] **Ball, A. Dwayne, Pedro S. Coelho, and Manuel J. Vilares**, "Service personalization and loyalty," *Journal of Services Marketing*, 2006, *20* (6), 391–403.

[5] **Blattberg, Robert C. and John Deighton**, "Interactive marketing: Exploiting the age of addressability," *Sloan Management Review*, 1991, *33* (1), 5–14.

[6] **Bounie, David, Antoine Dubus, and Patrick Waelbroeck**, "Selling Strategic Information in Digital Competitive Markets," 2018.

[7] **Brunton, Finn and Helen Nissenbaum**, *Obfuscation: A User's Guide for Privacy and Protest*, MIT Press, 2015.

[8] **Chellappa, Ramnath K. and Raymond G. Sin**, "Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma," *Information Technology and Management*, 2005, *6.*

[9] **Ching, Andrew T. and Fumiko Hayashi**, "Payment card rewards programs and consumer payment choice," *Journal of Banking & Finance*, 2010, *34* (8), 1773 – 1787. New Contributions to Retail Payments: Conference at Norges Bank (Central Bank of Norway) 14 and 15 November 2008.

[10] **Hauswald, Robert and Robert Marquez**, "Information Technology and Financial Services Competition," *The Review of Financial Studies*, 2003, *16* (3), 921–948.

[11] **Kahn, Charles M. and Jose M. Linares-Zegarra**, "Identity Theft and Consumer Payment Choice: Does Security Really Matter?," *Journal of Financial Services Research*, 2015, pp. 1–39.

[12] **Lacker, Jeffrey M.**, "The Economics of Financial Privacy: To Opt Out or Opt In?," *Federal Reserve Bank of Richmond Economic Quaterly*, Summer 2002, *88* (3), 1–16.

[13] **Linden, Greg, Brent Smith, and Jeremy York**, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet computing*, 2007, *7* (1), 76–80.

[14] **Mester, Loretta J., Leonard I. Nakamura, and Micheline Renault**, "Transactions Accounts and Loan Monitoring," *The Review of Financial Studies*, 2007, *20* (3), 529–556.

[15] **Mikians, Jakub, Laszlo Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris**, "Detecting price and search discrimination on the Internet," *HotNets-XI Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 2012.

[16] **Smith, H.Jeff, Sandra J. Milberg, and Sandra J. Burke**, "Information Privacy: Measuring Individuals' Concerns about Organizational Practices," *MIS Quaterly*, 1996, *20* (2), 167–196.

[17] **Taylor, Curtis R.**, "Consumer privacy and the market for customer information," *The RAND Journal of Economics*, 2004, *35* (4), 631–651.

[18] **Tsai, Janice Y., , Serge Egelman, Lorrie Cranor, and Alessandro Acquisti**, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research*, June 2011, *22*, 254–268.

[19] **Tucker, Catherine E.**, "Social networks, personalized advertising, and privacy controls," *Journal of Marketing Research*, 2014, *51*, 546–562.

[20] **Varian, Hal R.**, "Price discrimination and social welfare," *American Economic Review*, 1985, *75* (4), 870–875.

[21] **Villas-Boas, J.Miguel**, "Price Cycles in Markets with Customer Recognition," *The RAND Journal of Economics*, 2004, *35* (3), 486–501.

[22] **von Kalckreuth, Ulf, Tobias Schmidt, and Helmut Stix**, "Choosing and using payment instruments: evidence from German microdata," *Empirical economics*, 2014, *46*, 1019–1055.

CONTENTS

BIBLIOGRAPHY **30**

CHAPTER 1: GENERAL INTRODUCTION **35**

BIBLIOGRAPHY **44**

CHAPTER 2: ONLINE PRIVACY CONCERNS, PRIVACY-ENHANCING TECHNOLOGIES AND THE WILLINGNESS TO SHARE PERSONAL INFORMATION **46**

2.1 Introduction 46

2.2 Related literature 48

2.3 Data description 50

2.3.1 Source of the data 50

2.3.2 Willingness to share personal information. 51

2.4 Model and findings 52

2.4.1 The model 57

2.4.2 Findings 61

2.5 Conclusion 66

BIBLIOGRAPHY **67**

2.6 Appendix 1: OLS estimations 69

2.7 Appendix 2: Brant test 72

2.8 Appendix 3: Descriptive statistics 73

CHAPTER 3: PAYMENT INSTRUMENTS, FINANCIAL PRIVACY AND ONLINE PURCHASES **76**

3.1 Introduction 76

3.2 Related literature 78

3.3 Survey and data description 81

3.4 Econometric model 85

3.5 Estimation results 89

3.6 Conclusion 93

BIBLIOGRAPHY **95**

3.7 Appendix: descriptive statistics 98

CHAPTER 4: FINANCIAL PRIVACY, SCREENING AND PAYMENT INSTRUMENTS **100**

4.1 Introduction 100

4.1.1 Example 1: American Express 102

4.1.2 Example 2: OpenBanking 103

4.1.3 Example 3: FICO Score XD and FinTechs 103

4.2 Related literature 104

4.3 Model set-up 105

4.3.1 Borrowers 106

4.3.2 The bank 107

4.3.3 Expected utility functions 109

4.3.4 Timing 111

4.4 Case 1: Lending decisions according to a screening technology 111

4.4.1 Stage 2: lending decisions and payoffs 111

4.4.2 Stage 1: choice of payment instrument by borrowers 113

4.5 Case 2: Screening-based lending decisions with financial privacy concerns . . . 115

4.5.1 Stage 2: lending decisions and payoffs 115

4.5.2 Stage 1: choice of payment instrument by borrowers 116

4.6 Case 3: Lending decisions without payment data based screening 119

4.6.1 Stage 2: lending decisions and payoffs 119

4.6.2 Stage 1: choice of payment instrument by borrowers 120

4.7 Conclusion 122

BIBLIOGRAPHY **123**

CHAPTER 5: CONCLUSION **139**

BIBLIOGRAPHY **144**

CHAPTER 1

GENERAL INTRODUCTION

Consumers leave traces every time they browse the Internet. The use of these traces is at the core of the success of some of today's most well-known companies like Google or Facebook. Thanks to their large base of users[1] and their ability to collect information, Google and Facebook can transform personal data into advertising revenues. Revenues in 2017 were estimated around 109.65 billion US dollars for Google[2], and around 40.7 billion US dollars for Facebook.[3] In late October 2017, Alphabet (Google's parent company) reached 700 billion market capitalization in US dollars[4], higher than the estimated 2017 GDP of Switzerland.[5]

Firms have been gathering and using information for years in order to increase their revenues, for example by improving cost savings or their market understanding. This phenomenon has however reached an unprecedented level with the rise of the Internet,the reduction of the costs of data collection and the development of new analytical capabilities. Today, it is possible to transform consumers' digital footprints into profits. For example, data brokers collect personal information in order to sell it to firms, helping them to know more about their customers. In a 2014 study conducted by the US Federal Trade Commission (FTC), it was estimated that these data brokers own information '"on almost every U.S. household and commercial transaction".[6]

Firms can benefit from collecting and analyzing personal information in many ways. Using personal data can help firms increase their revenues by designing personalized products and services (Acquisti and Varian, 2005). Personal information can also generate cost savings. For example, in the case of the online advertising industry, personal data improves the matching between ads and consumers' needs (Blattberg and Deighton, 1991). Personal information can give firms competitive advantages as well, for example by making easier the implementation of price discrimination (Varian, 1985). Online vendors can, for instance, practice differential

[1] Google's share in the search engine market amounted to 91.74% in January 2018 (Statcounter - Search Engine Market Share Worldwide) while four of the five most used social media and messaging services in 2017 were owned by Facebook (F.Richter, Statista - "Facebook Inc. Dominates the Social Media Landscape"

[2] Statista - Annual revenue of Google from 2002 to 2017 (in billion U.S. dollars)

[3] Statista - Facebook's revenue and net income from 2007 to 2017 (in million U.S. dollars)

[4] J.C.Owens, Marketwatch "Alphabet heads for $700 billion market cap after earnings"

[5] IMF - World Economic Outlook A Shifting Global Economic Landscape

[6] Federal Trade Commission, 2014, "Data-brokers: A Call for Transparency and Accountability''.

pricing using location data (Mikians et al., 2012). Another competitive advantage that firms may get from using personal information is for example that personalizing services helps to strengthen consumers' loyalty as it increases switching costs (Ball et al., 2006). Proposing personalized services is made easier by personal data analysis as it helps firms to predict the preferences of consumers (Linden et al., 2003). This can also help firms to drive their innovation effort.

To gather personal information about Internet users, websites and firms use tools such as web trackers which allow them to collect browsing data. While firms have developed techniques that allow them to track consumers online, voluntary contribution of personal data still has its importance. The success of social networks relies for example on an active participation of consumers, mainly in the form of content exchange between users. In general, all industries that profit from user-generated contents (UGC) cannot properly work without consumers' willingness to contribute. The same goes for customer to customer platforms such as eBay, where the rating of sellers and buyers plays a crucial role.[7]

The willingness to share personal information is limited by privacy concerns. Four levels of privacy concerns have been highlighted by Smith et al. (1996): the volume of data collected, the unauthorized internally/externally secondary use (the secondary usage of personal data), the improper access (some data accessible whereas it should not be), and finally errors (errors in the collection of personal data, and the difficulty of repair). Privacy concerns among consumers have reached an important level in recent years. In 2013, the revelations by Edward Snowden shed a light on governments' global surveillance practices.[8] Data breaches have also become a major concern. For example, Equifax discovered in July 2017 that 145,500,000 of its consumer records had been stolen, about half the US population.[9]

The multiplication of such examples in news has deeply affected consumers' trust concerning personal data and privacy issues. In a 2014 US survey by the Pew Research Center, 91% of the respondents stated that they were worried about their loss of control over the way personal information is collected and used by firms.[10] Similarly, 85% of respondents in a 2017 French survey declared being worried about the protection of their personal data.[11]

[7] On such platforms, users are invited to give ratings in order to create a reputation system.
[8] G.Greenwald, The Guardian - "NSA collecting phone records of millions of Verizon customers daily"
[9] L.Mathews, Forbes - "Equifax Data Breach Impacts 143 Million American"
[10] Pew Research Center - "Public Perception of Privacy and Security in the post-Snowden era"
[11] CSA Research - "La protection des données personnelles"

Consumers are however more and more active with respect to privacy. In a 2013 study by the Pew Research Center, 86% of respondents stated that they had attempted to remove some of their personal information from the Internet in one way or another. 55% declared that they had adopted strategies or tools in order to avoid being tracked by specific individuals, firms or the government;[12] in a 2015 survey among representative French Internet users, 84% of respondents had taken privacy-enhancing steps online.[13] To feel safer online with respect to privacy, consumers have adopted several privacy-enhancing tools (PETs) and strategies. Such tools or strategies include deleting cookies regularly or installing privacy specific web browsing extensions.[14] Another solution for consumers is to practice obfuscation, which can be defined as " the deliberate addition of ambiguous, confusing, or misleading information to interfere with surveillance and data collection" (Brunton and Nissenbaum, 2015). Obfuscation strategies and the use of PETs may be a way for consumers to reduce the information asymmetry that often characterizes the use and collection by firms, as the nature of collected information and the purpose of this collection is not always clear or understandable.

Firms then face a trade-off: while it might be beneficial in the short run to extract profit from personal data analysis, it also could drive consumers away in the long run as they keep losing trust over privacy issues. Privacy concerns can have negative economic effects in several ways. For example, in the case of the online advertising industry, Tucker (2014) shows that personalized ads can be counterproductive if they result in consumers feeling concerned about their privacy. Privacy concerns may also reduce online spending by consumers (Akhter, 2012). Another example is how privacy preferences affect payment instruments choices: the possibility of remaining anonymous while using a payment instrument has a positive effect on its adoption (von Kalckreuth et al., 2014), partly because of the risk of identity theft (Kahn and Linares-Zegarra, 2015).

Our goal in this thesis is to look at the various ways the increasing control by consumers of the sharing of personal information affect the profitability of firms. Does this phenomenon result in a greater difficulty for firms to profit from personal data? Does it reflect a loss of trust so important that consumers prefer to renounce online activities such as e-commerce? Is there a way to reconcile information-based profit and consumers' privacy concerns?

[12] Pew Research Center - "Anonymity, Privacy, and Security Online"

[13] ACSEL and CDC - "Baromètre 2015 ACSEL-CDC de la Confiance des Français dans le numérique"

[14] These add-ons include communication anonymizers like virtual private networks (VPN) or extensions designed to block tracking online such as Ghostery or DoNotTrackMe.

A first possible outcome of the loss of consumer trust resulting from privacy concerns is an increased difficulty for firms to generate profit from personal information. Online advertising is an example of an industry that could suffer from this. In the digital economy, if some firms create profit from charging for various services, a high number of others build their business models around the selling of advertising spaces for targeted ads. The value of these slots rises with the audience and the ability of the firm to transform visits into personal data that is then bought and used to design personalized ads. This practice of selling personal data of users to advertising companies has allowed firms to propose free services. This has been true not only for data-driven firms like Google or Facebook, but also for media companies for example. Bounie et al. (2018) show how privacy concerns harm the efficiency of advertising algorithms, leading to an increase of advertising annoyance and ad blocking practices. As more and more consumers start installing ad blocking softwares on their Internet browsers, all these firms face an increasing difficulty to transform their audience into profit, and thus to propose free services. Consumers could also disengage more and more from online activities such as e-commerce for example. In general, all firms that base their ability to generate profit on the collection and analysis of personal information are threatened by the growing use of privacy-enhancing tools and strategies by consumers. This concerns the online advertising industry, social networks, online retailers as well as firms that use data to improve risk assessment like banks or insurance companies.

There is however another possible outcome if consumers are using privacy-enhancing technologies and strategies in order to have more control on privacy and data collection. While there might be negative effects for firms[15], it could also be a way to restore trust between firms and consumers. This would indicate that there is an economic value of privacy, for consumers as well as for firms. By analyzing how consumers seek privacy by increasing their control of the sharing of their information, this thesis studies the economic viability of this option.

We contribute to the economics of privacy by analyzing the economic implications of an active management of the sharing of personal information by consumers. With the development of the Internet and the analysis of consumers' personal information, the economic theory of privacy has attracted many scholars over the last years, tackling the various privacy trade-offs for both data holders and data subjects. This has led to fruitful work analyzing price discrimination,

[15]Consumers could for example use PETs to trick algorithms developed by firms. The use of PETs could also decrease the availability as well as the reliability of personal data.

the role of data intermediaries or the use of personal data in marketing techniques for example (Acquisti et al. 2015). Empirical work (including experiments) has addressed similar issues but has also analyzed the various determinants of privacy preferences. A subject of interest has for example been the valuation of privacy by consumers, showing for instance that some consumers may be willing to pay a price premium for purchasing from retailers who are respectful of privacy (Tsai et al. 2011). As of now, little work has been done about how the profitability of firms evolves according to consumers' willingness to share personal information.[16] Investigating this question is the main focus of this thesis.

This thesis consists of three chapters. Chapters 2 and 3 are based on French survey data.[17] Conducted in 2015, this survey includes questions about various aspects of online activity: privacy, e-commerce, online administration and so on. Chapter 4 presents a theoretical model. Overall, we find that it is economically valuable for firms to provide consumers with ways to have a better handle on the collection and the use of their personal information. Empowering consumers is a way for firms to resolve the trade-off they are facing between deriving profit from the use of personal data and the loss of consumer trust, as we show that empowered consumers are still ready to share personal information.

In Chapter 2 of this thesis, we empirically look at the global effect of PETs on the amount of personal information available to firms. Mainly, we want to know if there is a risk for firms that a growing use of PETs and obfuscation techniques will result in increasing difficulties in generating profit from personal data. Reflecting the various types of personal information, this question concerns many actors of the digital economy: social and professional networks, open source platforms, advertising firms and so on. Having simultaneously data about adoption of PETs and willingness to share personal information among representative French Internet users, we want to know how one affects the other. Investigating this issue allows us to understand if the ability of firms to collect and use data is threatened by consumers' increasing handling of personal information disclosure. If this phenomenon signals that it becomes more and more difficult to create profit from personal data, many industries should consider it with apprehension. This could happen because personal data would become more scarce or less reliable. Our study however finds that using PETs has positive effects on the willingness of consumers to share

[16] We should however mention Villas-Boas (2004) and Taylor (2004), who show how consumers can postpone purchases in order to avoid being identified and thus suffer from price discrimination. Strategically postponing purchasing allows consumers to access lower and non-discriminatory prices.

[17] ACSEL and CDC "Baromètre 2015 ACSEL-CDC de la Confiance des Français dans le numérique"

personal information.

This result suggests that the loss of trust resulting from privacy concerns that consumers have expressed may represent an economic opportunity for firms. It might indeed be beneficial for firms to develop solutions more respectful of consumers' privacy concerns, as it would not necessarily result in a decrease of the amount of data at their disposal. Chellappa and Sin (2005) find similarly that trust building activities makes easier for online vendors to acquire information about their customers and then propose personalized services. Moreover, as suggested by Tsai et al. (2011), firms may be able to use privacy protection as a valuable feature of their services. Overall, Chapter 2 highlights how improving privacy online can increase the participation of consumers in terms of personal information, indicating that it represents a way for firms to solve the trade-off they are facing between using personal data and antagonizing privacy concerns.

Another contribution of Chapter 2 is that it supports the regulatory approach of the European Union on personal information and privacy. In April 2016, the European Union adopted a regulation "on the protection of natural persons with regard to the processing of personal data and on the free movement of such data".[18] This "General Data Protection Regulation" (GDPR) will become enforceable on 25 May 2018. One of the key feature of the GDPR is that firms must have consumers' consent to collect and use their personal information: "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her". The unambiguous nature of personal data collection supposes that consent is clearly given by consumers. It also implies that the purpose of the collection is sufficiently clear, meaning that consumers are aware of what data is collected and for what purpose. We show in Chapter 2 that if this regulation is respected, it may be easier for firms to collect personal information.

Online advertising, online retailing and the personal data market are the industries that have interested the economics of privacy the most. While banks have been pioneers in generating profit from using customer information, few work has linked financial information (such as payment data and credit history) and privacy. Chapters 3 and 4 specifically focus on how an increasing control of consumers on privacy changes the economic landscape in the case of financial information. Tackling this issue allows us to link several strands of the economic

[18]The full text can be retrieved at the following address: http://data.europa.eu/eli/reg/2016/679/oj.

literature that have often been disconnected so far: privacy, payment instruments choice, online purchases and financial intermediation.

Banks and online retailers use financial information for various reasons, from fraud detection to proposing personalized offers (financial products for example). In chapters 3 and 4, we examine the notion of financial privacy, that is privacy related to financial data. Consumers may sometimes want to hide financial information, for example because they want to avoid being tracked by companies such as banks or insurers. Another reason might be that they want to hide some purchases from their relatives in the case of a shared account for example. 55% of consumers surveyed in a 2016 study by KPMG declared having renounced to buy something online because of privacy concerns[19]. Similarly, a survey conducted among 2000 representative French Internet users finds that 35% of them have renounced to a purchase at least once or change payment instrument for financial privacy reasons.[20] Overall, consumers sometimes value the anonymous nature of some payment instruments such as cash. In 2015 in the US, cash was still the most frequently used payment instrument according to a note from the Federal Reserve.[21]

In Chapter 3, we want to test whether privacy in the case of financial information may lead to positive economic effects. Specifically, we consider the use of various payment instruments as a way for consumers to improve financial privacy. Consumers can escape being tracked by banks by adopting non-bank payment instruments such as PayPal or Bitcoin. There is a growing number of payment instruments, some being proposed by telecommunication operators like Vodafone or digital companies such as Apple. This allows consumers to divide their purchases across various payment instruments, limiting the amount of data collected by actor. Overall, this could give consumers an incentive to increase their use of payment instruments online, including non-bank payment instruments. In other words, it could encourage consumers to purchase online.

Previous literature has found that privacy concerns negatively affect e-commerce (see Akhter (2012) for example). The vast majority of online purchases are paid with bank cards, which allow online retailers as well as banks to keep traces of these transactions. In the end, some

[19]Crossing the line - Staying on the right side of consumer privacy report.

[20]Chaire Valeurs et Politiques des Information Personnelles - Données personnelles et confiance : quelles stratégies pour les citoyens-consommateurs en 2017 ?

[21]The note "The State of Cash - Preliminary Findings from the 2015 Diary of Consumer Payment Choice" by Wendy Matheny, Shaun O'Brien, and Claire Wang can be retrieved here.

consumers may be reluctant to use their bank cards while shopping on the Internet. We show in Chapter 2 that the use of PETs induces higher willingness to share personal information. Considering this result in the context of e-commerce, one could expect that the adoption of non-bank payment instruments would improve the market participation of consumers.[22] Using the same data as in Chapter 2, we empirically find that consumers that use non-bank payment instruments tend to buy more online. Moreover, using MCMC Bayesian and instrumental variables regression techniques, we show that this use of non-bank payment instruments can in part be explained by financial privacy concerns.

As in Chapter 2, results of Chapter 3 indicate that consumers having more control on (financial) privacy could have positive effects. Indeed Chapter 3 suggests that developing an offer of payment instruments respectful of financial privacy would benefit e-commerce. This could correspond to increasing the number of payments instruments available to consumers when they purchase online. Conversely, not taking financial privacy concerns into account could lead to a decrease of online purchases. This would correspond to a loss of financial data for banks.

We focus on this very point in Chapter 4 where we investigate how financial privacy affects the profitability of banks as financial intermediaries. Indeed, in a context with growing financial privacy concerns and adoption of non-bank payment instruments adoption by consumers, mass financial data collection by banks may lead to increased adoption of non-bank payment instruments, which would ultimately represent a loss of data for banks.

We study this risk faced by banks using a theoretical model, in which a bank that offers credit wants to distinguish consumers according to their solvency. To do so, the bank uses a screening technology that is improved with the use of the bank card. Some consumers may then be tempted to use non-bank payment instruments in order to hide their true solvency level. Other consumers may not use the payment instrument of the bank because of a financial privacy cost.

Our model predicts that the profitability of the bank is negatively affected by financial privacy concerns. Firstly, the use of the payment instrument of the banks decreases, hence a lower profit on the payment instrument market. Secondly, the bank is less efficient in assessing creditworthiness, refusing credit to some solvent consumers. As a result, the bank may want to supply less intrusive payment instruments. Doing so would help the bank not to make type I mistakes,

[22] Interpreting such payments instruments as PETs with respect to banks and online retailers

that is confusing a solvent consumer with non-solvent ones. Again, we show that there is an economic incentive to be more respectful of consumers' privacy concerns.

The three chapters of this thesis overall find that improving privacy online has various positive effects. While improving privacy could be interpreted as a cost for firms, it could however be a way for firms to improve their ability to collect personal data, as shown in Chapter 2. Moreover, considering financial information and interpreting non-bank payment instruments as privacy-enhancing tools, we show in Chapter 3 that increasing financial privacy is beneficial for e-commerce. It would also toughen competition in the payment instrument market. As banks could risk a loss of financial information, they would have an incentive to also develop payment instruments more respectful of financial privacy. As a result, the overall state of financial privacy could be improved.

BIBLIOGRAPHY

[1] **Acquisti, Alessandro and Hal R. Varian**, "Conditioning prices on purchase history," *Marketing Science*, 2005, *24* (3), 1–15.

[2] **____ , Curtis R. Taylor, and Liad Wagman**, "The Economics of Privacy," *Journal of Economic Literature*, July 2015.

[3] **Akhter, Syed H.**, "Who spends more online? The influence of time, usage variety, and privacy concern on online spending," *Journal of Retailing and Consumer Services*, October 2012, *19*, 109–115.

[4] **Ball, A. Dwayne, Pedro S. Coelho, and Manuel J. Vilares**, "Service personalization and loyalty," *Journal of Services Marketing*, 2006, *20* (6), 391–403.

[5] **Blattberg, Robert C. and John Deighton**, "Interactive marketing: Exploiting the age of addressability," *Sloan Management Review*, 1991, *33* (1), 5–14.

[6] **Bounie, David, Antoine Dubus, and Patrick Waelbroeck**, "Selling Strategic Information in Digital Competitive Markets," 2018.

[7] **Brunton, Finn and Helen Nissenbaum**, *Obfuscation: A User's Guide for Privacy and Protest*, MIT Press, 2015.

[8] **Chellappa, Ramnath K. and Raymond G. Sin**, "Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma," *Information Technology and Management*, 2005, *6.*

[9] **Kahn, Charles M. and Jose M. Linares-Zegarra**, "Identity Theft and Consumer Payment Choice: Does Security Really Matter?," *Journal of Financial Services Research*, 2015, pp. 1–39.

[10] **Linden, Greg, Brent Smith, and Jeremy York**, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet computing*, 2007, *7* (1), 76–80.

[11] **Mikians, Jakub, Laszlo Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris**, "Detecting price and search discrimination on the Internet," *HotNets-XI Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 2012.

[12] **Smith, H.Jeff, Sandra J. Milberg, and Sandra J. Burke**, "Information Privacy: Measuring Individuals' Concerns about Organizational Practices," *MIS Quaterly*, 1996, *20* (2), 167–196.

[13] **Taylor, Curtis R.**, "Consumer privacy and the market for customer information," *The RAND Journal of Economics*, 2004, *35* (4), 631–651.

[14] **Tsai, Janice Y., , Serge Egelman, Lorrie Cranor, and Alessandro Acquisti**, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research*, June 2011, *22*, 254–268.

[15] **Tucker, Catherine E.**, "Social networks, personalized advertising, and privacy controls," *Journal of Marketing Research*, 2014, *51*, 546–562.

[16] **Varian, Hal R.**, "Price discrimination and social welfare," *American Economic Review*, 1985, *75* (4), 870–875.

[17] **Villas-Boas, J.Miguel**, "Price Cycles in Markets with Customer Recognition," *The RAND Journal of Economics*, 2004, *35* (3), 486–501.

[18] **von Kalckreuth, Ulf, Tobias Schmidt, and Helmut Stix**, "Choosing and using payment instruments: evidence from German microdata," *Empirical economics*, 2014, *46*, 1019–1055.

