

HEC MONTRÉAL

**Les comportements des usagers sur une plate-forme de
financement participatif**

**par
Nicolas Ricci**

**Sciences de la gestion
Option : Économie appliquée**

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

10 décembre 2014

©Nicolas Ricci, 2014

Résumé

Le financement participatif, ou *crowdfunding*, est un phénomène relativement nouveau permettant de financer des projets de petite et moyenne envergure par l'entremise de divers sites web. Ce type de financement gagne rapidement en popularité en ce qui concerne la création de biens numériques : films, musique, jeux vidéo, logiciels, etc. Dans un contexte où il devient de plus en plus difficile d'appliquer les lois relatives au droit d'auteur, le financement participatif peut-il prendre le relais et permettre de financer de tels biens ? Nous tentons de répondre à cette question à l'aide d'un modèle empirique évaluant l'impact des contributions obtenues par les créateurs de projets au moment T-1 sur les contributions obtenues au moment T. Ce modèle, basé sur celui employé lors d'une étude semblable dans le même domaine, utilise l'estimateur de système GMM permettant d'évaluer l'impact de variables indépendantes qui ne sont pas strictement exogènes. Nous interprétons nos résultats selon le cadre élaboré par la littérature existante, qui propose deux idéaux-types opposés de comportements des usagers du financement participatif : le resquillage et le grégarisme. Les résultats de notre analyse de régression nous permettent d'affirmer que le resquillage n'est généralement pas un problème en ce qui concerne le site étudié. Il existe, au contraire, un léger effet de renforcement positif qui relève des comportements grégaires.

Table des matières

1	Introduction	1
2	Revue de littérature	4
2.1	Description du phénomène	4
2.2	La plate-forme étudiée : Kickstarter.com	5
2.3	La relation entre biens numériques et biens publics	7
2.4	Deux types de comportements : resquillage et grégarisme	9
2.5	Resquillage, grégarisme et financement participatif	13
2.5.1	Modélisation	13
2.5.2	Données	15
2.5.3	Conclusions	16
3	Base de données	18
3.1	Collecte des données	18
3.2	Statistiques descriptives	19
3.2.1	Film & vidéo	20
3.2.2	Musique	21
3.2.3	Jeux vidéo	23
3.2.4	Logiciel	24
3.2.5	Autocorrélation et variance	25
4	Modélisation	27
4.1	Technique d'estimation	27
4.1.1	Données de panel dynamiques	27
4.1.2	Développement de l'approche instrumentale pour modèles dynamiques	27
4.1.3	Méthode des moments généralisée et estimateurs Arellano-Bond	29
4.2	Spécification du modèle	30
5	Analyse de régression	32
5.1	Statistiques de test	32
5.1.1	Considérations méthodologiques	32
5.1.2	Tests d'autocorrélation d'Arellano et Bond	34
5.1.3	Tests de Sargan-Hansen et Différence-en-Hansen	34

TABLE DES MATIÈRES

iii

5.2	Résultats principaux	35
5.2.1	Estimation SGMM	35
5.2.2	Autres tests de robustesse	37
6	Conclusion	39
6.1	Recherche empirique : caractéristiques des sites de crowdfunding	40
6.2	Approches théoriques	41
7	Bibliographie	43

Table des figures

1	Interface de Kickstarter.com	46
2	Statistiques de la catégorie Film & vidéo	47
3	Statistiques de la catégorie Musique	48
4	Statistiques de la catégorie Jeux vidéo	49
5	Statistiques de la catégorie Logiciels	50

Liste des tableaux

1	Variables dans Burtch et al. (2011)	15
2	Statistiques descriptives - Film	51
3	Statistiques descriptives - Musique	51
4	Statistiques descriptives - Jeux vidéo	51
5	Statistiques descriptives - Logiciels	51
6	Corrélations inter-temporelles	52
7	Décomposition de la variance	52
8	Résultats - Tests d'autocorrélation	53
9	Résultats - Tests de Sargan-Hansen	54
10	Résultats - analyse de régression - SGMM	55
11	Résultats - analyse de régression - MCO	56
12	Résultats - analyse de régression - SGMM (en retirant les <i>outliers</i>)	57

1 Introduction

Depuis quelques années, nous assistons à l'émergence d'un nouveau mode de financement en ligne : le *crowdfunding*, parfois traduit en français par le terme « financement participatif ». Cet outil, mis en place avec diverses variations par des dizaines de *start-ups* (ProFounder Financial, PeerBackers, Kiva.com, MyMajorCompany, Sellaband.com, Kickstarter.com, etc.) était à l'origine employé surtout pour financer des projets musicaux ou des œuvres caritatives. Aujourd'hui, il est utilisé par des centaines d'entrepreneurs proposant des projets de toutes catégories : musique, cinéma, bande dessinée, jeux vidéos, théâtre, etc. Si la plupart des sites qui offrent des plate-formes de *crowdfunding* tendent toujours vers le financement d'œuvres culturelles, certains permettent également de financer des logiciels ou des biens concrets, comme des synthétiseurs ou des outils professionnels de photographie. Les projets financés sont traditionnellement de petite envergure, mais l'ampleur que prend récemment ce modèle a permis à certaines entreprises d'atteindre un seuil critique de financement beaucoup plus important ; le concepteur de jeux vidéos Tim Schaffer a, par exemple, récolté plus de 3 millions de dollars US pour son dernier projet sur Kickstarter.com.

Ce mémoire entreprend d'étudier le phénomène du financement participatif dans un contexte où les biens numériques présentent de plus en plus des caractéristiques semblables à celles des biens publics. Comme noté précédemment, la plupart des projets entrepris sur les sites de financement participatif concernent des œuvres culturelles qui peuvent aisément être distribuées sous forme numérique : musique, films, logiciels, etc. Avec l'accès généralisé des consommateurs à Internet et la popularité croissante du partage de fichiers (par l'entremise des réseaux *peer-to-peer*¹), ce type de bien prend aujourd'hui des caractéristiques liées à la notion des biens publics ; c'est-à-dire qu'ils deviennent non-rivaux et non-exclusifs. Ces caractéristiques, selon la littérature économique, sont susceptibles de créer chez les consommateurs des comportements de passager clandestin (*resquillage*) – soit de les inciter à profiter du bien sans contribuer à son financement. Le partage de fichiers représente, de facto, un comportement de passager clandestin.

Le financement participatif constitue une forme de financement différente de celle

1. Parfois traduit directement en français par le terme « pair-à-pair » ou encore « poste-à-poste » (recommandé par l'OQLF), il s'agit d'un modèle de réseau informatique permettant à chaque client d'exécuter également des fonctions de serveur. Populaire dès 1999 avec le logiciel *Napster*, leur nature décentralisée permet le partage de fichiers de manière anonyme.

habituellement utilisée pour la production de biens culturels, soit la distribution au détail protégée par un système de propriété intellectuelle. Comme cette dernière méthode est menacée par la prolifération du partage de fichiers, il nous paraît important d'explorer d'autres façons de faire. Nous proposons donc d'effectuer une étude empirique des comportements des usagers d'un site de financement participatif, afin de déterminer si des problèmes de passager clandestin y sont observables. Nous explorons également la littérature concernant les « comportements grégaires »² afin de définir un comportement alternatif, opposé au *resquillage*, que peuvent adopter les usagers. L'objectif ultime de ce travail est de déterminer si, dans le contexte spécifique du financement participatif, l'un ou l'autre de ces comportements est observable, ou alternativement de déterminer lequel de ces deux comportements est prépondérant.

Pour répondre à ces questions, nous avons d'abord effectué une collecte de données sur le site de financement participatif Kickstarter.com. Cette collecte de données s'est échelonnée sur une période de quatre mois pendant lesquels, à chaque jour, à l'aide d'un logiciel conçu sur mesure, nous avons relevé le niveau de financement d'une panoplie de projets. Afin d'obtenir un ensemble de données compatible avec le cadre théorique proposé dans ce mémoire, les projets qui ont été suivis correspondent à la notion de biens culturels exposée ci-dessus, soit des biens pouvant aisément être transférés sous format numérique. La classification des projets sur Kickstarter a été utile en ce sens, chaque projet étant assigné dès le premier jour à une catégorie : musique, cinéma, etc. Pour chacune des catégories regroupant des biens de type culturel, la totalité des projets en cours a été soumise à la collecte de données - il s'agit d'une façon, pour nous, de répondre à la problématique de la sélectivité des données. Les résultats de cette collecte ont ensuite été compilés sous la forme de données de panel, chaque observation correspondant à un projet, à une dimension «temps» (nombre de journées jusqu'à la fin du projet) ainsi qu'à une dimension «financement». Différentes variables, incluant la contribution lors de la dernière journée ainsi que la contribution moyenne, ont été construites à partir de ces données brutes. Finalement, ces variables ont été utilisées afin de bâtir un modèle que nous testons à l'aide d'un estimateur de système GMM.

Nous employons un raisonnement relativement simple afin de déterminer s'il y a resquillage ou grégarisme. Si notre analyse de régression indique qu'une contribution au

2. Comportement où, au moment d'effectuer un choix, un individu tend à reproduire le comportement des autres individus qui ont fait face aux mêmes alternatives avant lui – soit par réciprocité, soit, suivant Banerjee (1992), par manque d'information concernant lesdites alternatives.

temps T-1 entraîne une diminution des contributions au temps T, nous en déduisons que les usagers considèrent que leurs contributions sont substituables. Dans une telle situation, ces derniers attendent que le bien soit financé afin d'en profiter sans payer (puisqu'il s'agit d'un bien numérique), et restreignent leurs propres contributions lorsqu'ils anticipent que le bien sera financé par d'autres usagers. Ce comportement correspond au resquillage. Si, au contraire, une contribution au temps T-1 entraîne une croissance des contributions au temps T, nous considérerons plutôt qu'il existe une tendance au grégarisme, la contribution étant alors un signal de valeur aux autres usagers. Nous incluons, afin d'obtenir des résultats plus précis, une variable représentant les sources externes de popularité d'un projet, ce qui nous permet d'isoler le signal envoyé par les contributions.

Nous obtenons des résultats indiquant que l'hypothèse du resquillage, dans le cas de l'ensemble de projets étudié, est fautive. Une contribution lors d'une période donnée entraîne une augmentation, plutôt qu'une diminution, des contributions lors des périodes subséquentes. Les meilleures spécifications du modèle indiquent que cet impact positif est modéré mais significatif. Nous trouvons par ailleurs que, parmi les variables disponibles, la contribution à la période précédente est celle qui produit l'effet le plus important.

2 Revue de littérature

2.1 Description du phénomène

Il y a peu d'ouvrages consacrés spécifiquement au phénomène du financement participatif, ce dernier ayant connu son essor dans les quelques dernières années seulement. Belleflamme et al. (2010) effectuent une première analyse empirique du phénomène ; avant cet ouvrage, on peut seulement trouver quelques études de cas portant sur des sites spécifiques (par exemple Brabham, 2008). Cette analyse note qu'il s'agit d'un marché jeune et peu développé, présent surtout dans les pays anglo-saxons. L'usage de plate-formes de financement participatif recouvre tout de même de nombreux domaines : journalisme, sport, culture, etc. L'objectif médian des entrepreneurs est de 100 000 euros ; le résultat médian réel est plutôt de 28 500 euros. 80% des sites offrent une sorte de récompense ou de rétribution aux usagers, dépassant donc la notion du simple don. L'objectif de tels projets n'est pas seulement d'obtenir du financement, mais également de générer de la publicité et, du même coup, de sonder le marché afin de connaître la demande potentielle pour le produit.

Schwienbacher et Larralde (2010) approfondissent ces notions. Ils notent entre autre que l'endettement est une source de fonds « stupide » (*stupid money*), c'est-à-dire qu'elle n'aide pas l'entrepreneur à mesurer l'enthousiasme qui existe réellement pour son projet. Le financement participatif permet donc à l'entrepreneur de sonder le public, et, en outre, de profiter de la « sagesse des foules » (Surowiecki, 2004). Ainsi, les consommateurs intéressés par le projet peuvent partager leurs impressions et permettre à l'entrepreneur d'améliorer ou d'ajuster le produit qu'il compte offrir. De plus, contrairement au capital-risque traditionnel, l'entrepreneur profite souvent des conseils de ses investisseurs sans être forcé de partager le contrôle de son projet avec eux. Ceux-ci restent généralement passifs ; rares sont les sites de financement participatif où les investisseurs obtiennent réellement une part dans le projet en question. Les auteurs notent également le créneau où opèrent la grande majorité des sites de *crowdfunding*, soit les entreprises de petite taille ; celles-ci ont beaucoup moins accès aux outils de financement traditionnels, comme le capital-risque et les offres publiques initiales. Le financement participatif vient donc desservir un marché autrefois délaissé, en rassemblant une clientèle qui peut parfois être éparse afin de financer des produits de niche.

Belleflamme et al. (2010) relèvent également des données au sujet des types et

des catégories de projet qui rencontrent le plus de succès. Ils notent que les entreprises qui réussissent à amasser des plus grandes sommes offrent généralement des produits tangibles, plutôt que des services. Au moment de leur analyse, les projets à buts non-lucratifs rencontraient beaucoup plus de succès que les firmes à but lucratif, soit environ 200% du financement que celles-ci réussissent à obtenir. Cela dit, les firmes sont plus nombreuses que les ONG sur les plate-formes de financement participatif, avec 35% contre 16% de l'échantillon ; 45% des données concernent des projets « uniques », lancés par des individus qui ne font partie ni d'une firme ou d'une ONG. 85% des répondants indiquent qu'ils s'adonnent au financement participatif partiellement pour des raisons de publicité, comme mentionné précédemment ; 60% souhaitent également obtenir des conseils de la part des usagers qui investissent dans leur projet. Ceci suggère qu'il existe une forte composante communautaire sur les plate-formes de *crowdfunding*.

De manière générale, nous pouvons donc noter que le financement participatif constitue un phénomène communautaire, social et interactif. Il sert autant à informer les consommateurs potentiels de l'existence de certains projets qu'à lever des fonds. Ce fait semble justifier notre approche, qui vise essentiellement à évaluer l'impact, sur le comportement des usagers, du contenu informationnel que véhicule le site que nous étudions, Kickstarter.com. En prenant compte des études citées, nous pouvons ajouter que le site est relativement innovateur, puisqu'il refuse tout projet caritatif – il impose par ailleurs que tout produit offert soit relativement marchandable, puisque l'auteur du projet doit être en mesure de récompenser ses investisseurs de manière tangible. Nous verrons plus précisément, dans la prochaine section, comment les projets que nous étudions s'apparentent à la littérature des biens publics.

2.2 La plate-forme étudiée : Kickstarter.com

Kickstarter.com, le site de financement participatif que nous étudions, a été fondé le 28 avril 2009 par Perry Chen, Yancey Strickler et Charles Adler, propriétaires de l'entreprise Kickstarter, Inc. Il ne s'agit pas de la première plate-forme de financement participatif, mais elle se distingue par le succès qu'elle a obtenu ainsi que par son mode de fonctionnement particulier. Chaque demande de financement sur Kickstarter doit correspondre à un projet ayant un échéancier fixe, c'est-à-dire un début et une fin - on ne peut demander du financement pour un projet sans objectif final (par exemple, une oeuvre caritative voulant lever des fonds de manière continue ne serait pas admissible).

Les projets concernent donc généralement des projets ayant pour but de produire un bien concret : un disque, un film, etc. Le site est également doté d'une structure de récompenses originale ainsi que d'un mode de financement de type «tout ou rien».