CHAPTER 2

ONLINE PRIVACY CONCERNS, PRIVACY-ENHANCING TECHNOLOGIES AND THE WILLINGNESS TO SHARE PERSONAL INFORMATION

2.1 Introduction

The Internet has become more and more integrated into the daily lives of many consumers in recent years. In the last decade, this integration has also increased with the widespread adoption of smartphones and other connected devices. As a result, it is now possible to record consumer activity at an unprecedented scale. This recording mainly takes the form of personal data collection. Firms are now able to track consumers online from website to website or from one smartphone application to another. All digital activities, being purchases, communications or contributions, are nowadays recorded and transformed into digital footprints.

These digital footprints are highly valuable for firms. Companies like Google and Facebook generate their profits from these footprints, by selling data to marketers and advertising companies in order to make online advertising more efficient. Firms can also use personal data to implement price discrimination, reduce innovation costs or to propose personalized services. This growing use of data in the economy has given rise to the data brokering industry. One of the nine data brokers investigated by the FTC in 2014 had information on 1.4 billion consumer transactions at the time. Another had around 3000 data points for nearly every U.S. consumer.[1] This secondary market of personal data has important consequences for competition, for example by allowing firms to determine the willingness to pay of consumers, and thus to apply price discrimination.

Consumers can, however, suffer negative externalities from the use of their personal information. We already have mentioned price discrimination, but one can also think of identity theft, spamming, data breaches and so on. As the use of personal data by firms was growing, more and more Internet users began to become aware about this exploitation of personal information online. For example, in a survey conducted in January 2014 by the Pew Research Center among representative US citizens, 91% of the respondents stated that they were concerned by their loss of control over the use of their personal data by companies[2]. A similar result can be

[1] Federal Trade Commission, 2014, "Data-brokers: A Call for Transparency and Accountability''.

[2] The complete report "Public Perceptions of Privacy and Security in the Post-Snowden Era" can be found here

found in a 2017 French study, where 91% of respondents stated that they wanted to control what firms can know about them online.[3].

While some Internet users have not taken any actions to deal with their concerns about privacy, others have decided to be more active. Indeed, more and more people use ad blockers and manage their browser privacy and history settings. Such tools and strategies include installing ad blocker, clearing browsing history and cookies, using digital currencies like Bitcoin, or using privacy-enhancing web browser extensions such as Ghostery and HTTPSEverywhere. In September 2013, the Pew Research Center published a survey where 86% of the respondents had taken measures in order to avoid surveillance by firms or other organizations[4], mainly clearing cookies and browsing history (64%), but also using temporary identifiers or remaining anonymous online. Consumers can also use obfuscation techniques, designed to derail online tracking by generating misleading information. For example, AdNauseam is a software linked to an ad-blocker that clicks on all blocked ads. As a result, AdNauseam makes online tracking for advertising purposes ineffective.

Two major aspects of the use of personal data in the digital economy are then at stake. First, this growing distrust of Internet users over the handling and the collection of their information could lead to a decreasing volume of personal data available for firms. Second, the developing use of privacy-enhancing technologies and obfuscation techniques could limit the reliability of the information gathered by firms, and create flawed algorithms for example. In the short run, firms may benefit from personal data collection, but may suffer from a loss of consumer trust in the long run. While some firms may not directly face this trade-off by relying on online tracking, other firms like social networks need consumers to share personal data.

The goal of this paper is to evaluate what is the effect of the use of privacy-enhancing technologies (PETs) on the willingness of consumers to disclose personal information. Namely, we want to study whether the growing use of PETs has a positive or a negative impact on the volume of personal data available to firms. A first outcome of the use of PETs is a decrease in the willingness of consumers to share personal information. A second outcome may occur: using PETs could bring consumers sufficient control, allowing them to keep disclosing personal information while accomodating their privacy concerns. This paper is based on a 2015

[3]"Etude sur les données personnelles" - Mediametrie and Chaire Valeurs et Politiques des Informations Personnelles

[4]The "Anonymity, Privacy, and Security Online" report can be retrieved here

cross-sectional survey designed to evaluate the level of trust that representative French Internet users have towards different Internet activities (including making online purchases, using online payment systems, online banking, blogs and so on) and actors (banks or social networks for example). An important feature of this survey is that it allows us to simultaneously capture the use of PETs and the willingness to share personal information.

Our contribution is an econometric result that may appear counter-intuitive: Internet users that have adopted solutions to protect their data are willing to share more personal data than the average. This suggest an eventual positive impact of PETs on the willingness of consumers to share personal data. Using PETs could indeed reconcile some Internet users with the collection of their information online, and even encourage data sharing. Moreover, this result is robust across four types of actors: State services, banks, online retailers and social networks. The general implication of the paper is that encouraging the use privacy-enhancing tools could help to solve the trust issue that is currently affecting the relationship between the majority of Internet users and the actors they are dealing with online.

Overall, our study finds that there is an economic value to privacy, and that improving privacy online could represent an economic opportunity. Specifically, consumers having more control of personal data sharing could be a solution to the trade-off between generating profit from personal information and loosing consumer trust.

The article proceeds as follows: section 2 reviews the related literature, section 3 describes the data, section 4 presents our model, estimation techniques and findings. Finally, section 5 concludes.

2.2 Related literature

This paper deals with the way consumers share personal information with respect to their privacy concerns. Privacy is an issue that has been discussed in economics for several years.

The analysis of privacy really began to be discussed intensively in Chicago, especially by Posner and Stigler [5] (Acquisti, 2010). In their framework, privacy is described as a set of information goods about an individual that is not shared with the the other agents in the market. Therefore privacy leads to markets with incomplete information, and consequently to market failures, as complete information is one of the basic assumptions of perfect competition (Posner

[5] See Stigler (1980) for example.

1978 and Posner 1981). Indeed, privacy is seen as a strategic concealment of personal information, that could increase total welfare if other agents were aware of the information. In the case of the hiring process for example, an individual with weak working skills could decide to hide negative characteristics during a job interview, thus leading to inefficiency. In the end, the costs of privacy are paid by the other agents in the market.

Firms can heavily benefit from gathering and analyzing personal information about their customers. They can for example improve their customer relationship management (CRM) as described by Richards and Jones (2008). Benefits of using personal data can also go to the extent of developing personalized offers for consumers, which can increase revenues for firms as pointed out by Acquisti and Varian (2005). Several firms propose nowadays recommendation algorithms.[6] Ball et al. (2006) show that when a firm uses personal information to design targeted services, it strengthens consumers' loyalty by increasing switching costs. Having personal information about consumers can also help to predict preferences (Linden et al., 2003). This can help to implement price discrimination as argued by Varian (1985). Knowing consumers' preferences can also facilitate the prediction of the evolution of aggregate demand for example. The advertising industry is another one that has benefited from the use of personal information, being able to design targeted ads, especially online. Blattberg and Deighton (1991) describe how the use of personal information in advertising may improve the match between the content of ads and the preferences of consumers.

Consumers can have incentives to share personal data. Varian (1996) argues that they could even suffer from too little exchange of personal information, in the case of health data for instance. Other benefits to disclosing personal information are the possibility of receiving personalized ads or money savings for example. In the case of online advertising, Goldfarb and Tucker (2010) argue that personalized ads might be less intrusive than non-targeted ones. Sharing personal information can also be a way for consumers to improve their online reputation, as in Spence's signalling model (Spence, 1973). Indeed, the reputation of an individual online is closely related to what information is available about that person. Not concealing too much personal information could allow an increase in the trust of others, for example on social networks or dating websites. In the world of digital marketplaces, Cabral and Hortascu (2010) show for example that a 1% increase of the number of negative reviews of a seller leads on average to a 7.5% decrease of prices. Similarly, Bounie et al. (2008) find that having good reviews on the

[6] See the example of a company like Netflix (Bennett and Lanning, 2007).

Amazon Marketplace can allow a seller to increase its prices up to 10%.

Although the use of personal information can be valuable for both firms and consumers, it might also yield negative effects. Disclosing personal information can indeed create negative externalities for consumers. Examples of such externalities are: excessive targeted advertising, spam, credit card hacking, identity theft and so on. Fudenberg and Tirole (1998) study the case of a monopolist in a durable good market where it may recognize past customers. Taylor (2004) finds that the surplus of consumers is captured by firms if consumers do not consider the possibility that merchants may use their personal information to implement price discrimination.

While consumers may have in part benefited from the use of their personal information by firms, the existence of negative externalities has triggered privacy concerns among Internet users. Negative externalities from the use of personal information by firms are at the source of the formation of privacy concerns. Four levels of privacy concerns have been highlighted by Smith et al. (1996). The first one is the concern about the collection level, that is a potential abuse in the volume of data collected. The second is the concern about unauthorized internal or external secondary use, meaning an unauthorized secondary usage of personal data. The third concern is the risk of improper access, that is the risk of some data being accessible when it should not be. Finally, the fourth concern is about errors that may appear during the process of data collection, and that may be difficult to repair. Because they feel that their privacy is not protected enough, consumers may start adopting privacy-enhancing strategies. Villas-Boas (2004) shows how consumers can postpone purchases in order to avoid being identified and thus suffer from price discrimination. Strategically postponing purchasing allows consumers to access lower and non-discriminatory prices.

2.3 Data description

2.3.1 Source of the data

We use a set of surveys that have been jointly conducted every 18 months from 2009 to 2015 by the Caisse des Dépôts and the ACSEL (Association pour le commerce et les services en ligne). These surveys has been conducted using online questionnaires, and the samples are representative of the French Internet population (with respect to age, sex, socioeconomic classification, urban areas and Internet use). They are cross-sectional studies and do not allow us to conduct a panel analysis.

The main objective of the survey is to measure the level of trust of Internet users in several online services (bank, administration, etc.). The survey is divided into several parts that deal with Internet access and use, e-commerce, payment instruments, online banking, online communication (chats, blogs, etc.), social networks, online administration, cloud services, Internet of things, security and authentication, personal data and privacy. Questions related to the use of privacy-enhancing technologies can only be found in the latest survey, the one we use in our estimations, which take the following form:[7]

$$\begin{aligned}(\text{Willingness to supply personal information})_i = \alpha &+ \Sigma_{j=1}^{3}\beta_j(\text{Privacy-enhancing tools} \\ &\quad \text{and strategies})_i \\ &+ \Sigma_{j=4}^{7}\beta_j(\text{Privacy-related risks})_i \\ &+ \Sigma_{j=8}^{9}\beta_j(\text{Online activity})_i \\ &+ \Sigma_{j=10}^{16}\beta_j(\text{Control variables})_i\end{aligned} \tag{2.1}$$

2.3.2 Willingness to share personal information.

We want to estimate what is the effect of the use of PETs on the willingness to share personal information. We construct the latter variable using one specific question from the survey.

The survey questions the respondents about their willingness to disclose six types of personal information (name, telephone number, ID number, bank card and account numbers, health information and personal information like tastes and preferences) with seven different types of actors that are: local territories, State services, banks, operators, online retailers, Internet actors like Google or Microsoft and finally social networks. Two dimensions can therefore be analyzed: whether the respondent wants to disclose to a specific actor or not, and whether the respondent is ready to disclose a specific information or not.

[7] As we use survey data, we have no way of observing the real behavior of respondents. Specifically, we cannot evaluate the gap between responses and actual behavior. This issue is not specific to this study, as it is a general problem with survey data. We believe that this is still relevant to study the relationship between trust over the collection of personal information and the use of privacy-enhancing technologies.

With whom are you ready to share these personal information?

	Local territories	State services	Banks	Operators	Online retailers	Internet actors	Social networks	I do not want to share this information
Name								
Adress								
Phone number								
ID card number								
Bank card or bank account number								
Health information								
Information about tastes or preferences								

We consider the willingness to supply personal information by an individual to be the sum of information he or she is willing to share with all the actors. Consequently, the amount of this willingness to share ranges from 0 to 49.

It is important to notice that this question only relates to the willingness to disclose personal information, and does not allow a measurement of the actual level of disclosure. For example, while it is impossible not to share banking details with the bank, a respondent could still prefer not to share this information with his bank. However, a survey conducted among 2000 representative French Internet users shows that there is little difference between the willingness t disclose and the actual disclosure.[8]

This table allows us to measure the willingness to disclose of all the respondents, whether this willingness translates into actual disclosure or not. This question has remained the same from 2009 to 2015, we are therefore able to study the evolution of the willingness to disclose over this period with cross-sections.

2.4 Model and findings

Before describing our estimation model, we conduct a descriptive analysis in order to guide us.

First, we take a look at how the supply of personal information evolved among representative French Internet users from 2009 to 2015.

[8]Chaire Valeurs et Politiques des Information Personnelles - Données personnelles et confiance : quelles stratégies pour les citoyens-consommateurs en 2017 ?

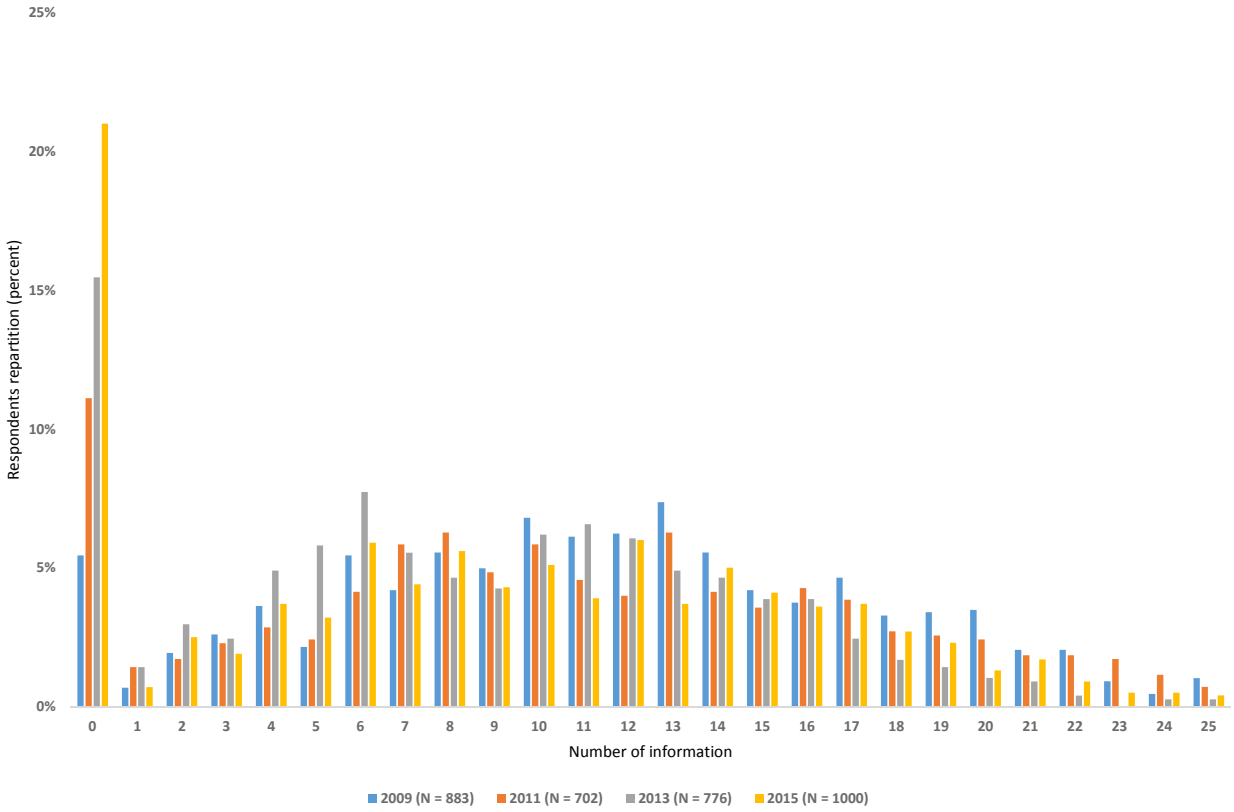


Figure 2.1 – Histogram of respondents by level of willingness to disclose

In 2009, a respondent was willing to make 12 disclosures of personal information in average. This level of disclosure decreased to around 10 in both 2011 and 2013, and finally 9 in 2015. However, the most striking feature of the following graph is indubitably the increase of the proportion of individuals who do not want to share any personal information between each survey. This proportion was equal to 5% in 2009, 11% in 2011, 15% in 2013, and more than 20% in the latest survey in 2015. Furthermore, this category becomes the most frequent from 2011 onwards.

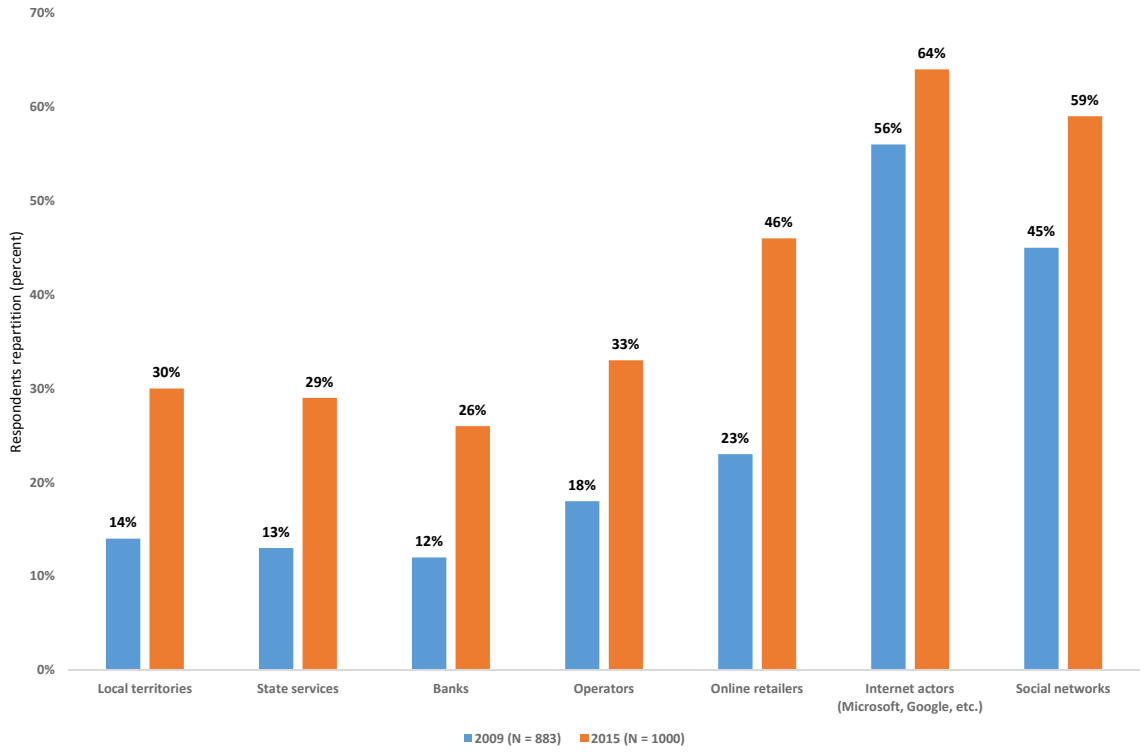


Figure 2.2 – Percent of respondents unwilling to disclose any personal information by actor

This stark decrease in the willingness to disclose personal information has impacted all the actors proposed in the surveys. In 2009, only the Internet actors and social networks seemed to trigger a high level of reluctance by Internet users to share personal data. While the highest percentages of refusal to disclose information are still associated with the Internet actors like Google or Microsoft and the social networks, they also are the actors that showed the smallest increases in the refusal to disclose personal information. This is likely explained by the fact that the percentages for both those type of actors were already high in 2009.

Notwithstanding an increasing reluctance to disclose personal information, the banks, the state services and the local territories are the three type of actors to which respondents are the most willing to disclose personal information in all four surveys.

The refusal to disclose personal information has also increased for all types of personal information. Health and personal (tastes and preferences) information are the most confidential: very few individuals are ready to share this information. The biggest increases in the refusal to disclose from 2009 to 2015 concern the taste and preferences of the individuals, their name,

and telephone number.

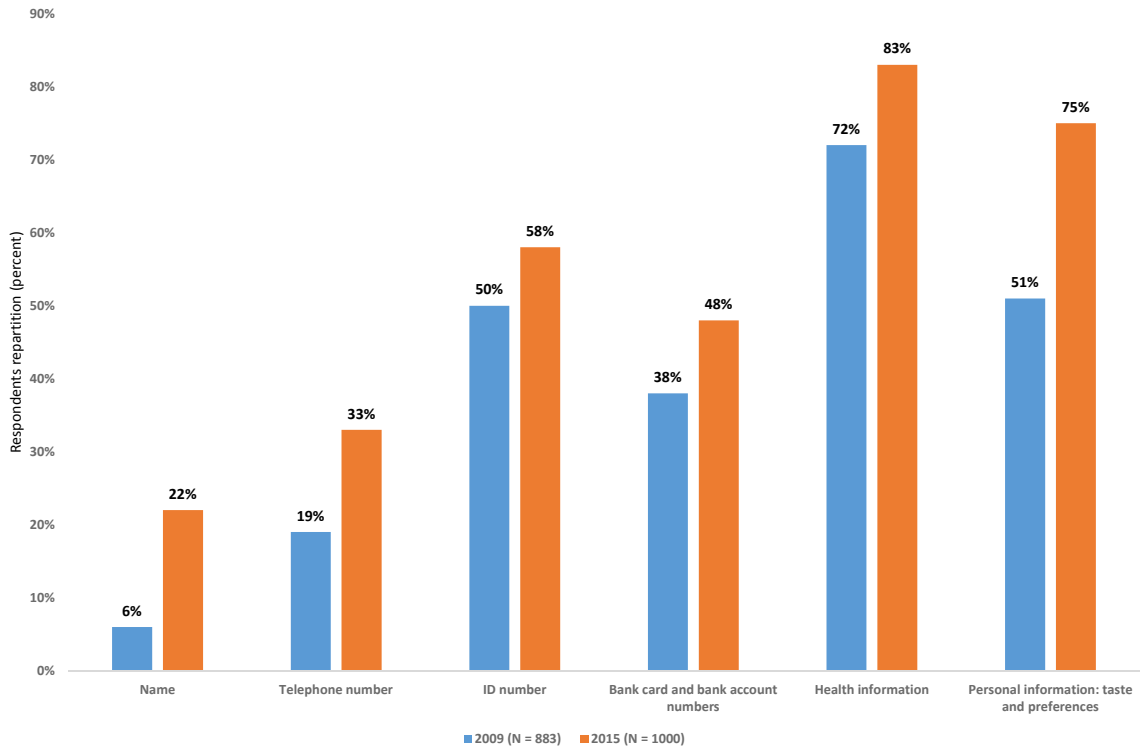


Figure 2.3 – Percent of respondents not willing to disclose by type of personal information

In the end, it is clear that the willingness to disclose personal information has decreased overall from 2009 to 2015, and this for all types of actors or information. This seems to suggest that Internet users have a growing distrust about the handling of their personal information online.

We are, however, mainly interested in the relation between the use of privacy-enhancing technologies and the willingness to disclose personal information. We focus on the survey of 2015 as we only have data on the use of PETs for that year.

An important part of our sample have used privacy protection tools and strategies at some point: 64% have cleared their browsing history, 71% clear their browsing cookies regularly and/or use privacy-enhancing web browser extensions, and finally 44% use ad blockers. These figures are similar to those of the Pew Research Center, as they found for example in 2013 that 64% of the people they interrogated had at some point cleared their cookies and/or their

browsing history.

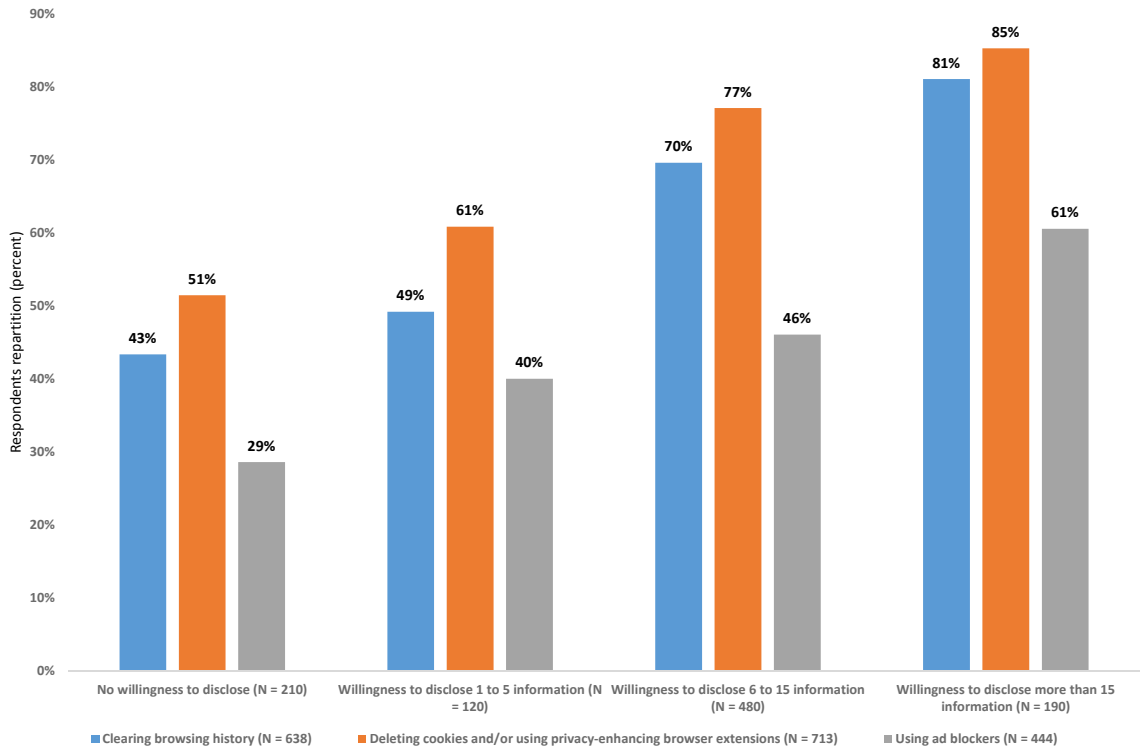


Figure 2.4 – Use of privacy-enhancing tools and strategies by level of willingness to disclose personal information

The proportion of individuals using privacy-enhancing tools and strategies is clearly increasing with the willingness to disclose personal information: while for example 43% of those who do not want to share any information have cleared their browsing history at some point, 81% have done so among those who are willing to disclose more than 15 personal information.

It then seems clear that there is no negative correlation between using PETs or privacy-enhancing strategies and the willingness to disclose personal information. Even more so, it appears that Internet users that have adopted ways of controlling their privacy online are ready to share more personal information. The following correlation confirms this statement:

Table 2.I – Correlation between the willingness to disclose and the use of PETs

	No disclosure	1 to 5 disclosures	6 to 15 disclosures	More than 15 disclosures	Clearing browsing history	Deleting cookies and/or using privacy-enhancing browser extensions	Using ad blockers
No disclosure	1						
1 to 5 disclosures	-0.1904	1					
6 to 15 disclosures	-0.4954	-0.3548	1				
More than 15 disclosures	-0.2497	-0.1788	-0.4653	1			
Clearing browsing history	-0.2196	-0.1124	0.1156	0.1739	1		
Deleting cookies and/or using privacy-enhancing browser extensions	-0.2265	-0.0854	0.1228	0.1495	0.4651	1	
Using ad blockers	-0.1643	-0.0327	0.0317	0.1572	0.1831	0.3045	1

It is worth noting that the use of privacy-enhancing tools and strategies is not correlated with the willingness to disclose personal information in a linear way.

2.4.1 The model

We want to test whether the use of privacy protection tools and strategies, when controlling for various other effects, positively influences or not the willingness to disclose personal information.

This willingness, our dependent variable, is categorized as follows: 0 for "no disclosure", 1 for ”1 to 5 disclosures", 2 for "6 to 15 disclosures" and 3 for "more than 15 disclosures". We categorize our dependent variable as the effect of the use of privacy-enhancing tools and strategies on the willingness to disclose is not linear. We provide an OLS estimation without this categorization in the appendix.[9]

Our equation is the following:

$$\begin{aligned}(\text{Willingness to disclose personal information})_i = \alpha &+ \Sigma_{j=1}^{3}\beta_j(\text{Privacy-enhancing tools} \\ &\text{and strategies})_i \\ &+ \Sigma_{j=4}^{7}\beta_j(\text{Privacy-related risks})_i \\ &+ \Sigma_{j=8}^{9}\beta_j(\text{Online activity})_i \\ &+ \Sigma_{j=10}^{16}\beta_j(\text{Control variables})_i\end{aligned} \tag{2.2}$$

[9]We also display the results of the Brant Test of parallel regression assumption in the appendix.

First, we use three binary variables that indicates whether an individual has a strategic approach over his or her supply of personal information (”Privacy-enhancing tools and strategies”): having cleared web browsing history or not, having cleared online cookies and/or used privacy-enhancing web browser extensions or not, and using ad blockers or not. These variables allow us to describe three dimensions of privacy concerns: privacy with respect to the entourage (clearing browsing history), privacy with respect to data collection (clearing cookies, using privacy-enhancing web browser extensions) and privacy with respect to advertising nuisance (using ad blockers). We expect to find that these variables have a positive impact on the willingness to disclose personal data.

Second, we use four binary variables to control for several risks of online information disclosure (”Privacy-related risks”): the perceived risk of family or friends accessing personal information on social networks, the perceived risk of governmental services keeping personal information indefinitely, the perceived risk of State surveillance when using governmental services and finally refusing to be geolocated or not. We expect these variables to have a negative impact on the willingness to disclose personal information.

Third, we use two binary variables to account for the level of online activity (”Online activity”): the frequency of use of Internet and the online purchasing frequency. We expect these two variables to have a positive effect on the willingness to share personal information.

Finally, we use several control variables: the highest diploma obtained (no diploma, high school, undergraduate and graduate studies), gender,the socio-economic category (1=employed, worker or farmers, 0 = other categories) and age category (15-24 y.o., 25-34 y.o., 35-49 y.o., more than 50 y.o.).

We use an ordered probit, as well as an ordinary least squares (OLS) regression for comparison’s sake, to estimate Equation (2.2).

The following tables show the correlation matrix of the explanatory variables.

Table 2.II – Cross-correlation table

Variables	Clearing browsing history	Deleting cookies and/or using privacy-enhancing browser extensions	Using ad blockers	Perceived risk of family or friends accessing personal information on social networks	Perceived risk of governmental services keeping personal information indefinitely
Clearing browsing history	1				
Deleting cookies and/or using privacy-enhancing browser extensions	0.4103	1			
Using ad blockers	0.1354	0.2625	1		
Perceived risk of family or friends accessing personal information on social networks	-0.0284	-0.0191	0.0261	1	
Perceived risk of governmental services keeping personal information indefinitely	0.0706	0.0575	0.0519	-0.0820	1
Perceived risk of State surveillance	0.0371	0.0341	0.0372	0.0793	0.0958
Refusing to be geolocated	0.0458	0.0515	0.0194	-0.0411	0.0673
Use of Internet frequency	0.2026	0.2288	0.1993	-0.0373	0.0153
Online purchasing frequency	-0.0006	0.0242	0.1003	-0.0142	0.0054
Education	0.0983	0.1227	0.1402	-0.0389	0.0198
Higher socioeconomic classification	0.0608	0.0969	0.1015	-0.0571	0.0272
Lower socioeconomic classification	-0.1138	-0.0774	-0.0780	0.1005	0.0143
Being between 15 and 24 years old	0.1282	0.0277	0.0659	-0.0369	0.0303
Being between 15 and 24 years old	-0.0032	0.0103	0.0069	0.0085	0.0030
Being more than 50 years old	-0.0617	-0.0261	-0.0571	0.0428	-0.0054

Cross-correlation table (continued)

Variables	Perceived risk of if State surveillance	Refusing to be geolocated	Use of Internet frequency	Online purchasing frequency	Education
Perceived risk of State surveillance	1				
Refusing to be geolocated	0.0514	1			
Use of Internet frequency	0.0094	-0.0506	1		
Online purchasing frequency	-0.0033	-0.1257	0.1407	1	
Education	-0.0218	-0.0044	0.2317	0.0997	1
Higher socioeconomic classification	-0.0440	-0.0035	0.0529	0.1336	0.3169
Lower socioeconomic classification	0.0099	0.0188	-0.0792	-0.0331	-0.1155
Being between 15 and 24 years old	0.0081	-0.0173	0.2218	0.0305	-0.0008
Being between 15 and 24 years old	-0.0610	-0.0316	0.0201	0.1129	0.1284
Being more than 50 years old	0.0776	-0.0230	-0.1212	-0.0854	-0.1231

Cross-correlation table (continued)

Variables	Higher socioeconomic classification	Lower socioeconomic classification	Being between 15 and 24 years old	Being between 25 and 34 years old	Being more than 50 years old
Higher socioeconomic classification	1				
Lower socioeconomic classification	-0.4659	1			
Being between 15 and 24 years old	-0.2118	-0.1245	1		
Being between 15 and 24 years old	0.0198	0.2283	-0.2211	1	
Being more than 50 years old	0.0090	-0.2451	-0.3749	-0.3537	1

2.4.2 Findings

The estimation results of Equation (2.2) are reported in Table 2.III. Table 2.IV displays the marginal effects.

Table 2.III – Result estimations

Dependent variable: Number of disclosures	Ordered probit	OLS
Privacy- enhancing tools and strategies		
Clearing browsing history	0.379***	0.319***
	(0.094)	(0.075)
Deleting cookies	0.292***	0.249***
and/or using privacy-enhancing browser extensions	(0.104)	(0.083)
Using ad blockers	0.266***	0.201***
	(0.083)	(0.067)
Privacy-related risks		
Perceived risk of family or friends	−0.261**	−0.197**
accessing personal information on social networks	(0.118)	(0.095)
Perceived risk of governmental services	−0.350**	−0.272***
keeping personal information indefinitely	(0.114)	(0.091)
Perceived risk of State surveillance	−0.224*	−0.186**
when using governmental services	(0.117)	(0.094)
Refusing to be geolocated	−0.388***	−0.312***
	(0.090)	(0.071)
Online activity		
Use of Internet frequency	0.175***	0.140***
	(0.065)	(0.052)
Online purchasing frequency	0.130**	0.111***
	(0.054)	(0.043)
Control variables	Yes	Yes
Constant cut 1	0.0223	
	(0.236)	
Constant cut 2	0.438	
	(0.237)	
Constant cut 3	2.003***	
	(0.244)	
Constant		0.857***
		(0.189)
Observations	811	811
Pseudo R^2 - R^2	0.0753	0.1659
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2.IV – Marginal effects

Dependent variable: Number of disclosures	No disclosure	1 to 5 disclosures	6 to 15 disclosures	More than 15 disclosures
Privacy- enhancing tools and strategies				
Clearing browsing history	-0.084^{***} (0.021)	-0.029^{***} (0.007)	0.014^{**} (0.006)	0.099^{***} (0.024)
Deleting cookies and/or using privacy-enhancing	-0.065^{***} (0.023)	-0.022^{***} (0.008)	0.011^{**} (0.005)	0.077^{***} (0.027)
browser extensions Using ad blockers	-0.060^{***} (0.018)	-0.020^{***} (0.006)	0.010^{**} (0.005)	0.070^{***} (0.021)
Privacy-related risks				
Perceived risk of family or friends accessing personal information on social networks	0.057^{**} (0.026)	0.020^{**} (0.009)	-0.009^{*} (0.005)	-0.068^{**} (0.031)
Perceived risk of governmental services keeping personal information indefinitely	0.077^{***} (0.025)	0.026^{***} (0.009)	-0.012^{**} (0.006)	-0.091^{***} (0.030)
Perceived risk of State surveillance when using governmental services	0.050^{*} (0.026)	0.017^{*} (0.009)	-0.008 (0.005)	-0.060^{*} (0.031)
Refusing to be geolocated	-0.388^{***} (0.090)	0.029^{***} (0.007)	-0.014^{**} (0.006)	-0.102^{***} (0.023)
Online activity				
Use of Internet frequency	-0.039^{***} (0.014)	-0.013^{***} (0.005)	0.006^{*} (0.003)	0.046^{***} (0.017)
Online purchasing frequency	-0.029^{**} (0.012)	-0.010^{**} (0.004)	0.005^{*} (0.002)	0.034^{**} (0.014)
Control variables	Yes	Yes	Yes	Yes
Note:	*p<0.1; **p<0.05; ***p<0.01			

We indeed find that, all other things being equal, using privacy-enhancing technologies positively influences the willingness to share personal information. The positive effects we find on the willingness to disclose personal data are strong and highly significant (at the 1 per cent level). Furthermore, they are consistent whether we use an ordered probit or an OLS regression. This is true for all three of our variables: clearing browsing history (0.379 for the ordered probit and 0.319 for the OLS), deleting cookies and/or using privacy-enhancing browser extensions (0.292 and 0.249) and using ad blockers (0.266 and 0.201).

These results tend to show that Internet users that have adopted privacy-enhancing technologies and other obfuscation techniques are willing to share more information on average. This suggests that using PETs helps Internet users to restore their trust in the handling of their personal information. While firms may consider improving privacy (for example requesting consent for data collection) as a cost, it could actually be a way to collect more information.

Other results are as we expected. All variables indicating privacy-related risks are negatively affecting the willingness to disclose personal information. For example, the effect associated with refusing to be geolocated is negative at the 1 per cent level (-0.388 for the ordered probit and -0.132 for the OLS). Also as expected, we find that the more Internet users are active online, the more they are willing to share personal data: using Internet more frequently is for example associated with a positive coefficient at the 1 per cent level (0.175 for the ordered probit and 0.140 for the OLS).[10]

We now estimate Equation (2.2) with respect to four different actors: State services, banks, online retailers and social networks. The willingness to share personal information with each of these actors ranges from 0 to 7.

We change the specification of privacy-related risks with respect to Equation (2.2). We adapt these risks for each estimation, depending on the actor:

- State services: perceived risk of governmental services keeping personal information indefinitely, perceived risk of State surveillance when using governmental services, refusing to be geolocated
- banks: perceived risk that financial data might be unfairly used when using online banking services, refusing to be geolocated
- online retailers: perceived risk of online retailers keeping personal information indefinitely, refusing to be geolocated
- social networks: perceived risk of family or friends accessing personal information on social networks, perceived risk of general loss of privacy on social networks, refusing to be geolocated

[10]For reason of convenience, we did not report results for ”control variables”: they are available upon request.

The rest of the specification remains the same as in Equation (2.2). We present the results for an ordered probit estimation in Table 2.[11]

Again, we find for each actor that using privacy-enhancing technologies have a positive impact on the willingness to share personal information. This seems especially true for State services, where all three variables are significant at the 1 per cent level.

The positive impact of clearing browsing history on the willingness to share personal information is the biggest for banks and online retailers (coefficients of 0.421 and 0.418). Online activity is still having a positive effect on the willingness to disclose personal data.

Overall, the use of PETs by consumers has positive effects for those who want to get access to personal information. With PETs, consumers seem more inclined to use services from their government, their banks, online retailers and social networks. All these actors which want to have access to consumers' personal information could then have an incentive to help the development of PETs or design services more respectful of privacy.

In the end, an increasing number of more active Internet users with respect to their privacy could be a good opportunity for the digital economy, from e-commerce to social networks.

[11] Results with an OLS estimation are available in the appendix.

Table 2.V – Result estimations (ordered probit)

Dependent variable: Number of disclosures	State services	Banks	Online retailers	Social networks
Privacy- enhancing tools and strategies				
Clearing browsing history	0.293*** (0.090)	0.421*** (0.081)	0.418*** (0.094)	0.263** (0.102)
Deleting cookies and/or using privacy-enhancing browser extensions	0.312*** (0.100)	0.292*** (0.089)	0.273*** (0.106)	0.140 (0.113)
Using ad blockers	0.240*** (0.078)	0.149** (0.071)	0.289** (0.081)	0.190** (0.087)
Privacy-related risks				
Perceived risk of family or friends accessing personal information on social networks				−0.289** (0.129)
Perceived risk of general loss of privacy on social networks				−0.204** (0.099)
Perceived risk of governmental services keeping personal information indefinitely	−0.240** (0.108)			
Perceived risk of State surveillance when using governmental services	−0.310*** (0.113)			
Perceived risk that financial data might be unfairly used when using online banking services		−0.122* (0.068)		
Perceived risk of online retailers keeping personal information indefinitely			−0.175** (0.082)	
Refusing to be geolocated	−0.163* (0.084)	−0.303*** (0.076)	−0.251*** (0.086)	−0.552*** (0.092)
Online activity	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Constant cut 1	0.257 (0.225)	0.265 (0.182)	0.465 (0.234)	0.289 (0.254)
Constant cut 2	0.728*** (0.226)	0.632*** (0.183)	1.108*** (0.236)	1.189*** (0.256)
Constant cut 3	1.389*** (0.228)	1.248*** (0.185)	1.849*** (0.240)	2.266*** (0.268)
Constant cut 4	2.006*** (0.231)	1.951*** (0.188)	2.755*** (0.251)	3.205*** (0.345)
Constant cut 5	2.684*** (0.238)	3.089*** (0.204)	3.751*** (0.333)	3.425*** (0.399)
Constant cut 6	3.347*** (0.258)	3.640*** (0.233)		
Observations	811	811	811	811
Pseudo R^2	0.0403	0.0537	0.0468	0.0695
Note:	*p<0.1; **p<0.05; ***p<0.01			

2.5 Conclusion

The study we have conducted in this paper finds what could be at first considered to be a counterintuitive result: the adoption of privacy-enhancing technologies and strategies is associated with an increasing willingness of consumers to share personal information. Indeed, we show using French survey data that consumers that have taken steps to improve their privacy online express a willingness to share more personal information than the average.

While privacy has classically been interpreted as a cost for firms, we find that it could on the contrary represent an economic opportunity. Specifically, improving privacy could resolve the trade-off that the data-driven economy increasingly faces between benefiting from data collection in the short run and loosing consumer trust in the long run. In a context dominated by a loss of trust of consumers toward actors collecting and handling their personal information, encouraging the existence and the adoption of privacy-enhancing technologies could help to restore this trust. Overall, the development of PETs could then have a positive impact on the digital economy, making consumers more at ease when disclosing their personal information.

Taking this paper into account, the adoption of the "General Data Protection Regulation" (GDPR) by the European Union represents an opportunity for firms.[12] One of the key feature of the GDPR is that firms must have consumers' consent to collect and use their personal information: "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her". Our study finds that this regulation may help firms to collect personal information and restore trust concerning privacy issues.

Our study does not imply however that regulatory authorities have no role to play. Adverse selection and trust manipulation are still possible outcomes, and authorities must insure that privacy preferences of consumers are actually respected.

[12]The full text can be retrieved at the following address: http://data.europa.eu/eli/reg/2016/679/oj.

BIBLIOGRAPHY

[1] **Acquisti, Alessandro**, "The Economics of Personal Data and the Economics of Privacy," *OECD Background Paper*, December 2010.

[2] ____ **and Hal R. Varian**, "Conditioning prices on purchase history," *Marketing Science*, 2005, *24* (3), 1–15.

[3] **Ball, A. Dwayne, Pedro S. Coelho, and Manuel J. Vilares**, "Service personalization and loyalty," *Journal of Services Marketing*, 2006, *20* (6), 391–403.

[4] **Bennett, James and Stan Lanning**, "The Netflix prize," *Proceedings of KDD Cup and Workshop*, 2007.

[5] **Blattberg, Robert C. and John Deighton**, "Interactive marketing: Exploiting the age of addressability," *Sloan Management Review*, 1991, *33* (1), 5–14.

[6] **Bounie, David, Marc Bourreau, Michel Gensollen, and Patrick Waelbroeck**, "Do Online Customer Reviews Matter? Evidence from the Video Game Industry," *Telecom ParisTech Working Paper No. ESS-08-02*, 2008.

[7] **Cabral, Luis and Ali Hortascu**, "The Dynamics of Seller Reputation: Theory and Evidence from eBay," *Journal of Industrial Economics*, 2010, *58.*

[8] **Fudenberg, Drew and Jean Tirole**, "Upgrades, Tradeins, and Buybacks," *The RAND Journal of Economics*, 1998, *29* (2), 235–258.

[9] **Goldfarb, Avu and Catherine E. Tucker**, "Privacy Regulation and Online Advertising," *Management Science*, 2010, *57* (1), 57–71.

[10] **Linden, Greg, Brent Smith, and Jeremy York**, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," *IEEE Internet computing*, 2007, *7* (1), 76–80.

[11] **Posner, Richard A.**, "The Right of Privacy," *Georgia Law Review*, 1978, *12* (3), 393–422.

[12] ____ , "The Economics of Privacy," *The American Economic Review*, 1981, *71* (2), 405–409.

[13] **Richards, Keith A. and Eli Jones**, "Customer relationship management: Finding value drivers," *Industrial Marketing Management*, 2008, *37* (2), 120–130.

[14] **Smith, H.Jeff, Sandra J. Milberg, and Sandra J. Burke**, "Information Privacy: Measuring Individuals' Concerns about Organizational Practices," *MIS Quaterly*, 1996, *20* (2), 167–196.

[15] **Spence, Michael**, "Job Market Signaling," *The Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

[16] **Stigler, George J.**, "An Introduction to Privacy in Economics and Politics," *The Journal of Legal Studies*, 1980, *9* (4), 623–644.

[17] **Taylor, Curtis R.**, "Consumer privacy and the market for customer information," *The RAND Journal of Economics*, 2004, *35* (4), 631–651.

[18] **Varian, Hal R.**, "Price discrimination and social welfare," *American Economic Review*, 1985, *75* (4), 870–875.

[19] ____ , "Economic Aspects of Personal Privacy," *Technical Report, University of California, Berkeley*, 1996.

[20] **Villas-Boas, J.Miguel**, "Price Cycles in Markets with Customer Recognition," *The RAND Journal of Economics*, 2004, *35* (3), 486–501.

2.6 Appendix 1: OLS estimations

Table 2.VI – Result estimations with uncategorized dependent variable (OLS)

Dependent variable: Number of disclosures	
Privacy- enhancing tools and strategies	
Clearing browsing history	2.441*** (0.527)
Deleting cookies and/or using privacy-enhancing browser extensions	1.834*** (0.584)
Using ad blockers	1.593*** (0.466)
Privacy-related risks	
Perceived risk of family or friends accessing personal information on social networks	−1.686** (0.666)
Perceived risk of governmental services keeping personal information indefinitely	−1.767*** (0.642)
Perceived risk of State surveillance when using governmental services	−1.642** (0.662)
Refusing to be geolocated	−1.958*** (0.500))
Online activity	
Use of Internet frequency	1.099*** (0.366)
Online purchasing frequency	0.335 (0.302))
Control variables	Yes
Constant	3.887*** (1.325)
Observations	811
R^2	0.1818
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2.VII – Result estimations - State and Banks (OLS)

Dependent variable: Number of disclosures	State services	Banks
Privacy- enhancing tools and strategies		
Clearing browsing history	0.405*** (0.124)	0.632*** (0.106)
Deleting cookies and/or using privacy-enhancing browser extensions	0.390*** (0.138)	0.415*** (0.116)
Using ad blockers	0.345*** (0.110)	0.238** (0.095)
Privacy-related risks		
Perceived risk of governmental services keeping personal information indefinitely	−0.316** (0.150)	
Perceived risk of State surveillance when using governmental services	−0.416*** (0.155)	
Perceived risk that financial data might be unfairly used when using online banking services		−0.160* (0.090)
Refusing to be geolocated	−0.224* (0.118)	−0.406*** (0.102)
Online activity	Yes	Yes
Control variables	Yes	Yes
Constant	0.640** (0.311)	1.077*** (0.207)
Observations	811	811
R^2	0.1285	0.1593
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 2.VIII – Result estimations - Online retailers and Social networks (OLS)

Dependent variable: Number of disclosures	Online retailers	Social networks
Privacy- enhancing tools and strategies		
Clearing browsing history	0.403***	0.163**
	(0.094)	(0.066)
Deleting cookies and/or using privacy-enhancing browser extensions	0.263**	0.075
	(0.105)	(0.074)
Using ad blockers	0.290***	0.128**
	(0.083)	(0.059)
Privacy-related risks		
Perceived risk of family or friends accessing personal information on social networks		−0.207**
		(0.083)
Perceived risk of general loss of privacy on social networks		−0.115*
		(0.065)
Perceived risk of online retailers keeping personal information indefinitely	−0.187**	
	(0.084)	
Refusing to be geolocated	−0.241***	−0.380***
	(0.089)	(0.063)
Online activity	Yes	Yes
Control variables	Yes	Yes
Constant	0.434*	0.616***
	(0.238)	(0.167)
Observations	811	811
R^2	0.1196	0.1377
Note:	*p<0.1; **p<0.05; ***p<0.01	

2.7 Appendix 2: Brant test

Table 2.IX – Brant test results

Variables	χ^2	P-value	Degree of freedom
All	34.89	0.247	30
Privacy- enhancing tools and strategies			
Clearing browsing history	0.54	0.765	2
Deleting cookies and/or using privacy-enhancing browser extensions	0.09	0.956	2
Using ad blockers	1.36	0.506	2
Privacy-related risks			
Perceived risk of family or friends accessing personal information on social networks	0.55	0.760	2
Perceived risk of governmental services keeping personal information indefinitely	0.38	0.827	2
Perceived risk of State surveillance when using governmental services	0.93	0.627	2
Refusing to be geolocated	0.44	0.802	2
Online activity			
Use of Internet frequency	0.94	0.624	2
Online purchasing frequency	1.00	0.607	2
Control variables	Yes		

2.8 Appendix 3: Descriptive statistics

Table 2.X – Statistics - Number of disclosures - Total

0 disclosure = 1	1 to 5 disclosures = 2	6 to 15 disclosures = 3	More than 15 disclosures = 4	N	Mean	St. Dev.
210	120	480	190	1000	1.65	0.048

Table 2.XI – Statistics - Number of disclosures - State services

0	1	2	3	4	5	6	N	Mean	St. Dev.
291	143	234	170	108	40	14	1000	1.837	0.05

Table 2.XII – Statistics - Number of disclosures - Banks

0	1	2	3	4	5	6	N	Mean	St. Dev.
259	115	216	218	165	20	7	1000	2.003	0.048

Table 2.XIII – Statistics - Number of disclosures - Online retailers

0	1	2	3	4	5	6	N	Mean	St. Dev.
457	203	200	114	24	2	0	1000	1.051	0.037

Table 2.XIV – Statistics - Number of disclosures - Social networks

0	1	2	3	4	5	6	N	Mean	St. Dev.
594	250	133	21	1	1	0	1000	0.588	0.026

Table 2.XV – Statistics - Binary variables

	N	Mean	St. Dev.
Privacy- enhancing tools and strategies			
Clearing browsing history	1000	0.638	0.0152
Deleting cookies and/or using privacy-enhancing browser extensions	1000	0.713	0.0143
Using ad blockers	1000	0.444	0.0157
Privacy-related risks			
Perceived risk of family or friends accessing personal information on social networks	1000	0.129	0.0106
Perceived risk of general loss of privacy on social networks	1000	0.255	0.0138
Perceived risk of governmental services keeping personal information indefinitely	1000	0.149	0.0113
Perceived risk of State surveillance when using governmental services	1000	0.137	0.0109
Perceived risk that financial data might be unfairly used when using online banking services	1000	0.516	0.0158
Perceived risk of online retailers keeping personal information indefinitely	1000	0.33	0.0149
Refusing to be geolocated	1000	0.737	0.0139
Individual variables			
Lower socioeconomic classification	1,000	0.294	0.0456
Higher socioeconomic classification	1,000	0.306	0.0146
Being between 15 and 24 years old	1,000	0.182	0.0386
Being between 25 and 34 years old	1,000	0.159	0.0366
Being more than 50 years old	1,000	0.397	0.0490

Table 2.XVI – Statistics - Frequency of purchase

Less often than once per month = 1	More than once once per month = 2	Once per week =3	Several times per week = 4	N	Mean	St. Dev.
316	394	72	29	811	1.77	0.75

Table 2.XVII – Statistics - Education

No diploma = 1	High school diploma = 2	Undergraduate and graduate studies = 3	N	Mean	St. Dev.
231	253	516	1000	2.285	0.82

Table 2.XVIII – Statistics - Use of Internet frequency

Less than several times per week = 1	Several times per week = 2	Several times per day = 3	N	Mean	St. Dev.
135	225	641	1000	2.506	0.023

CHAPTER 3

PAYMENT INSTRUMENTS, FINANCIAL PRIVACY AND ONLINE PURCHASES

3.1 Introduction

The protection of personal information has become a major concern for consumers in the digital economy (Acquisti et al., 2015). Personal *financial* information are no exception. As described by Lacker (2002), "a consumer's financial transactions give rise to a wealth of very personal data. Every credit card purchase, every ATM withdrawal, every loan payment, every paycheck deposit leaves an electronic trace at a person's bank." And with the advances in information technology, a variety of industries, including card networks, banks, retailers, and advertisers can easily track what consumers purchase, how much they spend, their location, their willingness to pay for specific items, etc., to a number of ends.