Tous les sites dédiés au financement participatif fonctionnent, à la base, de la même manière. L'entrepreneur qui désire obtenir du financement crée une page personnelle, où il décrit ce qu'il a l'intention de créer ainsi que ses besoins financiers ; les utilisateurs du site peuvent ensuite naviguer sur la page et décider s'ils désirent contribuer au projet. Il existe une caractéristique qui différencie le financement participatif des œuvres caritatives traditionnelles : l'entrepreneur doit promettre une récompense à l'individu qui contribue au financement de son projet. Certains sites, peut-être plus près des modes de financement et d'investissement traditionnels, proposent d'offrir une part ou un dividende à l'utilisateur qui contribue à un projet. La firme que nous étudions, Kickstarter.com, fonctionne d'une manière différente. Sur ce site, les entrepreneurs qui effectuent une demande de financement proposent diverses récompenses liées à l'achèvement du projet – par exemple, dans le cas d'un musicien voulant enregistrer un disque compact, on peut promettre aux participants un exemplaire du disque complété, mais également un mot personnel du musicien, une invitation à un concert, etc. De manière générale, plus la contribution est grande, plus la récompense est importante ; les entrepreneurs peuvent ainsi créer des catégories de récompenses afin de profiter de la plus grande volonté de payer de certains participants. Un système de récompenses à paliers multiples pourrait inclure, pour un projet de disque compact : un mot de remerciement pour les contributeurs ayant donné de 1 à 5 dollars ; une chanson à télécharger pour ceux ayant donné de 5 à 10 dollars ; de 10 à 20 dollars, un exemplaire du disque compact ; pour 20 dollars et plus, un exemplaire du disque compact, plus un billet de concert...

Notons donc qu'à partir d'un certain montant, le contributeur peut effectivement obtenir un exemplaire du produit auquel le financement est destiné. Ceci peut représenter un désavantage pour notre analyse dans la mesure où contribuer peut alors permettre à quelqu'un d'obtenir le produit plus rapidement ou avec plus de facilité qu'à travers le piratage. Le créateur du projet peut aussi faire des efforts supplémentaires afin de déjouer la tentation de resquiller - par exemple, en promettant une édition spéciale ou une dédicace à ses contributeurs, plutôt que la version grand public du produit. La décision de contribuer pourrait, en ce sens, provenir du désir d'obtenir une récompense plutôt que du désir d'obtenir le produit en soi. Il s'agit d'une possibilité dont nous ne pouvons pas tenir compte avec les données recueillies et qui pourrait affecter nos

résultats finaux.

Lors de leur visite sur Kickstarter.com, les usagers peuvent évidemment lire les descriptions des projets telles que fournies par les entrepreneurs, mais ils peuvent également relever certaines autres données intéressantes. Ce sont ces données – soit la somme totale demandée par l’entrepreneur, le niveau de financement déjà fourni par d’autres usagers et le nombre de jours restant avant la fin de la demande – que nous étudierons. Puisqu’elles sont accessibles aux usagers, elles peuvent potentiellement affecter leur comportement. Si ces derniers pensent qu’un projet est en voie d’être suffisamment financé, cela pourrait augmenter la probabilité qu’ils adoptent un comportement de passager clandestin. Notons par ailleurs qu’un projet n’obtient les sommes qui lui ont été promises que s’il atteint son objectif de financement, déterminé par l’entrepreneur au début du processus – et ce, dans les délais prévus (entre 1 et 60 jours, également déterminé par l’entrepreneur). Si, par exemple, un projet a comme objectif d’obtenir 5000\$ en 30 jours, mais qu’il n’obtient que 3500\$, alors le transfert d’argent entre les usagers et l’entrepreneur n’a simplement pas lieu et le projet échoue. C’est ce qu’on appelle un financement «tout ou rien» Tous les sites de financement participatif ne fonctionnent pas nécessairement de cette manière. Par exemple, l’article sur lequel nous nous sommes fondés pour effectuer notre analyse - Burtch et al. (2011) - étudie un site où les projets n’ont pas d’objectifs en tant que tel. Ainsi, les créateurs de projets récoltent la somme qui a été contribué après une période de temps prédéterminée, peu importe la taille de celle-ci. De telles considérations peuvent affecter la décision des usagers de contribuer à un projet ou non, et nous tenterons de mesurer leur impact.

2.3 La relation entre biens numériques et biens publics

La notion de « biens publics », telle qu’on l’entend aujourd’hui, a été développée d’abord par Samuelson (1954) ; elle a ensuite été approfondie par de nombreux auteurs, constituant un point de départ important pour l’étude économique des choix collectifs. Buchanan et Tullock (1958) ont, notamment, contribué de manière importante à la littérature ; leurs ouvrages ont donné naissance à une catégorie d’analyse consacrée à l’étude des choix collectifs dans une société constituée d’individus rationnels (*public choice*). Les biens publics sont définis, dans la littérature, par deux caractéristiques : la non-rivalité et la non-exclusivité. La caractéristique de non-rivalité signifie que la consommation du bien par un individu ne diminue pas la capacité de consommation

des autres individus – par exemple, la consommation d'un programme de télévision par un individu n'empêche pas sa consommation par un autre (du moins, tant que l'espace de visionnement n'est pas saturé). La non-exclusivité représente la difficulté d'empêcher un individu de profiter d'un bien, une fois celui-ci mis en marché ; les ondes radio, par exemple, peuvent être captées par n'importe quel individu doté d'un récepteur. Ces caractéristiques sont, comme le démontrent ces exemples, d'ores et déjà appropriées pour décrire les biens « informationnels » ; elles sont par ailleurs particulièrement bien adaptées à l'étude des biens numériques.

Rayna (2007) fait une analyse en profondeur des biens numériques et de leurs caractéristiques de biens publics. Si, comme noté précédemment, les biens comme la télévision et la radio ont toujours eu certaines caractéristiques de biens publics, les biens numériques contemporains correspondent presque totalement à la définition des biens publics. Cette catégorie de biens contient essentiellement des produits de divertissement – films, musique, jeux, livres – ainsi que certains outils professionnels, liés au développement ou à la bureautique. La considération de ces produits en tant que biens publics est intimement liée au développement des technologies de l'information et à Internet, mais aussi à la plus grande disponibilité de la bande passante, des connections haute-vitesse et de la capacité de stockage (disque dur).

L'état actuel des technologies de l'information permet en effet de transférer les biens numériques directement, à n'importe quel individu doté d'un ordinateur et d'une connexion Internet. De plus, il est possible de créer un nombre infini de copies d'un bien, sans jamais que la qualité du bien ne soit transformée, et sans qu'il n'y ait de coûts supplémentaires à l'achat d'un ordinateur. En termes économiques, ceci fait tendre le coût marginal d'une unité supplémentaire du produit vers zéro, dans un contexte où presque tous ont accès à la technologie nécessaire pour reproduire le bien en question. En bout de ligne, le produit prend des caractéristiques de non-rivalité, puisqu'il est possible pour n'importe qui de le reproduire à l'infini à un coût très bas. Le consommateur peut effectivement « donner » le bien à quelqu'un d'autre tout en le consommant lui-même. Il n'était pas impossible auparavant de créer des copies de films, de cassettes ou de disques compacts, mais de telles copies étaient souvent de qualité inférieure au produit original.

La plus grande différence entre les médias traditionnels et les biens numériques reste, cependant, au niveau de la non-exclusivité. Spécifiquement, la croissance d'In-

ternet représente une forte augmentation de la *difficulté*, pour les producteurs de biens numériques, d'exclure certains consommateurs. Lorsqu'un faussaire doit faire des efforts afin d'obtenir un produit en bonne qualité avant d'en faire une copie, on dit qu'il existe des « coûts de recherche » (*search costs*) élevés. De plus, dans un tel contexte, il est relativement simple d'identifier le faussaire comme étant la « source » unique des copies. On peut donc plus facilement appliquer les lois relatives à la protection de la propriété intellectuelle. La nature décentralisée des réseaux *peer-to-peer* et des sites BitTorrent³ contemporains rend une telle tâche beaucoup plus compliquée, car chaque consommateur devient lui-même une sorte de faussaire. Dès qu'il télécharge le produit, il devient une « source » lui-même, et la diffusion du produit progresse selon une suite géométrique. La difficulté d'appliquer le concept de propriété intellectuelle sur Internet contribue fortement à rendre les biens numériques plus conformes à la notion de bien public.

2.4 Deux types de comportements : resquillage et grégarisme

D'un côté, comme nous l'avons démontré, le phénomène que nous étudions se situe dans un contexte de financement de biens publics. Dans un tel contexte, la littérature économique – et, de manière plus large, celle des choix rationnels – considère que le problème principal sera lié au resquillage, ou « comportement de passager clandestin ». De l'autre côté, les études issues du domaine des « systèmes de l'information » ont, en ce qui concerne les réseaux de propagation de biens numériques par Internet, observé des comportements à l'opposé du resquillage. De nombreux exemples de travaux empiriques provenant de chaque domaine existent ; nous ne nous baserons pas sur ceux-ci spécifiquement⁴ mais leurs résultats peuvent nous donner certaines intuitions.

Selon la logique du *free-rider*, il est dans l'intérêt (au sens économique) d'un individu faisant partie d'un groupe de tenter de profiter des bienfaits de l'action collective sans y sacrifier ni temps ni ressources. On peut imaginer comme exemple un ouvrier profitant d'une hausse de salaire négociée par son syndicat sans qu'il n'ait payé de cotisations, ou, simplement, un individu qui profite des infrastructures publiques (routes,

3. Type spécialisé de protocole *peer-to-peer* particulièrement en vogue aujourd'hui parce qu'il permet de réduire la saturation des connections. Il réduit également la dépendance sur les connections à un usager spécifique en permettant à un fichier d'être téléchargé de diverses sources en même temps.

4. Voir plutôt la section suivante sur la formulation du modèle ; notre exemple principal sera Burtch et al. (2011).

etc.) tout en faisant de l'évasion fiscale. Cette notion a été formalisée par Mancur Olson dans *The Logic of Collective Action* (1965), où le problème de l'action collective au niveau politique est essentiellement représenté comme un problème de coûts diffus et de bienfaits concentrés. Becker (1974), dans la tradition de l'école de Chicago, affirme que le désir de reconnaissance sociale peut constituer une motivation rationnelle pour le financement volontaire de biens publics, alors que certains auteurs comme Roberts (1984) étudient le *crowding out*, ou remplacement, des contributions volontaires par le financement fourni par l'État. Andreoni (1990) développe un modèle qui mélange altruisme et intérêt personnel. Andreoni (1998) examine également un mode de financement similaire à celui que nous étudierons avec Kickstarter, c'est-à-dire où l'obtention des fonds est contingente à l'atteinte d'un objectif spécifique de financement. Cet ouvrage démontre une forte augmentation du resquillage dès le moment où le bien est financé, c'est-à-dire qu'il atteint son objectif. Il devient alors beaucoup plus difficile pour le producteur du bien d'obtenir du financement supplémentaire.

Plusieurs auteurs liés au domaine des systèmes d'information (*information systems*), à la croisée du management et de l'informatique, ont également contribué à la littérature concernant les biens numériques et le concept de biens publics. Asvanund et al. (2004) évaluent l'impact de la taille des réseaux *peer-to-peer* sur la tendance des usagers à adopter des comportements de passager clandestin ; ils trouvent que les réseaux de plus grande taille sont plus souvent sujets au resquillage. D'autres auteurs, par contre, obtiennent à la suite d'études empiriques des résultats à l'opposé des résultats obtenus traditionnellement par la science économique. En effet, la littérature issue des systèmes d'information se concentre essentiellement sur les « comportements grégaires », plutôt que sur le resquillage. Xia et al. (2011), ainsi que Gu et al. (2009), trouvent des exemples de renforcement positif en ce qui concerne les réseaux *peer-to-peer*. C'est-à-dire que les membres de ces réseaux contribuent de manière accrue – en partageant plus de fichiers, ou en contribuant davantage de bande passante – dès lors qu'il perçoivent qu'un nombre plus élevé de fichiers est disponible sur le réseau en question. Les usagers du service agiraient donc selon une certaine norme sociale de réciprocité. Nandi et Rochelandet (2008) opposent explicitement la réciprocité au resquillage, et trouvent le premier concept plus à même d'expliquer les comportements des membres de réseaux *peer-to-peer*.

Il est plausible, comme l'indique Rayna (2007), que ce type de comportement soit dû au coût très peu élevé d'une contribution additionnelle dans le cas des réseaux *peer-*

to-peer. En effet, la contribution d'un individu se limite dans ce cas à rendre disponible certains fichiers sur son ordinateur ; l'utilisateur sacrifie, dans le pire des cas, un peu de bande passante. Borck et al. (2006) étudient cependant un cas où les contributions financières sont bien réelles, et trouvent également que le sens de réciprocité joue un rôle important au moment où l'usager décide de la somme à contribuer. Ces résultats soulèvent cependant quelques questions, comme l'indiquent Burtch et al. (2011) : les usagers sont-ils réellement motivés par la réciprocité, ou font-ils du mimétisme ?

De fait, Banerjee (1992) décrit de tels comportements comme étant une recherche d'information. Les usagers ne connaissent pas vraiment le comportement « juste » à adopter, ou la somme juste à contribuer ; ils adoptent donc le comportement le plus commun parmi les autres usagers, ou encore le comportement qui leur est le plus familier, même jusqu'au point d'ignorer leurs propres préférences. Selon cette logique, l'individu qui choisit le premier, au sein d'un groupe, a une influence prépondérante sur les choix des autres membres du groupe. On suppose que le second individu à faire un choix est indifférent entre une option A et une option B, c'est-à-dire qu'il ne possède pas d'information lui permettant de départager les deux ; il s'en remet donc au choix du premier individu. Le troisième agit de la même sorte avec le choix du deuxième, et ainsi de suite, de telle sorte qu'une option A peut être choisie même si dans l'agrégat on aurait préféré l'option B. Ce genre de raisonnement constitue, après la réciprocité, une seconde alternative aux comportements de passager clandestin.

Certains aspects du grégarisme soulevés par Banerjee sont d'intérêt pour notre analyse. D'abord, la disponibilité d'information permettant de reproduire le comportement d'autrui est une condition préalable du grégarisme. Dans la plupart des cas, cette information n'est pas disponible de manière libre et gratuite aux participants - ceux-ci doivent payer pour y accéder. Dans notre cas cependant, Kickstarter fait office d'intermédiaire en permettant à ses utilisateurs de connaître de nombreuses informations relatives aux niveaux de financement des divers projets. En plus d'avoir un accès facile à la progression du financement des projets, les utilisateurs peuvent également consulter des listes de projets plus populaires - c'est-à-dire, des projets qui obtiennent le plus de contributions lors des derniers jours. La disponibilité de toutes ces différentes données faciliterait et encouragerait, selon le cadre proposé par Banerjee, les comportements grégaires.

Dans le cas étudié, soit Kickstarter, la question de la rationalité d'un comportement grégaire se pose. Les modèles de grégarisme, souvent appliqués au monde financier, font l'hypothèse que les participants cherchent à obtenir des avantages pécuniers en adoptant un tel comportement. Par exemple, un investisseur peut se laisser emporter par l'engouement pour un titre financier quelconque car il croit que les autres investisseurs possèdent de meilleures informations que lui. Dans le cas de Kickstarter, cependant, ce que le participant cherche à déterminer est la qualité des divers projets. Contrairement à un environnement où les participants cherchent à maximiser leurs profits, donc, la sélection d'un projet sur une plate-forme de financement participatif conserve un aspect subjectif. À l'inverse de l'aspect informationnel discuté précédemment, ceci pourrait contribuer à diminuer les comportements grégaires, puisque chaque contributeur potentiel est doté de préférences qui ne peuvent être totalement déterminées par les signaux d'autrui.

Il est difficile de mesurer le désir de réciprocité des individus participant à des projets de financement collectif – il faudrait peut-être sonder chaque participant afin de connaître ses « vraies intentions » au moment de l'investissement. La découverte des motivations psychologiques profondes des individus n'étant pas le but de cette étude, nous trouvons plus approprié d'adopter un cadre d'analyse plus près de celui proposé par Banerjee. Nous considérerons donc que les données obtenues sont simplement dotées d'un contenu informationnel – quel qu'il soit – qui a un impact sur le comportement des usagers des réseaux de financement participatif, la motivation ultime pouvant être de nature soit égoïste ou altruiste.