Banks for example can use payments data to cross-sell financial products such as credits (Mester et al., 2007), to detect fraudulent use of debit/credit cards, or again to create card-related reward programs (Ching and Hayashi, 2010). Lloyds Banking Group, the largest bank in the UK with 22 million current account customers, launched for instance in 2013 a program that allows customers to receive cashback offers from big brands (Starbucks, Hertz, etc.) based on where they actually spend their money (as identified by their debit and credit card purchase history). This program let people earn back between 5% and 15% when they shop at retailers (The Guardian, "Lloyds bank jumps on the cashback bandwagon", November 2, 2013). Card networks also compile audience segments (sporting goods store, specialty retailer, etc.), and sell them to online data brokers (Exelate) or ad companies to target audiences geographically. For example, "using the Mastercard data, a burger or pizza chain might use the system to push promotions to neighborhoods in which people spend more than the average at fast-food joints." (AdAge, "Mastercard, AmEx Quietly Feed Data to Advertisers", April 16, 2013.)

For all of these reasons, consumers can be reluctant to shop online using debit/credit cards issued by banks to avoid to be tracked, targeted or solicited.[1] This reticence may be even

[1] An article on the Belgium RTBF website in November 2015 relates the story of a researcher that saw a bank transfer from an NGO helping victims of the civil war in Syria blocked by her bank because the word 'Syria' appeared in the title of the transaction. It was a legitimate reimbursement of travel expenses, but the bank considered it as a potential attempt at money laundering or funding of terrorist activities. The article can be retrieved at the following address; last visit: 11/12/2015.

stronger when consumers want to buy sensitive goods such as medication, healthcare expenses, gifts, gambling, adult products and so on. For the same reasons, they may also want to hide information from relatives or other people (in the case of joint accounts), or from government-related institutions who may access bank statements.

To preserve their personal financial data, privacy-minded users may decide to use non-bank payment instruments delivered by firms that are not directly attached to their checking account.[2] For example, the transactions carried out with PayPal in France cannot be completely tracked by banks even though consumers use the cards issued by their bank: a transaction made with PayPal does not give rise to a similar writing on the consumer's checking account and the bank has neither information on the purchased item nor information on the retailer. Similarly, the use of electronic payments systems such as electronic currencies (Bitcoin and other cryptocurrencies) are completely anonymous and disconnected from bank accounts. Such payment services may therefore be used by consumers in online purchases to preserve personal data from banks and relatives.[3] This is exactly what is confirmed by the 2014-2015 Survey of Consumer Payment Choice (SCPC) from the Federal Reserve Bank of Boston (Schuh and Shy, 2016): about 20 per cent of the US consumers adopted virtual currencies because they distrust banks or sovereign currency or because they want to make payments anonymously.

This paper precisely investigates whether the use of non-bank payment instruments that preserve financial privacy from banks, relatives or government-related institutions may increase online purchases. The intuition is simple. Using bank payment instruments such as debit or credit cards to purchase privacy-sensitive goods may entail an extra cost for consumers as banks can clearly track their purchases. In the absence of non-bank payment instruments, a part of consumers may simply decide to stop shopping online in order to preserve their personal financial information. This line of reasoning is supported by the empirical privacy literature: privacy concerns negatively affect online purchases (Akhter, 2012). However, when consumers can use non-bank payment instruments, they can now buy online without sharing personal financial information with banks or relatives, resulting in a possible positive impact on online purchases. This intuition is confirmed by our empirical analysis. We use a survey conducted among a representative sample of 1,000 French Internet consumers in 2015, and analyze the

[2]The terminology "payment instrument" is commonly used in the literature to qualify a means of payment (Bagnall et al., 2016). This terminology should not be confused with 'instruments' or 'instrumental variables' used in econometrics and later on in Section 4.

[3]Bitcoins are not only used for transactional purposes but also for speculative ones (Bolt and van Oordt, 2016).

online purchasing decisions as well as the use of bank and non-bank payment instruments. Using an econometric model with two equations that we estimate by a two-step regression and a Bayesian Markov Chain Monte Carlo model to account for a potential endogeneity problem, we find evidence that the use of non-bank payment instruments positively influences consumers' online purchases.

This paper contributes to the economic literature on two points. First, previous contributions in the economic literature have separately focused on privacy, payments and online purchases. To the best of our knowledge, this paper is the first to merge these different strands of the literature and to analyze how financial privacy concerns may induce consumers to choose specific payment instruments for not disclosing personal financial information. As the review of the literature will show, financial privacy has only been studied from the viewpoint of the regulation of financial intermediaries. Our paper contributes therefore to the literature by showing that financial privacy is also a concern for consumers that should be taken into account to promote a sound development of the digital economy. Second, this paper proposes an original estimation method to test the model predictions. We use Bayesian econometrics to estimate an endogenous binary variable model that deals explicitly with the existence of a potential endogeneity issue if the use of a non-bank payment instrument is correlated with unobservable variables that influence online purchases.

The article proceeds as follows. Section 2 provides a discussion of the relevant literature. Section 3 describes the data, present the estimation strategy and discuss the estimation results. Section 4 concludes.

3.2 Related literature

This paper studies how consumers may adopt strategic behaviors when using payment instruments in online purchases to protect their financial privacy from banks and/or relatives. It is at the crossroads of different strands of the literature on the economics of privacy, the regulation of financial privacy, information sharing between banks, and the economics of payments.

First, two papers on the economics of privacy analyze the link between privacy issues and electronic commerce. Akhter (2012) analyzes survey data from a sample of 1,097 Internet subscribers in three Midwest states in the United States. The author finds that privacy concern has a negative and statistical significant influence on online spending. Similarly, Tsai et al. (2011) de-

sign an experimental study based on a search engine that displays the privacy policies of specific online shopping sites. They test whether participants presented with salient privacy information would be more likely to purchase from sites with privacy indicators than participants who did not see that information. They find that participants provided with salient privacy information took that information into consideration, making purchases from websites offering medium or high levels of privacy. Overall, these papers show that privacy concerns negatively affect online purchases. Our paper confirms in part their findings as we find that privacy-minded users purchase less online than the others. However, we also find evidence that online consumers may adopt strategies to protect their privacy from banks and relatives by using non-bank payment instruments, resulting in a positive influence on online purchases.

The literature on the economics of privacy has also considered the use of strategies by consumers to avoid price discrimination (Acquisti et al., 2015). For instance, Villas-Boas (2004) show that consumers can postpone their purchase to avoid being identified by the price-discriminating firm. Likewise, Conitzer et al. (2012) show that consumers can decide to remain anonymous from the retailer in order to avoid price discrimination. We extend this idea in our paper by showing that consumers may adopt strategies in order to avoid discriminatory practices based on their transaction history by financial institutions (scoring practices and other commercial use of personal data).

A number of other studies have also specifically focused on the effects of financial privacy regulations. In the U.S. for instance, the Gramm-Leach-Bliley Act (GLBA) allows a variety of financial institutions to collect, share and use personal information about their customers. The GLBA requires financial institutions to provide each consumer with a privacy notice explaining the exploitation of their personal data. The notice must also identify the consumer's right to opt out of the information being shared with unaffiliated parties. However, there are two main exemptions that authorizes information sharing despite objections from consumers: first it allows an institution to disclose personal information to affiliated institutions without providing notice of the disclosure and an opportunity to opt out; second, it allows an institution to disclose non-public personal financial information to non-affiliated third parties that jointly offer marketing with the original institution. If Lacker (2002) argues that the market for financial privacy has the characteristics that should yield efficient outcomes, Swire (2002) advances that the GLBA could lead financial institutions to review their data exploitation practices and to get rid of those with a low respect of privacy. The GLBA has also drawn criticisms concerning the level of

its privacy protection. Janger and Schwartz (2002) for example consider that consumers are not sufficiently informed and that this lack of information reduces their bargaining power with financial institutions. Concerning the impact of the GLBA, Sheng and Cranor (2006) find that the sharing of information about consumers between affiliates and non-affiliates has increased since the adoption of the GLBA. Similarly, Cranor et al. (2013) show that an important number of financial institutions shares consumers' data without allowing them to limit or to stop data disclosure.

The legislation regarding financial privacy in Europe is less permissive than in the U.S. Financial privacy is regulated by the Data Protection Directive 95/46/ec, which allows financial institutions to collect data about consumer identification data and products and services management, but not to sell or share them with non-affiliates. Financial institutions can share data among affiliates, but not without an authorization by both privacy regulation authorities and consumers. Jentzsch (2007) finds however that financial privacy regulation in Europe does not significantly reduce credit reporting practices. U.S. financial institutions have a more intensive use of credit reporting, but the difference is not imputable to differences in regulation between Europe and the U.S. Our analysis does not directly deal with the efficiency of financial privacy regulations but show however that consumers are sensitive to financial privacy concerns and that prohibiting banks to communicate financial information with non-affiliates allow precisely strategic consumers to shop more often online.

This paper is also related to the literature on information sharing in credit markets. Pagano and Japelli (199) showed that information sharing between lending institutions helps decreasing adverse selection, a decrease that can take the form of a lower amount of loans to risky borrowers (Hertzberg et al., 2011). Information sharing also yields an increase in the effort of potential borrowers (Padilla and Pagano, 1997 and 2000), but also lower competition in the market (Bouckaert and Degryse, 2006). Karapetyan and Stasescu (2014) show that this lower competition can lead to higher information acquisition in the credit market. Recently, Kim and Wagman (2015) argue that consumer information exchange in financial markets can lead to lower prices for consumers and to higher screening of financial products applicants, which induces an increase in ex-ante social welfare. Shy and Stenbacka (2015) analyze how firms make higher profits in a situation of weak privacy protection where firms can easily share consumer information than in a situation of strong privacy protection. Our paper differs from theses studies in that we analyze information acquisition by financial institutions, but from the perspective

of the consumers, that is their willingness to let financial institutions access their online purchases history.

Finally, our paper is also linked to the literature about the choice of payment instruments, a choice that can be affected by anonymity. While Markose and Loke (2003) suggest that there is a perfect substitution between cash and card payments, Drehmann et al. (2002) argue that the fact that cash preserves anonymity makes card payments not a perfect substitute. Using survey data about the German payment behavior, von Kalckreuth et al. (2014) find that the anonymity permitted by a payment instrument explains its adoption. Anonymity is also a key feature of payment instruments such as Bitcoin. According to Kahn and Linares-Zegarra (2015), this desire for anonymity could in part be explained by the risk of identity theft. Athey et al. (2017) illustrates how privacy related to the government, an intermediary or the public affects payment instruments choices. We show in this paper that consumers may want to choose payment instruments more respectful of their financial privacy for other reasons than identity theft, mainly for keeping banks and relatives from having access to their transaction history.[4]

3.3 Survey and data description

We use a survey conducted in May 2015 by ACSEL/Caisse des Dépôts on a sample of 1,000 French Internet users aged 15 years and older. The survey was drawn from access panels (directories of people willing to participate in surveys on a regular basis). The sample is representative of the French Internet population (in terms of age, sex, socioeconomic classification, urban areas and Internet use). The survey has been conducted using online questionnaires.

The main objective of the survey is to measure the level of trust of Internet users in several online services (bank, administration, etc.). The survey is divided into several parts that deal with Internet access and use, e-commerce, payment instruments, online banking, online communication (chats, blogs, etc.), social networks, online administration, cloud services, Internet of things, security and authentication, personal data and privacy. We focus our empirical study on the questions related to e-commerce (frequency of purchase, average monthly spending, trust in online retailers, security and privacy policy, etc.) and payment instruments. We now describe these questions in more detail.

[4]The social consequences of surveillance by relatives and friends are studied by sociologists such as Castells (2001).

French e-commerce is one of the most developed in Europe. In 2014, 34.7 millions of online consumers (79 percent of French Internet users) spent 57 billion euro with 164,000 online retailers (FEVAD, 2015). In our survey, 81 per cent of the respondents (811 respondents) in 2015 report to have made at least one online purchase during the last 12 months. Figure 3.1 displays the distribution of the frequencies of online purchases. 49 per cent of respondents purchase more than once per month but less than once per week. Overall, 62 per cent of the online consumers claim to make more than one purchase per month.

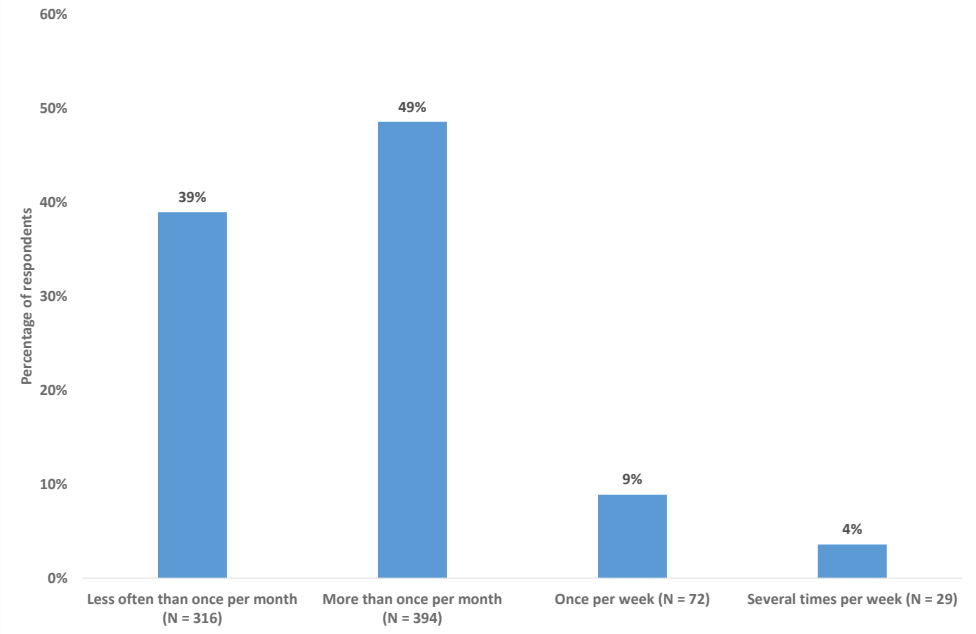


Figure 3.1 – Frequencies of purchases of online consumers

To make their online purchases, consumers can use several payment instruments such as payment cards, PayPal, checks, etc. (Figure 3.2) Payment cards provided by banks[5] and PayPal are by far the two most used payment instruments: 95 per cent declare using a payment card, and 36 per cent PayPal.[6] We also observe that 22 per cent of the respondents use other payment instruments. We know that 11 per cent of the respondents use gift cards that are redeemable only for purchases at retailers and that cannot be cashed out. The remaining 11 per cent of the respondents use other unknown payment instruments.

[5] 95 per cent of the payment cards issued in France are debit cards (see Bagnall et al., 2016).

[6] As from April 2015, 165 million of PayPal accounts are active with an average of 23 transactions by account in the first quarter of 2015. At the end of March 2015, 30 per cent of US online transactions were made using PayPal (retrieved from Paypal.com). No public data are provided for France.

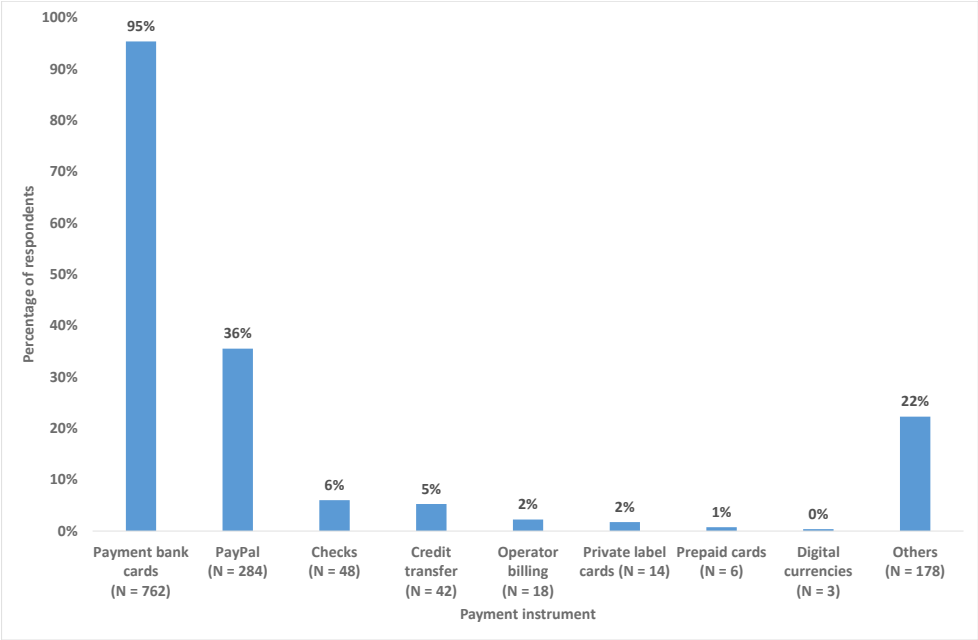


Figure 3.2 – Use of payment instruments

Among online consumers, 53 per cent only use one payment instrument, 29 per cent use two payment instruments, and 17 per cent use three or more payment instruments. Internet users who only use one payment instrument prefer payment cards provided by banks (94 per cent), while only two per cent of them prefer PayPal. People using two or more payment instruments prefer payment cards and PayPal (66 and 58 per cent, respectively). The use of the other payment instruments is limited. It is worth noting for our purpose that PayPal is not accepted by all retailers in France, which explains its limited rate of penetration. For instance, Amazon does not accept PayPal payments. However, it is crucial for the rest of the analysis to keep in mind that a retailer who accepts PayPal in France always accepts a payment card.[7] As a consequence, using PayPal is a real choice for a consumer as he could have used his bank card instead.

To summarize, consumers can use payment instruments provided by banks (payment cards, checks, and credit transfers) that are related to their bank account. They can also use payment instruments that are not attached to their checking account, i.e. payment instruments provided either by financial intermediaries such as PayPal or private companies such as operator billing, private label cards, prepaid cards, and digital currencies. We refer to these payment instruments

[7]PayPal transaction fees are on average 1.9 per cent higher than bank card fees.

as ”non-bank payment instruments”. In Figure 3.3, we note that 54 and 2 per cent of the respondents report that they exclusively use bank or non-bank payment instruments, and 24 per cent report that they use both bank and non-bank payment instruments. A significant proportion of French online consumers use therefore both bank and non-bank payment instruments.

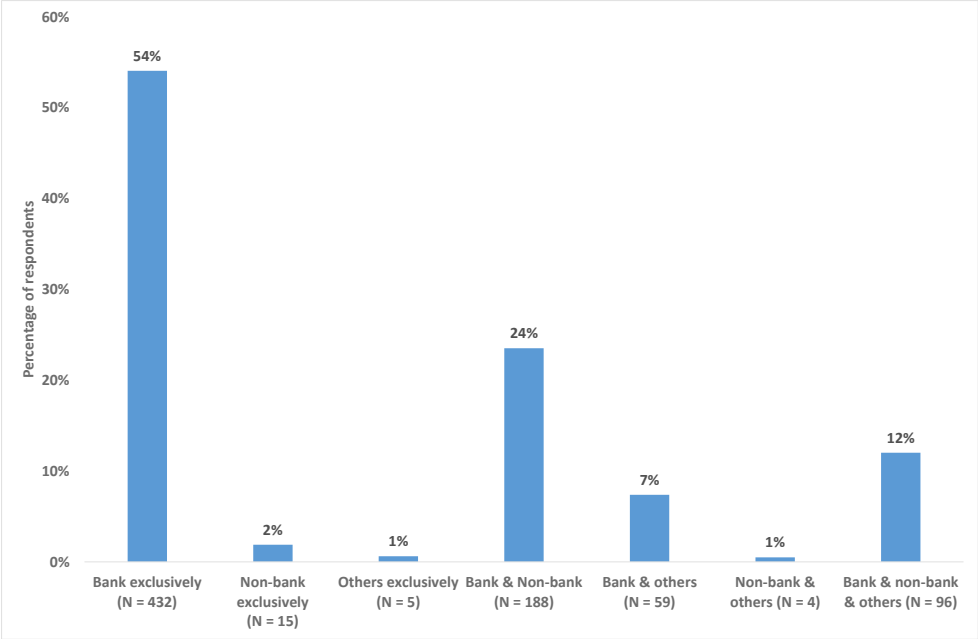


Figure 3.3 – Exclusive use of bank and non-bank payment instruments

These preliminary descriptive statistics raise the question of the motivation of consumers to use non-bank payment instruments for their online purchases. In this paper, we hypothesize that they use in part non-bank payment instruments to protect their financial privacy. Convenience, security and person-to-person transfers are further characteristics that could also explain the use of a non-bank payment instrument. We explain below how we control for these other dimensions in the econometric analysis.

Using a bank payment instrument leaves indeed a trace on the bank account such as the name of the retailer, the amount of the transaction, its date, etc. By contrast, using a non-bank payment instrument does not necessary leave a trace on the bank statement. For example, when using PayPal, Internet users can decide to credit their online PayPal account, which allows them to directly pay online. In that case, banks have absolutely no information on online purchases.[8]

[8]This situation can be compared to the use of cash at point-of-sale transactions. Cash is anonymous and cannot be tracked by banks.

Likewise, prepaid cards, virtual currencies and payment services provided by Internet providers or private companies (operator billing) cannot be tracked by banks.

Internet users are aware of the characteristics of the payment instruments provided by banks and non-banks. They may decide, for various reasons, to purchase online goods using non-bank payment services to protect personal information from further uses by banks. For example, Internet users may want to avoid to be profiled or targeted by financial commercial campaigns. Others may be simply unwilling to let relatives access financial information on a bank statement. Regardless of the reason, privacy-minded consumers may purchase online goods that they would not have purchased with a bank payment instrument if they have the option to use a non-bank payment instrument that protects their financial privacy.

3.4 Econometric model

In this section, we test whether the use of a non-bank payment instrument, when controlling for various other effects, positively influences the frequency of online purchases.

The dependent variable, "Frequency of online purchases", is a categorical variable with four categories: "less often than once per month", "more than once per month", "once per week", and "several times per week". Our main variable of interest w is *non-bank*, a binary variable indicating whether or not a consumer has used at least one non-bank payment instrument in online purchases (as described in Section 3). There is a potential endogeneity issue if the use of a non bank-payment instrument is correlated with unobservable variables that influence the frequency of online purchases y, such as cash used in online transactions or direct money transfers via online accounts for example. In this case, the coefficients estimated by Ordinary Least Squares (OLS) can be biased. We take into account the potential endogeneity by specifying a system of two equations, that explicitly models the use of a non-bank payment instrument for a consumer i:

$$w_i = \begin{cases} 0 & \text{if } w_i^* \leq 0, \\ 1 & \text{if } w_i^* > 0, \end{cases} \tag{3.1}$$

with:

$$w_i^* = X_i\beta + \varepsilon_i, \tag{3.2}$$

where w_i^* is a latent variable, β is of dimension k, and X_i is a set of control variables. The main equation of interest is the following:

$$y_i = w_i\delta + V_i\gamma + \eta_i, \tag{3.3}$$

where V_i is a set of exogenous explanatory variables that influences the frequency of online purchases, and δ is the parameter associated with the endogenous binary variable.

In Equations (3.2) and (3.3), we assume that $(\varepsilon_i, \eta_i)'$ is normally distributed with mean $(0,0)'$ and covariance Σ for $i = 1, \ldots, n$:

$$\Sigma = \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}. \tag{3.4}$$

The parameter ρ represents the correlation between the unobservable variables. If ρ is equal to 0, there is no endogeneity bias. For values of $\rho \neq 0$, estimating the effect of using a non-bank payment instrument on the frequency of online purchases could lead to biased results. Parameter σ^2 is the variance of η_i. As the variance in Equation (3.2) is not identified, we normalize it to 1. This is standard restriction in probit models (Wooldridge, 2006).

Endogeneity results from the correlation between the unobservable variables of the two equations. We deal with this issue in two ways. First, we use a two-step approach with instrumental variables. We estimate the first equation with a probit model, and we use the predicted value of this first step to estimate the effect of the use of a non-bank payment instrument on the frequency of online purchases. This approach solves the endogeneity problem provided that we have valid instruments, but does not allow us to estimate the correlation between coefficients. Second, we propose a Bayesian analysis that explicitly deals with the correlation between the unobserved variables of the two equations to account for the potential endogeneity of the choice of the type of payment instrument on the frequency of online purchases. We estimate the system of simultaneous Equations (3.2) and (3.3) using (3.4) by an MCMC method[9] that can deal with

[9]In principle, the parameters of the system of equations could also be estimated by Maximum Likelihood. However, Waelbroeck (2005) and others have shown that this method provides less reliable estimates of the elements of the covariance matrix of unobservable variables than the Bayesian method. See for a general introduction to numerical Bayesian tools Greenberg (2014). The identification of the endogenous effect in this model is guaranteed by the non-linearity of the probability to use a non-bank payment instrument. The estimation procedure can be carried on with the same set of explanatory variables in the two equations, contrary to instrumental variable method that requires some variables to be excluded in linear regression models.

the case $\rho \neq 0$.[10,11]

We specify Equation (3.3) as:

$$\begin{aligned} y_i = \alpha &+ \delta w_i \\ &+ \gamma_1(\text{Convenience and efficacy})_i \\ &+ \Sigma_{j=2}^{3}\gamma_j(\text{Risk})_i \\ &+ \Sigma_{j=4}^{7}\gamma_j(\text{Online activity})_i \\ &+ \Sigma_{j=8}^{14}\gamma_j(\text{Individual variables})_i \\ &+ \eta_i \end{aligned} \tag{3.5}$$

First, Internet users may prefer to use a non-bank payment instrument for other reasons than financial privacy such as ”Convenience and efficacy”. For example, some online retailers offer the possibility to store personal financial information on their websites to save time during the checkout process. This is especially important concerning PayPal, which allows consumers to pay online only by logging into their PayPal account. As PayPal is the most used non-bank payment instrument in the data set, we want to control for this effect in order to isolate a possible financial privacy effect. As a consequence we introduce a variable named "Financial info stored on e-commerce websites" that determines whether a consumer has stored financial information at online retailers.