Notons finalement qu'il existe plusieurs autres facteurs qui peuvent affecter la décision de contribuer. La variété de projets disponibles sur Kickstarter, par exemple, est un facteur duquel nous ne pouvons tenir compte avec les données que nous avons en main. La présence ou l'absence d'alternatives à un projet aura certainement un impact sur sa performance en termes de financement. Cette variable peut aussi être interprétée suivant notre cadre d'analyse : la présence d'un plus grand nombre de projets, par exemple, pourrait amoindrir l'effet des comportements grégaires en diminuant l'attrait de chaque projet individuel. En d'autres mots, un projet populaire qui attirerait normalement des contributions de type grégaire connaîtra un engouement moins important s'il doit rivaliser avec un grand nombre d'autres projets pour l'attention des contributeurs. Ainsi, même si l'hypothèse du grégarisme s'avère correcte, on pourrait néanmoins observer un effet moins important du phénomène dans les cas où moins de

projets existent, et vice versa. Inversement, dans le cas où l'hypothèse du resquillage s'avère correcte, on pourrait imaginer que ce comportement serait accentué par un plus grand nombre de projets, puisque l'acte de financer tous les projets que l'on trouve intéressants deviendrait d'autant plus difficile.

2.5 Resquillage, grégarisme et financement participatif

La formulation de notre modèle s'inspire essentiellement du travail de Burtch et al. (2011), qui effectuent une étude semblable à la nôtre mais pour un secteur différent. Cet ouvrage explore également la question des comportements des usagers de sites de financement participatif – l'analyse effectuée par les auteurs concernant cependant un site de journalisme en-ligne plutôt qu'un site consacré à la production de biens numériques, comme Kickstarter. Les auteurs partent du postulat que des comportements de passager clandestin ainsi que des comportements « grégaires » ont été observés dans la littérature, et qu'il est donc nécessaire d'effectuer une étude empirique afin de déterminer quel de ces deux effets sera le plus important dans le contexte du *crowdfunding*.

2.5.1 Modélisation

Burtch et al. construisent donc un modèle empirique, avec pour objectif de tester l'impact des contributions précédentes – c'est-à-dire lors d'une période $t - 1$ – sur les contributions actuelles (à la période t). Ils commencent en décrivant l'interprétation qu'ils feront de leurs résultats : la direction du coefficient obtenu sur les contributions précédentes indiquera s'ils ont à faire à des comportements de passager clandestin ou à des comportements grégaires. En effet, si le coefficient de cette variable explicative est positif, cela signifie que les contributions précédentes ont un effet de *renforcement* et que l'on a donc principalement à faire à des comportements grégaires ; si, au contraire, son coefficient est négatif, alors il y a principalement un effet de *substitution*⁵ et l'on considère que le resquillage prédomine. La direction du coefficient indique donc aux auteurs sur le type de comportement observé, alors que sa taille les renseigne sur le niveau de signification et l'impact global du dit comportement sur les contributions au moment t .

5. Dans le sens où les individus considèrent que les contributions d'autres usagers se substituent aux leurs, et qu'ils n'ont donc plus besoin de contribuer.

En élaborant plus précisément leur modèle empirique, les auteurs définissent cependant plus de variables explicatives que la seule « contribution précédente ». En fait, l'effet entendu par ce terme est réparti dans deux variables indépendantes, soit la contribution à $t - 1$ et ce que les auteurs appellent la fréquence des contributions (également à $t - 1$). La variable de contribution décrit simplement le montant recueilli par un projet i au cours d'une période de temps t , soit une journée. La *fréquence des contributions*, elle, décrit le montant total d'argent recueilli par un projet divisé par le nombre de jours depuis le début de ce projet. Les auteurs font le postulat que cette variable a un effet sur le comportement de l'utilisateur, ce dernier considérant par exemple qu'un projet qui récolte en moyenne 30\$ par jour aura plus de chance d'aboutir qu'un projet qui récolte 10\$ par jour. Pour les auteurs, il s'agit essentiellement d'incorporer à leur analyse l'information additionnelle que détient l'utilisateur au sujet de la durée du projet.

Les auteurs créent également quelques autres variables explicatives, en plus des données liées aux niveaux de contributions. La variable de visionnement de pages, évaluée à l'instant t , correspond au nombre d'individus ayant visionné la page web d'un projet au cours d'une journée. Cette variable est construite à partir de statistiques additionnelles obtenus auprès du gestionnaire du site web en question par les auteurs. Elle doit servir à mesurer l'intérêt général qu'il existe pour un projet à un temps donné. La variable de budget restant (ou manque à gagner), dont l'impact est plus difficile à prédire, doit servir à tenir compte du niveau de financement restant avant qu'un projet soit complété. Les auteurs supposent qu'un manque à gagner important au niveau de financement pourrait inciter les utilisateurs à « serrer les rangs » et à contribuer davantage, ou, alternativement, il pourrait indiquer que la tâche est insurmontable et donc décourager les contributions. Finalement, les auteurs disposent également de deux variables instrumentales afin d'assurer que leur modèle soit identifié : les références et les tendances de recherche. Ces deux variables servent à mesurer diverses sources exogènes de popularité des projets. La première est construite à partir de données fournies par le gestionnaire du site web, alors que la deuxième provient de données recueillies par le moteur de recherche Google au sujet des divers « mots-clés » présents dans les titres des projets. Dans le cas de ce deuxième instrument, les auteurs ont recueilli manuellement les mots-clés des projets et s'en sont servi afin de recueillir des données de domaine public.

TABLE 1 – Variables dans Burtch et al. (2011)

Variable	Définition
1. contribution_freq	Contributions recueillies divisé par le nombre de jours écoulés.
2. contribution	Nombre de dollars recueillis en une journée.
3. remaining_budget	Part du budget qu'il reste à contribuer afin de compléter le projet.
4. page_views	Nombre de fois que la page d'un projet a été consultée.
5. search_trends	Indique si le projet est recherché sur Google.
6. referrals	Nombre de visionnements provenant de sites extérieurs.

Le problème technique principal de leur analyse, tel que décrit par les auteurs, est l'endogénéité de la contribution à $t - 1$, utilisée comme variable indépendante, par rapport à la contribution à t , qui est la variable dépendante. Cette situation est décrite dans la littérature sous l'appellation de modèle de données de panel dynamique. Tel que décrit dans le chapitre Modélisation, de nombreuses tentatives ont été faites au cours des années afin de trouver un estimateur permettant de pallier à ce problème ; celle qui a eu le plus de succès a mené à l'élaboration de l'estimateur SGMM. Cet estimateur a pour particularité de permettre l'utilisation des divers retards de la variable de contribution en tant que variables instrumentales, permettant ainsi d'effacer l'aspect endogène du retard utilisé comme variable déterminante. Cet estimateur a donc été conçu spécifiquement pour le type de modèle empirique proposé par Burtch et al., et son utilisation s'est imposée. Les spécifications techniques de cet estimateur, importantes aussi pour notre analyse, sont présentées dans une section subséquente.

2.5.2 Données

Les auteurs ont créé un logiciel simple permettant d'extraire les données pertinentes du site web étudié dans leur article. Des données supplémentaires ont été récoltées, comme mentionné précédemment, auprès du gestionnaire du site web, ainsi qu'auprès de Google Analytics, une plate-forme de données publiques entretenue par Google. Leurs données s'étendent de mars 2009 à janvier 2011 mais, puisque Google compile des données sur les termes de recherche seulement depuis le 26 août 2009, les auteurs choisissent d'effectuer leur analyse seulement à partir de cette date. Notons par ailleurs que les données récoltées dans cet article ne concernent que les projets qui ont réussi à obtenir du financement ; ce fait introduit un certain biais de sélectivité avec lequel les

auteurs doivent composer, et auxquels ils répondent en effectuant des manipulations de données supplémentaires. Comme nos données comprennent tous les projets en existence au cours d'une certaine période, nous n'aurons pas à adresser de tels problèmes de sélectivité.

La base de données des auteurs comprend un total de 154 projets lancés. Seulement 100 de ceux-ci ont réussi à compléter le processus de financement et sont donc inclus dans l'analyse finale. Les montants recueillis par ces projets totalisent 66,016.21\$, ces montants étant répartis de manière relativement égale sur l'ensemble des projets (65% des projets reçoivent 52% du financement total). Les auteurs disposent de 5,874 observations pour effectuer leur analyse, les projets ayant été complétés selon divers échéanciers – en moyenne, un projet a été financé en 54,5 jours, avec un minimum de 1 jour et un maximum de 294 jours. Il est évident que, sur le site web étudié par les auteurs, la rigidité de l'échéancier joue beaucoup moins que dans le cas de Kickstarter, où il est impossible de dépasser 90 jours. Les budgets obtenus ne dépassent pas 7,685\$, ce qui est également beaucoup moins que ce le maximum observé sur Kickstarter.

2.5.3 Conclusions

Les auteurs estiment leur modèle à l'aide de la méthode SGMM, en utilisant donc les retards de la variable de contribution comme instruments. Des instruments additionnels - références, tendances de recherche - sont également inclus. Au moment d'estimer les coefficients, les auteurs trouvent que la fréquence des contributions a un effet négatif significatif, avec un coefficient de -0,31 à $p < 0,01$. Cet estimé tient la route avec ou sans l'inclusion de variables instrumentales. Comme les variables sont log-transformées, ce résultat indique qu'une augmentation de 1% de la fréquence des contributions est associée à une baisse de 0,31% de la contribution actuelle. Les auteurs observent donc qu'il existe un processus de substitution chez l'utilisateur, ce qui suggère la présence de comportements rationnels. L'effet est toutefois partiel, puisque le coefficient est inférieur à 1. Il n'y a donc pas un *crowding out* complet des contributions. La contribution à la période précédente produit également un effet significatif, mais contraire à celui observé pour la fréquence des contributions. La simple variable de contribution a un coefficient positif de 0,29 pour $p < 0,01$, ce qui indique qu'une augmentation des contributions à la période précédente encourage les contributions à la période actuelle. Les auteurs proposent que ce résultat indique une certaine inertie des contributions. Le niveau de

contributions actuel est également affecté de manière positive par le nombre de visionnements de la page du projet, ce qui est assez logique. Les autres variables, dont celle représentant le manque à gagner au niveau du financement, n'ont pas d'impact significatif.

Les auteurs terminent en concluant qu'il existe, en ce qui concerne leur analyse, un phénomène de substitution dans le financement participatif, c'est-à-dire que la contribution d'un dollar aura un effet de *crowding out* sur les contributions à venir. Il est difficile de dire s'il s'agit de la bonne conclusion, puisque, comme mentionné dans le paragraphe précédent, il est possible d'observer des effets contraires pour différentes définitions de la « situation » financière d'un projet. Les auteurs notent tout de même que ce phénomène de substitution s'applique plus spécifiquement à cette notion de « contribution frequency », et qu'une plus grande concentration des contributions mène donc, peut-être contre-intuitivement, à des résultats inférieurs. Même si la variable de *contribution* vient contredire partiellement la conclusion vers laquelle tendent les auteurs, le résultat est suffisamment intéressant pour qu'on le retienne afin de le comparer à celui que nous obtiendrons.

Nous croyons que le contexte de l'étude de Burtch et al. est suffisamment semblable au nôtre pour que nous puissions nous inspirer de leur cadre d'analyse, et en particulier du modèle empirique qu'ils ont élaboré, mais que les données que nous étudions sont suffisamment originales pour justifier notre effort. La plate-forme Kickstarter couvre une plus grande variété de produits numériques, la plupart desquels progressent, depuis quelques années, vers une nature qui s'apparente de plus en plus à des biens publics, tout en restant sujets à la propriété intellectuelle. Notre analyse constitue donc une première tentative d'identifier s'il est plausible d'établir un mode de financement différent pour ces produits. De plus, le fonctionnement de Kickstarter diffère en certains points de la plate-forme étudiée par Burtch et al. Les utilisateurs de Kickstarter obtiennent notamment certaines récompenses lorsqu'ils contribuent à un projet et que celui-ci est complété. Finalement, les échéanciers – qui, comme nous l'avons vu, avaient peu d'impact sur la décision de contribuer dans l'article de Burtch et al. – sont plus rigides et restreints sur Kickstarter, ce qui ajoute un élément de stratégie supplémentaire que doivent tenir en compte à la fois les utilisateurs et les propriétaires des projets. Sans tenter de mesurer spécifiquement les effets de ces diverses considérations, nous considérons néanmoins qu'elles modifient l'environnement des contributeurs de Kickstarter par rapport à ceux étudiés par Burtch et al.

3 Base de données

3.1 Collecte des données

Nous avons effectué, pour vérifier les hypothèses présenté dans notre modèle, une collecte de données à partir du site web de Kickstarter. Ce site est organisé, comme mentionné précédemment, en une douzaine de catégories permettant aux auteurs ainsi qu'aux usagers du site de parcourir les projets selon la « discipline » à laquelle ils appartiennent. Ainsi, les catégories sont énumérées dans un menu à la droite de la page principale : *Art, Danse, Mode, etc.* Le site permet également la navigation selon certains critères, comme la popularité, le financement recueilli, etc. Comme nous proposons d'étudier les projets les plus vulnérables au resquillage, nous avons isolé les catégories *Film & vidéo, Musique, Jeux vidéos* et *Logiciels*. Lorsque l'utilisateur clique sur l'une de ces catégories, il atterrit sur une page contenant une quinzaine de projets, résumés par une boîte contenant leur titre, une image, le pourcentage de complétion du financement, le nombre de dollars recueillis et le nombre de jours jusqu'à l'échéance du projet. Il est possible d'accéder à plus de projets en cliquant sur le lien « Page suivante », ce qui affiche quinze autres projets, etc. Nous avons choisi de recueillir les données disponibles sur ces pages. Afin d'obtenir les données mentionnées – nom, pourcentage de complétion, dollars recueillis et jours jusqu'à l'échéance – nous avons choisi de procéder par une technique dénommée *web scraping*. Cette technique permet d'extraire n'importe quelle information présente dans le code HTML d'une page web, pourvu que celle-ci ne soit pas générée de manière dynamique. Après évaluation des outils disponibles pour compléter cette tâche, nous avons choisi d'utiliser le langage informatique *Python* ainsi que le logiciel *Scrapy*, ce dernier étant conçu précisément pour faciliter le *web scraping*. Nous avons, de plus, utilisé le plug-in *Firebug* du navigateur internet *Firefox* afin de pouvoir explorer le code HTML de Kickstarter et d'identifier la location précise des informations que nous recherchions. Le programme de recherche que nous avons écrit à la suite de ce processus a été mis en marche à chaque jour, les données recueillies étant placées de manière automatique dans un fichier CSV. Ces données ont ensuite été triées et nettoyées. Le programme utilisé n'étant pas tout à fait au point, certaines données manquaient occasionnellement à l'appel, et l'ensemble de données utilisé pour l'analyse peut donc être considéré « imbalancé ». La technique d'estimation que nous utilisons, heureusement, permet d'effectuer une analyse de régression même avec certaines données manquantes.

Rappelons que nous avons également proposé d'utiliser la popularité des projets, telle que comptabilisée par la popularité sur Facebook, en tant qu'instrument exogène (ou du moins, aussi exogène qu'il est possible de trouver). Cette information, bien que clairement visible sur Kickstarter, n'était pas disponible dans le code HTML du site. Elle était plutôt générée de manière dynamique par du code JavaScript intégré à même le code HTML de base. Ceci a posé problème, vu l'incapacité de notre programme à lire le contenu généré de manière dynamique. Ainsi nous avons décidé d'adopter une méthode alternative, soit l'automatisation d'un navigateur internet. Cette tâche, bien qu'également effectuée en langage *Python*, a nécessité l'usage d'un programme différent. Nous avons cette fois opté pour le programme *Selenium*, qui permet à la fois d'automatiser un navigateur internet et d'extraire des informations des pages web ainsi générées. Avec cette technique, le navigateur procède à la page désirée, attend que le contenu dynamique soit généré, et extrait ensuite le dit contenu à un fichier CSV. Cette méthode a comme désavantage d'être passablement plus lente que le *web scraping* (quelques heures contre quelques minutes), en plus de nécessiter une vigilance constante vu le grand nombre de données à recueillir et la tendance du programme à s'éteindre lors du chargement de certaines pages. Nous n'avons donc fait qu'un seul passage afin de récolter les données – notre instrument représente donc la popularité d'un projet à un seul moment dans le temps, soit au moment où le projet prend fin.

3.2 Statistiques descriptives

Au total, 234 421 observations ont été collectées lors de l'exercice, et ce sur une période de quatre mois (123 jours) allant du 29 février au 1er juillet 2012. Ces données sont réparties dans quatre catégories représentant des types de produits différents, tel que décidé par les administrateurs de Kickstarter : *Film & vidéo*, *musique*, *jeux vidéo* et *logiciels*. Les deux dernières catégories sont, par ailleurs, d'une taille beaucoup moins importante que les premières – les catégories *Film & vidéo* et *Musique* contiennent respectivement 110 823 et 110 825 observations, alors que la catégorie *Jeux vidéo* en contient 11 068 ; la catégorie *Logiciels* n'en contient que 1 645. Ces données représentent près de 10 000 projets lancés sur Kickstarter au cours de la période étudiée. Nous ne possédons pas l'intégralité des données pour chacun de ces projets. Il manque, par exemple, une partie des données pour les projets présents au tout début de la collecte, ainsi qu'une autre partie pour ceux incomplets à la toute fin de la collecte. Plusieurs projets n'ont pas réussi à recueillir de fonds, mais la plupart ont obtenu la promesse

de certaines sommes, aussi minimales soient-elles. Dans cette section, nous effectuons un survol des données recueillies dans les quatre catégories différentes. Nous avons, par ailleurs, construit trois variables dichotomiques pour mieux comprendre ces données. Une variable d'achèvement indiquant si le projet en question a été comptabilisé du début à la fin ; une variable indiquant s'il a été un succès (c'est-à-dire, s'il a obtenu 100% ou plus de son objectif de financement) ; et une variable servant à identifier la dernière journée où le projet était en financement. Cette dernière variable nous permet de regrouper toutes les données de fin-de-projet afin de mieux connaître les résultats finaux typiques obtenus sur Kickstarter. Des variables d'objectif et de limite de temps ont également été construites à partir des données disponibles.

3.2.1 Film & vidéo

La catégorie *Film & vidéo* contient, outre les projets n'ayant réussi à attirer aucun financement quel qu'il soit, 95 404 observations. Ces observations sont réparties sur 3 914 projets ; nous excluons ici 556 projets qui n'ont attiré aucun financement. De ce nombre, 1 470 ont réussi à atteindre leur objectif de financement, soit une proportion d'environ 37%. Remarquons d'abord que la somme maximale atteinte par un projet dans cette catégorie est de 322 022\$. Ce projet a été un succès, puisqu'il a obtenu 5,36 fois son objectif de financement de 60 000\$; il s'agit du projet le plus important de cette catégorie ayant été financé lors de la période étudiée⁶. La plupart des autres projets ont obtenu des sommes bien moindres, soit en moyenne 6 433\$ à la conclusion de la période de financement. Ce chiffre s'élève à 11 692\$ pour ceux ayant atteint leur objectif. En terme de pourcentages, le projet moyen obtient 0,63% de son objectif ; le projet ayant le mieux performé a obtenu plus de 26 fois le montant demandé, soit 157 485\$ pour un objectif de 6 000\$⁷. Les projets réussis ont en moyenne un plus petit objectif que ceux qui ont échoué, soit 9 316\$ contre 20 058\$ - il s'agit d'une différence assez importante qui semble confirmer l'intuition que les plus petits projets ont une meilleure chance de survie. Notons que les écarts-type sont relativement importants en ce qui concerne les montants de dollars recueillis – 15 764\$ en général, 22 856\$ pour les projets réussis – ce qui indique que les observations sont plutôt dispersées et qu'il existe donc une vaste gamme de projets. On remarque également une dispersion des données

6. Il s'agit de « BronyCon : The Documentary » un film documentaire long-métrage au sujet d'un congrès d'amateurs de la série télévisée *My Little Pony : Friendship is Magic*.

7. « Tropes and Women in Video Games », un projet de documentaire explorant les stéréotypes féminins dans les jeux vidéo.

en ce qui concerne les objectifs de financement, mais dans ce cas les écarts-types sont moindres lorsqu'il s'agit de projets réussis (15 992\$ contre 41 852\$), ce qui indique que ceux-ci sont regroupés d'une manière passablement plus dense.

Le plus long délai qu'un auteur peut attribuer à son projet sur Kickstarter est de 60 jours (cette limite a récemment été réduite – elle était auparavant de 90 jours). 202 projets, soit environ 5%, ont exactement un délai de 60 jours. Le projet moyen a un délai de 30 jours accompagné d'un écart-type de 13 jours. Les projets réussis ont un délai moyen quasiment identique, soit de 29 jours, avec un écart-type de 12 jours.

Voir Figure 2 en annexe.

Le tableau précédent nous permet d'explorer davantage les relations entre les variables. Notons premièrement que les points des graphiques représentent les projets en fin de financement ; chaque graphique contient donc 3 914 points. Dû au grand nombre d'observations ainsi qu'à l'existence de certains *outliers* importants, la direction de la droite de régression n'est pas toujours visible à l'oeil nu. La relation entre les variables de limite de temps et de pourcentage de financement accumulé, par exemple, semble très faiblement négative ; en réalité, la plupart des valeurs de pourcentages se situent sous la barre des 500%, et la droite passe de 80% pour des délais très courts à 45% pour les projets de plus de 60 jours. La relation entre la limite de temps et les dollars représente sensiblement la même chose, et a une pente semblable. On remarque également que les projets avec de plus grands objectifs s'accordent généralement des délais un peu plus longs, allant jusqu'à une moyenne d'environ 30 000\$ pour les projets de 60 jours.

Les autres relations sont plus évidentes et plus informatives. Ainsi on voit que les projets avec de plus grands objectifs atteignent la fin de leur financement avec des pourcentages moindres. Cette relation est cependant balancée par celle entre les objectifs et les dollars, qui indique que les projets avec de plus grands objectifs obtiennent certainement des plus grandes sommes d'argent. Avec les informations obtenues précédemment, nous avons une idée du fonctionnement des projets de différentes tailles sur Kickstarter : les projets avec de plus grands objectifs obtiennent des promesses plus importantes en terme de dollars, mais ils ont tout de même plus de difficulté à progresser vers leur objectif et à l'atteindre.

3.2.2 Musique

Les données contenues dans la catégorie *Musique* ressemblent en plusieurs points

à celles de la catégorie précédente. Elle contient 98 773 observations dans 4020 projets ; 700 projets n'ayant obtenu aucun financement sont exclus. Notons premièrement que les sommes impliquées dans cette catégorie sont moindres, comparativement à celles de la catégorie précédente ; et, peut-être par conséquence, les succès sont plus nombreux. Les projets récoltent en moyenne 3 638\$ à la fin de la période de financement, ou 5 792\$ pour les projets réussis. Ces chiffres sont d'environ 50% inférieurs à ceux obtenus pour des projets de la catégorie *Film & vidéo*. De fait, on compte 1 904 réussites et 2 116 échecs, soit un ratio de réussite de 47%, supérieur de 10% à celui de la catégorie précédente. Malgré ces moyennes inférieures, le projet le plus important de cette catégorie a obtenu un financement de 1 089 032\$⁸, *largement supérieur au plus grand projet de la catégorie Film & vidéo*. Ce projet a obtenu 10,89 fois son objectif de financement (qui était donc de 100 000\$), et ce en seulement 32 jours. Les projets réussis démontrent par ailleurs un écart-type plus large que la moyenne dans cette catégorie ; somme toute cependant, les écarts-type (et donc la dispersion des observations) ressemblent largement à ceux de la catégorie précédente.

En termes d'objectifs, les sommes demandées sont aussi généralement plus petites que celles de la catégorie précédente : le projet moyen a un objectif de 5 710\$, avec une moyenne de seulement 4 238\$ pour les projets réussis et de 7 036\$ pour les projets échoués. Ici, la relation est moins importante mais elle tient quand même – les projets avec un objectif plus élevé échouent plus souvent que ceux avec un objectif moindre. La moyenne des projets pour le pourcentage de complétion du financement est de 0,78%, nettement plus élevée que pour la catégorie recouvrant les films et vidéos. Les observations de pourcentage sont relativement denses, et les *outliers* les plus frappants dans cette catégorie – outre le projet ayant obtenu plus d'un million de dollars – représentent des projets avec de très petits objectifs (5 à 100\$ pour ceux ayant atteint plus de 1000% de complétion).

Finalement, 224 projets de la catégorie *Musique* atteignent la limite de temps de 60 jours. La moyenne se situe, comme pour la catégorie précédente, à 30 jours ; les projets réussis sont ici aussi légèrement moins longs que les projets échoués, avec une moyenne de 28 jours contre 32.

Voir Figure 3 en annexe.

8. « Amanda Palmer : New Record, Art Book & Tour » ; il s'agit d'un projet avec plusieurs facettes.

Les données présentées dans ces graphiques sont organisées de manière très semblable à celles de la catégorie *Film & vidéo*. Parmi les différences principales, on note que les observations liées à la variable de dollars sont regroupées plus densément autour de plus petites valeurs, ce qui change passablement les pentes des droites mettant cette variable en relation avec les objectifs et les pourcentages (mais pas leur direction). Notons également que la forme plus écrasée des nuages de points est également influencée par l'*outlier* extrême qui a récolté plus de 1 000 000 de dollars. Si on prend en compte ces quelques variations, les conclusions que l'on peut tirer de ces graphiques sont en tous points similaires à celles de la catégorie précédente.

3.2.3 Jeux vidéo

Cette catégorie est considérablement plus petite que les deux premières – en excluant les 25 projets n'ayant reçu aucune attention, il nous reste 9624 observations réparties dans 492 projets. Le ratio de réussite est plus faible que dans les autres catégories, soit 119 réussites et 373 échecs pour un ratio de 24%. Bien qu'il y ait considérablement moins d'observations dans cette catégorie, certains projets ont tout de même obtenu des sommes importantes. On note un financement maximal de 2 947 771\$ pour un projet de grande envergure (objectif de 1 000 000\$). Nous ne possédons pas la totalité des données pour ce projet puisqu'il a débuté avant le début de notre collecte de données ; cependant, le deuxième projet le plus important – pour lequel nous avons toutes les observations – a obtenu un financement assez proche, avec 2 830 110\$ (3,14 fois le montant souhaité). Si les jeux vidéo sont moins nombreux sur Kickstarter, ces deux projets démontrent néanmoins une forte présence des amateurs sur le site, et la possibilité d'y financer des jeux d'une envergure assez importante⁹ Entraînée entre autres par ces deux projets, la moyenne de financement pour cette catégorie est de 128 001\$, avec un écart-type de 420 461\$, ce qui dépasse de loin les statistiques équivalentes pour les deux catégories précédentes. La moyenne pour les projets ayant échoué est moins extrême, à 8 717\$. Cette catégorie semble donc favoriser quelque peu les « méga-projets » - moins de réussites, mais des résultats plus éclatants pour les projets qui atteignent leur objectif.

Les objectifs en tant que tel sont également considérablement plus élevés dans cette catégorie, soit 52 837\$ pour les projets réussis et 54 551\$ pour ceux qui ont échoué.

9. Il s'agit respectivement de « Double Fine Adventure », un projet du concepteur de jeux vidéos Tim Schaffer, et d'un *remake* de « Wasteland », jeu culte des années 1980 se déroulant dans un environnement post-apocalyptique.

Remarquons cependant que les écarts-type sont de plus de 100 000\$, ce qui indique une dispersion relativement importante des objectifs. Ici aussi, la durée moyenne d'un projet est d'environ 30 jours; nous pouvons constater qu'il y a une distribution assez égale et identique des projets sur Kickstarter. 15 projets ont la durée la plus longue, soit 60 jours.

Voir Figure 4 en annexe.

Ces tableaux nous indiquent, encore une fois, que les données semblent être organisées à peu près de la même façon que dans les autres catégories. Une différence peut nous intéresser : la relation entre les pourcentages et les objectifs est moins fortement négative que pour les deux catégories précédentes, ce qui semble confirmer que les projets de plus grande taille sont comparativement favorisés dans celle-ci. Ces projets semblent, à tout le moins, être à peu près aussi aptes que les petits projets à obtenir de bonnes proportions du financement qui leur est nécessaire. Notons aussi que la pente de la droite pour les objectifs et les limites de temps illustre une relation positive plus forte entre ces deux variables dans cette catégorie – il est donc plus évident ici que les projets avec de grands besoins de financement s'accordent des délais plus longs.

3.2.4 Logiciel

Cette dernière catégorie est de loin la plus petite des quatre que nous avons étudiées. Elle contient 1182 observations réparties sur seulement 54 projets. Seulement 12 de ceux-ci ont été des réussites, ou 22% - c'est le ratio de réussite le plus bas des quatre catégories, tout en étant relativement proche de celui de la catégorie *Jeux vidéo*. Le projet avec le financement le plus important a obtenu des sommes totalisant 302 373\$¹⁰. Son objectif était de 200 000\$; les projets d'envergure (relativement) importante sont donc aussi présents dans cette catégorie malgré son faible volume. Chose intéressante, la moyenne en dollars promis par les usagers pour les projets réussis est de 41 527\$, alors que celle pour les projets échoués est de 1 937\$; pourtant, les objectifs de financement sont essentiellement les mêmes, soit 25 819\$ et 22 745\$ respectivement. Les usagers dans cette catégorie ont donc tendance à converger vers certains projets spécifiques en délaissant presque totalement les autres. On observe cependant, comme d'habitude, des écarts-type élevés en ce qui concerne les projets réussis – la moyenne est donc certainement tirée vers le haut par quelques projets dominants.

10. « Light Table », un nouvel environnement de développement intégré.

Nous pouvons noter quelques autres statistiques intéressantes : les projets qui réussissent obtiennent en moyenne 2,09 fois leur objectif de financement (un projet a obtenu 7,89 fois son objectif de 3 000\$); le plus petit projet à être financé a obtenu 570\$; et la durée moyenne des projets se situe aussi, pour cette catégorie, autour de 30 jours.

Voir Figure 5 en annexe.

Les graphiques de cette catégorie sont caractérisés par leurs intervalles de confiance très larges, dû petit nombre d'observations. Néanmoins, ils ont à peu près la même forme générale que les graphiques des autres catégories. Comme pour la catégorie *Jeux vidéo*, la relation positive entre la taille des objectifs et la durée des projets est mise en évidence. On reconnaît aussi distinctement une relation positive entre les sommes récoltées en termes de dollars et la durée des projets, alors que la droite représentant cette relation était plutôt plate dans les trois autres cas.

3.2.5 Autocorrélation et variance

Comme nous l'avons noté précédemment, Clark et al. (2009) calculent les corrélations inter-temporelles des variables explicatives afin de justifier l'utilisation de l'estimateur de système GMM. Rappelons que cet estimateur est particulièrement approprié dans le cas où certaines variables ressemblent à des marches aléatoires, c'est-à-dire que la réalisation de la valeur à T dépend de la valeur à T-1. Dans ce cas, l'estimateur DGMM est inefficace et il est préférable d'y ajouter des informations additionnelles, ce qui le transforme en SGMM. Nous calculons donc les auto-corrélations des variables *contribution* et *contribution_frequency* pour chaque catégorie de projet, et obtenons les résultats suivants :

Voir Tableau 6 en annexe.

Dans chacun des cas, notons que la fréquence des contributions est fortement corrélée avec son propre lag, ce qui n'est pas surprenant puisqu'elle constitue une sorte de moyenne mobile des contributions. En ce qui concerne la variable de contribution, elle est toujours corrélée de manière positive avec son lag, mais avec une importance plus ou moins forte dépendant de la catégorie. Les contributions pour les catégories musique et jeux vidéo sont fortement corrélées, alors que celles des catégories film et logiciel le sont considérablement moins. Comme toutes les catégories sont dotées de variables à faible

variance inter-temporelle, nous considérons que le choix de l'estimateur SGMM est correct. Cependant, l'importance moindre des auto-corrélations pour certaines catégories est intéressante à noter. Par exemple, la catégorie des films présente une variation de ses contributions, dans le temps, beaucoup plus marquée.