Second, we include two variables related to the risks of electronic commerce ("Risk"): the perceived risk of banking details being hacked on electronic commerce websites and the perceived risk of banking details being consulted by third parties on online banking websites. This type of variables has previously been used by Akhter (2012) and Tsai et al. (2011). The authors find evidence of a negative impact of privacy concerns on online purchases. We expect these variables to be negatively correlated with the frequency of online purchases.

[10]Bayesian MCMC methods are based on a simulation of a Markov process in the parameters space (including the latent variables), which converges to a limiting distribution that it is exactly the posterior distribution of the parameters. Each parameter is simulated sequentially from its posterior conditional distribution with the Gibbs algorithm. When the exact posterior conditional distribution is not available, the Metropolis-Hastings algorithm simulates a draw by using a probabilistic acceptance method. The details of this procedure are available upon request and are also described in the context of an ordinal endogenous variable model by Bounie et al. (2016).

[11]The idea is to use ”data augmentation” to simulate observations for the unobservable endogenous latent variable w_i of Equation (3.2) and to draw simulations from the full posterior distribution of β, δ, γ, ρ, σ, and w_i.[12] Indeed, once we observe or simulate w_i in Equation (3.2), the system of Equations (3.2) and (3.3) is a simple Seemingly Unrelated Regression model (SUR).

Third, we control for the level of online activities ("Online activity") to capture online experience with a binary variable that indicates whether an individual connects to the Internet every day and a binary variable for the number of passwords used to secure online accounts (1 = more than 10, 0 = less than 10). We also use average online spending (less than euro 50, between euro 50 and euro 250) to determine the link between online spending and purchase frequency.

Finally, we control for various individual variables ("Individual variables"): whether the respondent has children or not, the socio-economic category (1=employed, worker or farmers, 0 = other categories), the age category (15-24, 25-34, 35-49, 50+) and the level of education (no diploma, high school, undergraduate and graduate studies).

We also specify Equation (3.2) as:

$$\begin{aligned} w_i^* = \alpha &+ \beta_1(\text{Convenience and efficacy})_i \\ &+ \Sigma_{j=2}^{3}\beta_j(\text{Risk})_i \\ &+ \Sigma_{j=4}^{7}\beta_j(\text{Online activity})_i \\ &+ \Sigma_{j=8}^{14}\beta_j(\text{Individual variables})_i \\ &+ \Sigma_{j=15}^{17}\beta_j(\text{Instruments})_i \\ &+ \varepsilon_i \end{aligned} \tag{3.6}$$

In addition to control variables in Equation (3.5), we use three instruments that allow us to have a consistent estimation despite the risk of endogeneity between the use of a non-bank payment instrument and the frequency of online purchases. In order to have robust results, our instruments must be correlated with our endogenous variable (the use of a non-bank payment instrument) and exogenous with our main dependent variable (the frequency of online purchases).

First, people may use several payment instruments when shopping online as they purchase to various online retailers. A consumer might then not have the choice but to use a non-bank payment instrument at some point. We control for this effect with a variable that indicates the number of payment instruments used in online purchases.

Second, we use a binary variable that indicates whether an Internet user is ready to share personal information with the bank : "Willingness to disclose personal information to the bank". This variable is used to highlight a potential financial privacy effect with respect to the use of

a non-bank payment instrument. We expect to find that people who are not willing to share personal information with their bank are more likely to choose a non-bank payment instrument during their online purchases.

Third, we include a variable related to the empowerment of Internet users: the use of ad-blockers and/or privacy enhancing web browser extensions ('Ghostery' or 'HTTPS everywhere' for example). We expect this variable to be positively correlated with the use of a non-bank payment instrument, as people who are more aware of the potential negative incidences of privacy issues (such as tracking, profiling, solicitation, etc.), may start adopting privacy enhancing tools. The use of a non-bank payment instrument is an example of such a tool, so again we expect a positive correlation between these two variables.

3.5 Estimation results

Estimation results are reported in Table 3.I.[13] Parameters of Equations (3.2) and (3.3) are estimated using two econometric methods: a two-step regression model in columns (1) and (2),[14] and by Bayesian MCMC methods in columns (3) and (4).

As indicated in Table 3.I, we find a statistical significant positive effect at the 1 per cent level of the use of a non-bank payment instrument on online purchases (0.190 for the two-step regression and 0.223 for the Bayesian MCMC). These estimation results confirm that the use of a non-bank payment instrument positively affects the frequency of online purchases. The mean of the dependent variable is 1.77 (see Appendix A2) which corresponds to a frequency between "less often than once per month" and "more than once per month". As an illustration, if we assume that 1.77 corresponds to one purchase every two months, and the value 2 is at least once per month, the use of a non-bank payment instrument doubles the frequency of online purchases, which is economically significant.

The use of instrumental variables appears to be crucial in our estimations, all three of them being significant in predicting the use of a non-bank payment instrument. We find that Internet users who are willing to disclose personal information about their tastes or preferences to their bank are less likely to use a non-bank payment instrument, which supports our hypothesis that

[13] For reason of convenience, we did not report the control variables related to "Online activity" and "Individual variables"; the results are available upon request.

[14] We also estimated equation (3.2) with a linear probability regression using OLS. Results are similar and available upon request.

financial privacy plays a role in the choice of payment instruments. Similarly, our results show that people who delete their cookies regularly and/or use privacy enhancing browser extensions tend to have more usage of a non-bank payment instrument, which subsequently turns into more purchases. As commented in Section 3.2, the existing literature suggests that people who are more concerned with their online privacy are likely to purchase less. Our result suggests that the relation is more subtle: as people become more familiar with privacy issues and adopt privacy enhancing tools, they start to use non-bank payment instruments which allows them to trust e-commerce more and also to purchase more. Finally, we unsurprisingly find that the number of payment instruments is importantly correlated with the use of a non-bank payment instrument.

Using the Bayesian MCMC, we find that the parameter ρ is not significantly different from 0. This illustrates the validity of our instrumental variables. Note that when ρ is equal to 0, Equations (3.2) and (3.3) can be estimated separately. The two methods provide robust estimates of the effect of using a non-bank payment instrument on the frequency of online purchases.

It is important to outline that we control for various other factors that might increase the use of a non-bank payment instrument. Amongst them, the convenience of using online payment instruments (measured by the variable that determines whether a consumer has stored financial information online) is strongly positive, which shows that it was therefore important to capture this (efficacy) dimension of online payment instruments that is not related to financial privacy.

As expected, variables related to risks associated to unauthorized uses of personal data (the perceived risk of banking details being hacked on e-commerce or being consulted by third parties on online banking websites) have a negative impact, which confirms the results found in the existing literature (Acquisti et al., 2015). Comparing this to the fact that privacy-related variables have an indirect positive effect on the frequency of online purchases seems to indicate that there are probably two types of Internet users. On the one hand, there are people who are afraid of making online transactions and who reduce in turn their online purchases. On the other hand, there are Internet users who adopt protection technologies and who trust e-commerce more, resulting in an increase of their online purchases. These two effects should be investigated further.

Concerning online activity, we find that Internet users who spend less than 50 euros per month on average purchase significantly less. We also show that people who have many passwords (more than ten) have a higher purchase frequency. Our individual variables highlight that

Internet users aged from 15 to 34 years old have a higher use of electronic commerce, as well as people with children.

Table 3.I – Estimation results

	Two-step regression		MCMC	
	Step 1: Use of a non-bank payment instrument	Step 2: Frequency of online purchases	Step 1: Use of a non-bank payment instrument	Step 2: Frequency of online purchases
Payment instruments				
Non-bank		0.190***		0.223***
		(0.066)		(0.061)
Convenience and efficacy				
Financial information stored on e-commerce websites	0.068	0.175***	0.059	0.203***
	(0.020)	(0.059)	(0.177)	(0.055)
Risk				
Perceived risk of banking details being hacked on e-commerce websites	0.355	−0.316***	0.385	−0.253***
	(0.310)	(0.090)	(0.272)	(0.083)
Perceived risk of banking details being consulted by third parties on online banking websites	-0.225	−0.171***	-0.100	−0.145***
	(0.195)	(0.057)	(0.101)	(0.057)
Online activity	Yes	Yes	Yes	Yes
Individual variables	Yes	Yes	Yes	Yes
Instruments				
Willingness to disclose personal information to the bank	−0.440*		−0.541*	
	(0.242)		(0.292)	
Use of privacy web browser extensions and deleting cookies	0.427*		0.406*	
	(0.258)		(0.235)	
Number of payment instruments	3.065***		3.03***	
	(0.207)		(0.176)	
Constant	−5.832***	2.193***	−4.931***	1.934***
	(0.821)	(0.211)	(0.491)	(0.153)
σ		0.679***		0.679***
				(0.018)
ρ				0.064
				(0.088)
Observations	715	715	715	715
Pseudo R^2 - R^2	0.72	0.24		
Prob > $\chi 2$	0.000	0.000		

Note: *p<0.1; **p<0.05; ***p<0.01

3.6 Conclusion

This paper has investigated whether the use of non-bank payment instruments by consumers has a positive impact on their online purchases. Using bank payment instruments such as debit or credit cards to purchase privacy-sensitive goods (medication, healthcare items, gifts, gambling, adult products, etc.), may indeed induce a cost for consumers as banks can clearly track their purchases. In the absence of alternative non-bank payment instruments, a part of consumers may simply decide not to shop online so as not to disclose personal financial information. However, when consumers can use non-bank payment instruments, they can now buy online without fear of sharing information with banks or relatives. The use of non-bank payment instruments can therefore lead privacy-minded consumers to increase their online purchases.

Our empirical results confirm this intuition but also show that the relationship between privacy and online purchases is more subtle. On the one hand, as the existing literature has shown, general privacy concerns tend to decrease online spending. On the other hand, some Internet users do not necessarily want to be identified by the bank during all their online activities. They are ready to take actions to hide some of their purchases from the bank or relatives. Using what we have called non-bank payment instruments such as PayPal and virtual currencies allows these consumers to purchase privacy-sensitive goods and to increase their online purchases. These findings have two main implications.

First, there is a recent trend by large financial institutions and 'Fintech startups' to propose more personalized payment solutions such as personalized coupons. We believe that financial intermediaries should think about enlarging their portfolio of online payment instruments to account for different levels of privacy: secure identification for the payment of local or federal taxes, anonymous transactions for gifts and for the purchase of privacy-sensitive goods. This approach could be paralleled to the European EIDAS regulation on electronic transactions that acknowledges the concept of federation of identities and leaves the door open for pseudonymized transactions.

Second, payment instruments and bank accounts provide useful and ongoing information on consumers' financial statements (overdrafts, revenues, etc.). By analyzing these traces, banks can evaluate the creditworthiness of consumers and the potential risks of lending money. They can also monitor borrowers to mitigate losses due to bad debt. Following Mester et al. (2007), providing deposit-taking and lending jointly is a capital advantage for banks in the competition

with non-bank institutions that do not manage checking accounts. However, the development of non-bank institutions in the domain of payments could seriously affect this competitive advantage. If a significant fraction of consumer payments are carried out with non-bank institutions, banks will have in return less transactions to manage, and hence less information to make lending decisions. In the end, the profitability of banks could be affected.

BIBLIOGRAPHY

[1] **Acquisti, Alessandro, Curtis R. Taylor, and Liad Wagman**, "The Economics of Privacy," *Journal of Economic Literature*, July 2015.

[2] **Akhter, Syed H.**, "Who spends more online? The influence of time, usage variety, and privacy concern on online spending," *Journal of Retailing and Consumer Services*, October 2012, *19*, 109–115.

[3] **Athey, Susan, Christian Catalini, and Catherine Tucker**, "The Digital Privacy Paradox: Small Money, Small Costs, Small Talk," 2017.

[4] **Bagnall, John, David Bounie, Kim Huynh, Anneke Kosse, Scott Schuh, Tobias Schmidt, and Helmut Stix**, "Consumer Cash Usage and Management: A Cross-Country Comparison with Diary Survey Data," *International Journal of Central Banking*, December 2016.

[5] **Bolt, Wilko and Maarten van Oordt**, "On the value of virtual currencies," Technical Report, Bank of Canada Staff Working Paper, 2016-42 2016.

[6] **Bouckaert, Jan and Hans Degryse**, "Entry and Strategic Information Display in Credit Markets," *Economic Journal*, 2006, *116*, 702–720.

[7] **Bounie, David, Abel François, and Patrick Waelbroeck**, "Debit Card and Demand for Cash," *Journal of Banking & Finance*, 2016, *73*, 55–66.

[8] **Castells, Manuel**, *The Internet Galaxy: Reflections on the Internet, Business and Society*, Oxford University Press, 2001.

[9] **Ching, Andrew T. and Fumiko Hayashi**, "Payment card rewards programs and consumer payment choice," *Journal of Banking & Finance*, 2010, *34* (8), 1773 – 1787. New Contributions to Retail Payments: Conference at Norges Bank (Central Bank of Norway) 14 and 15 November 2008.

[10] **Conitzer, Vincent, Curtis R. Taylor, and Liad Wagman**, "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases," *Marketing Science*, March 2012, *31* (2), 277–292.

[11] **Cranor, Lorrie F., Kelly Idouchi, Pedro G. Leon, Manya Sleeper, and Blase Ur**, "Are They Actually Any Different? Comparing Thousands of Financial Institutions Privacy Practices," *The Twelfth Workshop on the Economics of Information Security (WEIS 2013)*, June 2013.

[12] **Drehmann, Mathias, Charles Goodhart, and Malte Krueger**, "The challenges facing currency usage: will the traditional transaction medium be able to resist competition from new technologies?," *Economic Policy*, 2002, *17*, 193–228.

[13] **Greenberg, E.**, *Introduction to Bayesian Econometrics*, Cambridge University Press, 2014.

[14] **Hertzberg, Andrew, Jose M. Liberti, and Daniel Paravisini**, "Public Information and Coordination: Evidence from a Credit Registry Expansion," *Journal of Finance*, 2011, *66*, 379–412.

[15] **Janger, Edward J. and Paul M. Schwartz**, "The Gramm-Leach-Bliley Act, Information Privacy, and the Limits of Default Rules," *Minnesota Law Review*, 2002, *86*, 1219.

[16] **Jentzsch, Nicola**, *Financial Privacy - An International Comparison of Credit Reporting Systems*, Springer, 2007.

[17] **Kahn, Charles M. and Jose M. Linares-Zegarra**, "Identity Theft and Consumer Payment Choice: Does Security Really Matter?," *Journal of Financial Services Research*, 2015, pp. 1–39.

[18] **Karapetyan, Artashes and Bogdan Stasescu**, "Information Sharing and Information Acquisition in Credit Markets," *Review of Finance*, 2014, *18* (4), 1583–1615.

[19] **Kim, Jin-Huyk and Liad Wagman**, "Screening incentives and privacy protection in financial markets: a theoretical and empirical analysis," *RAND Journal of Economics*, Spring 2015, *46*, 1–22.

[20] **Lacker, Jeffrey M.**, "The Economics of Financial Privacy: To Opt Out or Opt In?," *Federal Reserve Bank of Richmond Economic Quaterly*, Summer 2002, *88* (3), 1–16.

[21] **Markose, Sheri M. and Ying J. Loke**, "Network effects on cash card substitution in transactions and low interest rate regimes," *The Economic Journal*, 2003, *113*, 456–476.

[22] **Mester, Loretta J., Leonard I. Nakamura, and Micheline Renault**, "Transactions Accounts and Loan Monitoring," *The Review of Financial Studies*, 2007, *20* (3), 529–556.

[23] **Padilla, Jorge A. and Marco Pagano**, "Endogenous Communication among Lenders and Entrepreneurial Incentives," *Review of Financial Studies*, 1997, *10*, 205–236.

[24] **Pagano, Marco and Tullio Japelli**, "Information Sharing in Credit Markets," *The Journal of Finance*, December 1993, *48* (5), 1693–1718.

[25] **Schuh, Scott and Oz Shy**, "U.S. Consumers' Adoption and Use of Bitcoin and other Virtual Currencies," April 2016. Paper presented at the DeNederlandsche bank, Conference entitled "Retail payments: mapping out the road ahead".

[26] **Sheng, Xinguang. and Lorrie F. Cranor**, "An Evaluation of the Effect of US Financial Privacy Legislation Through the Analysis of Privacy Policies," *I/S: A Journal of Law and Policy for the Information Society*, Fall 2006, *2* (3), 934–979.

[27] **Shy, Oz and Rune Stenbacka**, "Customer Privacy and Competition," 2015.

[28] **Swire, Peter P.**, "The Surprising Virtues of the New Financial Privacy Law," *Minnesota Law Review*, 2002, *86*, 1263.

[29] **Tsai, Janice Y., , Serge Egelman, Lorrie Cranor, and Alessandro Acquisti**, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research*, June 2011, *22*, 254–268.

[30] **Villas-Boas, J.Miguel**, "Price Cycles in Markets with Customer Recognition," *The RAND Journal of Economics*, 2004, *35* (3), 486–501.

[31] **von Kalckreuth, Ulf, Tobias Schmidt, and Helmut Stix**, "Choosing and using payment instruments: evidence from German microdata," *Empirical economics*, 2014, *46*, 1019–1055.

[32] **Waelbroeck, Patrick**, "Computational Issues in the Sequential Probit Model," *Computational Economics*, 2005, *26* (2), 141–161.

[33] **Wooldridge, Jeffrey M.**, *Introductory Econometrics: A Modern Approach*, Michigan State University, 2006.

3.7 Appendix: descriptive statistics

Table 3.II – Statistics - Binary variables

	N	Mean	St. Dev.	Min	Max
Payment instruments					
Use of non-bank payment instruments	811	0.453	0.017	0	1
Convenience and efficacy					
Financial information stored on electronic commerce websites	811	0.449	0.498	0	1
Risk					
Perceived risk of banking details being hacked on electronic commerce websites	1,000	0.860	0.347	0	1
Perceived risk of banking details being consulted by third parties on online banking websites	1,000	0.458	0.498	0	1
Online activity					
Using Internet several times every day	1,000	0.641	0.480	0	1
More than 10 passwords	1,000	0.170	0.376	0	1
Average monthly spending in electronic commerce inferior to euro 50	723	0.408	0.492	0	1
Average monthly spending in electronic commerce between euro 50 and euro 250	723	0.519	0.500	0	1
Individual variables					
Having children	1,000	0.262	0.440	0	1
Lower socioeconomic classification	1,000	0.294	0.456	0	1
Inactives	1,000	0.40	0.49	0	1
Being between 15 and 24 years old	1,000	0.182	0.386	0	1
Being between 25 and 34 years old	1,000	0.159	0.366	0	1
Being more than 50 years old	1,000	0.397	0.490	0	1
Instruments					
Willingness to disclose personal information to the bank	1,000	0.042	0.006	0	1
Use of privacy web browser extensions and deleting cookies	1,000	0.713	0.014	0	1

Table 3.III – Statistics - Frequency of purchase

Less often than once per month = 1	More than once once per month = 2	Once per week =3	Several times per week = 4	N	Mean	St. Dev.
316	394	72	29	811	1.77	0.75

Table 3.IV – Statistics - Number of payment instruments

One = 1	Two = 2	Three or more = 3	N	Mean	St. Dev.
316	394	72	811	1.77	0.75

Table 3.V – Statistics - Education

No diploma = 1	High school diploma = 2	Undergraduate and graduate studies = 3	N	Mean	St. Dev.
231	253	516	1000	2.285	0.82

CHAPTER 4

FINANCIAL PRIVACY, SCREENING AND PAYMENT INSTRUMENTS

4.1 Introduction

As data becomes increasingly valuable in economic activities, it is worthwhile to remember that banks have a vast amount of data at their disposal. This data stems from every financial transaction: purchases, payments, transfers and so on. For years, banks haven been using this type of information to make better decisions. For example, analyzing transactional data can facilitate credit card fraud detection or marketing messages customization. Banks also look at data as a way to significantly reduce credit risk, as data analytics could help to have a better assessment of solvency.

Nowadays, it is becoming more difficult for banks to collect data, as an increasing number of non-bank actors are entering the payment instruments market, like telecommunication operators (Orange or Vodafone for example) or tech companies for instance (Apple, PayPal and so on). The fact that the competition for the access to consumer data is increasing can explain in part why banks seem to push for the disappearance of cash: it does not generate any useful information for them.[1] Mobile payment is an example of service designed and offered by banks to replace cash.[2]

Cash does not however seem to be an endangered species: it still was the most frequently used payment instrument in 2015 in the US according to a note from the Federal Reserve.[3] In India, Prime Minister Narendra Modi's decision to demonetize 500 and 1000 rupee notes in November 2016 led to a massive cash shortage which triggered chaotic scenes across the country. [4] As the banking sector has a growing interest in collecting and using financial data, cash appears to be a way for consumers to maintain financial privacy.[5] Some consumers indeed

[1] Other reasons include the cost of handling cash or dealing with money laundering and other illegal activities

[2] As an example, several French banks such as BNP Paribas, Crédit Agricole or Société Générale have conjointly developed a mobile payment instrument called Paylib.

[3] The note "The State of Cash - Preliminary Findings from the 2015 Diary of Consumer Payment Choice" by Wendy Matheny, Shaun O'Brien, and Claire Wang can be retrieved here.

[4] 50 days of pain: What happened when India trashed its cash by Rishi Iyengar (CNN).

[5] For example, 93% of respondents to the report Americans' Attitudes About Privacy, Security and Surveillance by Mary Madden and Lee Rainie (PewResearch Center) declare that controlling what information is available to whom is important.

do value privacy in the case of financial data. For example, 55% of consumers surveyed in 2016 by KPMG declared having declined to buy something online because of privacy concerns[6]. Similarly, a survey conducted among 2000 representative French Internet users finds that 35% of them have declined to a purchase or changed payment instrument for financial privacy reasons at least once.[7] In this context, mass financial data collection and analysis by banks may lead to a growing adoption of non-bank payment instruments, which would ultimately represent a loss of data for banks.

The question we want to answer to is how the profitability of banks as financial intermediaries is challenged by the concerns of some of their customers for financial privacy. In the market for the supply of payment instruments, financial privacy concerns affect the payment instrument choice of consumers: they may turn to a non-bank payment instrument that allows them to preserve financial privacy. Concerning the lending decision process, the bank relies on payment data to have efficient screening. Financial privacy may decrease the amount of payment data available to the bank, leading to inefficient lending decisions.

We design a model where a bank uses probability to estimate the solvency of borrowing consumers according to a screening technology that is based on payment data analytics. The more a consumer uses the payment instrument proposed by the bank, the more the screening process of the bank is precise. Few papers link the supply of payment instruments by banks and lending strategy.

Consumers choose their payment instruments: for a given level of consumption, they decide the proportion that is going to be paid with the bank's payment instrument. The rest of that consumption is paid using an alternative instrument that does not generate data collected by the bank (cash, non-bank payment instruments like PayPal or Bitcoin and so on). The choice of payment instruments may be influenced by financial privacy concerns, the cost of the payment instruments and the desire of consumers of to be granted credit. By correctly determining the cost of using its payment instrument, the bank make efficient lending decisions. Efficient lending decisions means granting credit to all solvent borrowers and none to non-solvent ones. In order to do so, the bank must however take the presence of alternative payment instruments into account.

[6]Crossing the line - Staying on the right side of consumer privacy report.

[7]haire Valeurs et Politiques des Information Personnelles - Données personnelles et confiance : quelles stratégies pour les citoyens-consommateurs en 2017 ?

Financial privacy concerns may however increase competition in the supply of payment instruments, making the bank obligated to consider this factor. The fact that some borrowers may choose not to use the bank payment instrument because of these concerns can damage the efficiency of lending by the bank, as it has less data to improve its screening process.

Thus, we show how it is imperative for the bank to remember that its customers can have financial privacy concerns, as it could affect the efficiency of its loan attribution. A way to deal with this issue is to lower the usage fee of its payment instrument, in order to better compete with the alternatives. In other words, banks should not act as if they face no competition regarding the supply of payment instruments. Another threat is a loss of market power in lending, especially when there is an increasing number of possibilities to get credit without the banks (peer-to-peer lending, digital currencies, PayPal Working Capital and so on). It could then be better for banks to slow down on data analytics. Proposing a payment instrument weakly intrusive with respect to financial data may be a way to keep financial privacy minded customers by their side in such a context.

Taking financial privacy concerns into consideration when analyzing lending decisions and the market for payment instruments could also be crucial for regulators.

We now illustrate the importance of payment data analysis for financial intermediaries with three examples.

4.1.1 Example 1: American Express

On January 31^{th} 2009, the New York Times published an article on how American Express was looking at its customers' spending habits in order to improve the precision of scoring.[8] American Express for example explained the decision to cut some clients' credit lines by stating: "other customers who have used their card at establishments where you recently shopped, have a poor repayment history with American Express." By association, several solvent American Express customers were then deemed not worthy of credit. Even upon request, American Express did not specify which retailers were considered to assess whether a customer is creditworthy or not.

[8] R.Lieber American Express Kept a (Very) Watchful Eye on Charges - New York Times

4.1.2 Example 2: OpenBanking

During the end of the year 2017, several UK banks informed their customers that they could accept the sharing of their financial data with third parties, based on a program called "Open Banking".[9] Created by the UK's Competition and Markets Authority (CMA), the goal of the "Open Banking Implementation Entity"[10] is to work with UK banks, fintechs, other third parties and consumers' associations in order to develop Application Programme Interfaces (APIs). These APIs are designed to help the rise of new services, facilitate credit applications and so on. However, according to an Accenture survey[11], two-thirds of UK consumers do not want to share financial data with third parties. Moreover, the head of the UK's Financial Inclusion Centre Mick McAteer stated that "Open Banking" could worsen the situation of low incomes individuals, leading to more financial exclusion.