Nous présentons également un tableau de décomposition des variances pour chacune des catégories. Celui-ci décompose les écarts-types en catégories *within* et *between*, qui représentent respectivement les variations à l'intérieur des projets – soit la variation inter-temporelle – et les variations entre les projets – soit la variation sectionnelle. Clark et al. (2009) notent que l'estimateur SGMM se fie sur la présence de variations inter-temporelles pour fonctionner ; pour notre cas, nous observons que ce type de variation est présent pour chacune des catégories. Pour la variable de contribution, les écarts-types *within* sont tous plus importants que les écarts-types *between* ; pour les fréquences de contribution, les résultats sont partagés, mais la variation inter-temporelle reste comparativement importante dans chaque catégorie. Finalement, en ce qui concerne la variable de budget, la variation sectionnelle est largement plus importante, ce qui est logique – les sommes varient plus entre les projets qu'entre les différentes périodes d'un même projet. Notons également que les catégories Film et Musique présentent des données relativement semblables, alors que la catégorie Jeux vidéo présente une variation beaucoup plus importante à tous les niveaux.

Voir Tableau 7 en annexe.

4 Modélisation

4.1 Technique d'estimation

4.1.1 Données de panel dynamiques

Tel que discuté dans la revue de littérature, l'analyse que nous proposons (avec une même variable, à des temps différents, utilisée à la fois comme variable déterminante et comme variable déterminée) correspond à un type de modèle empirique nommé modèle de données de panel dynamique. Le problème principal d'un tel modèle est l'endogénéité de la variable déterminante - soit, dans notre cas, la variable de contribution au temps $t - 1$ - et de la variable déterminée, soit la contribution au temps t . Les modèles de données de panel statiques, qui ne sont pas soumis à une telle problématique, peuvent (en l'absence d'autres problèmes) être estimés par MCO - c'est ce qu'on appelle l'estimateur *pooled*. Les données de panel dynamiques, cependant, présentent des particularités, dues à l'endogénéité, qui empêchent l'estimation par MCO. Comme l'indique Wooldridge (2002), le problème d'endogénéité des modèles dynamiques est associé conceptuellement aux problèmes d'effets non-observés. Les effets non-observés se rapportent généralement à des caractéristiques propres aux individus que l'on suit dans le temps. Ces caractéristiques varient donc entre les individus mais restent fixes à travers toutes les périodes. Plus précisément, en ce qui concerne notre analyse, les caractéristiques spécifiques aux projets individuels, qui affectent la récolte de financement mais ne sont pas directement incluses dans notre modèle, peuvent être considérées comme des effets non-observés. Tel que mentionné par Wooldridge, des solutions aux problèmes d'effets non-observés existent, mais celles-ci sont inadaptées dans le cas des données de panel dynamiques. Comme il est donc impossible de résoudre les divers problèmes d'endogénéité avec les outils traditionnels, nous devons nous tourner vers l'utilisation de variables instrumentales. Pirotte (2011) cite premièrement Balestra et Nerlove (1966) qui, procédant par analogie avec les modèles d'équations simultanées, adoptent premiers l'approche par variable instrumentale afin d'estimer un modèle dynamique.

4.1.2 Développement de l'approche instrumentale pour modèles dynamiques

Balestra et Nerlove (1966) proposent un estimateur par variables instrumentales en raisonnant que, s'il est possible de trouver un instrument valide pour la variable

dépendante retardée, alors il sera possible d'estimer le modèle correctement – tant et aussi longtemps que cet instrument respecte les conditions d'un estimateur à effets aléatoires, soit l'exogénéité de l'instrument par rapport à l'effet non-observé ainsi qu'au terme d'erreur (exogénéité stricte). Pirotte note que ce type d'estimateur est peu employé puisque, dans les faits, il est difficile de trouver un instrument qui soit à la fois valide et non-corrélé avec l'effet non-observé. Afin de résoudre ce problème, Anderson et Hsiao (1982) proposent d'éliminer l'influence des effets non-observés en raisonnant plutôt avec un modèle en première différence. Dans ce cas, le modèle est conçu de la manière suivante :

$$\begin{aligned} y_{it} - y_{i,t-1} &= \beta_1(y_{i,t-1} - y_{i,t-2}) + \beta_2(x_{it} - x_{i,t-1}) + (u_{it} - u_{i,t-1}) \\ \Delta y_{it} &= \beta_1 \Delta y_{i,t-1} + \beta_2 \Delta x_{it} + \Delta u_{it} \end{aligned} \quad (1)$$

Bien qu'elle permette d'éviter le problème que rencontre l'estimateur Balestra-Nerlove, la méthode de Anderson et Hsiao nécessite tout de même l'utilisation de variables instrumentales. Comme l'indique Pirotte, il existe « une autre corrélation entre la variable $\Delta y_{i,t-1}$ et la perturbation $\Delta u_{i,t}$ par l'intermédiaire de la perturbation $u_{i,t-1}$. En effet, il est évident que $\Delta y_{i,t-1}$ est fonction de $\Delta u_{i,t-1}$, donc de $u_{i,t-1}$. » À ce moment, les auteurs proposent deux types d'instruments qui sont chacun corrélés avec $\Delta y_{i,t-1}$ mais pas avec $\Delta u_{i,t}$ – soit l'instrument en différence, $\Delta y_{i,t-2}$, et l'instrument en niveau, $y_{i,t-2}$. Lorsqu'il est question de maximiser le nombre de données disponibles pour l'estimation, Pirotte et Roodman notent chacun qu'il est préférable d'employer l'instrument en niveau puisque celui-ci est disponible à $t = 3$ alors que l'instrument en différence n'est disponible qu'à partir de $t = 4$ (le premier instrument disponible dans ce cas étant $y_{i,2} - y_{i,1} = \Delta y_{i,2}$). Nous avons noté qu'il n'était pas réaliste, en ce qui concerne les modèles dynamiques, d'accepter l'hypothèse d'exogénéité stricte. En effet, pour ce type de modèle, le terme d'erreur est – par définition – incorporé aux valeurs explicatives à partir de $x_{i,t+1}$. Wooldridge propose donc simplement d'abandonner une partie de l'hypothèse, pour ne conserver que l'exogénéité séquentielle – c'est-à-dire que nous permettons à $u_{i,t}$ d'être corrélé avec les valeurs futures de la variable explicative, soit $(x_{i,t+1}, \dots, x_{i,T})$. Cette condition s'exprime donc de la manière suivante :

$$E(u_{it} | x_{it}, x_{i,t-1}, \dots, x_{i1}, c_i) = 0 \quad (2)$$

C'est cette condition d'orthogonalité plus faible qui nous permet, en l'absence de variables instrumentales adéquates, de proposer l'utilisation comme instrument de la variable dépendante elle-même – comme le font Anderson et Hsiao avec leurs instruments en différence et en niveau. À l'aide de ces instruments, il est donc possible de procéder par 2SLS (moindres carrés en deux étapes) afin d'estimer le modèle.

4.1.3 Méthode des moments généralisée et estimateurs Arellano-Bond

Si l'estimateur de Anderson et Hsiao est consistant, il est toutefois relativement imprécis pour diverses raisons. Pirotte cite notamment l'utilisation d'un nombre inadéquat de variables instrumentales, l'autocorrélation des termes d'erreurs ainsi que la corrélation potentiellement faible entre la variable endogène retardée et les instruments employés. Arellano et Bond (1991) ont donc entrepris la tâche de le rendre plus efficace. Comme Roodman le souligne, il serait possible d'améliorer l'efficacité de l'estimateur Anderson-Hsiao en utilisant davantage de valeurs retardées de la variable dépendante en tant qu'instruments. Cette approche cause cependant une perte d'information, puisqu'il faut alors éliminer les observations pour lesquelles les valeurs retardées en question sont absentes ; ainsi, plus on ajoute des instruments, plus on risque de perdre des observations. Arellano et Bond, suivant Holtz-Eakin, Newey et Rosen (1988), proposent de remplacer le vecteur d'instruments utilisé pour un 2SLS par une matrice d'instruments, effectuant ainsi la transition vers la méthode des moments généralisée (GMM). Ceci permet d'utiliser toutes les valeurs retardées (lags) valides de la variable en tant qu'instruments, sans pour autant perdre d'observations. Arellano et Bond fournissent également une version robuste de l'estimateur ainsi développé, qui permet de procéder même lorsque l'hypothèse – généralement non-réaliste – de l'homoskedasticité des termes d'erreur n'est pas valide. Avec ces modifications, nous arrivons à l'estimateur différence GMM (DGMM) de Arellano et Bond. Cet estimateur a ensuite été développé davantage par Arellano et Bover (1995) ainsi que par Blundell et Bond (1998).

Blundell et Bond soulignent un problème avec l'estimateur DGMM : lorsque la variable dépendante ressemble à une marche aléatoire, « les instruments en niveaux sont faiblement corrélés avec les variables du modèle en différence première » (Pirotte). Donc, même en décuplant le nombre d'instruments tel que recommandé par Arellano et Bond, on obtient finalement une estimation assez faible. Afin de pallier à ce problème, Blundell et Bond développent une approche proposée d'abord par Arellano et Bover

(1995). Alors que l'approche Arellano-Bond originale estime un modèle en différence avec des instruments en niveaux, Arellano et Bover proposent plutôt de transformer les instruments eux-mêmes afin de les rendre exogènes aux effets fixes – soit en prenant la première différence, soit en retirant la moyenne de toutes les observations futures disponibles (méthode de « l'écart à la moyenne » ou orthogonal deviation). Ces instruments sont ensuite utilisés pour estimer le modèle en niveaux. En bout de ligne, l'estimateur « système » de Blundell et Bond fonctionne en dédoublant l'ensemble original de données, afin de produire un nouvel ensemble contenant à la fois les données originales et les données transformées (Roodman). On obtient donc un modèle « empilé » de la manière suivante :

$$\begin{pmatrix} \Delta y \\ y \end{pmatrix} = \beta_1 \begin{pmatrix} \Delta y_{-1} \\ y_{-1} \end{pmatrix} + \beta_2 \begin{pmatrix} \Delta X \\ X \end{pmatrix} + \begin{pmatrix} \Delta u \\ \varepsilon \end{pmatrix} \quad (3)$$

On estime ensuite ce modèle selon la méthode des moments généralisée, en utilisant pour la partie supérieure les instruments de type Arellano-Bond et pour la partie inférieure les instruments de type Arellano-Bover. On arrive de cette manière à l'estimateur de système GMM (SGMM) de Blundell et Bond – c'est cet estimateur que nous utiliserons afin d'estimer notre modèle.

4.2 Spécification du modèle

Les données que nous récoltons ont largement les mêmes caractéristiques que celles amassées par Burtch et al. Nous proposons donc d'adapter leur modèle à notre base de données. Nous récupérons également leur méthode d'estimation, qui semble la plus apte à tenir compte de certains problèmes qu'il serait possible de rencontrer dans une base de données telle que la nôtre (grand N , risques d'auto-corrélation dû à un T aussi assez grand).

Étant donné la composition de notre base de données, nous pouvons développer sensiblement les mêmes variables que l'ont fait Burtch et al. Nous disposons de suffisamment d'information pour identifier la variable de *contributions* – pour un projet donné, la différence entre son financement à une période t et son financement à $t - 1$ – ainsi que celle de fréquence des contributions, soit, pour chaque période donnée,

la totalité du financement accumulé divisé par le nombre de jours depuis le début du projet. Nous avons également les données nécessaires pour créer la variable de manque à gagner pour chaque projet, à chaque période. Malheureusement, les données au sujet du visionnement des pages de projets par les utilisateurs ne nous sont pas disponibles, tout comme celles concernant les références à partir d'autres sites web ou celles des tendances recherche Google.

Les deux variables auxquelles nous n'avons pas accès permettaient de tenir compte de sources exogènes de popularité des divers projets. Nous considérons que l'absence de ces variables peut être partiellement palliée par l'utilisation d'une variable supplémentaire dans notre modèle. Nous proposons donc d'utiliser, en remplacement, un indicateur de la popularité des divers projets créé par les gestionnaires du site Kickstarter et observable par le public. Il est possible, par exemple, d'obtenir de la page de chaque projet – une fois celui-ci terminé – le nombre de mentions sur Facebook du projet (« likes »). Ceci pourrait nous servir d'indicateur de popularité tout en demeurant un facteur externe au financement. Avec ces quelques modifications, nous obtenons donc un modèle assez semblable à celui mentionné précédemment, soit :

$$\begin{aligned} contribution_{it} = & \beta_0 + \beta_1 * contribution_frequency_{i,t-1} + \\ & \beta_2 * contribution_{i,t-1} + \beta_3 * fb_like_i + \\ & \beta_4 * remaining_budget_{it} + \alpha_i + \varepsilon_{it} \end{aligned} \quad (4)$$

Nous n'incluons pas, dans notre analyse, des vecteurs permettant de tenir compte des périodes de temps (comme la saisonnalité, par exemple). Puisque notre analyse est effectuée sur une période de temps plus courte que celle étudiée par Burtch et al., nous ne nous attendons pas à des effets de saisonnalité très importants - nous n'avons pas à faire, comme dans leur étude, à des événements marquants comme le temps des impôts, la saison des fêtes, etc.

5 Analyse de régression

5.1 Statistiques de test

5.1.1 Considérations méthodologiques

Dès lors que l'on effectue la régression avec la totalité des instruments, le logiciel indique la présence d'un problème. Comme noté dans la section portant sur la modélisation, il s'agit du problème de surabondance des instruments par rapport au nombre total d'observations. Roodman (2009) note ce problème, indiquant qu'il s'agit d'une des préoccupations principales liées à l'estimation SGMM. Il décrit cette problématique de la façon suivante : l'estimateur SGMM, utilisé dans notre analyse, a comme particularité de créer de nouveaux instruments pour chaque période de temps T - ainsi, la matrice d'instruments croît de manière quadratique par rapport à T , ce qui peut générer un très grand nombre d'instruments lorsque les données étudiées s'étalent sur plusieurs périodes. C'est effectivement le cas de notre banque de données, où T peut atteindre 60, ce qui est considérablement plus large que les problèmes grand N , petit T auxquels l'estimateur SGMM était consacré à l'origine.

Roodman note par ailleurs les symptômes principaux d'un trop grand nombre d'instruments : surapprentissage (*overfitting*) des variables endogènes, imprécision des estimés de la matrice de pondération et biais subséquent des écarts-type, et faiblesse du test diagnostic de Hansen. Un modèle qui est *overfitted* sera biaisé - les études citées par Roodman indiquent qu'il est possible de réduire le biais en réduisant le nombre d'instruments, et qu'il est donc souhaitable d'évaluer la robustesse des résultats obtenus face à une telle réduction des instruments. En ce qui concerne les biais des écarts-type, Roodman recommande la correction de Windmeijer (2005), aujourd'hui incorporée dans la plupart des logiciels statistiques (dont la commande *xtabond2* dont nous faisons usage). Finalement, en ce qui concerne les tests de Hansen, la prolifération d'instruments provoque un risque de faux-positifs. Ces tests, qui évaluent l'exogénéité des instruments employés, peuvent être faussés par un trop grand nombre d'instruments et ainsi donner des résultats « parfaits » (valeur-p de 1.000).

Afin de pallier à ces problèmes, Roodman recommande de réduire le nombre d'instruments utilisés lors de l'estimation. Il démontre qu'il existe deux techniques valides permettant d'accomplir ceci : on peut soit utiliser un groupe d'instruments *collapsed*

– où l'on restreint la taille horizontale de la matrice d'instruments – ou encore limiter le nombre de retards employés. Ces deux techniques ont pour effet de rendre la matrice d'instruments linéaire en T , ce qui l'empêche de croître de manière incontrôlable. Roodman suggère donc de les employer soit individuellement ou en combinaison, et d'évaluer l'impact de la réduction du nombre d'instruments sur les tests de Hansen. Il recommande par ailleurs d'adopter une attitude relativement conservatrice par rapport à ceux-ci; comme la prolifération d'instruments peut donner des faux-positifs, il est préférable d'obtenir une statistique de Hansen proche des seuils de signification (0,05 ou 0,1).

Nous effectuons une analyse en quatre parties, chaque partie correspondant à une des quatre catégories Kickstarter pour lesquelles nous avons récolté des données. Chacune de ces catégories possède un nombre variable d'observations; puisque le nombre d'observations détermine, pour l'estimateur SGMM, le nombre d'instruments, et que ce nombre d'instruments détermine également la validité de l'estimé, nous devons donc utiliser la méthode préconisée par Roodman afin de choisir le nombre d'instruments approprié pour chaque catégorie. Comme il s'agit d'une manière plus ou moins arbitraire de procéder, nous effectuons divers tests afin de vérifier la robustesse des résultats. Nous estimons d'abord les coefficients de chaque catégorie avec l'ensemble d'instruments «optimal»; ensuite, nous effectuons des estimations supplémentaires en variant le nombre d'instruments ainsi qu'en retirant certaines variables du modèle. Pour chaque estimation, nous obtenons les statistiques pour le test d'autocorrélation d'Arellano et Bond, le test de Sargan-Hansen ainsi que le test de différence-en-Hansen. À titre de rappel, le test d'autocorrélation Arellano-Bond identifie la présence d'autocorrélation dans le terme d'erreur – par exemple, si le test AR(1) est significatif, il y a présence d'autocorrélation dans les premiers lags et nous ne pouvons donc les utiliser comme instruments. La logique est la même pour AR(2) et AR(3). Nous verrons que tous les tests de niveau AR(3) sont non-significatifs, ce qui signifie qu'il est toujours possible d'utiliser les lags de 3 (et plus) en tant qu'instruments. Le test de Sargan-Hansen, lui, vérifie la validité des instruments dans leur totalité, tandis que les tests de Différence-en-Hansen vérifient la validité des groupes d'instruments individuels (GMM et IV classique). Notons finalement que nous appliquons, dans tous les cas, la correction de Windmeijer, tel que recommandé par Roodman (2006) afin d'obtenir des écarts-type robustes.

5.1.2 Tests d'autocorrélation d'Arellano et Bond

Nous fournissons dans le **tableau 8** les résultats des tests d'autocorrélation pour certaines spécifications du modèle, soit avec des instruments à partir des 3e, 4e et 5e lags. Dans trois catégories sur quatre, le test Arellano-Bond indique la présence d'autocorrélation du 1er degré ou AR(1), ce qui est à prévoir puisque nous estimons un modèle dynamique où le premier lag de la variable dépendante est comprise parmi les variables indépendantes. Pour nos deux catégories les plus importantes, soit Film et Musique, le test indique également la présence d'un processus AR(2), ce qui interdit l'utilisation des lags du second degré en tant qu'instruments. Pour ces catégories, seuls les lags de 3 et plus peuvent être utilisés comme instruments – en effet, pour les processus AR(3), le test présente toujours un résultat non-significatif ne permettant pas de rejeter l'hypothèse nulle. La catégorie Jeux Vidéo permet l'utilisation de lags de 2 et plus, et il ne semble même pas y avoir d'autocorrélation d'ordre 1 dans la catégorie Logiciels ; cependant cette catégorie présente certains problèmes dont nous discutons plus bas.

5.1.3 Tests de Sargan-Hansen et Différence-en-Hansen

Les résultats des divers tests de Hansen sont présentés, pour quelques spécifications du modèle, dans le **tableau 9**. Remarquons qu'il est difficile d'obtenir des résultats solides avec le modèle tel que spécifié initialement – certains ensembles d'instruments produisent des statistiques de Hansen qui semblent au moins acceptables (retard de 4 pour Film, 5 pour Musique, ensemble d'instruments collapsed pour Jeux Vidéo), mais vu l'avertissement de Roodman de se méfier des résultats qui passent la barre «de justesse», nous avons décidé de refaire l'expérience avec quelques autres spécifications. Lorsque nous modifions le modèle en retirant carrément certaines variables indépendantes de l'ensemble des instruments GMM, nous notons une amélioration importante des différentes statistiques de Hansen. Par exemple, en retirant la fréquence de contribution, certaines statistiques des catégories Film et Musique qui peinaient à dépasser le seuil de signification de 10% semblent à présent devenir clairement non-significatives. Ces spécifications produisent également des coefficients statistiquement significatifs, comme nous le verrons dans la section suivante. Le retrait total d'un des deux ensembles d'instruments semble donc être une solution possible au problème de la surabondance d'instruments. Notons finalement qu'il semble impossible d'obtenir des résultats valides à partir des données de la catégorie Logiciels. Nous soupçonnons

que ceci est dû au nombre restreint de projets dans cette catégorie – les estimés des coefficients pour celle-ci ne seront donc pas rapportés, puisque nous doutons de leur validité.

5.2 Résultats principaux

5.2.1 Estimation SGMM

Suivant la section précédente, il semble clair que les tests statistiques employés valident certains ensembles d'instruments plutôt que d'autres. Tel que mentionné précédemment, il faut choisir la meilleure spécification du modèle suivant les résultats de nos tests de Hansen. Selon Roodman (2009), le problème principal de l'estimateur SGMM est une surabondance d'instruments pouvant causer un surajustement des variables instrumentées. Ce surajustement tend à biaiser les résultats vers ceux obtenus par l'estimateur des MCO. Le test de Hansen nous permet de résoudre ce problème en évaluant différents ensembles d'instruments ; les meilleurs peuvent ensuite être choisis pour l'estimation en tant que telle. Roodman indique que, pour le test de Hansen, on doit préférer des résultats supérieurs (l'hypothèse nulle étant que les instruments sont valides) mais proches des seuils de signification - 0,05 et 0,10. Comme on peut le voir dans le tableau 9, la meilleure spécification varie selon l'ensemble de données analysé. La restriction des instruments au 4ème retard semble préférable pour les films, au 5ème retard préférable pour la musique, et un ensemble d'instruments collapsed - donc très petit - semble préférable pour les jeux vidéos. Ce sont les résultats correspondant à ces spécifications que nous privilégierons pour nos conclusions.

Nous rapportons néanmoins les résultats de toutes les spécifications essayées, afin de démontrer l'impact du choix de la spécification sur le coefficient estimé. Le **tableau 10** démontre que les estimés varient à l'intérieur d'une fourchette relativement large selon la spécification choisie. Cependant, les variations ne sont pas énormes ; le signe des coefficients demeure dans tous les cas le même, et nos résultats sont assez précis pour nous permettre d'identifier l'effet de chacune des variables indépendantes. Seule l'ampleur exacte de cet effet reste incertaine.

Rappelons d'abord l'interprétation que nous pouvons faire de chacune de nos variables indépendantes. La variable de contribution représente le montant de dollars contribués à un projet à la période $T-1$. Ce montant peut être observé par le contributeur potentiel en tout temps. Nous faisons l'hypothèse que ceci aura un effet sur

le comportement du contributeur au temps T : s'il diminue sa contribution lorsqu'il observe une contribution additionnelle au temps $T-1$, alors il fait du resquillage, et le coefficient de la variable contribution sera négatif; s'il l'augmente, alors on observe un comportement grégaire, et le coefficient sera positif. L'interprétation de la fréquence de contribution est semblable, mais elle prend en compte le temps écoulé – le contributeur prend alors en compte la contribution moyenne depuis le début du projet lorsqu'il prend sa décision. Ensuite, la variable de budget restant (manque à gagner) indique si le contributeur augmente ou diminue sa contribution par rapport aux besoins du projet qu'il observe. S'il la diminue lorsque les besoins du projet diminuent et qu'il l'augmente lorsqu'ils augmentent, cela peut indiquer qu'il tente de «substituer» les contributions des autres pour les siennes; si au contraire la relation est inverse, cela indique qu'il a tendance à adopter un comportement grégaire peu importe les besoins du projet. Finalement, nous incluons un indicateur externe de la popularité du projet basé sur le nombre de *likes* obtenus sur Facebook.

Les coefficients de la variable de contribution sont, pour chacune des catégories que nous avons conservé (Film, Musique et Jeux Vidéo), positifs. Pour la catégorie Film, le coefficient varie entre 0.36 et 0.53, dépendant de l'ensemble d'instruments choisi – les statistiques de test les plus fiables sont associées à un coefficient de 0.36. Pour la catégorie Musique, le coefficient associé au meilleur test statistique est de 0.57, tandis que pour la catégorie Jeux Vidéo il est considérablement plus large, soit de 0.81. La signification générale de ces résultats est claire : une augmentation des contributions au temps $T-1$ conduit à une augmentation correspondante des contributions au temps T . L'effet semble être moyennement important pour les deux catégories les plus importantes – les coefficients indiquent qu'une contribution d'un dollar au temps $T-1$ entraîne par la suite une contribution moindre, soit entre 0.35 et 0.56 dollars. La situation est un peu différente pour la catégorie Jeux Vidéo, où une contribution d'un dollar entraîne presque une contribution identique voir même supérieure. Selon le cadre d'analyse que nous avons proposé, ces résultats indiquent que les comportements des utilisateurs de Kickstarter sont largement grégaires face aux contributions totales qu'ils observent à la période précédant leur décision de contribuer.

L'effet de la fréquence de contribution est moins certain. Bien que le coefficient soit solidement positif – mais moins important que celui de contribution à 0.15 – pour la catégorie Film, il est équivalent à 0 pour la catégorie Musique et est carrément négatif pour la catégorie Jeux Vidéo (-0.13). Ceci indique que, face à la contribution moyenne

du projet observé, l'utilisateur fait moins preuve de grégarisme qu'en ce qui concerne la contribution totale. En termes concrets, cela signifie qu'un projet qui amasse rapidement du financement n'en obtiendra pas nécessairement davantage lors de la période suivante – tout dépendant de la catégorie où il se trouve. Ce fait souligne la difficulté pour un projet d'établir du momentum dans son financement – il correspond à l'observation, faite plus haut, qu'un projet qui amasse un dollar à T-1 obtient subséquemment moins de un dollar en contributions «grégaires» à T. Notons par ailleurs que le niveau de significativité de cette variable est plus incertain que celui de la variable de contribution. Dans plusieurs cas - surtout en ce qui concerne la catégorie Film - les statistiques relatives à la significativité sont relativement faibles.

La variable de manque à gagner, qui dénote le montant de financement qu'un projet doit obtenir avant d'atteindre son objectif, a dans tous les cas un effet peu important sur la variable dépendante – et dans plusieurs cas cette variable est carrément non-significative. Le coefficient est dans tous les cas faiblement négatif, ce qui semble indiquer qu'un projet qui a de plus grands besoins aura plus de difficultés à obtenir du financement. Quant à la variable de popularité sur Facebook, elle est encore plus souvent non-significative, et également d'une force très faible; fait toutefois intéressant, elle est significative et moyennement importante pour certaines spécifications du modèle en ce qui concerne les données sur les jeux vidéo. Par exemple, avec un ensemble d'instruments collapsed, on obtient un coefficient significatif de 0.27. On peut en déduire une présence plus importante de ces projets – et de leurs contributeurs éventuels – sur les réseaux sociaux, alors que ces derniers n'auraient qu'un impact négligeable sur les projets dans les catégories Film et Musique.

5.2.2 Autres tests de robustesse

Nous effectuons quelques tests supplémentaires afin de vérifier la robustesse de nos résultats, et afin d'examiner l'effet de certaines modifications aux spécifications de notre modèle. Le **tableau 11** permet de comparer nos résultats avec ceux issus d'une simple estimation par moindres carrées ordinaires. Les coefficients obtenus par MCO ne sont pas énormément différents de ceux obtenus avec l'estimateur SGMM. Pour le premier retard de la variable de contribution, nous obtenons des coefficients de 0.21 pour la catégorie Film, 0.38 pour la catégorie Musique et 0.76 pour la catégorie Jeux Vidéo; chacun de ces coefficients est légèrement moins important que celui obtenu

par SGMM. La fréquence de contribution passe à 0.13 pour la catégorie Film, 0.09 pour Musique et -0.06 pour Jeux Vidéo; des résultats encore une fois semblables à ceux obtenus par SGMM. Les variables subséquentes présentent également des résultats comparables. Il y a cependant dans certains cas une perte de significativité lorsque nous utilisons l'estimateur MCO, par exemple dans le cas de la variable de contribution dans la catégorie Film.

Nous ré-estimons ensuite notre modèle en retirant certaines observations aberrantes (outliers), soit en retirant pour chaque catégorie le 1er et le 99e percentile des projets en termes d'objectifs. Ceci nous permet de déterminer, par exemple, si certains projets de très grande envergure causent une variation démesurée dans les coefficients. Cela semble être le cas pour certaines catégories, mais pas toutes; par exemple, pour la catégorie Film, le coefficient pour la variable de contribution passe de 0.35 à 0.32, alors qu'elle passe de 0.56 à 0.18 dans le cas de la catégorie Musique. Pour la catégorie Jeux Vidéo, cette même variable passe de 0.80 à 0.68. Nous voyons donc qu'en certains cas – nommément les projets de Musique pour cette banque de données – l'effet de la variable de contribution peut être relativement ambigu, bien qu'il demeure positif. Les autres résultats de cette ré-estimation peuvent être consultés dans le **tableau 12**; de manière générale, ils ne diffèrent pas significativement de ceux obtenus lors de la première estimation.

6 Conclusion

Bien que l'étude de Burtch et al. relève certains faits suggérant que les contributions, dans le financement participatif, peuvent souffrir d'un effet de substitution – c'est-à-dire que des sommes contribuées à la période T-1, ou encore un accroissement rapide du financement, pouvaient entraîner une diminution à court-terme des nouvelles contributions – nous ne découvrons aucun effet de ce genre dans notre analyse. Tous les coefficients qui pourraient être révélateurs de comportements de resquillage sont positifs, et tendent donc plutôt vers le gréganisme, dans le sens où un dollar contribué à un projet entraînera des contributions additionnelles. Notons cependant qu'il n'est pas possible d'identifier un réel effet d'entraînement, où seul les comportements grégaires auraient de l'importance : les coefficients, s'ils sont positifs, ont généralement seulement une influence modérée. Une hypothèse de gréganisme plus fort, où une contribution enverrait un signal aux usagers et entraînerait une vague croissante de contribution, eût été plus crédible si ces coefficients étaient supérieurs à 1. Pour nos résultats tels qu'ils sont, nous ne pouvons qu'avancer une hypothèse de gréganisme faible ou modéré, où la contribution d'un dollar entraîne des contributions subséquentes de manière décroissante.

Notons que, en ce qui concerne les modèles formels de comportements grégaires – par exemple celui développé par Banerjee (1992) – le gréganisme est plus précisément décrit comme étant une situation où les agents abandonnent leur préférences personnelles pour adopter le choix «grégaire», et ce même si ce dernier est moins avantageux. Nous avons tenté de distinguer les choix des usagers sur Kickstarter de leurs préférences «réelles» en introduisant les *like* sur Facebook en tant qu'instruments externes ; malheureusement, dans plusieurs cas, le coefficient de cet instrument s'est avéré non-significatif, ce qui a réduit le montant d'information que nous avons pu en retirer. Malgré tout, les divers tests fournis dans le cadre de l'estimateur SGMM indiquent que notre instrument est exogène, et qu'il permet donc, partiellement au moins, d'effectuer la tâche souhaitée. En trouvant des instruments additionnels de ce type, il devrait être possible de distinguer davantage les vraies préférences des usagers de l'aspect «grégaire» de leurs contributions et donc d'isoler la composante grégaire du financement sur Kickstarter et sur les autres sites de financement participatif. Ainsi faisant nous pourrions identifier les plate-formes caractérisées par un niveau plus ou moins élevé de gréganisme (Kickstarter) ou de resquillage (l'article de Burtch et al.). Ensuite, l'étape subséquent

sera de déterminer les facteurs entraînant les divers comportements observables sur les différentes plate-formes de crowdfunding.

La croissance de l'industrie du *crowdfunding* pourrait fournir une piste intéressante en ce qui concerne la recherche sur les comportements des utilisateurs. Chaque nouveau site qui voit le jour constitue en effet une nouvelle expérience naturelle (*natural experiment*) dont les caractéristiques propres affectent d'une manière ou d'une autre les comportements de ses usagers. L'existence de ces expériences naturelles pourrait nous permettre d'accumuler deux types distincts de connaissances au sujet du financement participatif : premièrement, des connaissances empiriques au sujet des caractéristiques des plate-formes de financement et des comportements des usagers de celles-ci ; ensuite, le développement d'une approche théorique permettant de modéliser et de prédire les comportements en se basant sur les sus-dites caractéristiques.