4.1.3 Example 3: FICO Score XD and FinTechs

The US data analytics company FICO provides most American banks with its FICO score, used to assess the creditworthiness of consumers. Recently, FICO has developed a solution called "FICO Score XD", designed to improve financial inclusion.[12] This scoring adds landline, cable and mobile payments to data traditionally used in credit scoring (public records, data from credit bureaus, etc.). This solution is aimed at developing countries, where many consumers are not banked, but have an important use of mobile payments. Similarly, many FinTechs are starting to base scoring on mobile data, creating what can be called 'mobile scoring"'.[13].

This paper proceeds as follows: section 2 reviews the related literature, section 3 presents a model where the bank uses payment data to improve its lending decisions through borrowers' screening, section 4 explores the effect of borrowers' financial privacy concerns on the relationship between payment instruments, credit and screening. Finally, section 5 concludes

[9] K.Peachey Why banks will share your financial secrets - BBC

[10] Open Banking

[11] I.Withers Two thirds of consumers don't want to share financial data in blow to 'open banking' revolution - The Telegraph

[12] FICO Score XD

[13] See for example companies like Tala, Finca or BigDataScoring.

4.2 Related literature

This paper contributes to the literature on banks' screening practices by illustrating how financial privacy concerns (privacy over financial data) can ultimately influence the efficiency of lending. More specifically, it analyzes the link between credit screening and the choice of payment instruments as well as how the latter is influenced by financial privacy concerns. We discuss papers on payment instruments, lending decisions and privacy concerns in this section.

Following the seminal paper by Broecker (1990), many articles have tackled the issue of screening by banks (see Mukminov (2015) for an extensive survey on the disclosure). Broecker (1990) designs a set-up where screening practices use imperfect independent information to perform costless creditworthiness tests. We also consider costless screening with imperfect information, but in a case where the information results from the use of payment instruments. In our model, screening can help to avoid what Kanniainen and Stenbacka (1998) call type I and type II errors: type I is not attributing a loan to a low-risk borrower, type II is granting a credit to a high-risk borrower. In a model of spatial bank competition, Hauswald and Marquez (2006) show that the bigger the distance between the bank and the borrower, the less efficient becomes the process of granting credit. Here, we consider the distance as being the use of the bank card: the more it is used by borrowers, the "closer" they are to the bank.

A useful source of information for banks are bank-customer relationships. They can be defined in a broader sense than just lending relationships: savings, checking accounts, personal relationship with bank employees and so on. We consider bank-customer relationship in our paper as the use of a payment instrument proposed by the bank, which is then able to gather data on customers' consumption practices. In that regard, we should mention the paper by Mester et al. (2007) which shows how analyzing transactions accounts to get data on borrowers' activities makes monitoring practices easier for banks. Some papers have argued that the same case can be made for screening practices. For instance, Vissing-Jorgensen (2011) shows using panel data that some consumption patterns may indicate high credit risk.Along these lines, Hibbeln et al. (2015) argue that consistent analysis of checking accounts can reduce credit risk, especially if this data is combined with data on credit card accounts. Schoar (2012) shows that increasing personal contact with borrowers helps to diminish moral hazard, and facilitates repayment. Finally, Puri et al. (2013) illustrate how access to credit is improved by valuable information provided by bank-customer relationships.

Our paper is also linked to the literature about the choice of payment instruments. This choice can be affected by different factors: consumer characteristics such as age, education (see Stavins (2001) on the subject) or habits (see van der Cruijsen et al. (2015) for instance), payment instruments characteristics such as convenience or cost (see Schuh and Stavins (2010) for example), or incentives designed by merchants (see Arango et al. (2015) on the matter). We consider in our model that this choice is influenced by financial privacy concerns, for example between card and an alternative such as cash payments. Drehmann et al. (2002) show that card payments are not a perfect substitute for cash, as the latter preserves anonymity. Along this line, von Kalckreuth et al. (2014) find using German payment behavior survey data that a payment instrument's anonymity is a positive determinant of its adoption. Athey et al. (2017) illustrates how privacy related to the government, an intermediary or the public affects payment instruments choices. We build on this literature in two ways. First, we consider how the efficiency of bank's credit granting strategy is impacted by choices of payment instruments. Second, we discuss how these choices are influenced by financial privacy concerns.

In this paper, we begin by considering a model where the bank uses payment data analytics to design its credit granting policy through the screening of borrowers. In a second case, we introduce financial privacy concerns for a fraction of borrowers, which makes makes it harder for the bank to assess credit risk.[14].

4.3 Model set-up

We design a model inspired by Hauswald and Marquez (2003), which studies how progress in information technology affects the competition between financial intermediaries. We use the model in the paper as a basis because it gives us a good way to explore the links between information processing and lending decisions. Hauswald and Marquez (2003) do not however describe the nature of the information used by banks. In our model, we characterize this information as being payment instrument data. Analyzing such data can provide crucial information to the bank on credit risk. Another key element in our model differs from Hauswald and Marquez (2003): as the information available to the bank is provided by borrowers' consumption, it allows borrowers to control what the bank knows about them. We do not consider screening as only being a technological investment. Finally, as our focus in this paper is on the relationship

[14]We also design in the appendix a model without financial privacy concerns and where the bank is not able to use payment data to infer the type of borrowers.

between a bank and potential borrowers, we do not consider competition on the credit market.

Our model describes the relationship between a bank and borrowers.

Borrowers simultaneously apply for a loan at a bank and consume a basket of goods. They pay for their consumption using either a payment instrument delivered by the bank or alternatives (cash, a card from another bank, etc.). When borrowers use the payment instrument of the bank, they improve its screening ability. While solvent borrowers may have an incentive to use this instrument to signal themselves as creditworthy consumers, others might use alternative payment instruments to trick the bank into granting them credit. Indeed, without enough payment data, the bank may make some mistakes during its screening process.

The bank has a monopoly on credit, it faces competition for its payment instrument offer.

4.3.1 Borrowers

Each borrower is characterized by its type $\theta \in \{l,h\}$, each type corresponding to a solvency probability p_θ with $p_h > p_l$.

There is a proportion q of type h consumers (conversely a proportion $(1-q)$ of type l consumers), q being known both by the bank and all borrowers.

Borrowers have a consumption of size 1 which brings them a utility of size v. They choose the proportion $\alpha \in [0,1]$ of their consumption that they are going to pay for using the payment instrument offered by the bank. The bank determines the cost f_b of this payment instrument.

On top of this fee f_b, a proportion ω of borrowers is subjected to an additional privacy cost c when using the payment instrument of the bank. Financial privacy concerns are independent from the likelihood to reimburse.

This financial privacy cost c is exogenous, and the total cost of using the bank's payment instrument is equal to $f_b + c$ for borrowers with financial privacy concerns.

Proportions q and ω are known by all the players in the game. The bank is however not able *ex-ante* to distinguish between the four types of borrowers: low-risk financial privacy minded borrowers (proportion ωq of the total population), high-risk financial privacy minded borrowers (proportion $\omega_(1-q)$), low-risk borrowers that do not have financial privacy concerns (proportion $(1-\omega)q$) and high-risk borrowers that do not have financial privacy concerns (proportion $(1-\omega)(1-q)$).

The remaining proportion $(1-\alpha)$ of the consumption is paid using alternative payment instruments (cash, non-bank payment instruments, etc.). We suppose that the cost of using these alternative payment instruments is equal to s. We also suppose that this cost s is exogenous. When all payment instruments have the same cost (i.e. $f_b = s$), borrowers use the payment instrument of the bank.[15]

We denote α_θ^P the payment instrument choice of borrowers of type θ who are sensitive to financial privacy, and α_θ^{NP} the choice of borrowers not sensitive to it.

4.3.2 The bank

Simultaneously to their consumption, borrowers apply for a credit of amount normalized to 1. We normalize the amount of the credit because we want borrowers as similar as possible, solvency and financial privacy concerns excepted. The bank decides to grant a borrower a credit or not according to a signal $\varepsilon \in \{l,h\}$. This signal results from the screening of each borrower. This screening process is costless for the bank.

This credit of value 1 generates a cash flow of an amount with value I. The credit is reimbursed with probability p_θ for a θ borrower. In that case, a fraction δ of I is kept by the borrowers, while the bank gets $(1-\delta)I$, which covers the amount of the credit and the associated interests.

ASSUMPTION 1

The outside option when there is default is equal to 0 for both borrowers and the bank.[16]

The expected profit derived from the bank by granting credit to a type θ borrower can be written as follows:

$$p_\theta(1-\delta)I + (1-p_\theta)*0 - 1$$

$$\Leftrightarrow p_\theta(1-\delta)I - 1$$

The expected gain of providing a credit to a type h borrower is positive, while it is negative for a type l borrower:

[15] This allows the model to be simpler without affecting the results.

[16] This hypothesis allows the model to be simpler without affecting the results.An illustration of it would be the following: in case of default, the bank manages to extract all the cash flow I but looses the same amount in recovering and insurance cost. In the end, an absence of reimbursement yields no loss for the bank, but neither any profit. The same goes for borrowers.

$$p_l(1-\delta)I < 1 < p_h(1-\delta)I$$

This represents an incentive for the bank to implement screening of potential borrowers. The bank only grants credit when it receives a signal $\varepsilon = h$.

It is *ex-ante* efficient to authorize lending,even without perfect screening.[17] This is expressed by:

$$(qp_h + (1-q)p_l)(1-\delta)I > 1$$

Only borrowers are aware of their type, and the bank has beliefs about it according to the signal ε. We denote the precision of this signal ϕ. ϕ is the probability that the signal is correct (conversely, $(1-\phi)$ is the probability that the signal is incorrect).as

There is a direct link between payment data analytics and screening precision. In other words, the quality ϕ of this signal for each borrower depends on the proportion α of consumption that this borrowers has paid with the bank's payment instruments, with $\phi = \phi(\alpha)$. Analyzing payment data brings additional information to the bank in its screening process: $\phi'(\alpha) > 0$. When a consumer exclusively uses the payment instrument of the bank, then the bank knows its type with certainty: $\phi(1) = 1$. Having $\phi'(\alpha) > 0$ ensures that adding payment data into the screening technology is always beneficial. When a borrower does not use the payment instrument of the bank, the precision is equal to $\frac{1}{2}$: $\phi(0) = \frac{1}{2}$.[18]

ASSUMPTION 2

We assume that $\phi''_\theta(\alpha) = 0$.

The amelioration of the screening technology is linear in α.

[17] This allows to have lending when the bank has no information.

[18] The value of $\phi(0)$ is not important, it only provides a basic precision of the bank.

The bank receives each signal with probability:

$$\mathbb{P}(\varepsilon = h) = q[\phi(\alpha_h^P)\omega + (1-\omega)\phi(\alpha_h^{NP})] + (1-q)[(1-\phi(\alpha_l^P))\omega + (1-\omega)(1-\phi(\alpha_l^{NP}))]$$

$$\mathbb{P}(\varepsilon = l) = q[(1-\phi(\alpha_h^P))\omega + (1-\omega)(1-\phi(\alpha_h^{NP}))] + (1-q)[\phi(\alpha_l^P)\omega + (1-\omega)\phi(\alpha_l^{NP})]$$

Upon receiving the signal ε, the bank then knows that this signal is correct with probability:

$$\mathbb{P}(\theta = h|\varepsilon = h) = \frac{\mathbb{P}(\varepsilon = h|\theta = h)\mathbb{P}(\theta = h)}{\mathbb{P}(\varepsilon = h)} \equiv H$$

$$\mathbb{P}(\theta = h|\varepsilon = h) = \frac{q[\phi(\alpha_h^P)\omega + (1-\omega)\phi(\alpha_h^{NP})]}{q[\phi(\alpha_h^P)\omega + (1-\omega)\phi(\alpha_h^{NP})] + (1-q)[(1-\phi(\alpha_l^P))\omega + (1-\omega)(1-\phi(\alpha_l^{NP}))]}$$

$$\mathbb{P}(\theta = l|\varepsilon = l) = \frac{\mathbb{P}(\varepsilon = l|\theta = l)\mathbb{P}(\theta = l)}{\mathbb{P}(\varepsilon = l)} = \equiv L$$

$$\mathbb{P}(\theta = l|\varepsilon = l) = \frac{(1-q)[\phi(\alpha_l^P)\omega + (1-\omega)\phi(\alpha_l^{NP})]}{q[(1-\phi(\alpha_h^P))\omega + (1-\omega)(1-\phi(\alpha_h^{NP}))] + (1-q)[\phi(\alpha_l^P)\omega + (1-\omega)\phi(\alpha_l^{NP})]}$$

Conversely,

$$\mathbb{P}(\theta = l|\varepsilon = h) \equiv 1 - H$$

$$\mathbb{P}(\theta = h|\varepsilon = l) \equiv 1 - L$$

4.3.3 Expected utility functions

Borrowers maximize their expected utility $\mathbb{E}(U|\theta)$ with respect to their frequency of use of the payment instrument proposed by the bank, denoted α. All things considered, we have the following expected utilities for financial privacy conscious consumers:

Low-risk (type h borrower)

$$\begin{aligned} \max_{\alpha} \quad & \mathbb{E}(U|\theta = h) = v - \alpha(f_b + c) - (1-\alpha)s \\ & + \mathbb{P}(\varepsilon = h/\theta = h) p_h \delta I \end{aligned} \tag{4.1}$$

High-risk (type l borrower)

$$\begin{aligned} \max_{\alpha} \quad & \mathbb{E}(U|\theta = l) = v - \alpha(f_b + c) - (1-\alpha)s \\ & + \mathbb{P}(\varepsilon = h/\theta = l) p_l \delta I \end{aligned} \tag{4.2}$$

A proportion $(1-\omega)$ of borrowers have a financial privacy cost c equal to zero, and the following expected utilities:

Low-risk (type h borrower)

$$\begin{aligned} \max_{\alpha} \quad & \mathbb{E}(U|\theta = h) = v - \alpha f_b - (1-\alpha)s \\ & + \mathbb{P}(\varepsilon = h/\theta = h) p_h \delta I \end{aligned} \tag{4.3}$$

High-risk (type l borrower)

$$\begin{aligned} \max_{\alpha} \quad & \mathbb{E}(U|\theta = l) = v - \alpha f_b - (1-\alpha)s \\ & + \mathbb{P}(\varepsilon = h/\theta = l) p_l \delta I \end{aligned} \tag{4.4}$$

Expected utilities of borrowers are linear in alpha because of Assumption 2. At the equilibrium, borrowers set $\alpha = 0$ or $\alpha = 1$: they exclusively use either the payment instrument of the bank or alternatives.

Expected profit of the bank

The bank determines the cost f_b of the use of its payment instrument, and grants credit when it receives a signal $\varepsilon = h$ from borrowers.

We denote $\pi(f_b)$ the profit the bank derives from the use of its payment instrument.

The total expected profit of the bank is:

$$\begin{aligned}\max_{f_b} \quad & \mathbb{E}(\Pi_B) = \mathbb{P}(\varepsilon = h)\Big[\mathbb{P}(\theta = h|\varepsilon = h)p_h(1-\delta)I \\ & + (1-\mathbb{P}(\theta = l|\varepsilon = h))p_l(1-\delta)I - 1\Big] + \pi(f_b)\end{aligned} \tag{4.5}$$

4.3.4 Timing

The timing of the model is as follows:

1. The bank sets the cost f_b of its payment instrument
2. Borrowers choose the proportion α of their consumption that is paid with the payment instrument of the bank. Using its screening technology, the bank decides to grant each borrower a credit or not. Borrowers that have been granted a credit then reimburse it or not according the probability p_θ.

We resolve the model by determining the equilibrium cost of using the payment instrument of the bank f_b^*, which induces levels of use of this instruments by borrowers (i.e. α^*). In a first case, we look at a benchmark equilibrium, that is to say without considering financial privacy concerns. We introduce these concerns in a second case, in order to determine their impact on the expected profit of the bank. We finally consider a case where the bank commits not to use payment data during its screening process, which means that the value of α chosen by borrowers does not affect the precision of the bank's screening.

4.4 Case 1: Lending decisions according to a screening technology

We do not consider financial privacy concerns in this first case: the financial privacy cost c is equal to 0 for all borrowers, regardless of their type. The cost of using the payment instrument of the bank is then reduced to its fee f_b. Accordingly, the proportion ω of financial privacy minded borrowers is also reduced to 0. This case represents the ideal setting for the bank, as using payment data for screening purposes does not generate financial privacy costs among its customers.

4.4.1 Stage 2: lending decisions and payoffs

The expected utility functions of borrowers are linear in α: they set either $\alpha = 0$ or $\alpha = 1$. In other words, they either exclusively use the payment instrument of the bank, or they never

use it.

We denote α_θ the choice of type θ borrowers. Borrowers play a Nash equilibrium game, conditional on the value of f_b:

		Type h borrowers	
		$\alpha_h = 0$	$\alpha_h = 1$
Type l borrowers	$\alpha_l = 0$	$(Credit, Credit)$	$(No credit, Credit)$
	$\alpha_l = 1$	$(No credit, Credit)$	$(No credit, Credit)$

The two situations where type h borrowers set $\alpha_h = 1$ constitute Nash equilibrium, as does the situation where no borrowers use the payment instrument of the bank ($\alpha_l = \alpha_h = 0$). However, the choice of the value of α does not only depend on access to credit or not, but also on the value of f_b. Hence, all four situations of payment instrument choices by borrowers are theoretically possible. These choices correspond to the following expected profits for the bank:

$$(\alpha_l = 0, \alpha_h = 0) : \mathbb{E}(\Pi_b) = \frac{1}{2}[qp_h(1-\delta)I + (1-q)(p_l(1-\delta)I) - 1]$$

$$(\alpha_l = 1, \alpha_h = 0) : \mathbb{E}(\Pi_b) = (1-q)f_b + q[p_h(1-\delta)I - 1]$$

$$(\alpha_l = 0, \alpha_h = 1) : \mathbb{E}(\Pi_b) = qf_b + q[p_h(1-\delta)I - 1]$$

$$(\alpha_l = 1, \alpha_h = 1) : \mathbb{E}(\Pi_b) = f_b + q[p_h(1-\delta)I - 1]$$

The only situation where the bank has no information on the true type of borrowers is the first one, that is when no borrowers ($\alpha_l = \alpha_h = 0$) use its payment instrument and the precision of the screening technology is only $\phi(0) = \frac{1}{2}$.

Otherwise, if one type of borrowers uniformly sets $\alpha = 1$, then the bank can perfectly infer the type of all borrowers. This is possible for two reasons. First, the bank observes both the signal sent by consumption and the value of α. Second, there are only two types of borrowers in this setting, all without financial privacy concerns: by knowing with certainty the behavior of one type of borrowers, the bank can perfectly infer which borrowers are from the the other type.

The bank then only needs to have one type of borrowers using exclusively its payment instrument. Type h borrowers have an obvious interest to do so, as it allows them to be sure of being granted credit. They set $\alpha_h = 1$ if their expected utility (Equation 4.3) is higher than if they never use the payment instrument of the bank ($\alpha_h = 0$).

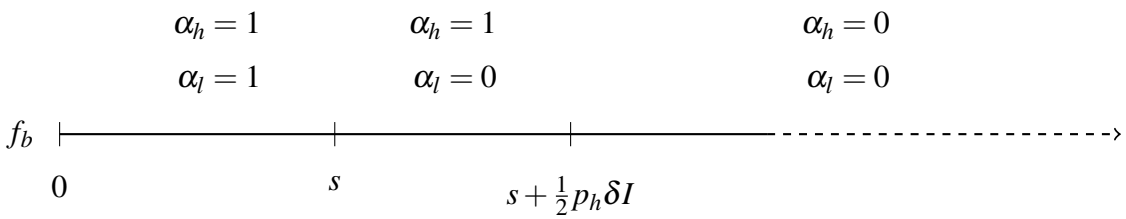
To get credit, type l borrowers may want to hide their true type to the bank by setting $\alpha_l = 0$. This is however useless if type h borrowers have chosen $\alpha_h = 1$. Type l borrowers then can only fool the bank by exclusively using the alternative if type h borrowers also have decided not to use the payment instrument of the bank. If that is not the case, type l borrowers use whatever payment instrument is the cheapest (Equation 4.4).

LEMMA 1

Choosing a value of f_b which brings no use of its payment instrument is always a dominated strategy for the bank. [19]

4.4.2 Stage 1: choice of payment instrument by borrowers

The goal for the bank is to face a situation where all type h borrowers use exclusively its payment instrument ($\alpha_h = 1$). Borrowers choose which payment instrument they are going to use by comparing costs (f_b with s) and their possibility of being granted credit or not. For a given f_b^*, borrowers compare their expected utilities between setting $\alpha = 1$ and $\alpha = 0$. Knowing the value of the alternative s, the bank then faces three different tarification strategies, each corresponding to different payment instrument choices by borrowers:[20]



From Lemma 1, we know that the bank is always going to set the cost f_b of its payment instrument inferior to $s + \frac{1}{2} p_h \delta I$, in order to have at least a part of borrowers that uses exclusively this instrument. Indeed, all type h borrowers set $\alpha = 1$, as they have a higher expected utility in that case than if they choose $\alpha = 0$.

[19] The best situation that the bank can face is the one where at least one type of borrowers uses its payment instrument, as it provides all information needed for making perfect lending decisions. In that case, the bank makes no losses on the credit market.

[20] A detailed resolution of the demands for the bank payment instrument can be found in the Annex A.

Setting $f_b^* < s + \frac{1}{2} p_h \delta I$ allows the bank to infer the type of all borrowers. Knowing with certainty which borrowers are low-risk ones, the bank makes perfect lending decisions, which prevents the bank of losing money on the credit market.

At the equilibrium, the bank has two strategies: either it chooses $f_b^* = s + \frac{1}{2} p_h \delta I$ and only type h borrowers set $\alpha = 1$ (**Strategy A**), or it chooses $f_b^* = s$ and all borrowers use exclusively its payment instrument (**Strategy B**). We rewrite Equation 4.5 accordingly:

$$\textbf{Strategy A:}\quad \Pi_A = \mathbb{E}(\Pi_b / f_b^* = s + \frac{1}{2} p_h \delta I) = q((s + \frac{1}{2} p_h \delta I) + q[p_h(1-\delta)I - 1] \tag{4.6}$$

$$\textbf{Strategy B:}\quad \Pi_B = \mathbb{E}(\Pi_b / f_b^* = s) = s + q[p_h(1-\delta)I - 1] \tag{4.7}$$

PROPOSITION 1

If the proportion q of type h borrowers is inferior to $\overline{q} = \frac{s}{s + \frac{1}{2} p_h \delta I}$, the bank sets $f_b^* = s$ (strategy B, Equation 4.7). If $q \geq \overline{q}$, $f_b^* = s + \frac{1}{2} p_h \delta I$ (strategy A, Equation 4.6).

With respect to the proportion of low-risk borrowers q, the bank either decides to have a monopoly on the use of payment instruments by setting $f_b^* = s$ or decides to set $f_b^* = s + \frac{1}{2} p_h \delta I$ and only attract type h borrowers towards its payment instrument. While the profit made on the credit market remains the same, these two pricing strategies do not yield the same profit on the payment instrument supply market:

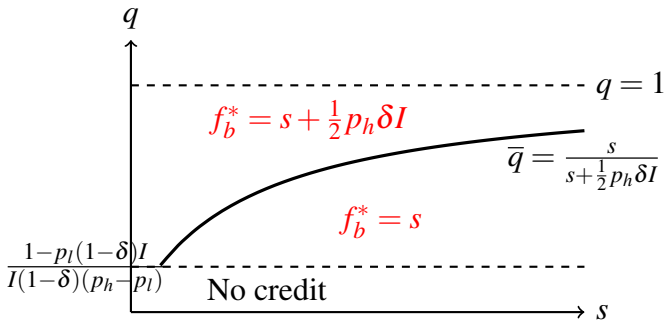


Figure 4.1 – $f_b{}^*$ with respect to the proportion q of type h borrowers and the cost s of using an alternative payment instrument

At the equilibrium, the fact that the bank can use payment instrument data for screening purposes generates optimal lending decisions: all type h borrowers are granted credit, while no

type l borrowers are. In a context where the credit market is characterized by low risk (high proportion q of low-risk borrowers), the bank has also enough leverage to propose a payment instrument more costly than any alternative due to its monopoly on credit.

It is however imperative for the bank to take the value of s into account to make perfect lending decisions. This exogenous cost s could be interpreted as the ease of access to alternative payment instruments (cash, digital currencies and so on). The current context with the development of many new non-bank payment instruments could correspond in our model as a decrease of the value of s,. Our results indicate that the most appropriate response of the bank is to lower the fee f_b of its payment instrument. Because the bank needs payment information to conduct its credit activity properly, s puts a competitive pressure on the value f_b the bank can determine.

4.5 Case 2: Screening-based lending decisions with financial privacy concerns

In this second case, we introduce financial privacy concerns among a proportion ω of borrowers. Borrowers that are concerned about financial privacy support an additional cost c when using the payment instrument of the bank. We analyze what consequences do these financial privacy concerns have on the choice of payment instrument by borrowers, the bank's pricing strategy and the efficiency of its lending decisions.

With financial privacy concerns, the main difficulty for the bank is that it is not able to precisely identify the reason why a borrower does not use its payment instrument. It might be a type l borrower that wants to hide its type to the bank, or it could be a type h borrower who refuses to use the payment instrument of the bank because of financial privacy concerns. In this second case, the bank may make type I mistakes by refusing to lend to a type h borrower.

4.5.1 Stage 2: lending decisions and payoffs

Depending on the choice of payment instrument made by borrowers, information available to the bank about the true type of borrowers varies. We denote α_θ^P the payment instrument choice of borrowers of type θ who are sensitive to financial privacy, and α_θ^{NP} the choice of borrowers not sensitive to it.