6.1 Recherche empirique : caractéristiques des sites de crowdfunding

La première piste à explorer est celle des caractéristiques propres à chaque plate-forme (site) de financement participatif. Si, dans le cas du site étudié par Burtch et al., le resquillage pouvait être à redouter, nous démontrons cependant que ce n'est pas le cas pour tous les sites de financement participatif. Nos résultats illustrent, à tout le moins, qu'il existe potentiellement divers comportements liés à ce type de financement. Il est probable, comme le soulèvent Burtch et al., que les caractéristiques spécifiques à chaque site de financement participatif peuvent influencer les comportements de leurs usagers. Nous identifions trois différents types de caractéristiques permettant de classer les plate-formes selon une hiérarchie précise. D'abord, le modèle de financement, c'est-à-dire la façon selon laquelle la plate-forme étudiée permet aux auteurs d'obtenir des fonds pour faire progresser leur projet ; ensuite, le type d'industrie ou de domaine économique où opère la plate-forme (par exemple, le journalisme pour Burtch et al., ou les diverses industries numériques pour Kickstarter) ; et finalement, la structure de récompenses prévue par la plate-forme.

Divers modèles de financement existent actuellement : notons par exemple la distinction entre les sites de type *All or Nothing* («tout ou rien») où les créateurs de projets n'obtiennent du financement que s'ils atteignent un objectif pré-établi, et

les sites de type *Keep It All* qui permettent au créateurs de projets de garder tout le financement qu'ils obtiennent, peu importe leur objectif. Le premier type de site – Kickstarter en est un exemple – encourage certainement les usagers à contribuer sans prendre en considération la situation financière d'un projet qui les intéresse, puisque leur argent leur sera retourné si le projet n'atteint pas son objectif. Le second type de site, en contrepartie, encourage plutôt les usagers à être plus soucieux de la progression d'un projet avant d'y contribuer. De nombreux sites faisant partie de cette seconde catégorie ont récemment vu le jour, comme par exemple FundAnything, IndieGogo ou RocketHub.

Il est également possible d'imaginer que des usagers agiraient de manière différente dépendamment de l'industrie spécifique où la plate-forme de financement évolue : spectacles de musique, dons charitables, etc. Nos résultats sont relativement uniformes, mais chacune des trois catégories pour lesquelles nous obtenons des résultats présente des caractéristiques propres légèrement différentes des autres. Par exemple, la popularité sur Facebook semble jouer un rôle plus important dans le cas des projets impliquant des jeux vidéo. Comme la popularité croissante du financement participatif a entraîné la création de plate-formes consacrées à des types de projets spécifiques, il devient possible d'évaluer la propension au resquillage dans des secteurs mieux définis même que ceux que nous avons choisi. Et comme les secteurs desservis par les plate-formes de financement participatif correspondent plus ou moins, dépendant des cas, à des biens publics, une telle analyse permettrait de voir si le resquillage est surtout présent dans le cas de financement de biens publics.

Enfin, comme nous l'avons mentionné, le type de compensation offerte aux investisseurs affecte certainement leur comportement. Si Kickstarter offre divers types de compensations de nature plus ou moins «concrète», d'autres sites offrent moins – par exemple, de simples remerciements – inclus dans le produit financé – alors que d'autres sites offrent plus, allant jusqu'à proposer aux usagers d'acheter des parts dans des compagnies ainsi financées.

6.2 Approches théoriques

Une fois les caractéristiques empiriques du financement participatif établies, l'étape suivante sera de tenter d'associer un modèle théorique à ces connaissances. De nombreux modèles portant sur les biens publics existent, mais certains correspondent plus

spécifiquement au contexte du financement participatif. Dans cette situation, nous tenterions d'expliquer pourquoi il n'existe pas de resquillage sur Kickstarter, en dépassant la simple présence de comportements grégaires. Il serait alors intéressant de déterminer si les caractéristiques et la conception des sites de financement participatif permettent carrément d'éliminer le resquillage, et ce de manière systématique. Une telle analyse entrerait dans le cadre de la littérature portant sur les contrats d'assurance (*assurance contracts*). Cette littérature, largement fondée par Bagnoli et Lipman (1989), décrit certains mécanismes issus de la théorie des jeux permettant de surmonter le problème de l'approvisionnement privé des biens publics. L'article de Tabarrok (1996) portant sur les contrats d'assurance «dominants» introduit dans le modèle un entrepreneur rationnel qui résout le problème d'approvisionnement en créant un contrat avec certaines caractéristiques. Il pourrait être intéressant de voir si les diverses caractéristiques des sites de crowdfunding – mentionnées plus haut – permettent à ces sites de ressembler plus ou moins au type de contrat décrit par Tabarrok – et de voir si les données que l'on peut récupérer au sujet des comportements des utilisateurs soutiennent cette thèse. Si c'est le cas, ce fait pourrait expliquer pourquoi on n'observe pas de resquillage sur Kickstarter dans notre analyse – la conception de ce site pourrait effectivement représenter un exemple de contrat d'assurance dominant. Dans un cadre plus large, l'expérimentation dans le domaine des plate-formes de financement participatif présente l'opportunité de découvrir des mécanismes d'approvisionnement privés de biens publics applicables à toutes sortes de sphères de l'économie, et non pas seulement aux biens numériques.

7 Bibliographie

Anderson, T. W. et Hsiao, C. (1982) Formulation and estimation of dynamic models using panel data, *Journal of Econometrics*, Elsevier, 18, 1, janvier, 47-82.

Andreoni, J. (1990) Impure Altruism and Donations to Public Goods : A Theory of Warm-Glow Giving, *The Economic Journal*, 100, 401, 464-477.

Andreoni, J. (1998) Toward a Theory of Charitable Fund-Raising, *The Journal of Political Economy*, 106, 6, 1186-1213.

Arellano, M., et Bond, S. (1991) Some Tests of Specification for Panel Data : Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, Wiley Blackwell, 58, 2, avril, 277-97.

Arellano, M. et Bover, O. (1995) Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics*, Elsevier, 68, 1, juillet, 29-51.

Asvanund, A., Clay, K., Krishnan, R., et Smith, M.D. (2004) An Empirical Analysis of Network Externalities in Peer-to-Peer Music-Sharing Networks, *Information Systems Research* , 15, 2, 155-174

Balestra, N. et Nerlove, M. (1966) Pooling Cross-Section and Time Series Data in the Estimation of Dynamic Models : The Demand for Natural Gas, *Econometrica*, juillet 1966.

Banerjee, A. (1992) A Simple Model of Herd Behavior, *The Quarterly Journal of Economics* , 107, 3, 797-817.

Becker, G. (1974) A Theory of Social Interactions, *Journal of Law and Economics*, 19, 2, 245-48.

Belleflamme, P., Lambert, T., et Schwienbacher, A. (2010) Crowdfunding : Tapping the Right Crowd, SSRN Working Paper.

Blundell, R. et Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, Elsevier, août, 87, 1, 115-143.

Borck, R., Frank, B., et Robledo, J. (2006) An Empirical Analysis of Voluntary Payments for Information Goods on the Internet, *Information Economics and Policy*, 18, 2, 229-239.

Burtch, G., Ghose, A. et Wattal, S. (2012), An Empirical Examination of the Antecedents and Consequences of Investment Patterns in Crowd-Funded Markets, *Information Systems Research*, Forthcoming.

Brabham, Daren. "Moving the crowd at iStockphoto : The composition of the crowd and motivations for participation in a crowdsourcing application" *First Monday* [Online], Volume 13 Numéro 6 (21 Mai 2008)

Clark, R., Doraszelski, U., Draganska, M. (2009) The Effect of Advertising on Brand Awareness and Perceived Quality : An Empirical Investigation using Panel Data, *Quantitative Marketing and Economics*, 7, 2, 207-236.

Gu, B., Huang, Y., Duan, W., et Whinston, A. (2009) Indirect Reciprocity in Online Social Networks – a Longitudinal Analysis of Individual Contributions and Peer Enforcement in a Peer-to-Peer Music Sharing Network, *McCombs Research Paper Series*

Nandi, T., et Rochelandet, F. (2008) The Incentives for Contributing Digital Contents over P2p Networks : An Empirical Investigation, *Review of Economic Research on Copyright Issues*, 5, 2, 19-35.

Olson, M. (1965). *The Logic of Collective Action : Public Goods and the Theory of Groups*. Harvard University Press.

Pirotte, A. (2011). *Économétrie des données de panel*. Economica.

Rayna, T. (2007) *Digital Goods as Durable Public Goods*. Thèse, PhD. Aix-Marseille.

Roberts, R. (1984) A Positive Model of Private Charity and Public Transfers, *The Journal of Political Economy*, 92, 1, 136-148.

Roodman, D. (2006) How to do xtabond2 : An introduction to "difference" and "system" GMM in Stata (Working Paper No. 103). Obtenu sur le site du Center for Global Development : www.cgdev.org/content/publications/detail/11619.

Roodman, D. (2008) A short note on the theme of too many instruments (Working Paper No. 125). Obtenu sur le site du Center for Global Development : www.cgdev.org/content/publications/detail/14256.

Samuelson, P. (1954) The Pure Theory of Public Expenditure, *The Review of Economics and Statistics*, 36, 4, 387-389.

Schwienbacher, A., et Larralde, B. (2010) Crowdfunding of Small Entrepreneurial Ventures, SSRN Working Paper.

Surowiecki, J. (2004). *The wisdom of crowds : Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York : Doubleday.

Wooldridge, Jeffrey M. (2001). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Xia, M., Huang, Y., Duan, W., et Whinston, A. (2011) To Continue Sharing or Not to Continue Sharing? An Empirical Analysis of User Decision in Peer-to-Peer Sharing Networks, *Information Systems Research*, Forthcoming.

Annexe

FIGURE 1 – Interface de Kickstarter.com



FIGURE 2 – Statistiques de la catégorie Film & vidéo

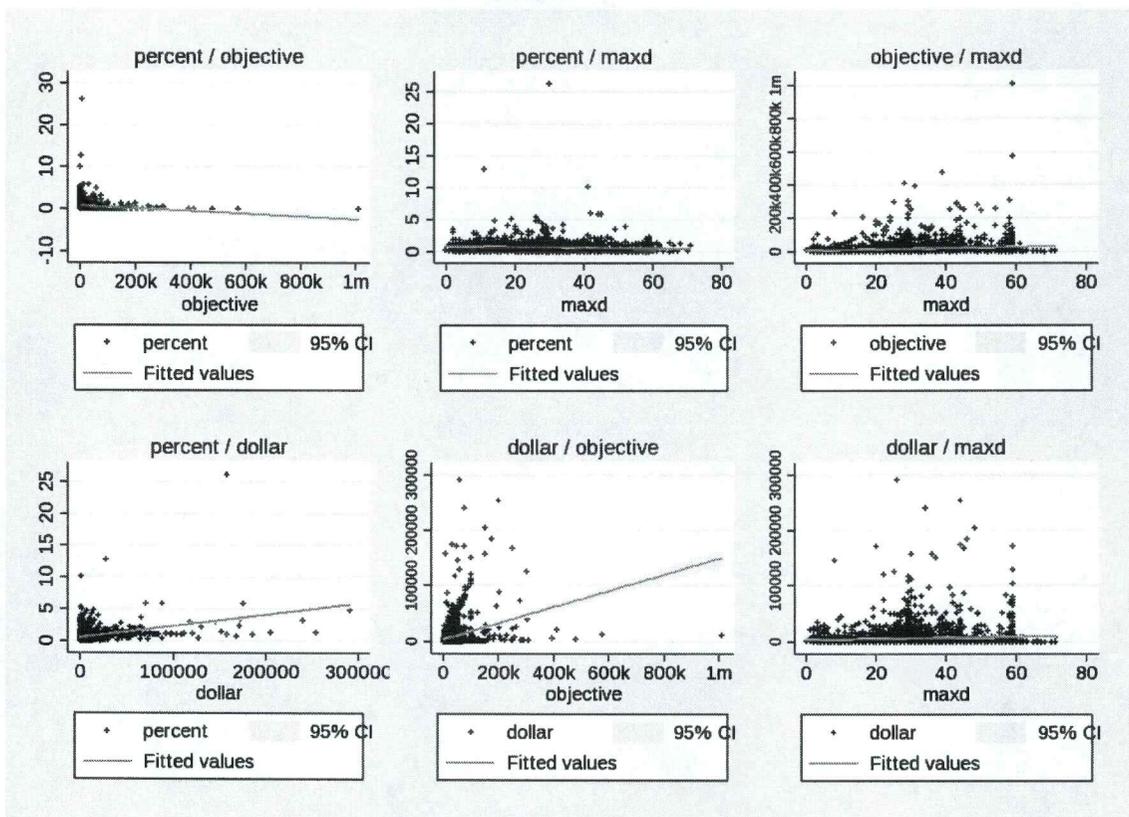


FIGURE 3 – Statistiques de la catégorie Musique

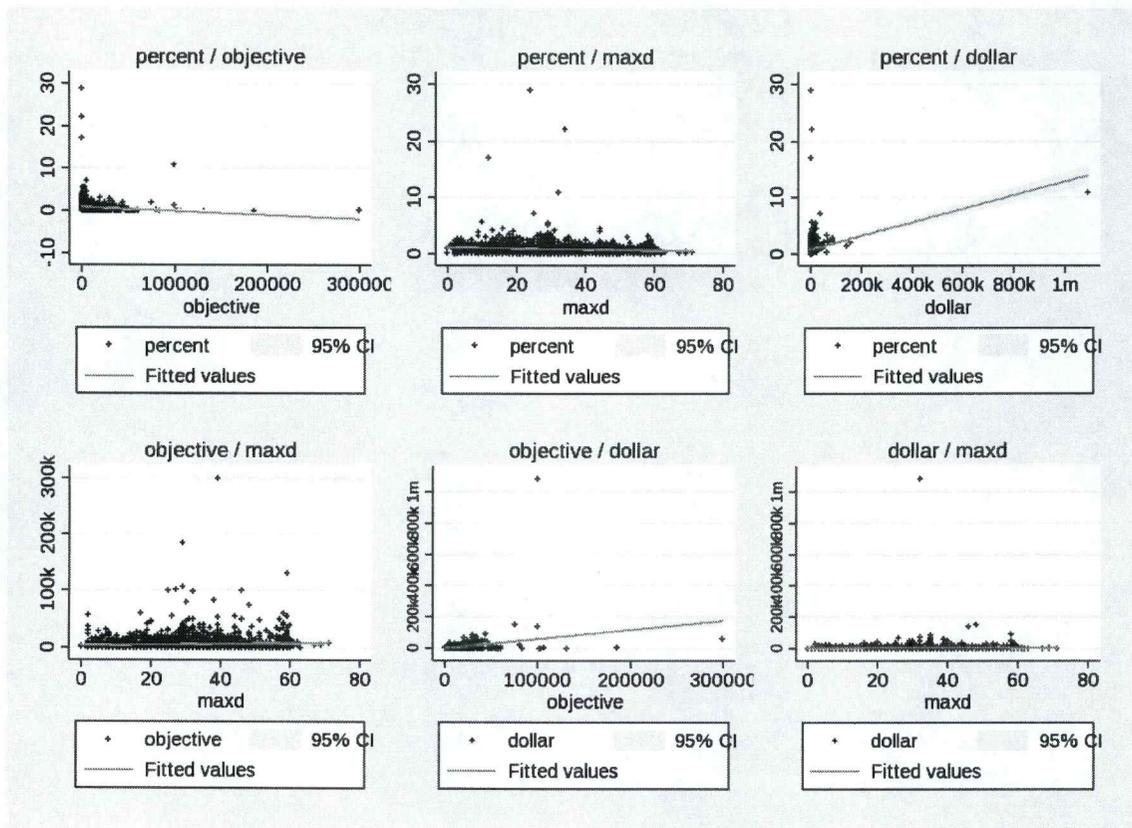


FIGURE 4 – Statistiques de la catégorie Jeux vidéo

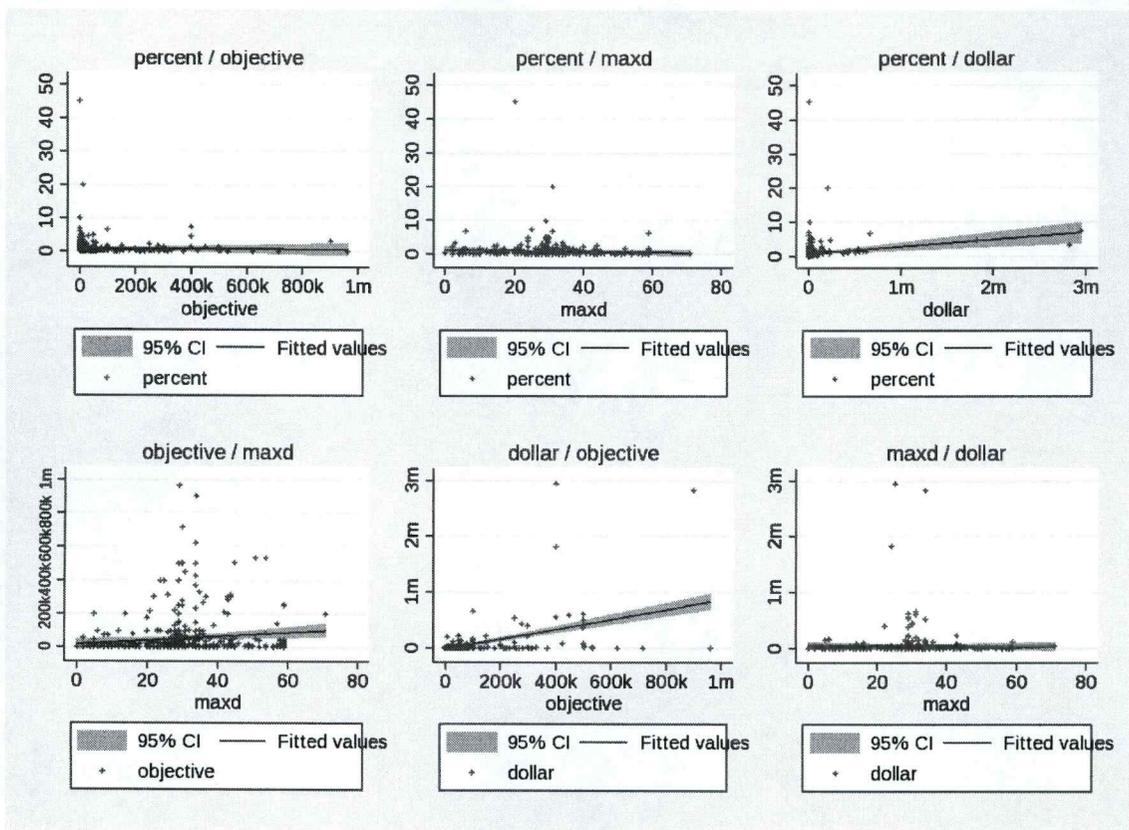


FIGURE 5 – Statistiques de la catégorie Logiciels

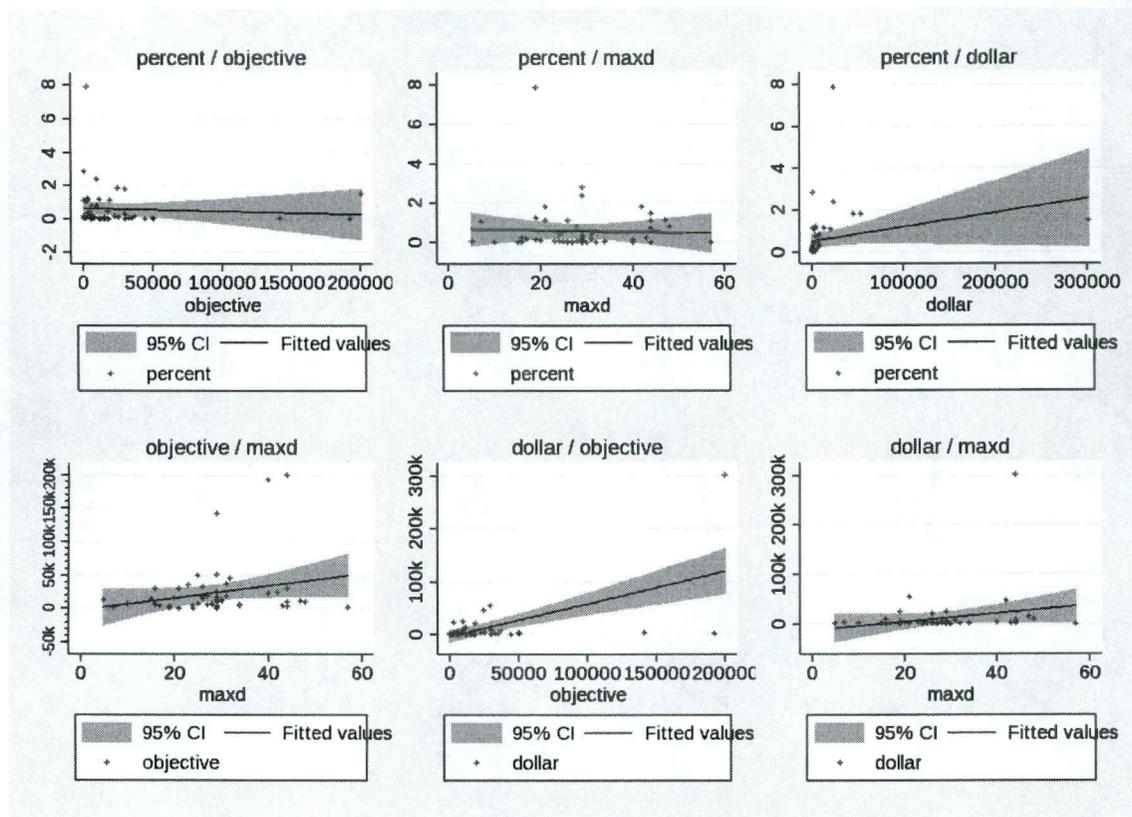


TABLE 2 – Statistiques descriptives - Film

	Obs.	Moyenne	Écart-type	Min	Max
contribution	87587	217.961	1088.723	-72253	75445
contribution_frequency	95394	407.2906	1728.893	0	154416
remaining_budget	95394	13294.18	34225.37	-261994	999900
fb_like	77556	989.7952	13108.22	0	384376

TABLE 3 – Statistiques descriptives - Musique

	Obs.	Moyenne	Écart-type	Min	Max
contribution	91106	112.4124	826.5548	-10330	118864
contribution_frequency	98770	267.7061	1963.717	0	329205
remaining_budget	98770	3558.911	14699.42	-989029	298467
fb_like	90296	848.9783	14447	0	385134

TABLE 4 – Statistiques descriptives - Jeux vidéo

	Obs.	Moyenne	Écart-type	Min	Max
contribution	8595	1518.732	9093.604	-6302	409384
contribution_frequency	9624	3750.798	35544.97	0	1957208
remaining_budget	9624	28987.94	154650.3	-2547764	953073
fb_like	7988	1121.44	4660.128	0	56806

TABLE 5 – Statistiques descriptives - Logiciels

	Obs.	Moyenne	Écart-type	Min	Max
contribution	1091	404.0394	1828.357	-2571	28643
contribution_frequency	1182	951.3417	3647.888	0	62482
remaining_budget	1182	14252.26	32672.62	-102204	191989
fb_like	1141	331.6564	911.6524	0	4883

TABLE 6 – **Corrélations inter-temporelles**

	contribution	contribution_frequency
Film	0.2603	0.9392
Musique	0.7262	0.9403
Jeux vidéo	0.8300	0.9689
Logiciels	0.4106	0.9088

TABLE 7 – **Décomposition de la variance**

	contribution		cont_freq		rem_budg	
	within	between	within	between	within	between
Film	945.81	596.03	1199.16	1984.356	4799.957	33207.27
Musique	611.26	510.70	1320.67	1617.16	3926.34	12696.06
Jeux vidéo	6529.00	5538.62	27429.98	20497.68	43291.82	140918.3
Logiciels	1415.58	899.449	2684.74	2003.99	12856.32	32867.73

TABLE 8 – Résultats - Tests d'autocorrélation

Film

Restrictions sur les instruments : collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
AR(1)	0.00	0.00	0.00	0.01	0.02
AR(2)	0.01	0.02	0.00	0.04	0.08
AR(3)	0.19	0.18	0.15	0.21	0.24

Musique

Restrictions sur les instruments : collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
AR(1)	0.03	0.00	0.00	0.03	0.00
AR(2)	0.23	0.09	0.06	0.23	0.07
AR(3)	0.35	0.32	0.31	0.35	0.32

Jeux vidéo

Restrictions sur les instruments : collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
AR(1)	0.00	0.00	0.00	0.00	0.00
AR(2)	1.00	0.93	0.96	0.92	0.85
AR(3)	0.90	0.97	1.00	0.94	0.77

^a En retirant la contribution comme v.i.

^b En retirant la fréquence de contribution comme v.i.

TABLE 9 – Résultats - Tests de Sargan-Hansen

<i>Film</i>									
Restrictions sur les instruments : collapsed 3e retard 4e retard 5e retard 4e retard ^a 4e retard ^b									
	0.11	0.15	0.20	0.11	0.48	0.00			
Hansen test of overid restrictions	0.11	0.15	0.20	0.11	0.48	0.00			
Gmm – excluding group	0.10	0.75	0.67	0.60	0.86	0.67			
Gmm – difference	0.67	0.02	0.06	0.03	0.14	0.00			
IV – excluding group	0.11	0.16	0.22	0.11	0.47	0.00			
IV – difference	0.38	0.14	0.13	0.21	0.50	0.04			

<i>Musique</i>									
Restrictions sur les instruments : collapsed 3e retard 4e retard 5e retard 4e retard ^a 4e retard ^b									
	0.00	0.00	0.02	0.07	0.28	0.18			
Hansen test of overid restrictions	0.00	0.00	0.02	0.07	0.28	0.18			
Gmm – excluding group	0.00	0.05	0.31	0.18	0.13	0.17			
Gmm – difference	0.00	0.00	0.01	0.12	0.62	0.34			
IV – excluding group	0.00	0.00	0.03	0.07	0.33	0.17			
IV – difference	0.49	0.02	0.05	0.95	0.09	0.98			

<i>Jeux vidéo</i>									
Restrictions sur les instruments : collapsed 3e retard 4e retard 5e retard 4e retard ^a 4e retard ^b									
	0.88	0.13	0.09	0.08	0.00	0.52			
Hansen test of overid restrictions	0.88	0.13	0.09	0.08	0.00	0.52			
Gmm – excluding group	0.91	0.98	0.96	0.94	0.75	0.81			
Gmm – difference	0.14	0.00	0.00	0.00	0.00	0.21			
IV – excluding group	0.87	0.13	0.26	0.25	0.01	0.53			
IV – difference	0.63	0.31	0.00	0.00	0.13	0.24			

^a En retirant la contribution comme v.i.^b En retirant la fréquence de contribution comme v.i.

TABLE 10 – Résultats - analyse de régression - SGMM

Film

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	-0.07 (0.16)	0.54*** (0.13)	0.36* (0.17)	0.52*** (0.14)	0.52*** (0.09)	0.41*** (0.15)
contribution_frequency	0.06 (0.07)	0.09* (0.04)	0.15* (0.06)	0.11* (0.05)	0.08* (0.04)	0.13 (0.08)
remaining_budget	-0.07 (0.04)	-0.01 (0.01)	-0.01 (0.11)	-0.01 (0.01)	0.00 (0.11)	-0.01 (0.01)
fb_like	0.02 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00* (0.00)	0.00 (0.00)	0.00 (0.00)

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	0.33** (0.09)	0.51*** (0.01)	0.47*** (0.01)	0.57*** (0.01)	0.46*** (0.01)	0.53*** (0.01)
contribution_frequency	0.09** (0.04)	0.02*** (0.00)	-0.04** (0.01)	0.00 (0.01)	-0.03*** (0.01)	0.01 (0.01)
remaining_budget	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)
fb_like	-4.62 (0.10)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

*Musique**Jeux vidéo*

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	0.81*** (0.10)	0.95*** (0.09)	0.92*** (0.09)	0.99*** (0.10)	0.99*** (0.11)	1.25*** (0.14)
contribution_frequency	-0.13*** (0.06)	-0.09*** (0.03)	-0.12** (0.04)	-0.10*** (0.03)	-0.10 (0.03)	-0.11*** (0.10)
remaining_budget	-0.02*** (0.03)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02** (0.00)	-0.01*** (0.00)
fb_like	0.27** (0.48)	0.10 (0.10)	0.18 (0.09)	0.08 (0.08)	0.09 (0.13)	-0.12 (0.20)

^a En retirant la contribution comme v.i.^b En retirant la fréquence de contribution comme v.i.

*** p < 0.001

** p < 0.01

* p < 0.05

() Écarts-types entre parenthèses

TABLE 11 – Résultats - analyse de régression - MCO

Film

Variable explicative	Coefficient	Écart-type
contribution	0.21	(0.12)
contribution_frequency	0.13**	(0.04)
remaining_budget	0.00	(0.00)
fb_like	0.00	(0.00)

Musique

Variable explicative	Coefficient	Écart-type
contribution	0.38***	(0.08)
contribution_frequency	0.09*	(0.03)
remaining_budget	-0.01**	(0.00)
fb_like	0.00	(0.00)

Jeux vidéo

Variable explicative	Coefficient	Écart-type
contribution	0.76***	(0.05)
contribution_frequency**	-0.06	(0.02)
remaining_budget	-0.01***	(0.00)
fb_like	0.27***	(0.03)

^a En retirant la contribution comme v.i.

^b En retirant la fréquence de contribution comme v.i.

TABLE 12 – Résultats - analyse de régression - SGMM (en retirant les outliers)

Film

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	-0.09 (0.09)	0.50*** (0.13)	0.32 (0.17)	0.48*** (0.15)	0.49*** (0.10)	0.40* (0.16)
contribution_frequency	0.01 (0.04)	0.04 (0.04)	0.10 (0.09)	0.07 (0.05)	0.03 (0.05)	0.13 (0.07)
remaining_budget	-0.05 (0.03)	-0.02 (0.13)	-0.02 (0.02)	-0.02 (0.01)	-0.03* (0.01)	-0.01 (0.02)
fb_like	0.02* (0.01)	0.01 (0.00)	0.01 (0.00)	0.01* (0.00)	0.01* (0.00)	0.00* (0.00)

Musique

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	0.07*** (0.02)	0.24*** (0.07)	0.11 (0.14)	0.18 (0.09)	0.13 (0.16)	-0.09 (0.24)
contribution_frequency	0.04* (0.02)	0.03 (0.03)	0.05* (0.03)	0.09 (0.03)***	0.03 (0.03)	-0.06 (0.10)
remaining_budget	0.02 (0.02)	-0.02*** (0.02)	-0.02* (0.01)	-0.01 (0.01)	-0.02 (0.01)	-0.04** (0.16)
fb_like	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Jeu vidéo

Restrictions sur les instruments :	collapsed	3e retard	4e retard	5e retard	4e retard ^a	4e retard ^b
contribution	0.68*** (0.09)	0.87*** (0.08)	0.84*** (0.07)	0.87*** (0.16)	0.88*** (0.11)	1.13*** (0.17)
contribution_frequency	-0.09** (0.03)	-0.09** (0.03)	-0.11** (0.04)	-0.12* (0.06)	-0.10** (0.04)	-0.04 (0.11)
remaining_budget	-0.03*** (0.01)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.01 (0.01)
fb_like	0.30* (0.13)	0.18 (0.00)	0.23*** (0.04)	0.23*** (0.05)	0.20** (0.07)	-0.14 (0.18)

^a En retirant la contribution comme v.i.^b En retirant la fréquence de contribution comme v.i.