Lemma 1 still applies in presence of financial privacy concerns: if at least one type θ of borrowers uniformly chooses to set $\alpha = 1$, then the bank is able to infer perfectly the type of all

borrowers, as in the previous case. This corresponds to the following expected profits:

$$(\alpha_l = 1, \alpha_h = 0) : \mathbb{E}(\Pi_b) = (1-q)f_b + q[p_h(1-\delta)I - 1]$$

$$(\alpha_l = 0, \alpha_h = 1) : \mathbb{E}(\Pi_b) = qf_b + q[p_h(1-\delta)I - 1]$$

$$(\alpha_l = 1, \alpha_h = 1) : \mathbb{E}(\Pi_b) = f_b + q[p_h(1-\delta)I - 1]$$

Borrowers of the same type may however make different payment instrument choices. A situation may arise where a part of both types of borrowers never uses the payment instrument of the bank, setting $\alpha = 0$. In this situation, some type h borrowers may be confused with type l borrowers.

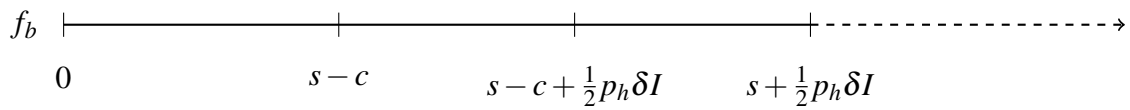
4.5.2 Stage 1: choice of payment instrument by borrowers

As in the first case, the bank must take into account the presence of alternative payment instruments in setting the fee f_b. In addition to the previous case however, the bank must now also consider the presence of financial privacy concerns: the financial privacy cost c and the proportion ω of borrowers that supports it. We suppose that $c > \frac{1}{2}p_h\delta I$, meaning that the financial privacy cost is superior to the gain of applying for credit for type h borrowers.[21]

[21] This hypothesis allows the model to be simpler without affecting the results. Considering the case $c < \frac{1}{2}p_h\delta I$ only gives to the bank another possible strategy, which does not bring more insights about the impact of financial privacy on the profitability of the bank. Moreover, this additional strategy only appears at the equilibrium for restrictive values of the parameters of the model. The resolution of the model when $c < \frac{1}{2}p_h\delta I$ can be found in Annex E.

The bank has four different strategies (Π_1 to Π_4), each corresponding to choices of payment instrument made by borrowers with respect to the value of f_b:[22]

Π_1	Π_2	Π_3	Π_4
$\alpha_h^{NP}=1$	$\alpha_h^{NP}=1$	$\alpha_h^{NP}=1$	$\alpha_h^{NP}=0$
$\alpha_h^{P}=1$	$\alpha_h^{P}=1$	$\alpha_h^{P}=0$	$\alpha_h^{P}=0$
$\alpha_l^{NP}=1$	$\alpha_l^{NP}=0$	$\alpha_l^{NP}=0$	$\alpha_l^{NP}=0$
$\alpha_l^{P}=1$	$\alpha_l^{P}=0$	$\alpha_l^{P}=0$	$\alpha_l^{P}=0$



Here are the different expected profits faced by the bank depending on the value of f_b and the choice of payment instruments by borrowers:

$$\textbf{Strategy } \Pi_1 \ (f_b = s-c)\textbf{:} \quad \Pi_1 = (s-c) + q[p_h(1-\delta)I - 1] \tag{4.8}$$

$$\textbf{Strategy } \Pi_2 \ (f_b = s-c+\frac{1}{2}p_h\delta I)\textbf{:} \quad \Pi_2 = q[s-c+\frac{1}{2}p_h\delta I + p_h(1-\delta)I - 1] \tag{4.9}$$

$$\begin{aligned}\textbf{Strategy } \Pi_3 \ (f_b = s+\frac{1}{2}p_h\delta I)\textbf{:} \quad \Pi_3 &= (s+\frac{1}{2}p_h\delta I)q(1-\omega) \\ &+q(1-\frac{\omega}{2})(p_h(1-\delta)I-1) + (\frac{1-q}{2})(p_l(1-\delta)I-1)\end{aligned} \tag{4.10}$$

$$\textbf{Strategy } \Pi_4 \ (f_b > s+\frac{1}{2}p_h\delta I)\textbf{:} \quad \Pi_4 = \frac{1}{2}[q(p_h(1-\delta)I) + (1-q)(p_l(1-\delta)I) - 1] \tag{4.11}$$

[22]The detailed resolution of the choice of payment instruments can be found in the appendix B.1.

PROPOSITION 2

At the equilibrium, the bank chooses: [23]

- **strategy** Π_1 **(**$\alpha_h^P = \alpha_h^{NP} = \alpha_l^P = \alpha_l^{NP} = 1$**) with** $f_b^* = s - c$ **when** $\omega > \underline{\omega}$ **and** $q < \overline{q}$
- **strategy** Π_2 **(**$\alpha_h^P = \alpha_h^{NP} = 1$ **and** $\alpha_l^P = \alpha_l^{NP} = 0$**) with** $f_b^* = s - c + \frac{1}{2} p_h \delta I$ **when** $\omega > \overline{\omega}$ **and** $q \geq \overline{q}$
- **strategy** Π_3 **(**$\alpha_h^{NP} = 1$ **and** $\alpha_h^P = \alpha_l^P = \alpha_l^{NP} = 0$**) with** $f_b^* = s + \frac{1}{2} p_h \delta I$
  - **when** $\omega < \underline{\omega}$ **and** $q < \overline{q}$
  - **when** $\omega < \overline{\omega}$ **and** $q \geq \overline{q}$

When a part of borrowers are sensitive to financial privacy, the bank face a trade-off between lowering the fee f_b of its payment instrument and making losses on the credit market. To make perfect lending decisions, the bank must decide to take into account the cost c of financial privacy. Lowering f_b, however, leads to smaller profits on the payment instrument market.

Propostion 2 is fairly intuitive: when the proportion ω of borrowers sensitive to financial privacy is too important, the bank decreases the cost of using its payment instrument, and thus chooses either strategy Π_1 or strategy Π_2. If the proportion ω of financial privacy sensitive borrowers is low, the bank accepts to lose money on the credit market, as these losses are compensated by making more profit on the payment instrument market.

The decision to decrease f_b not only depends on the proportion ω of borrowers sensitive to financial privacy: it also depends on the proportion q of type h borrowers. The less there are type h borrowers, the quicker the bank is to lower the fee f_b as bad lending decisions are more costly when q is low.

We now compare the equilibrium expected profits of the bank in the presence of financial privacy concerns (case 2) with the case where there are none (case 1). Without financial privacy concerns, the bank has two equilibrium strategies, choosing strategy Π_A (Equation (4.6)) or choosing strategy Π_B (Equation (4.7)). In case 2, the bank chooses between three strategies: strategy Π_1 (Equation (4.8)), strategy Π_2 (Equation (4.9)), strategy Π_3 (Equation (4.10)).

PROPOSITION 3

$min(\Pi_A, \Pi_B) > max(\Pi_1, \Pi_2, \Pi_3)$ [24]

[23] See Annex B.2 for proof.

[24] See Annex B.3 for proof.

At the equilibrium, the banks always makes more profit in case 1 than in case 2, as the existence of financial privacy concerns has important consequences on the profitability of the bank.

Firstly, if the cost f_b is unchanged, less borrowers use the payment instrument of the bank with respect to a situation without financial privacy concerns. If the bank wants to maintain the demand for its payment instrument, the bank has no choice but to set a lower f_b. As a result, the profit derived by the bank from the use of its payment instrument decreases due to the existence of financial privacy concerns. Competition in the payment instrument market is made stronger by the presence of financial privacy concerns.

Secondly, it might be optimal for the bank to accept making bad lending decisions in case 2. These mistakes also lower the expected profit at the equilibrium with respect to the case without financial privacy concerns.

4.6 Case 3: Lending decisions without payment data based screening

As seen in the previous case, the expected profit of the bank decreases when a part of borrowers is sensitive to financial privacy concerns. In this third case, we investigate whether the bank has an incentive to disentangle payment and credit. Such a commitment would then reduce the financial privacy cost c to 0 for all borrowers.

In this case, the bank does not have a screening technology to improve its lending decision: the use of its payment instrument by potential borrowers does not provide additional information on the precision of the signal about the type. The precision of the signal received by the bank is then $\phi = \frac{1}{2}$ regardless of α.

4.6.1 Stage 2: lending decisions and payoffs

The bank agrees to loan to a borrower only if it receives the signal $\varepsilon = h$. It knows that there is only a probability $\phi = \frac{1}{2}$ that this signal is correct. As screening is not improved by consumption data, the profit the bank derives from credit is not related to the profit it derives from the use of its payment instrument. Denoting $\pi(f_b)$ the profit on the payment instrument market, the expected profit $\Pi_F P$ of the bank in that case is the following:

$$\textbf{Strategy } \Pi_{FP}\textbf{:} \quad \pi(f_b) + \frac{1}{2}[q(p_h(1-\delta)I) + (1-q)(p_l(1-\delta)I) - 1] \tag{4.12}$$

4.6.2 Stage 1: choice of payment instrument by borrowers

Potential borrowers choose the share α of their consumption that they pay using the payment instrument of the bank. As their expected utility functions are linear in α, they choose either $\alpha^* = 0$ or $\alpha^* = 1$.

In this case, the choice of payment instrument does not affect lending decisions by the bank: borrowers use the payment instrument that is the least costly for them. No borrowers have financial privacy concerns ($c = 0$ for all), the cost of the payment instrument of the bank is then f_b for all borrowers, and using an alternative costs s.

LEMMA 2

Borrowers use exclusively one type of payment instrument: the bank's instrument if it is the cheapest (i.e. $f_b \leq s$), and only alternatives otherwise ($f_b > s$).[25]

PROPOSITION 4

At the equilibrium, the bank sets $f_b^* = s$.

The bank derives the same profit from credit regardless of the value of f_b. It chooses then the greatest f_b which allows it to capture all the demand of the use of payment instruments, that is $f_b = s$. At the equilibrium, the expected profit of the bank is then:

$$\textbf{Strategy } \Pi_{FP} \textbf{ } (f_b = s)\textbf{:} \quad s + \frac{1}{2}[q(p_h(1-\delta)I) + (1-q)(p_l(1-\delta)I) - 1] \tag{4.13}$$

All potential borrowers choose $\alpha^* = 1$. Their expected utility functions are:

$$\mathbb{E}(U|\theta = h, \alpha^* = 1) = v - s + \frac{1}{2}p_h\delta I$$

$$\mathbb{E}(U|\theta = l, \alpha^* = 1) = v - s + \frac{1}{2}p_l\delta I$$

As the bank does not have the ability to improve its lending decisions by analyzing payment data, it looses money on the credit market. The bank makes both type I and type II errors: some type h borrowers are not granted a credit while some type l ones are.

[25] If borrowers are indifferent between payment instruments ($f_b = s$), they use the bank's instrument.

The bank may however still have an interest in not using payment data in its credit granting strategy, as doing so makes the payment instrument of the bank more attractive.

PROPOSITION 5

When q is inferior to some values, the bank has an incentive to renounce the use of payment data in its screening process:[26]

- $\Pi_{FP} > \Pi_1$ **if** $q < q_1 = \frac{(1-\delta)p_l I + 2c - 1}{(1-\delta)(p_h - p_l)I + 2}$
- $\Pi_{FP} > \Pi_2$ **if** $q < q_2 = \frac{2s + (1-\delta)p_l I - 1}{2 - 2c + 2s + (1-\delta)(p_h - p_l)I}$
- $\Pi_{FP} > \Pi_3$ **if** $q < q_3 = \frac{2s}{(1-\omega)(p_h I + 2s + 1)}$

Overall, there exists situations where the bank has an incentive to assure borrowers that it is not going to make use of any consumption data in its screening technology. While the bank loses money on the credit market front, this loss is compensated by the growth of the demand for its payment instrument. This strategy is profitable when q is low, meaning when there are few type h borrowers and the credit market is risky. Under certain levels of q, the bank is better off by focusing most of its activity on the payment instrument market.

Values of q under which it is profitable for the bank to give up on linking payment data and screening (q_1, q_2 and q_3) increase with the weight of financial privacy:

$$\frac{\partial q_1}{\partial c} > 0$$

$$\frac{\partial q_2}{\partial c} > 0$$

$$\frac{\partial q_3}{\partial \omega} > 0$$

The bigger the importance of financial privacy (either by the value of the financial privacy cost c or the proportion ω of borrowers sensitive to it), the more the bank has an incentive to stop using payment data in its screening technology.

While having payment data at its disposal is a great asset for the bank when making lending decisions, it might come at the cost of rising financial privacy concerns among its customers. In

[26] See Annex D for proof.

the case where too many consumers have those concerns, it can be economically valuable for the bank to develop payment instruments that are more respectful of financial privacy.

4.7 Conclusion

The goal of this paper is to question whether the profitability of banks is challenged by the rise of financial privacy concerns. Investigating this issue is relevant as more and more consumers are sensitive to the collection and the use of their financial information. Financial privacy concerns could then lead some consumers to use non-bank payment instruments. This phenomenon would affect not only the profit derived by banks on the payment instruments market, but also their ability to make efficient lending decisions.

We find that financial privacy concerns toughens competition in the market for payment instruments, forcing the bank to lower the cost of its instrument if it wants to keep full coverage of the market. This has a direct consequence on lending: if the bank has less payment data at its disposal, the bank makes mistakes during its lending decision process. We find that it is optimal in some cases for the bank to decline to incorporate payment data in its screening process.

Another solution for banks is to improve their respect for financial privacy. One way of doing so is to propose payment instruments that allow consumers to control what transactional data is at their bank's disposal. Another way is to develop financial privacy preserving credit scoring. Tobback and Martens (2017) propose privacy-friendly credit scoring algorithms that are based on payment data.

BIBLIOGRAPHY

[1] **Arango, Carlos, Kim P. Huynh, and Leonard Sabetti**, "Consumer payment choice: Merchant card acceptance versus pricing incentives," *Journal of Banking & Finance*, 2015, *55*, 130–141.

[2] **Athey, Susan, Christian Catalini, and Catherine Tucker**, "The Digital Privacy Paradox: Small Money, Small Costs, Small Talk," 2017.

[3] **Broecker, Thorsten**, "Credit-Worthiness Tests and Interbank Competition," *Econometrica*, 1990, *58* (2), 429–452.

[4] **Drehmann, Mathias, Charles Goodhart, and Malte Krueger**, "The challenges facing currency usage: will the traditional transaction medium be able to resist competition from new technologies?," *Economic Policy*, 2002, *17*, 193–228.

[5] **Hauswald, Robert and Robert Marquez**, "Information Technology and Financial Services Competition," *The Review of Financial Studies*, 2003, *16* (3), 921–948.

[6] ____ **and** ____ , "Competition and Strategic Information Acquisition in Credit Markets," *The Review of Financial Studies*, 2006, *19* (3).

[7] **Hibbeln, Martin Thomas, Lars Norden, Piet Usselmann, and Marc GÃ¼rtler**, "Informational Synergies in Consumer Credit," *Paris December 2015 Finance Meeting EUROFIDAI - AFFI*, 2015.

[8] **Kanniainen, Vesa and Rune Stenbacka**, "Lending market structure and monitoring incentives," *Working paper*, 1998.

[9] **Mester, Loretta J., Leonard I. Nakamura, and Micheline Renault**, "Transactions Accounts and Loan Monitoring," *The Review of Financial Studies*, 2007, *20* (3), 529–556.

[10] **Mukminov, Rinat**, "Deposit Markets, Lending Markets and Bank Screening Incentives." PhD dissertation, Hanken School of Economics 2015.

[11] **Puri, Manju, Joerg Rocholl, and Sascha Steffen**, "What kinds of bank-client relationships matter in reducing loandefaults and why?," *Working Paper*, 2013.

[12] **Schoar, Antoinette**, "The Personal Side of Relationship Banking," *Working Paper, MIT*, 2012.

[13] **Schuh, Scott and Joanna Stavins**, "Why are (some) consumers (finally) writing fewer checks? The role of payment characteristics," *Journal of Banking & Finance*, 2010, *34* (8), 1745–1758.

[14] **Stavins, Joanna**, "Effect of consumer characteristics on the use of payment instruments," *New England Economic Review*, 2001, *3*, 20–31.

[15] **Tobback, Ellen and David Martens**, "Retail credit scoring using ne-grained payment data," *Working Paper*, 2017.

[16] **van der Cruijsen, Carin, Lola Hernandez, and Nicole Jonker**, "In love with the debit card but still married to cash," *DNB Working Paper*, 2015.

[17] **Vissing-Jorgensen, Annette**, "Consumer Credit: Learning Your Customer's Default Risk from What (S)he Buys," *Working Paper*, 2011.

[18] **von Kalckreuth, Ulf, Tobias Schmidt, and Helmut Stix**, "Choosing and using payment instruments: evidence from German microdata," *Empirical economics*, 2014, *46*, 1019–1055.

Annex A: Case 1 - Lending decisions according to a screening technology

Payment instrument demands

Consumers of each type choose to use the bank payment instrument or not depending on the respective resulting expected utilities. These utilities are however also affected by the choice made by the other type of consumers. We have the following choices:

Users of type h

In the case where no type l consumers uses the payment instrument of the bank ($\alpha_l = 0$), type h consumers set $\alpha_h = 1$ if :

$$\mathbb{E}(U|\theta = h, \alpha_h = 1) > \mathbb{E}(U|\theta = h, \alpha_h = 0)$$

$$\Leftrightarrow v - f_b + (p_h(1-\delta)I) > v - s + \frac{1}{2}(p_h \delta I)$$

$$\Leftrightarrow f_b < s + \frac{1}{2}(p_h \delta I)$$

Conversely, type h consumers set $\alpha_h = 1$ in the case where type l consumers already use the bank payment instrument if:

$$\mathbb{E}(U|\theta = h, \alpha_h = 1) > \mathbb{E}(U|\theta = h, \alpha_h = 0)$$

$$\Leftrightarrow v - f_b + (p_h \delta I) > v - s + (p_h \delta I)$$

$$\Leftrightarrow f_b < s$$

Users of type l

In the case where no type h consumers uses the payment instrument of the bank ($\alpha_h = 0$), type l consumers set $\alpha_h = 1$ if :

$$\mathbb{E}(U|\theta = l, \alpha_l = 1) > \mathbb{E}(U|\theta = l, \alpha_l = 0)$$

$$\Leftrightarrow v - f_b > v - s + \frac{1}{2}(p_l \delta I)$$

$$\Leftrightarrow f_b < s - \frac{1}{2}(p_l \delta I)$$

Conversely, type l consumers set $\alpha_l = 1$ in the case where type h consumers already use the bank payment instrument if:

$$\mathbb{E}(U|\theta = l, \alpha_l = 1) > \mathbb{E}(U|\theta = l, \alpha_l = 0)$$

$$\Leftrightarrow v - f_b > v - s$$

$$\Leftrightarrow f_b < s$$

Annex B: Case 2 - Screening-based lending decisions with financial privacy concerns

Annex B.1: Payment instrument demands

Users of type h

A first case arises when at least a proportion of users of type l is not using the bank payment instrument(α_l^{NP} or $\alpha_l^P \neq 1$). In this case, the bank cannot distinguish between between a users of type h which would choose $\alpha_h^P = 0$ and a user of type l also choosing $\alpha_l = 0$. Hence, no matter the choice of users of type h, the bank will not be able to identify them.

Users of type h who are sensitive to financial privacy will choose $\alpha = 1$ if:

$$\mathbb{E}(U|\theta = h^{\omega}, \alpha_h^P = 1) > \mathbb{E}(U|\theta = h^{\omega}, \alpha_h^P = 0)$$

$$\Leftrightarrow v - (f_b + c) + p_h \delta I > v - s + \frac{1}{2}(p_h \delta I)$$

$$\Leftrightarrow f_b < \frac{1}{2}\left[p_h \delta I\right] + (s - c)$$

Users of type h not sensitive to financial privacy will choose $\alpha_h^{NP} = 1$ if:

$$\mathbb{E}(U|\theta = h^{1-\omega}, \alpha_h^{NP} = 1) > \mathbb{E}(U|\theta = h^{1-\omega}, \alpha_h^{NP} = 0)$$

$$\Leftrightarrow v - f_b + (p_h \delta I) > v - s + \frac{1}{2}(p_h \delta I)$$

$$\Leftrightarrow f_b < \frac{1}{2}\left[p_h \delta I\right] + s$$

A second case arises when all the uses of type l are using the bank payment instrument, $\alpha_l^{NP} = \alpha_l^P = 1$, and therefore are perfectly identified by the bank.

In that case, users of type h who are sensitive to financial privacy will choose $\alpha_h^P = 1$ if:

$$\mathbb{E}(U|\theta = h^{\omega}, \alpha_h^P = 1) > \mathbb{E}(U|\theta = h^{\omega}, \alpha_h^P = 0)$$

$$\Leftrightarrow = v-(f_b+c)+(p_h\delta I) > v-s+(p_h\delta I)$$

$$\Leftrightarrow f_b < s-c$$

Users of type h who are not sensitive to financial privacy will choose $\alpha_h^{NP}=1$ if:

$$\mathbb{E}(U|\theta = h^{1-\omega}, \alpha_h^{NP}=1) > \mathbb{E}(U|\theta = h^{1-\omega}, \alpha_h^{NP}=0)$$

$$\Leftrightarrow v-f_b+(p_h\delta I) > v-s+(p_h\delta I)$$

$$\Leftrightarrow f_b < s$$

Users of type l

A first case arises when at least a proportion of users of type h is not using the bank payment instrument(α_h^{NP} or $\alpha_h^P \neq 1$). As for users of type h, the bank cannot distinguish between between a users of type h which would choose $\alpha_h^P=0$ and a user of type l also choosing $\alpha_l=0$. Hence, no matter the choice of users of type l, the bank will not be able to identify them.

In that case, users of type l who are sensitive to financial privacy will choose $\alpha_l^P=1$ if:

$$\mathbb{E}(U|\theta = l^{\omega}, \alpha_l^{P}=1) > \mathbb{E}(U|\theta = l^{\omega}, \alpha_l^{P}=0)$$

$$\Leftrightarrow v-(f_b+c) > v-s+\frac{1}{2}(p_l\delta I)$$

$$\Leftrightarrow f_b < (s-c)-\frac{1}{2}(p_l\delta I)$$

Users of type l who are not sensitive to financial privacy will choose $\alpha_l^{NP}=1$ if:

$$\mathbb{E}(U|\theta = l^{1-\omega}, \alpha_l^{NP}=1) > \mathbb{E}(U|\theta = l^{1-\omega}, \alpha_l^{NP}=0)$$

$$\Leftrightarrow v - f_b > v - s + \frac{1}{2}(p_l \delta I)$$

$$\Leftrightarrow f_b < s - \frac{1}{2}(p_l \delta I)$$

A second case arises when all the uses of type h are using the bank payment instrument, $\alpha_h^{NP} = \alpha_h^P = 1$, and therefore are perfectly identified by the bank.

In that case, users of type l who are sensitive to financial privacy will choose $\alpha_l^P = 1$ if:

$$\mathbb{E}(U|\theta = l^{\omega}, \alpha_l^P = 1) > \mathbb{E}(U|\theta = l^{\omega}, \alpha_l^P = 0)$$

$$\Leftrightarrow v - (f_b + c) > v - s$$

$$\Leftrightarrow f_b < s - c$$

Users of type l who are not sensitive to financial privacy will choose $\alpha_l^{NP} = 1$ if:

$$\mathbb{E}(U|\theta = l^{1-\omega}, \alpha_l^{NP} = 1) > \mathbb{E}(U|\theta = l^{1-\omega}, \alpha_l^{NP} = 0)$$

$$\Leftrightarrow v - f_b > v - s$$

$$\Leftrightarrow f_b < s$$

Annex B.2: Proof of proposition 2

It is trivial to see that strategy Π_4 is always dominated at the equilibrium.

Strategy Π_1

Comparing strategies Π_1 and Π_2, we have:

$$\Pi_1 > \Pi_2$$

$$\Leftrightarrow (s-c)+q[p_h(1-\delta)I-1] > q[s-c+\frac{1}{2}p_h\delta I+p_h(1-\delta)I-1]$$

$$\Leftrightarrow q < \overline{q} = \frac{s-c}{s-c+\frac{1}{2}p_h\delta I}$$

Comparing strategies Π_1 and Π_3, we have:

$$\Pi_1 > \Pi_3$$

$$\Leftrightarrow \omega > \underline{\omega} = \frac{2c+qp_h\delta I-2s(1-q)+(1-q)(p_l(1-\delta)I-1)}{2qs+qp_h\delta I+q(p_h(1-\delta)I-1)}$$

Strategy Π_2

Comparing strategies Π_2 and Π_3, we have:

$$\Pi_2 > \Pi_3$$

$$\Leftrightarrow \omega > \overline{\omega} = \frac{2qc+(1-q)(p_l(1-\delta)I-1)}{2qs+qp_h\delta I+2q(p_h(1-\delta)I-1)}$$

We also have:

$$\underline{\omega} < \overline{\omega}$$

$$\Leftrightarrow q < \frac{s-c}{s-c+\frac{1}{2}p_h\delta I} = \overline{q}$$

Annex B.3: Proof of Proposition (3)

Strategy A

We compare the expected profit of the bank using **strategy A** (Equation (4.6)) in the case without financial privacy concerns with any equilibrium expected profit of case 2.

Here is **strategy A**, as well as the strategies between which the bank chooses at the equilibrium in case 2:

$$\textbf{Strategy A: } \mathbb{E}(\Pi_b / f_b^* = s + \frac{1}{2} p_h \delta I) = q((s + \frac{1}{2} p_h \delta I) + q[p_h(1-\delta)I - 1]$$

$$\textbf{Strategy 1: } \Pi_1 = \mathbb{E}(\Pi_B | f_b^* = s - c) = (s - c) + q[p_h(1-\delta)I - 1]$$

$$\textbf{Strategy 2: } \Pi_2 = \mathbb{E}(\Pi_B | f_b^* = s - c + \frac{1}{2} p_h \delta I) = q[s - c + \frac{1}{2} p_h \delta I + p_h(1-\delta)I - 1]$$

$$\begin{aligned} \textbf{Strategy 3: } \Pi_3 = \mathbb{E}(\Pi_B | f_b^* = s + \frac{1}{2} p_h \delta I) &= (s + \frac{1}{2} p_h \delta I) q(1-\omega) \\ &+ q(1 - \frac{\omega}{2})(p_h(1-\delta)I - 1) + (\frac{1-q}{2})(p_l(1-\delta)I - 1) \end{aligned}$$

Comparing **strategy A** and **strategy 1**, we have:

$$\textbf{Strategy A} > \textbf{Strategy 1}$$

$$\Leftrightarrow q((s + \frac{1}{2} p_h \delta I) + q[p_h(1-\delta)I - 1] > (s - c) + q[p_h(1-\delta)I - 1]$$

$$\Leftrightarrow q > \frac{s - c}{s + \frac{1}{2} p_h \delta I}$$

We now that the bank chooses **strategy A** for $q > \frac{s}{s + \frac{1}{2} p_h \delta I} > \frac{s-c}{s + \frac{1}{2} p_h \delta I}$. This means that **strategy A** always dominates **strategy 1** in a case where **strategy A** is an equilibrium strategy.

Comparing **strategy A** and **strategy 2**, we have:

Strategy A > Strategy 2

$$\Leftrightarrow q((s+\frac{1}{2}p_h\delta I)+q[p_h(1-\delta)I-1]>q[s-c+\frac{1}{2}p_h\delta I+p_h(1-\delta)I-1]$$

$$\Leftrightarrow s+\frac{1}{2}p_h\delta I>s-c+\frac{1}{2}p_h\delta I)$$

which is always true.

Comparing **strategy A** and **strategy 3**, we have:

Strategy A > Strategy 3

$$\Leftrightarrow q((s+\frac{1}{2}p_h\delta I)+q[p_h(1-\delta)I-1]>(s+\frac{1}{2}p_h\delta I)q(1-\omega)$$
$$+q(1-\frac{\omega}{2})(p_h(1-\delta)I-1)+(\frac{1-q}{2})(p_l(1-\delta)I-1)$$

$$\Leftrightarrow (s+\frac{1}{2}p_h\delta I)q\omega+\frac{\omega}{2}(p_h(1-\delta)I-1)-(\frac{1-q}{2})(p_l(1-\delta)I-1)>0$$

which is always true, all parts of this sum being positive (as $p_h(1-\delta)I<1$).

Strategy B

We now compare the expected profit of the bank using **strategy B** (Equation (4.7)) in the case without financial privacy concerns with any equilibrium expected profit of case 2.

Here is **strategy B**, as well as the strategies between which the bank chooses at the equilibrium in case 2:

$$\textbf{Strategy B: } \mathbb{E}(\Pi_b / f_b^* = s) = s + q[p_h(1-\delta)I - 1]$$

$$\textbf{Strategy 1: } \Pi_1 = \mathbb{E}(\Pi_B | f_b^* = s - c) = (s-c) + q[p_h(1-\delta)I - 1]$$

$$\textbf{Strategy 2: } \Pi_2 = \mathbb{E}(\Pi_B | f_b^* = s - c + \frac{1}{2}p_h\delta I) = q[s - c + \frac{1}{2}p_h\delta I + p_h(1-\delta)I - 1]$$

$$\begin{aligned}\textbf{Strategy 3: } \Pi_3 = \mathbb{E}(\Pi_B | f_b^* = s + \frac{1}{2}p_h\delta I) &= (s + \frac{1}{2}p_h\delta I)q(1-\omega) \\ &+ q(1 - \frac{\omega}{2})(p_h(1-\delta)I - 1) + (\frac{1-q}{2})(p_l(1-\delta)I - 1)\end{aligned}$$

Comparing **strategy B** and **strategy 1**, we have:

$$\textbf{Strategy B} > \textbf{Strategy 1}$$

$$\Leftrightarrow s + q[p_h(1-\delta)I - 1] > (s-c) + q[p_h(1-\delta)I - 1]$$

$$\Leftrightarrow s > s - c$$

which is obviously always the case.

Comparing **strategy B** and **strategy 2**, we have:

$$\textbf{Strategy B} > \textbf{Strategy 2}$$

$$\Leftrightarrow s+q[p_h(1-\delta)I-1] > q[s-c+\frac{1}{2}p_h\delta I+p_h(1-\delta)I-1]$$

$$\Leftrightarrow q < \frac{s}{s-c+\frac{1}{2}p_h\delta I}$$

We now that the bank chooses **strategy B** for $q < \frac{s}{s+\frac{1}{2}p_h\delta I} < \frac{s}{s-c+\frac{1}{2}p_h\delta I}$. This means that **strategy B** always dominates **strategy 2** in a case where **strategy B** is an equilibrium strategy.

Comparing **strategy B** and **strategy 3**, we have:

Strategy B > **Strategy 3**

$$\Leftrightarrow s+q[p_h(1-\delta)I-1] < (s+\frac{1}{2}p_h\delta I)q(1-\omega)+q(1-\frac{\omega}{2})(p_h(1-\delta)I-1)+(\frac{1-q}{2})(p_l(1-\delta)I-1)$$

$$\Leftrightarrow q < \frac{s+\frac{1}{2}(p_l(1-\delta)I-1)}{(s+\frac{1}{2}p_h\delta I)(1-\omega)-\frac{\omega}{2}(p_h(1-\delta)I-1)-\frac{1}{2}(p_l(1-\delta)I-1)}$$

which is always true for a case where **strategy B** is chosen (*i.e* $q < \frac{s}{s+\frac{1}{2}p_h\delta I}$).

Annex D: Proof of Proposition (5)

When the bank does not link payment data and screening, its expected profit is the following:

$$\mathbb{E}(\Pi_b)^* = s + \frac{1}{2}[q(p_h(1-\delta)I) + (1-q)(p_l(1-\delta)I) - 1]$$

If the bank decides to use payment data for screening purposes, there are three possible equilibrium strategies:

$$\textbf{Strategy 1: } \Pi_1 = \mathbb{E}(\Pi_B | f_b^* = s-c) = (s-c) + q[p_h(1-\delta)I - 1]$$

$$\textbf{Strategy 2: } \Pi_2 = \mathbb{E}(\Pi_B | f_b^* = s-c+\frac{1}{2}p_h\delta I) = q[s-c+\frac{1}{2}p_h\delta I + p_h(1-\delta)I - 1]$$

$$\begin{aligned}\textbf{Strategy 3: } \Pi_3 = \mathbb{E}(\Pi_B | f_b^* = s+\frac{1}{2}p_h\delta I) &= (s+\frac{1}{2}p_h\delta I)q(1-\omega) \\ &+q(1-\frac{\omega}{2})(p_h(1-\delta)I-1) + (\frac{1-q}{2})(p_l(1-\delta)I-1)\end{aligned}$$

Disentangling payment and credit is profitable for the bank with respect to **Strategy 1** when:

$$\mathbb{E}(\Pi_b)^* > \Pi_1$$

$$\Leftrightarrow q < q_1 = \frac{(1-\delta)p_l I + 2c - 1}{(1-\delta)(p_h - p_l)I + 2}$$

Disentangling payment and credit is profitable for the bank with respect to **Strategy 2** when:

$$\mathbb{E}(\Pi_b)^* > \Pi_2$$

$$\Leftrightarrow q < q_2 = \frac{2s + (1-\delta)p_l I - 1}{2 - 2c + 2s + (1-\delta)(p_h - p_l)I}$$

Disentangling payment and credit is profitable for the bank with respect to **Strategy 3** when:

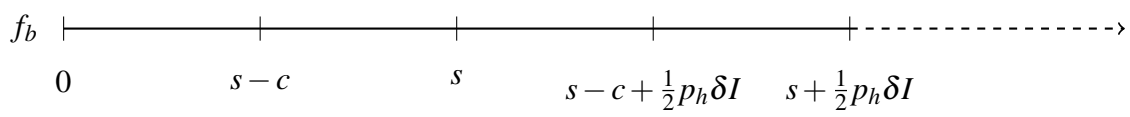
$$\mathbb{E}(\Pi_b)^* > \Pi_3$$

$$\Leftrightarrow q < q_3 = \frac{2s}{(1-\omega)p_h I + 2s(1-\omega) + (1-\omega)}$$

Annex E: Equilibrium strategies for $c < \frac{1}{2} p_h \delta I$

The bank may face situation where the cost related to financial privacy is low $c < \frac{1}{2} p_h \delta I$. In that case, knowing the choice of payment instruments by borrowers with respect to f_b, the bank have an additional strategy with respect to the high privacy cost case:[27]

Π_1	Π_5	Π_2	Π_3	Π_4
$\alpha_h^{NP} = 1$	$\alpha_h^{NP} = 1$	$\alpha_h^{NP} = 1$	$\alpha_h^{NP} = 1$	$\alpha_h^{NP} = 0$
$\alpha_h^{P} = 1$	$\alpha_h^{P} = 1$	$\alpha_h^{P} = 1$	$\alpha_h^{P} = 0$	$\alpha_h^{P} = 0$
$\alpha_l^{NP} = 1$	$\alpha_l^{NP} = 1$	$\alpha_l^{NP} = 0$	$\alpha_l^{NP} = 0$	$\alpha_l^{NP} = 0$
$\alpha_l^{P} = 1$	$\alpha_l^{P} = 0$	$\alpha_l^{P} = 0$	$\alpha_l^{P} = 0$	$\alpha_l^{P} = 0$



We denote this new situation strategy Π_5:

$$\textbf{Strategy 5: } \Pi_5 = \mathbb{E}(\Pi_B | f_b^* = s) = s(1-\omega(1-q)) + q[p_h(1-\delta)I - 1] \tag{4.14}$$

Comparing strategies Π_1 and Π_5, we have:

$$\Pi_5 > \Pi_1$$

$$\Leftrightarrow s(1-\omega(1-q)) + q[p_h(1-\delta)I - 1] > (s-c) + q[p_h(1-\delta)I - 1]$$

$$\Leftrightarrow q > \widetilde{q} = 1 - \frac{c}{s\omega}$$

Comparing strategies Π_5 and Π_2, we have:

$$\Pi_5 > \Pi_2$$

$$\Leftrightarrow s(1-\omega(1-q)) + q[p_h(1-\delta)I - 1] > q[s - c + \frac{1}{2} p_h \delta I + p_h(1-\delta)I - 1]$$

[27] See Annex B.1

$$\Leftrightarrow q < \frac{s(1-\omega)}{s(1-\omega) - c + \frac{1}{2} p_h \delta I} < \overline{q}$$

Comparing strategies Π_1 and Π_3, we have:

$$\Pi_1 > \Pi_3$$

$$\Leftrightarrow \omega > \widetilde{\omega} = \frac{q p_h \delta I + (1-q)(p_l(1-\delta)I - 1) - 2s(1-q)}{2s(2q-1) + q p_h \delta I + p_h(1-\delta)I - 1}$$

At the equilibrium, the bank then chooses:

- strategy Π_1 ($\alpha_h^P = \alpha_h^{NP} = \alpha_l^P = \alpha_l^{NP} = 1$) with $f_b^* = s - c$ when $\omega > \underline{\omega}$ and $q < \overline{q} < \widetilde{q}$
- strategy Π_2 ($\alpha_h^P = \alpha_h^{NP} = 1$ and $\alpha_l^P = \alpha_l^{NP} = 0$) with $f_b^* = s - c + \frac{1}{2} p_h \delta I$ when $\omega > \overline{\omega}$ and $q \geq \overline{q}$
- strategy Π_3 ($\alpha_h^{NP} = 1$ and $\alpha_h^P = \alpha_l^P = \alpha_l^{NP} = 0$) with $f_b^* = s + \frac{1}{2} p_h \delta I$
  - when $\omega < \widetilde{\omega} < \underline{\omega}$ and $\widetilde{q} < q < \overline{q}$
  - when $\omega < \overline{\omega}$ and $q \geq \overline{q}$
- strategy Π_5 ($\alpha_l^P = 0$ and $\alpha_h^P = \alpha_l^P = \alpha_l^{NP} = 1$) with $f_b^* = s$
  - when $\widetilde{q} < q < \overline{q}$ and $\omega > \underline{\omega}$
  - when $\widetilde{q} < q < \overline{q}$ and $\omega < \widetilde{\omega} < \underline{\omega}$

CHAPTER 5

CONCLUSION

In a famous paper, Stigler (1961) wrote that "knowledge is power", a statement that firms have increasingly embraced during the last years. With the simultaneous decrease of data collection costs and the rise of new analytical capabilities such as machine learning, firms are incorporating more and more data in their business. This has especially been the case for personal information. Several reasons have pushed firm to collect personal data: the possibility of proposing personalized services or an increase of advertisement relevancy for example. The development of online social networks, social media or user generated content (UGC) has also transformed many consumers into producers of personal information (Acquisti et al., 2015). This unprecedented importance of data has affected numerous industries (online retailing, travel agencies, etc.) and allowed the rise of new ones such as the data brokering industry or the online advertising ecosystem (Evans, 2009). The growing intensity of firms' data collection practices and the opacity that often characterizes their purposes have however generated rising privacy concerns among consumers. In the 2017 State of Consumer Privacy and Trust survey, 68% of respondents stated that they did not trust brands to handle their personal information appropriately.[1]

While there might benefit from the use of their personal data, it may indeed also come at a cost. Examples of data disclosure externalities are: data breaches, spam, credit card hacking, identity theft and so on. Furthermore, some firms could use data collection in order to extract more consumer surplus, for example through price discrimination (Odlyzko, 2003). Overall, consumers are increasingly losing trust in firms with respect to the collection and the use of personal information. The loss of consumer trust is a major threat for the digital economy: in a 2016 UK study, a third of respondents declared refusing to several online services due to privacy concerns.[2] Firms increasingly face a trade-off: collecting personal data is profitable, but it comes at the cost of a loss of consumers' trust. Importantly, this loss of trust occurs in a context where personal data is at the core of many firms' business models, especially in the digital economy. For example, most free websites generate their revenues by selling slots for

[1] Gigya - "The 2017 State of Consumer Privacy and Trust", survey conducted with 4000 adults, one-half in the US and one-half in the UK

[2] W.Ashford, ComputerWeekly - "Trust issues stifling digital economy growth, study shows"

targeted advertising. This practice is made more difficult with the loss of consumer trust, and it becomes increasingly problematic for free websites to remain profitable.

Some consumers have taken steps to actively manage the sharing of their personal information because of their privacy concerns. It is an increasingly common behavior as more and more Internet users are adopting various privacy-enhancing technologies (PETs) or strategies. Indeed, an increasing number of Internet users are regularly deleting their cookies or browsing history, installing advertisement avoidance technologies (AATs) or privacy-related browser extensions for example. Overall, consumers are increasingly able to control how their information is collected. This may be worrisome for firms as it could for example hurt their ability to collect personal information, limiting the amount of data at their disposal. The use of PETs could also affect the reliability of data, making algorithms less efficient for example by creating bias.

Previous academic work has mainly considered the economic incentives to use personal information and consumers' attitudes and behaviors regarding privacy (Acquisti et al., 2015). Outside price discrimination avoidance (Villas-Boas (2004), Taylor (2004)), few papers have however focused on strategic consumers with respect to privacy. Analyzing this type of strategic behavior, we find a possible solution to the loss of consumer trust that threatens the digital economy: consumers may be willing to share personal information if they have tools to manage the disclosure of data at their disposal. Not being able to choose their level of identification or tracking is a cost for consumers. This cost can translate in a decreasing willingness to share personal data. Providing solutions such as PETs may be a way to reconcile the use of personal information by firms and consumers' privacy concerns.

Following this path, firms could encourage the sharing of personal information by consumers, and even have access to more precise and reliable data. In other words, proposing PETs or services more respectful of privacy may represent a business opportunity for firms. Investing in privacy could for example lower the direct and indirect costs (such as reputation loss) of data breaches. It could also be a competitive asset, as several papers have shown that consumers have a willingness to pay for privacy (Tsai et al. (2011) or Savage and Waldman (2013) for example). These incentives could be especially important concerning information deemed sensitive by consumers such as health information. In a 2017 international survey by Gemalto on the Internet of things (IoT), 60% of respondents were concerned about data leakage and 54% about hackers accessing personal information.[3]Our result in Chapter 2 supports the idea that

[3]Gemalto - "The State of IoT Security"

firms have economic incentives to improve privacy as we empirically show that the use of PETs is positively correlated with willingness to share personal information.

We find a similar result concerning financial information in Chapter 3 where we make the link between financial privacy and payment instruments. Using multiple payment instruments (bank cards, PayPal, Bitcoin, etc.) is a way for consumers to have more control over their financial privacy. Consumers may want to hide their transactions for several reasons: because they are buying goods they consider sensitive (medication, gambling, etc.) or because they want to avoid being solicited too much. Assimilating non-bank payment instruments to PETs, we find that their adoption is beneficial for e-commerce. Indeed, we show that using non-bank payment instruments has positive effects on consumers' online purchases frequency. We also show that financial privacy concerns positively influences the use of non-bank payment instruments.

This result suggests that there are economic incentives to encourage the development of non-bank payment instruments. Mainly, if consumers have multiple payment instruments at their disposal, they will increasingly turn to e-commerce as consumers are then able to reconcile online purchasing and financial privacy. Not taking consumers' desire for financial confidentiality could slow down the rise of online retailing.

A growing use of non-bank payment instruments would however mean a loss of financial data for banks. In Chapter 4, we investigate how this loss of information affects banks' profitability. Indeed, when consumers have financial privacy concerns, they have an increasing use of payment instruments that prevent tracking from financial institutions (mainly non-bank payment instruments and cash). In this chapter, we design a theoretical model that links payment instrument choices of consumers with a bank's ability to screen loan applicants. We find that a lower use of the payment instrument of the bank severely affects its profits.

One reason for this decrease in profit is a lower use of banks' payment instruments as the competition gets tougher on the payment instruments market. Another reason is that banks' grant credit less efficiently as they have fewer information at their disposal. The main risk for banks is to refuse credit to solvent borrowers that have a low use of bank payment instruments because of their financial privacy concerns. Banks may then confuse these borrowers with risky ones, that also have little use of bank payment instruments, but due to their desire to trick financial institutions about their financial situation.

In line with the results of Chapters 2 and 3, a solution for banks might be to propose to

consumers a payment instrument that is respectful of their financial privacy concerns. This could give a powerful incentive for consumers to use that instrument and then make banks more competitive on the payment instrument market. Moreover, it could bring banks sufficiently enough information to improve their screening abilities, while matching consumers' financial privacy preferences.

Overall, it seems that proposing tools, being PETs or payment instruments, that take consumers' (financial) privacy preferences into account may help gain their trust. Not considering those preferences leads to a loss of trust by consumers, and could ultimately represent a loss of data for firms. In that regard, trust increasingly becomes an economic asset. In general, if firms adapt their data collection practices, they could regain consumer trust, and simultaneously keep access to consumers' personal information. If this access is the result of consumers' consent, this could mean having data in bigger volume and with better reliability.

The European Union has taken measures concerning the loss of consumer trust faced by the digital economy. The recently adopted GDPR is designed to clarify the way personal information is collected and used by firms. With that regulation, the EU defines opt-in as the standard for measuring consumers' consent on data collection: there must be an explicit acceptance on consumers' side. It has been argued that consumers' decision to opt-in or opt-out can be analyzed using Coase's theorem (1960) (Lacker, 2002): whether the standard is opt-in or opt-out, control over data will efficiently go to the actor that gives it the most value. Coase's theorem however only works if there are no transactions costs, a situation that is arguably made difficult by a loss of trust. In that context, the use of PETs can be interpreted as a way of decreasing transaction costs.

Some firms may still be tempted to take advantage of the trust of consumers. This might take the form of adverse selection, like in the case of the TRUSTe label. TRUSTe is a company that gives privacy certifications to websites. It was for shown that websites that had been certified as trustworthy with the TRUSTe label were actually more likely to be untrustworthy regarding privacy (Edelman, 2011). Another example is Ghostery: while it blocks tracking from third parties, it still collects data about its users to third parties, playing the role of a data broker. It has also been suggested that increased perceived control over the sharing of personal information is sufficient to trick consumers into disclosing more sensitive information (Brandimarte et al., 2010). In that regard, providing PETs could be a way to trick consumers into sharing more data.

Collecting anonymous data is also not enough to insure consumers' privacy. It is indeed fairly easy to process to the de-anonymization of a database: only four points of reference are for example needed to uniquely identify individuals in cellphone data of over 1.5 million users (de Montjoye et al., 2013). The promise of an anonymization process is thus not enough to insure privacy, as "there is no such thing as anonymous online tracking".[4] Consumers could be conviced to share personal information given the promise of anonymization of data, while this process has in reality little effect. Regulation thus still has an important role to play, making sure that consumer trust is not manipulated. The role of regulation also concerns the competitive landscape. Indeed, generalizing opt-in consent could lead to more concentrated markets, especially in the case of priceless markets (Campbell et al., 2015). If regulatory agencies are not vigilant enough, an undesired effect of the GDPR could then be a decrease of competition.

Still, an increased empowerment of consumers (through the use of PETs or an opt-in requirement) could turn out to be an opportunity, both for consumers and firms. Mainly, it might at least partly resolve the trade-off between the personalization of goods or services and privacy costs. While it may be a cost for firms to improve privacy online, it may also be an investment. This investment would encourage consumers to share personal information, allowing firms to have more data to generate revenues from. Moreover, it could improve the reliability of data as less consumers use obfuscation techniques. In the end, these effects could be crucial for the adoption of the Internet of things or the development of algorithms.

Privacy concerns and consumer trust are also relevant topics in developing countries. Data analysis is an important foundation of financial inclusion, for example in the case of mobile scoring. Banks, operators, credit bureaus and fintechs are increasingly relying on data (mobile payment data, telecom invoices, etc.) to form their lending decisions. Insuring that privacy concerns are sufficiently taken into account would help to build consumer trust, and thus improve financial inclusion.

[4] A.Narayanan - "There is no such thing as anonymous online tracking", The Center for Internet and Society, Standford Law School

BIBLIOGRAPHY

[1] **Acquisti, Alessandro, Curtis R. Taylor, and Liad Wagman**, "The Economics of Privacy," *Journal of Economic Literature*, July 2015.

[2] **Brandimarte, Laura, Alessandro Acquisti, and George Loewenstein**, "Misplaced Confidences: Privacy and the Control Paradox," *Ninth Annual Workshop on the Economics of Information Security (WEIS)*, 2010.

[3] **Campbell, James, Avi Goldfarb, and Catherine Tucker**, "Privacy regulation and market structure," *Journal of Economics & Management Strategy*, 2015, *24* (1), 47–73.

[4] **Coase, Ronald H.**, "The Problem of Social Cost," *Journal of Law and Economics*, 1960, *3*, 1–44.

[5] **de Montjoye, Yves-Alexandre, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel**, "Unique in the Crowd: The privacy bounds of human mobility," *Scientific Report*, 2013, *3*.

[6] **Edelman, Benjamin**, "Adverse selection in online âtrustâ certifications and search results," *Electronic Commerce Research and Applications*, 2011, *10* (1), 17–25.

[7] **Evans, David S.**, "The Online Advertising Industry: Economics, Evolution, and Privacy," *Journal of Economic Perspectives*, 2009.

[8] **Lacker, Jeffrey M.**, "The Economics of Financial Privacy: To Opt Out or Opt In?," *Federal Reserve Bank of Richmond Economic Quaterly*, Summer 2002, *88* (3), 1–16.

[9] **Odlyzko, Andrew**, "Privacy, Economics, and Price Discrimination on the Internet," *Proceedings of the 5th International Conference On Electronic Commerce (ICEC 2003)*, 2003.

[10] **Savage, Scott and Donald M. Waldman**, "The Value of Online Privacy," *Working paper*, 2013.

[11] **Stigler, George J.**, "The Economics of Information," *The Journal of Political Economy*, 1961, *69* (3), 213–225.

[12] **Taylor, Curtis R.**, "Consumer privacy and the market for customer information," *The RAND Journal of Economics*, 2004, *35* (4), 631–651.

[13] **Tsai, Janice Y., , Serge Egelman, Lorrie Cranor, and Alessandro Acquisti**, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research*, June 2011, *22*, 254–268.

[14] **Villas-Boas, J.Miguel**, "Price Cycles in Markets with Customer Recognition," *The RAND Journal of Economics*, 2004, *35* (3), 486–501.

CONTRIBUTIONS A L'ECONOMIE DE LA VIE PRIVEE ET DES DONNEES FINANCIERES

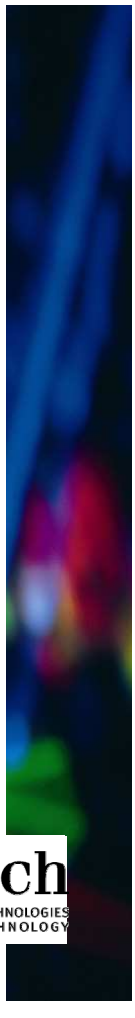
Yann BALGOBIN

RESUME : De nombreuses firmes collectent et utilisent des informations sur les consommateurs afin d'augmenter leurs profits. Ces entreprises pourraient toutefois avoir de grandissantes difficultés à générer des profits à partir des données personnelles. Premièrement, de plus en plus de consommateurs sont inquiets au sujet du respect de leur vie privée en ligne. Deuxièmement, un nombre croissant d'outils permet de contrôler la collecte de données personnelles. Nous montrons dans la thèse que l'usage de tels outils ont un effet positif sur la disposition des consommateurs à partager leurs informations personnelles, permettant ainsi une collecte de données plus respectueuse de la vie privée. Ces conclusions s'appliquent également au sujet des données financières. Encourager l'usage de moyens de paiement non-bancaires pourrait être bénéfique au commerce en ligne. Enfin, dans un contexte où les consommateurs sont inquiets quant au respect de leur vie privée, les banques pourraient bénéficier du fait de rendre leurs pratiques de *screening* moins intrusives.

MOTS-CLEFS : Economie numérique - données personnelles - banque - commerce en ligne - paiement

ABSTRACT : Many firms collect and use information about consumers to increase their revenues. Firms may face greater difficulty to generate profit from personal data. Firstly, because consumers are increasingly concerned about their privacy. Secondly, because more and more privacy-enhancing technologies (PETs) become available. We find in the thesis that the use of PETs could positively influence consumers' willingness to share personal information, enabling a data collection that takes privacy concerns into account. We make similar conclusions in the case of financial information. Developing the use of non-bank payment instruments could benefit e-commerce. Finally, in a context where consumers are concerned with their privacy, banks may benefit from making screening less intrusive.

KEY-WORDS : Digital economy - privacy - banking - e-commerce - paiement



TELECOM ParisTech

EDiTE

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY