

HEC MONTRÉAL

**Sensibilité et exposition des
consommateurs aux sources
d'information externes:
une étude spatiale et longitudinale
du cinéma québécois**

Par

Thomas Boquet

**Sciences de la gestion
(Marketing)**

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

Novembre 2013

© Thomas Boquet, 2013

Sommaire

La littérature en marketing a exploré les effets de la recherche d'information, du type de sources considérées, de la valence de ces sources d'information et de l'expertise des consommateurs sur la consommation de films et d'autres produits. En effet, de nombreux auteurs ont reconnu le caractère marqué de l'évolution du partage de l'information à travers le temps et son importance sur le choix de consommation (Hennig-Thurau et al 2010, Debenedetti 2006). Dans tous les cas, le degré avec lequel les consommateurs ont accès à l'information n'a pas été intégré pleinement dans les recherches en association avec des données longitudinales et géographiques. Il y a donc un intérêt certain à explorer dans quelles mesures l'exposition aux sources d'information externes joue un rôle sur l'effet de leur valence et influence le succès commercial des films sortis en salle.

Nous examinons plus particulièrement la façon avec laquelle l'exposition des consommateurs aux critiques cinématographiques modère l'effet de la cote des critiques et influence la consommation de ces biens expérientiels à travers le temps. En utilisant plusieurs bases de données assemblées pour les besoins de l'étude et des modèles linéaires mixtes, nous estimons d'une nouvelle manière les effets des variables étudiées.

Nous trouvons que l'exposition à l'information accentue l'effet de sa valence avec une évolution entre les périodes considérées. La cote a un impact moins important la première semaine de sortie par rapport aux semaines d'exploitation restantes. De plus, il existe un seuil d'informa-

tion minimum dont a besoin le consommateur afin de clarifier sa compréhension de l'expérience comme décrite par les critiques. C'est lorsque l'asymétrie d'information est forte (la première semaine) que les consommateurs ont besoin d'accumuler un certain niveau d'information avant de considérer sa valence. À partir de ce point, plus le consommateur est exposé à l'information et plus la valence des critiques a un effet important sur les revenus du box-office. Après la première semaine de sortie, nous observons que le consommateur n'a plus nécessairement besoin d'être exposé aux critiques de films pour que la cote ait un effet positif sur leur succès commercial. L'accès à des sources d'informations telles que le bouche-à-oreille et les autres types de recommandation émises par les consommateurs explique ce comportement. L'effet de l'exposition a tout de même une influence positive sur les revenus du box-office. Notre travail nous permet aussi de confirmer la présence d'une structure de corrélation complexe qu'il est nécessaire de prendre en compte pour une analyse de sensibilité ou pour la prédiction du box-office des films.

Ces résultats empiriques peuvent servir de guide aux gestionnaires qui voudraient mieux contrôler la diffusion d'information sur leurs produits. Les nouveaux modèles estimés sont quant à eux des outils puissants pour prédire la réaction du box-office lorsque les acteurs de l'industrie cinématographique doivent fixer les niveaux des variables qu'ils contrôlent. Ils apportent une précision accrue afin de mettre en place des stratégies de diffusion de l'information au niveau géographique permettant d'optimiser les revenus des films.

Mots-clés : industrie cinématographique, modèles linéaires mixtes, critiques cinématographiques, asymétrie d'information, sources d'information externes

Remerciements

Premièrement, je tiens à remercier mes directeurs de mémoire, les professeurs Denis Larocque et Renaud Legoux, pour leurs conseils précieux qui m'ont permis de réaliser ce travail de recherche. C'est grâce à leur soutien et aux nombreux échanges que nous avons eu que j'ai pu découvrir comment mener à bien un projet de recherche tout en travaillant sur un sujet passionnant. J'adresse aussi des remerciements à tous les professeurs et les étudiants que j'ai rencontré lors de ma maîtrise, car c'est aussi grâce aux discussions que nous avons eu que j'ai pu réaliser ce mémoire. Évidemment, je remercie mes amis, ma famille et toutes les autres personnes qui ont joué un rôle, de près ou de loin, dans ce travail. Je voudrais particulièrement remercier Patricia, Sonia, Anaïs, Philippe, Marie-Pier et tous les amis de la maison du 3355.

Table des matières

Sommaire	i
Remerciements	iii
1 Introduction	1
2 Revue de la littérature	5
2.1 Biens expérientiels	6
2.2 Recherche d'information	8
2.3 Sources d'information	10
2.3.1 Cadre de l'industrie cinématographique	11
2.3.2 Critiques cinématographiques	12
2.3.3 Effet de prédiction et effet d'influence	13
2.4 Autres sources d'informations	17
2.4.1 Publicité	17
2.4.2 Bouche-à-oreille	18
2.5 Caractéristiques des films	19
2.5.1 Visa d'exploitation	19
2.5.2 Star power	21
2.5.3 Saisonnalité	22
2.5.4 Distributeur	22
2.5.5 Genre	23
2.5.6 Durée	24
2.5.7 Synthèse de la revue de littérature	25
3 Hypothèses de recherche	26
4 Méthodologie	29

4.1	Cadre conceptuel	29
4.2	Outils d'analyse et de manipulation des données	29
4.3	Données	31
4.3.1	Base de données CINÉAC	32
4.3.2	Enquête Sociale Générale Cycle 19 : Emploi du temps	38
4.4	Assemblage des bases de données	41
4.4.1	Base de données sur mesure et Cansim	41
4.4.2	Assemblage final	42
4.5	Variables utilisées	44
4.5.1	Variables dépendantes	44
4.5.2	Variables d'intérêt	44
4.5.3	Covariables	46
4.6	Modèles	52
4.6.1	Structure des données	53
4.6.2	Modélisation des effets fixes et des effets aléatoires	54
4.6.3	Estimation avec le logiciel R	60
4.6.4	Algorithmes et techniques d'estimation utilisées	62
4.6.5	Vecteurs de variables de contrôle	62
5	Analyses et résultats	64
5.1	Statistiques descriptives	64
5.2	Paramètres de variance-covariance	71
5.2.1	Sélection de la structure	71
5.3	Effets fixes	73
5.3.1	box-office composé et interprétation graphique	76
5.3.2	Analyse comparative des modèles estimés	80
5.3.3	Comparaison avec la littérature	81
5.3.4	Nombre d'écrans à la première semaine	87
5.3.5	Synthèse de l'analyse	87
6	Conclusion	89
6.1	Objectifs de recherche	89
6.2	Apports du modèle	90
6.3	Limites et avenues de recherche	91
A	Modèles estimés	93

B Tableaux des paramètres estimés

96

Table des figures

2.1	Visa d'exploitation MPAA	20
4.1	Cadre conceptuel	30
4.2	Base de données ACCESS pour les films	33
4.3	Base de données ACCESS pour les cinémas	34
4.4	Structure des données disponibles	40
4.5	Cinémas retenus et minutes de lecture de journaux	43
4.6	Structure de corrélation	57
5.1	Éléments de marque	65
5.2	Caractéristiques des films	67
5.3	Saisonnalité et cote	69
5.4	Variables continues	70
5.5	Effets marginaux du nombre de minutes de lecture pour les semaines subséquentes	77
5.6	Effets marginaux du nombre de minutes de lecture pour la première semaine	78

Liste des tableaux

4.1	Semaine d'avant première - Première condition	37
4.2	Semaine d'avant première - Deuxième condition	38
5.1	Distribution des modalités des variables concernant les éléments de marque	66
5.2	Distribution des modalités de la variable visa	66
5.3	Distribution des modalités de la variable genre	67
5.4	Distribution des modalités de la variable durée	68
5.5	Distribution des modalités de la variable distributeur	68
5.6	Distribution des modalités de la variable mois	69
5.7	Choix de la structure de varance-covariance- Critères d'information	72
5.8	BOC - Test des effets de type III (Satterthwaite)	74
5.9	BOC - Effets fixes	75
5.10	Comparatif des effets fixes	80
B.1	BOC - Effets fixes	96
B.2	BOC - Simple Coding	99
B.3	BOC - Estimateurs des variances	102
B.4	BOPS - Estimateurs des variances	102
B.5	BOPS - Tests des effets de type III	103
B.6	BOPS - Effets fixes	104
B.7	BOSS - Estimateurs des variances	105
B.8	BOSS - Tests des effets de type III	105
B.9	BOSS - Effets fixes	106
B.10	BOT - Estimateurs des variances	107
B.11	BOT - Tests des effets de type III	107
B.12	BOT - Effets fixes	108

Chapitre 1

Introduction

Le rôle et l'effet des sources d'information externes sur le succès commercial de différents produits et services ont suscité un vif intérêt dans la littérature de nombreuses sphères des sciences humaines (Debenedetti, 2006). Ces sources d'information externes sont consultées par les consommateurs dans les situations où ils doivent choisir de consommer un ou plusieurs produits disponibles dans un ensemble de choix mis à leur disposition. Les personnes influentes qui émettent des critiques concernant les biens et services sont donc prises en compte dans le processus de choix du consommateur. En effet, ces sources d'informations sont des reflets de l'opinion d'une personne considérée comme experte du marché et des produits en question. Les critiques sont donc des éléments qui peuvent être déterminants dans le choix de consommation. Elles constituent un des vecteurs d'information sur les choix disponibles dans de nombreuses industries (Debenedetti, 2006). Ces sources d'information sont généralement indépendantes du circuit de production et de distribution et plus réputées que les sources comme le bouche-à-oreille (Debenedetti, 2006). L'augmentation de la taille de l'ensemble de choix du consommateur a motivé l'apparition de sources d'information externes de plus en plus efficaces afin de diriger chaque personne vers une alternative spécifique.

L'industrie cinématographique offre un terrain d'étude propice à la mesure de l'impact de ces sources d'information sur le succès commercial des produits disponibles sur le marché. En effet, les films ont un court cycle de vie, sont des produits uniques et ont des caractéristiques esthétiques qui sont complexes à évaluer *ex-ante* consommation. Cette industrie a généré à elle seule des revenus de 1,6 milliard de dollars en 2010 (Statistique Canada 2012). La marge bénéficiaire d'exploitation n'a cessé d'augmenter entre 2007 et 2010 au Canada et au Québec (Statistique Canada, 2012). Il est donc pertinent de proposer de nouvelles dimensions à exploiter afin d'en améliorer la compréhension. Il existe un avantage certain à considérer de nouvelles méthodes pour soutenir la rentabilité de ce domaine des arts.

La littérature en marketing a pu se développer autour de l'industrie cinématographique étant donné la disponibilité des données et leur fréquence de publication régulière. De nombreux auteurs ont tiré des conclusions concernant l'effet des critiques cinématographiques sur le box-office des films (Reinstein et Snyder 2005 , Gazley, Gemma et Sinha 2011, Desay et Basuroy 2005 , Eliashberg et Shugan 1997, Basuroy, Chatterjee et Ravid 2003). Ils décomposent l'effet de ces sources d'information en deux composantes : l'effet d'influence qui est propre à la propriété des critiques d'influencer la décision des consommateurs et l'effet de prédiction qui se définit comme étant la propriété des critiques de refléter les préférences des consommateurs. La littérature a cherché à attribuer l'un ou l'autre des effets aux critiques cinématographiques et l'on trouve surtout des déclinaisons de l'effet des facteurs d'intérêt par rapport au temps.

Malgré les nombreux articles scientifiques qui sont parus, les recherches passées considèrent les critiques comme étant diffusée d'une manière homogène à travers la population. Sous cette hypothèse, les consommateurs reçoivent le même niveau d'information et la dimension de l'exposition

aux critiques n'est pas prise en compte.

Il est possible que certains groupes de consommateurs soient plus au moins exposés aux critiques selon la fréquence à laquelle ils ont accès à cette information. Il existe de grandes chances pour que cette dimension soit importante dans la modélisation de la demande de films.

La problématique de ce mémoire est de quantifier dans quelles mesures l'exposition des consommateurs aux critiques cinématographiques modère les composantes de l'effet de la valence des critiques sur le box-office.

Nous analysons l'effet de prédiction et l'effet d'influence en prenant en compte l'exposition à ces sources d'informations externes à travers le temps. Nous introduisons des variables reflétant l'exposition des consommateurs aux critiques cinématographiques en exploitant le caractère géographique des données disponibles. L'éventail de mesures de l'exposition aux critiques est en effet associé à des régions administratives spécifiques. Les modèles linéaires mixtes que nous proposons sont comparés aux alternatives utilisées dans la littérature et nous mettons en opposition les différents résultats.

Nos résultats théoriques viennent renforcer la littérature en marketing qui s'intéresse au comportement du consommateur en approchant la recherche d'information au niveau temporel et au niveau spatial. Nous réexaminons ainsi la dichotomie entre l'effet de prédiction et l'effet d'influence des critiques cinématographiques avec un cadre plus flexible et aussi plus précis. Nos résultats empiriques permettent de mesurer l'impact de différentes intensités d'exposition au niveau géographique et à travers les temps. Les modèles estimés permettent de prédire la réaction du box-office selon la région où est diffusé un film en utilisant la note que les critiques lui attribuent. Ils apportent une précision accrue afin de mettre en place des stratégies de diffusion de l'information au niveau géographique permettant d'optimiser les revenus des films.

Pour y parvenir, nous commençons par explorer l'arborescence de la littérature appliquée à la recherche d'information et à l'industrie cinématographique. Nous allons comprendre les mécanismes sous jacents au processus de consommation de bien expérientiels puis nous irons documenter la recherche et la construction de notre base de données finale qui a servi à l'analyse. Par la suite, nous nous pencherons sur la modélisation théorique de la demande de films qui nous servira à tester nos hypothèses de recherche. Nous analyserons ensuite les résultats de la modélisation en prenant soin de comparer l'effet que nous estimons avec notre nouveau modèle aux effets estimés avec les modèles utilisés couramment dans la littérature. Nous viendrons clôturer les conclusions tirées de l'analyse comparative par une discussion sur la portée et les améliorations souhaitables au cadre théorique que nous avons emprunté et au modèle statistique estimé.

Chapitre 2

Revue de la littérature

De nombreuses études ont approché les facteurs déterminant le succès commercial des biens expérientiels comme les films cinématographiques. Cependant, peu d'auteurs ont utilisé un cadre théorique simple et une manière de représenter visuellement les résultats permettant aux analyses d'amener un levier considérable dans l'explication des phénomènes sous-jacents aux effets trouvés (Chang et Ki, 2005). Cet exercice est un atout précieux qui permet d'exprimer les résultats de la recherche d'une manière intuitive. Alors que Chang et Ki (2005) et Reddy et al. (1998) ont commencé à développer un schéma théorique autour des biens expérientiels, nous l'enrichissons grâce à une base de données et une manne de facteurs et de variable de contrôle plus précis.

Afin de bien approcher la recherche d'information d'un consommateur *ex-ante* consommation d'un produit, nous commencerons par introduire la nature et les attributs des biens expérientiels comme les produits artistiques. En effet, la recherche d'informations que les consommateurs doivent acquérir quand ils sont face à un choix concernant ce type de biens est centrale dans notre analyse. Nous présenterons parallèlement les chemins empruntés par la littérature en marketing appliquée à l'industrie

cinématographique afin de nous diriger vers les sources qui permettent au consommateur de trouver de l'information. Ces sources d'information prendront une place importante, premièrement en tant que critère de choix d'un produit et deuxièmement en tant qu'indicateur des préférences des consommateurs et donc de la performance commerciale de ce produit. Nous continuerons graduellement à introduire des éléments du contexte de l'industrie cinématographique afin de cadrer le rôle des sources d'informations et plus particulièrement des critiques cinématographiques. Après avoir opposé l'effet d'influence et l'effet de prédiction de celles-ci, nous recenserons les facteurs propres aux films qui déterminent leur succès au box-office et que nous contrôlons lors de notre analyse.

2.1 Biens expérientiels

Dans un premier temps, nous nous penchons sur les origines de l'appréciation de ce type particulier de produit pour comprendre plus loin dans cette revue de littérature quels sont les processus sous-jacents à leur consommation. Les biens expérientiels sont des objets difficiles à étudier en marketing car leur évaluation par le consommateur est complexe. Les premières réflexions sur la consommation de tels types de produits ont été élaborées dans un contexte de recherche exploratoire. Dans les deux premières études publiées, Levy (1959) explore le comportement des consommateurs au niveau de la partie symbolique des produits tandis que Berlyne (1971) réalise une expérience concernant l'esthétique qui se réfère à la conception du « beau » par les sens. La littérature contemporaine (Hirschman et Holbrook 1982 , Holbrook et Hirschman 1982 , Eliashberg et Sawhney 1994) nous permet d'arriver à une description précise de l'appréciation de ces produits. Les biens expérientiels développent chez le consommateur des aspects émotifs et fantaisistes à travers des sens-

tions multiples qui sont apparentées à l'aspect hédonique des produits.

Les produits artistiques comme les oeuvres cinématographiques sont soumis à une dynamique de choix particulière de la part du public. Selon Hirschman et Holbrook (1982), la perception idéale de la réalité par le consommateur, qui est la manière subjective dont celui-ci se représente le produit, est un moteur de ce choix. Le consommateur est donc mentalement plus sollicité lors de ce processus, car il prend en compte la signification symbolique des caractéristiques non tangibles des objets comme la beauté, la profondeur ou la qualité (Holbrook et Hirschman, 1982). Les films l'engagent dans l'amusement, les loisirs et le plaisir qui y sont associés et qui définissent l'expérience qu'il vit. Les deux auteurs proposent que le consommateur choisisse d'allouer son capital mental parmi les différentes alternatives qui s'offrent à lui ou qu'il puisse également le conserver pour l'utiliser dans le futur. Selon Eliashberg et Sawhney (1994), « la consommation hédonique est une interaction entre les individus et le produit expérientiel » (traduction libre). Dans leur étude, les auteurs modélisent l'expérience du public visionnant un film. Plus précisément, ils définissent l'appréciation de l'expérience comme dépendante de l'interaction de l'humeur temporelle des consommateurs, des différences individuelles (personnalité) et du contenu émotionnel de celle-ci. L'élément de maximisation du profit économique est vu comme n'étant qu'une partie non dominante dans le choix des individus alors que l'expérience de consommation constitue la partie la plus importante (Reddy et al., 1998 ; Chang et Ki, 2005).

La deuxième caractéristique notable des biens expérientiels réside dans la difficulté pour un consommateur de leur attribuer une valeur avant de les avoir consommés (Shapiro et Varian, 1999). De ce fait, des éléments extérieurs à la relation entre le consommateur et les biens expérientiels peuvent jouer un rôle dans le choix de consommation. La dynamique de

l'expérience de consommation, qui est immédiate, peut être influencée par une attente concernant l'appréciation du film relative à une exposition à des sources d'influences externes telles que la publicité, le bouche-à-oreille et les critiques (Eliashberg et Sawhney, 1994). Ces sources d'influence externes peuvent être rencontrées simplement lors de la consommation de produits complémentaires, mais aussi lors de la recherche d'information par les consommateurs. Cette étape, qui précède l'expérience, y joue un rôle non négligeable lors du processus complet de consommation et nous la développons dans la prochaine partie.

2.2 Recherche d'information

Un consommateur a généralement besoin d'information quand il projette d'acheter un bien ou un service. Pour retracer son parcours lors de sa recherche d'information, nous nous penchons sur les types d'informations qui sont liés à la consommation et sur l'intensité de la recherche d'information en reprenant les travaux effectués en économie et en marketing. Murray (1991) part de la littérature existante (Bettman 1979, Leigh et Rethans 1984, Lynch et Srull 1982, Hansen 1972, Jacoby, Chestnut, et Silverman 1977, van Raaij 1977) et définit deux types de sources d'information. La recherche d'information interne réfère à une recherche d'expérience passée du consommateur dans sa mémoire. La recherche d'information externe est définie comme une motivation consciente du consommateur d'enrichir sa connaissance de la consommation d'un produit et de son environnement. Nelson (1970), avance que l'information concernant un produit peut être acquise soit en cherchant de l'information le concernant, soit en ayant eu une expérience permettant d'inférer sa qualité sur un produit similaire.

Ces deux activités de recherche d'information du consommateur sont

motivées par sa volonté de réduire le risque de consommer un produit. En effet, le risque perçu dépend de l'évaluation qu'il fera d'un produit par rapport aux autres alternatives en prenant en compte le coût de chaque alternative (Murray, 1991). C'est le niveau de risque perçu qui va influencer le besoin d'information du consommateur et ainsi définir indirectement le volume de recherche et le type de sources qu'il va préférer (Cox 1967 p. 604, Capon et Burke 1977, Locander et Hermann 1979, Lutz et Reilly 1973). Dans le cas des biens expérientiels, le consommateur est confronté à des biens hétérogènes et la plupart du temps nouveaux. Il est donc dans un contexte où son seul choix est d'obtenir de l'information qui provient de sources externes. De plus, sa recherche d'information sera plus grande et il va préférer les sources d'information expérientielles car il n'a jamais pu avoir l'expérience de consommer ce produit et son risque perçu est élevé (Lutz et Reilly 1973). Ces éléments pourront alors forger la connaissance du consommateur et par apprentissage lui faciliter l'acquisition et la compréhension de nouvelles informations concernant le produit (Hoyer et MacInnis 2001).

Nelson (1970) reprend les bases de la théorie de recherche d'information de Stigler (1961) afin de pouvoir poser un cadre théorique pour expliquer comment un individu peut optimiser sa recherche d'information. Selon lui, un consommateur peut trouver le nombre de recherches optimales qu'il devra faire. Il devra alors simplement arrêter de chercher quand le bénéfice marginal de chercher une fois supplémentaire excédera son coût. Dans un autre contexte, Alba et Hutchinson (1987) définissent l'arrêt de la recherche d'information par l'accumulation suffisante de connaissances sur le produit. Les caractéristiques généralement utilisées lorsque l'on s'intéresse à la recherche d'information sont la valence et le volume. La valence fait référence à la capacité d'une information de synthétiser la qualité du produit d'une manière positive ou négative tandis que le

volume correspond à la quantité d'information disponible sur un produit sur une période donnée. Selon ces différentes manières de considérer la recherche d'information, nous pouvons dégager plusieurs renseignements pertinents pour la présente recherche. Sachant que le prix d'un billet de cinéma n'est pas très élevé et homogène entre les films de même qualité technique, il est possible que les consommateurs ne s'engagent pas dans de nombreuses recherches et sélectionnent peu de sources d'information. Selon ce raisonnement, nous nous attendons à ce que les consommateurs recherchent de l'information par des sources qui sont facilement accessibles, comme les journaux quotidiens. La recherche d'information nécessite un contact avec des sources d'information qui peuvent avoir différentes caractéristiques. Pour comprendre les implications du type de sources d'information sur le box-office, nous explorons par la suite des résultats empiriques importants de la littérature.

2.3 Sources d'information

Les consommateurs désireux d'acquérir de l'information sur les biens expérientiels disposent de diverses sources qui ont été mises au jour dans la littérature en marketing. Nous allons premièrement nous pencher sur des résultats confirmés par plusieurs études avant d'analyser les sources d'information séparément. Faber et O'Guinn (1984) affirment que les éléments importants lors de la recherche d'information sur un film sont les extraits, les aperçus (*previews*) et les bandes-annonces de films (*trailers*). Ces éléments qui font partie de la stratégie publicitaire sont plus importants que le bouche-à-oreille (*word-of-mouth*) et les critiques selon les individus interrogés dans leur sondage.

Levene (1992) montre que les critiques positives sont des sources d'information qui sont utilisées par les consommateurs. Dans son expérience

sur des étudiants, cette variable est positionnée derrière le jeu d'acteur, l'intrigue, les commentaires en provenance d'un ami (bouche-à-oreille). Chen, Liu, et Zhang (2011) confirment l'existence d'un effet positif des sources d'information externes sur les revenus du box-office des films. Dans le prolongement de ces résultats, Murray (1991) arrive à la conclusion qu'en général, les stratégies visant à simuler les sources personnelles d'information comme le bouche-à-oreille et les recommandations d'un cercle de connaissance sont bénéfiques aux services.

Globalement, nous observons dans la littérature que trois grandes sources d'information qu'il est possible d'exploiter ressortent. En effet, la mesure d'éléments comme la qualité et le nombre de diffusions en salle de *trailers* ou de bandes-annonces étant compliqués à mesurer, nous utilisons les éléments connus dans la littérature. Nous considérons donc la publicité, le bouche-à-oreille et les critiques cinématographiques. Pour comprendre comment la littérature en marketing s'est développée autour des différentes sources d'information et de l'industrie cinématographique, nous arborons les grandes lignes de la recherche déjà effectuée dans la prochaine partie de cette revue de littérature.

2.3.1 Cadre de l'industrie cinématographique

Dans la littérature en marketing appliquée à l'industrie cinématographique, le comportement du marché et des acteurs ont été analysés principalement sous deux angles différents. Premièrement, la littérature s'intéresse directement au comportement des consommateurs lors de leur processus de choix en ayant comme unité d'observation les individus (Wyatt et Badger 1984, Desai et Basuroy 2005, D'Astous et Touil 1999, Vézina 1997, D'Astous et Colbert 2002, Gemma et Sinha 2011). Les chercheurs se placent donc à un niveau microscopique afin de déterminer quels sont les facteurs influençant directement les choix des consommateurs.

La deuxième approche, que nous utilisons dans le cadre de ce mémoire, consiste à analyser à un niveau macroscopique l'industrie du cinéma en prenant comme unité d'observation les films (Litman 1983 , Hirschman et Pieros 1985, Litman et Kohl 1998, Wallace, Seigerman et Holbrook 1993, Sochay 1994, Eliashberg et Shugan 1997, Ravid 1999). Ce sont ici les variables composant la demande de films qui sont mises en relation et analysées afin d'estimer leurs effets sur la performance des films au box-office. La prochaine partie vient ancrer l'élément central de ce mémoire et développe les caractéristiques inhérentes aux critiques cinématographiques.

2.3.2 Critiques cinématographiques

Les premières ébauches de recherche concernant les critiques cinématographiques (non professionnels) apparaissent dans l'ouvrage de Katz et Lazarsfeld (1955, p. 307). Les deux auteurs essaient de comprendre qualitativement quelles sont les caractéristiques des individus qui sont considérés comme des critiques et quelle est leur influence sur les personnes qui y sont exposées. Plus tard, une étude formelle de Litman (1983) conclura que les critiques constituent l'une des trois variables les plus importantes dans la prédiction du box-office agrégé. Dans le prolongement, Wyatt et Badger (1984) essaient d'identifier si l'influence des critiques a un effet sur le public et tentent de mesurer cet effet. Ils conduisent une expérience sur des étudiants afin de dériver leurs résultats. Ceux-ci indiquent que la direction de la critique a un effet significatif sur l'intérêt porté au film et sur son évaluation. Plus récemment, Wallace, Seigerman et Holbrook (1993) mettent en avant le fait que la variable concernant les critiques incluse dans leur régression a un effet quadratique sur le revenu de location d'un film. Un mauvais film aurait ainsi intérêt à ne pas avoir une note positive de sa critique, mais plutôt à avoir une critique la plus mauvaise

possible alors qu'un bon film aurait besoin d'avoir une excellente critique pour mieux performer.

2.3.3 Effet de prédiction et effet d'influence

Dans la littérature, deux rôles ont été attribués aux critiques. D'une part, elles sont considérées comme influençant les décisions des consommateurs quant au choix d'un film et indirectement la performance commerciale des films (Wyatt et Badger 1984, Wyatt et Badger 1990, D'Astous et Touil 1999). D'autre part, elles sont décrites comme ayant un pouvoir de prédiction du box-office des films (Litman et Kohl 1989, Prag et Casavant 1994, Eliashberg et Shugan 1997, Zuckerman et Kim 2003, Chang et Ki 2005, Terry, Butler et De Armond 2005). La littérature est divisée entre ces deux effets en ce qui concerne les critiques cinématographiques et la plupart des études se positionnent dans l'une ou l'autre des directions. Cependant, Basuroy, Chatterjee et Ravid (2003) estiment que les deux effets peuvent coexister. Nous relatons consécutivement les différentes études qui ont un lien direct avec l'effet de prédiction et l'Effet d'influence.

Effet de prédiction

Dans leur étude empirique sur le rôle des critiques, Eliashberg et Shugan (1997) examinent le potentiel des critiques de pouvoir prédire ou influencer le box-office. Ils remarquent que les critiques cinématographiques sont, à un niveau agrégé, des indicateurs à moyen terme de la performance des films plutôt que des leaders d'opinion. Leurs résultats ne soutiennent pas l'effet d'influence. L'effet le plus marqué des critiques n'est pas observé dans la semaine qui suit la publication de ces sources d'information. Il ne semble donc pas exister de lien causal directement engendré par un effet des critiques marqué entre deux périodes. C'est plutôt en tant qu'in-

dicateurs qu'ils considèrent les critiques, car elles ne sont pas corrélées fortement avec le box-office des premières semaines, mais plutôt avec le box-office total cumulé. Cet indicateur reflète donc à quel point le public de films va l'apprécier.

Considérées ainsi, les critiques ont des caractéristiques particulières. Premièrement, elles peuvent s'adapter au goût du public à travers le temps afin de devenir et rester populaires. Deuxièmement, les magazines, journaux et sites internet peuvent choisir les critiques en fonction de leur opinion qui correspondrait aux opinions de leur lectorat. Dans les deux cas, cette congruence d'opinion entre les critiques et le lectorat de magazines fait partie du schéma théorique pouvant expliquer l'effet de prédiction de celles-ci.

Effet d'influence et sensibilité du consommateur aux critiques cinématographique

Il est possible de penser que les critiques cinématographiques influencent les revenus du box-office. Le critique peut être vu comme un leader d'opinion ou une personne influente. Cette personne est considérée par un groupe ou par d'autres personnes comme ayant une expertise ou un savoir supérieur sur un sujet particulier. Debenedetti (2006) décrit la critique comme « une source d'information ou de prescription ». La critique est donc vue comme étant un élément de « réassurance » au même titre que les bandes-annonces et les extraits. En tant que source d'information externe pour le consommateur, la critique a un rôle indépendant, car elle n'a pas de portée commerciale à l'inverse de la publicité. Cette indépendance est importante, car elle place l'information à un niveau qui peut être considéré comme authentique puisque l'appât du gain n'apparaît pas dans la sélection et le jugement des oeuvres. De plus, elle traduit une expérience vécue par une personne, car son auteur s'efforce de décrire les

aspects intangibles du produit. Cette traduction de concepts abstraits à une description plus complète est composée d'éléments appartenant au moins à trois catégories (Shrum 1991). La « prescription de fait » amène des éléments descriptifs et informatifs. La « prescription technique » pose un cadre contextuel afin de pouvoir analyser et comprendre l'oeuvre en question. La « prescription de jugement » permet d'évaluer positivement ou négativement l'oeuvre.

Le degré d'implication du public vis-à-vis du cinéma, qu'il est possible de mesurer par l'intensité de la fréquentation des cinémas, peut être une bonne base pour évaluer la consultation des critiques. Vézina (1997), avance aussi que les critiques ont plus d'influence que les proches si l'implication est plus marquée. En opposition, d'Astous (1999) et d'Astous et Colbert (2002) découvrent une relation négative entre la consultation des critiques et l'implication mesurée avec l'échelle de Zaichkowsky (1985). La connaissance et l'expertise des individus concernant le cinéma sont des aspects importants affectant la consultation des critiques. Vézina (1997) et d'Astous et Colbert (2002) montrent que ces deux caractéristiques sont liées à la considération de la critique cinématographique. La relation qui est positive décrit un comportement de « consommation addictive » endogène dans le sens où la lecture de critique rend l'individu un peu plus expert et son expertise lui permet de retirer une plus grande utilité de cette lecture.

Un autre axe à considérer part de la théorie de la comparaison sociale de Festinger (1954). Selon lui, un individu qui est en situation d'incertitude va avoir tendance à consulter des sources d'information qui peuvent être représentées par un ou plusieurs individus qui auront le même degré d'expertise que lui. Dans le cadre décisionnel du consommateur qui fait face à un ensemble de films et un ensemble de sources d'information données, nous pouvons penser que l'individu peu connaisseur ou qui prétend

faire partie du « grand public » aura tendance à se référer au bouche-à-oreille. Les individus qui se considèrent comme connaisseurs vont quant à eux s'intéresser à des sources qui proviennent d'experts, tels que les critiques. En effet, les critiques de cinéma professionnels ne représentent pas nécessairement les goûts de la population qui se propagent par le bouche-à-oreille selon Chakravarty, Liu et Mazumdar (2010).

En opposition, la fréquence des visites des cinéphiles au cinéma peut influencer la prise en compte de ces deux vecteurs d'information. D'Astous (1999), découvre qu'il peut exister une relation concave entre la consultation des critiques et le niveau d'expertise. Debenedetti (2006), cadre cet effet dans le sens du transfert de connaissances du prescripteur au consommateur discuté par Hatchuel (1995). Cette relation peut être justifiée par le fait qu'à partir d'un certain niveau, le consommateur a acquis assez d'information, en se documentant sur le cinéma via des sources considérées comme expertes, qu'il peut se faire plus facilement sa propre idée sur une oeuvre. Plus en détail, d'Astous et Touil (1999) trouvent qu'une critique va avoir un impact significatif si elle est en contradiction avec le jugement qu'un critique a l'habitude d'émettre.

Elberse et Eliashberg (2003) observent que l'effet entre les critiques et les revenus du box-office est significatif la première semaine. La traduction de cet effet au niveau méthodologique est que les critiques positives devraient avoir un impact positif sur les revenus du box-office de la première semaine. Les revues négatives devraient au contraire avoir un impact négatif sur les revenus du box-office de la première semaine. Gazley, Gemma et Sinha 2011 confirment l'effet d'influence de la critique sur la décision d'achat d'un billet de cinéma. L'effet qu'ils estiment est révélé comme étant plus faible que l'effet du bouche-à-oreille.

Conclusion entre les deux approches

La littérature est divisée entre l'effet d'influence et l'effet de prédiction des critiques cinématographiques et cette division semble aussi distincte par rapport à l'approche utilisée, le niveau d'agrégation et la taille de l'échantillon disponible. Nous pouvons trouver des explications plus précises sur les phénomènes sous-jacents à ces deux effets. Il existe très probablement des mécanismes non observables dans la prise en compte des critiques par le consommateur. Ces deux points de vue sont à considérer, car il est possible que les critiques puissent à la fois refléter les préférences du public et qu'elles puissent aussi influencer le choix des individus dépendamment de leur niveau d'exposition (Terry, Butler, D'Armond, 2011).

2.4 Autres sources d'informations

2.4.1 Publicité

La publicité ressort comme étant un outil important dans la promotion de films. Elberse et Eliashberg (2003) trouvent une relation statistiquement significative entre la publicité et les revenus. Ils décrivent la publicité comme étant une variable influençant indirectement les revenus du box-office à travers son impact sur la décision du distributeur de diffuser le film dans plus ou moins de salles (allocation d'écrans). La variable a donc un effet de prédiction du succès financier d'un film et du nombre d'écrans pour la semaine d'ouverture (movie marketability). Prag et Casavant (1994), dans une recherche sur le marché des États-Unis, montrent que l'investissement en marketing peut modérer les effets des variables incluses dans l'analyse. Généralement, ils remarquent que les coûts de production, la note du film telle que décrite par la MPAA, la présence de star(s) sont importants seulement si la variable reflétant les dépenses

en publicité n'est pas incluse. Il est donc possible que l'effet de ces variables soit en réalité dû à la publicité. L'effet de la publicité est selon leur étude un facteur ayant un impact important sur le nombre de places vendues. Il est supérieur à l'effet du budget et supérieur à l'effet des stars et des récompenses. Zufryden (1996) utilise quant à lui un modèle liant les dépenses en publicité à la sensibilisation, à l'intention d'aller voir un film et aux projections des ventes pour confirmer l'impact notoire que la publicité apporte sur le box-office.

2.4.2 Bouche-à-oreille

Dans le cadre de l'industrie cinématographique, le bouche-à-oreille, reflète la capacité du film à retenir un public à travers sa durée de diffusion. Le bouche-à-oreille est incarné par les commentaires des consommateurs vis-à-vis d'un film. Ces commentaires peuvent être retrouvés en ligne ou des espaces de partage et des systèmes de votes (par exemple sur la plateforme IMDb) permettent aux consommateurs de s'exprimer et de partager leur opinion sur un film. Le bouche-à-oreille est considéré dans la littérature comme une source d'information privilégiée par les consommateurs et peut donc être représenté de plusieurs manières. Le volume (le nombre de recommandations) du bouche-à-oreille a un impact sur le nombre de consommateurs qui connaîtront le produit. La valence a un lien avec l'effet de persuasion que les revues en ligne peuvent avoir sur les consommateurs. Liu (2006) confirme l'impact du volume sur les revenus du box-office des films, mais la valence ne semble pas avoir d'impact avec la configuration de sa base de données et les outils statistiques utilisés.

2.5 Caractéristiques des films

2.5.1 Visa d'exploitation

Le visa d'exploitation est inclus dans notre modèle, car il peut donner de l'information (un signal pour les consommateurs) sur des éléments concernant les préférences ou la sensibilité du public de films (Prag et Casavant, 1994). Selon cette logique, les films ayant un visa tout public auront donc une plus grande facilité à générer des entrées qu'un film ayant une restriction au niveau de l'âge. Ce mécanisme amène une possible corrélation entre le visa d'exploitation et le box-office. Les études antérieures sont basées sur une évaluation établie premièrement en 1968 par la Motion Picture Association of America (MPAA) et qui a pour but d'informer le public et les parents sur les contenus sensibles des films. Les visas d'exploitation de la MPAA sont décrits à la figure 2.1.

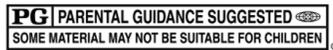
Dans la littérature, cette variable a été globalement confirmée comme étant un facteur prédictif du box-office (Litman 1983, Prag et Casavant 1994, Sochay 1994, Litman et Ahn 1998, Ravid 1999 ; Ravid et Basuroy 2004, Chang et Ki 2005, Brewer, Kelly, Jozefowicz 2009). Plus en détail, les recherches précédentes rapportent des effets ambigus pour les visas G, PG et R. Le visa PG13 a quant à lui l'effet le plus marqué et ressort comme étant la catégorie qui tend à générer des revenus plus élevés. L'explication de cet effet dans la littérature découle du constat de Prag et Casavant, qui expliquent que le public plus âgé dispose d'un revenu discrétionnaire plus élevé leur permettant de consommer ce type de film. De plus, ce public dispose d'une fenêtre temporelle plus large permettant aux cinémas de présenter les films avec un visa PG13 plus tard en soirée (Brewer, Kelly, Jozewicz, 2009). Chang et Ki (2005) concluent que c'est surtout l'étendue du public qui va varier avec le visa. Au Québec, c'est la régie du cinéma du Québec qui effectue le classement des films et nous

FIGURE 2.1 – Visa d’exploitation MPAA

GENERAL AUDIENCES



PARENTAL GUIDANCE SUGGESTED



PARENTS STRONGLY CAUTIONED



RESTRICTED



NO ONE 17 AND UNDER ADMITTED



repreons donc ce classement qui diffère peu de celui de la MPAA.

2.5.2 Star power

Dans le contexte où les acteurs présents dans les films peuvent être plus ou moins connus du public et peuvent refléter la qualité d'un film, nous allons décrire la place que prend cet élément de marque au sein de l'industrie cinématographique. Elberse (2007) mesure le retour sur investissement de la participation d'acteurs et d'actrices considérés comme étant des stars sur un marché d'action réel (Hollywood Stock Exchange) et sur un marché d'actions simulées. L'auteur découvre que les prix du HSX répondent significativement à l'annonce des acteurs auditionnés, soutenant directement l'hypothèse que l'implication des stars affecte positivement les revenus d'un film. Une star apporterait, selon les estimations obtenues, 3 millions de dollars de revenus supplémentaires en moyenne. La réputation des acteurs et leurs succès passés sont des variables influençant positivement leur impact sur le succès financier des films. Il existe de plus, un lien entre la popularité et le succès entre tous les acteurs d'un film impliquant une dynamique particulière au niveau du recrutement et de la rémunération des stars (Caves 2000, 2003). La présence d'un réalisateur jouissant d'une bonne réputation artistique aurait aussi pour effet d'influencer positivement la qualité perçue d'un film (Boor, 1992). D'un autre côté, plusieurs auteurs soutiennent que la présence d'une ou plusieurs stars (Austin, 1984; Sochay, 1994; Litman et Ahn, 1998) ne garantit pas la rentabilité du long métrage. C'est donc le film, considéré comme un tout avec certaines vedettes, qui semblent affecter leur rentabilité (Hennig-Thurau et al., 2007).

2.5.3 Saisonnalité

Nous emploierons le terme saisonnalité pour faire référence à la période de l'année durant laquelle est lancé le film. Sochay (1994) identifie la période de Pâques (mois de mars et avril), la période estivale (mois de mai à août) et la période des fêtes de fin d'année (mois de novembre et décembre). Litman (1983) avait déjà montré que la période des fêtes de fin d'année semblait appropriée pour la sortie d'un film tandis qu'Izod (1988) confirme ce résultat en ajoutant que la période estivale est aussi une période assurant un retour plus élevé lors du lancement d'un film. Simonoff et Sparrow (2000) identifient quant à eux les revenus sont bien plus élevés pour le troisième quartile en saison comparé aux revenus hors-saison.

De nombreuses études soutiennent que les cinémas ont un taux de fréquentation plus élevé durant les saisons mentionnées plus haut. L'effet saisonnier introduit ainsi une plus grande compétition lors de la sortie des films en saison, étant donné le nombre limité de salles disponibles. Simonoff et Sparrow (2000) mettent en avant que durant les fins de semaine en saison, les studios doivent avoir une gestion de leur logistique de diffusion efficace. Étant donné cette concurrence, ils émettent des signaux à la concurrence parfois une année avant la sortie afin de tenter de réserver une place à leur futur blockbuster potentiel.

Nous allons donc inclure dans cette étude des variables binaires qui traiteront l'effet saisonnier afin de le contrôler quand nous estimerons l'effet des variables auxquelles nous nous intéressons.

2.5.4 Distributeur

Étant donné la configuration particulière de l'industrie cinématographique, il est nécessaire de se pencher sur le rôle des distributeurs sur le marché.

Chang et Ki (2005), identifient un effet de la variable liée au pouvoir de marché des distributeurs sur la durée de présentation des films. Cette variable est définie comme le nombre de films de l'année en cours inclus dans le top 100 des films produits par un distributeur l'année précédente. Gemser, Oostrum et Leenders (2006) établissent une dichotomie entre les distributeurs majeurs et les autres distributeurs. Dans leur étude sur l'impact des critiques sur la performance au box-office des films d'auteur comparés aux films issus de la culture populaire (*mainstream movies*), la variable est significative et a un effet positif sur le box-office. Ce résultat est soutenu par les études antérieures de Litman et Kohl (1989) qui trouvent que les films distribués par un distributeur majeur génèrent des revenus plus élevés. D'un autre côté, Sochay (1994) et Litman et Ahn (1998) ne trouvent pas d'évidence statistique au niveau de la corrélation entre la taille du distributeur et les revenus liés à ses films. Sochay (1994) suppose qu'il pourrait exister une liaison entre le budget du film et le pouvoir des distributeurs à bénéficier de meilleures salles. Cet élément de distribution pourrait quant à lui avoir un impact direct sur le box-office. Selon les configurations des modèles utilisés, la littérature ne suggère pas une avenue claire concernant l'inclusion de facteurs représentant les caractéristiques des distributeurs. Nous préférons donc utiliser toutes les informations disponibles sur la distribution pour contrôler des relations possibles entre cette variable et le box-office.

2.5.5 Genre

La littérature concernant le genre des films est divisée et de nombreuses méthodes ont été utilisées pour déterminer si ce facteur pouvait être discriminant quant à la performance des films au box-office. Le genre s'est révélé être une variable ayant un pouvoir explicatif du box-office lors de son intégration dans les modèles utilisés dans la littérature en marketing.

En effet, pour Litman et Kohl (1989), Wyatt (1991), Sochay (1994) et Vézina (1994), le genre apparaît comme étant un facteur déterminant le succès financier des films. Par exemple, les films d'action et aventure, horreur et comédie ressortent comme ayant une capacité à générer un box-office plus élevé au Royaume-Uni (Elliot et Simmons, 2008). Les drames sont quant à eux les films les moins populaires aux États-Unis et au Royaume-Uni (Litman et Kohl, 1989 ; Chang et Ki, 2005). Jedidi et al. (1998) ne trouvent cependant pas d'évidence concernant le rôle du genre sur le succès commercial des films dans la catégorisation des films sélectionnés dans leur recherche. Chacun de leurs quatre groupes a des films de genres mixtes alors que leurs parts de marchés évoluent clairement différemment au cours du temps. Ravid (1999), Sochay (1994) et Ravid et Basuroy (2004) ouvrent cependant une discussion sur la classification stricte des films et l'aspect subjectif de la perception de ce classement par les consommateurs. Le film *Life of Pi* (2012) peut par exemple être classé dans la catégorie action et aventure, drame ou peut encore être considéré comme un film fantastique (classement IMDb). La classification par genre peut donc engendrer des problèmes de mesures qui dépendent de la subjectivité des personnes qui classent les films et du nombre de thèmes principaux y étant abordés.

2.5.6 Durée

Dans la littérature, la variable concernant la durée des films a rarement été explorée. Elle peut cependant faire référence à la concurrence sur le marché et notamment auprès des distributeurs qui doivent choisir d'allouer une fenêtre temporelle à chaque film au sein de leur établissement. Ainsi, la longueur d'un film peut contraindre un distributeur à ne pouvoir le diffuser qu'un certain nombre de fois en une journée (Bonnet, 2009). Elle peut aussi restreindre complètement la diffusion dans un éta-

blissement. La diffusion fait directement référence à la disponibilité géographique et temporelle du produit et la durée totale d'un film pourrait donc avoir un impact indirect sur le box-office.

2.5.7 Synthèse de la revue de littérature

Lors de cette revue de la littérature, nous avons répertorié les variables ayant une influence sur les revenus du box-office. Dans la prochaine partie, nous organisons les effets que nous voulons mettre au jour grâce à nos bases de données.

Chapitre 3

Hypothèses de recherche

Les différents résultats dans la littérature nous ont amené à exprimer d'une nouvelle manière les relations entre la disponibilité de l'information et le succès commercial des biens expérientiels. Premièrement, nous testons si l'effet de modération est différent d'une période à l'autre en prenant soin de suivre la littérature concernant le type d'aggrégation temporelle utilisée. Dans un deuxième temps, nous testons la présence d'un effet d'influence et/ou d'un effet de prédiction des critiques cinématographiques dans notre nouveau modèle. Nous utilisons une mesure de la cote de critiques, une mesure d'exposition aux critiques et un indicatif entre la première semaine et les semaines subséquentes (trois variables détaillées dans la partie méthodologique de cette étude) dans nos modèles. Nous posons les hypothèses suivantes pour identifier si les critiques partagent les préférences esthétiques des consommateurs ou si ce sont plutôt les consommateurs qui sont influencés par l'opinion des critiques.

H1 : L'effet des critiques cinématographiques est accentué par le degré d'exposition à celles-ci et différent entre les semaines subséquentes à la première semaine et la première semaine.

En effet, nous nous attendons à ce que les consommateurs qui sont plus

exposés à l'information aient une réaction relativement plus marquée au niveau de la cote par rapport aux consommateurs moins exposés à l'information. Afin de traiter cet effet d'une manière plus complète, nous vérifions si les réactions des consommateurs ne sont pas différentes selon la période où le film est présenté. Entre les semaines subséquentes et la première semaine, il serait possible que le contexte informationnel soit radicalement différent comme de nouvelles sources d'information sont disponibles après la sortie d'un film (bouche-à-oreille, nouveaux articles parus dans les journaux, etc.).

H2 : Les critiques influencent les revenus du box-office si les effets sont significatifs seulement pour la première semaine et non significatifs pour les semaines subséquentes à la première.

S'il n'existe qu'un effet d'influence, il est logique de dire que la relation entre nos variables d'intérêt et le box-office n'est présente que pour la première semaine. Ce serait ainsi d'autres variables qui détermineraient le succès des films après la première semaine.

H3 : Les critiques prédisent les revenus du box-office si les effets sont significatifs pour les semaines subséquentes.

D'un autre côté, s'il n'existe qu'un effet de prédiction, il est logique de dire que la relation entre les critiques et le box-office n'est présente qu'après la première semaine. Ce serait donc d'autres variables qui déterminerait le succès des films la première semaine.

H4 : Les critiques ont un effet d'influence et un effet de prédiction si les effets sont significatifs pour la première semaine et pour les semaines subséquentes.

Enfin, s'il y a présence d'un effet d'influence et d'un effet de prédiction, nous devrions observer un effet significatif de nos variables d'intérêt la première semaine, pour les semaines subséquentes à la première.

Afin de tester ces hypothèses de recherche nous utiliserons des bases de données et des outils que nous détaillons dans la prochaine section.

Chapitre 4

Méthodologie

4.1 Cadre conceptuel

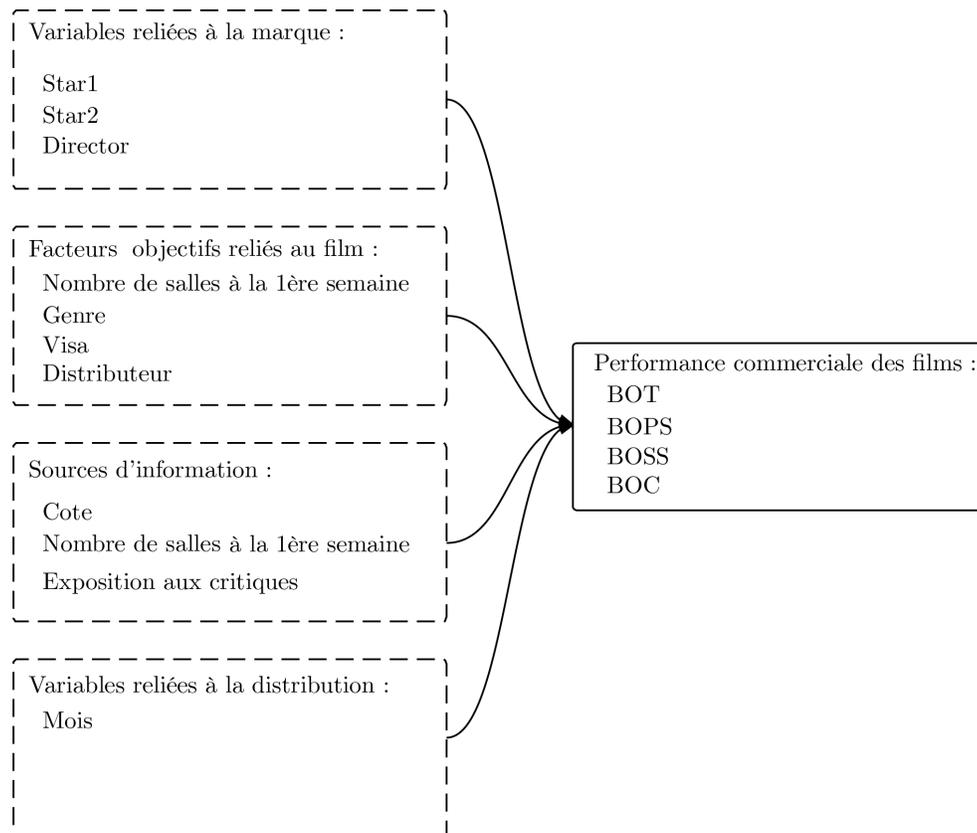
Nous enrichissons la représentation et la méthodologie de Chang et Ki (2005) afin d'intégrer de nouvelles variables en utilisant une base de données plus précise. Notre cadre conceptuel (figure 4.1) intègre donc quatre éléments clés :

- les variables reliées à la marque
- les facteurs objectifs liés aux films et des variables proxys
- les sources d'informations
- les variables reliées à la distribution

4.2 Outils d'analyse et de manipulation des données

Nous avons utilisé le logiciel SAS pour l'import et la manipulation des bases de données tandis que l'analyse a été effectuée avec le logiciel R version 3.0.0 (2013-04-03), 'Masked Marvel'. Nous utilisons en particulier

FIGURE 4.1 – Cadre conceptuel



les packages R suivants

- **lme4** (version 1.1-0 en développement, téléchargée sur Github le 01/10/2013), qui nous permet d’ajuster nos modèles linéaires mixtes
- **lmerTest** (version 2.0-0) qui nous permet d’accéder au calcul des valeurs-p en permettant la correction de Satterthwaite pour les degrés de liberté
- **pbkrtest** (version 0.3-7) qui nous permet d’accéder au calcul des valeurs-p en permettant la correction de Kenward-Roger pour les degrés de liberté et l’estimation des valeurs-p par bootstrap
- **ggplot2** (version 0.9.3.1) qui nous permet de représenter graphiquement des statistiques descriptives et les diagnostiques des régressions
- **xtable** (version 1.7-1) qui nous permet de convertir des sorties R en tableaux Latex
- **OpenStreetMap** (version 0.3.1) qui nous permet d’utiliser les cartes libres d’Open Street Map

4.3 Données

La problématique à laquelle nous voulons répondre introduit une nouvelle dimension qui permet d’évaluer l’effet des critiques. Plus précisément, nous mettons en relation l’exposition à l’information et la valence des critiques cinématographiques pour expliquer la performance des films au box-office. Afin de pouvoir tester l’hypothèse principale de ce mémoire, nous utilisons une base de données complexe, assemblée spécifiquement pour nos besoins. Nous reprenons les résultats des travaux de Bonnet (2009) et Belmati (2011) pour l’assemblage des bases de données et nous l’enrichissons de variables géographiques et de la variable spécifique concernant l’exposition à l’information représentée par le nombre de minutes de lecture de journaux quotidiens. Nous avons donc recours

à trois bases de données secondaires dans le cadre de ce mémoire. Premièrement, nous utilisons la base de données CINEAC qui couvre la population des films sortis dans tous les cinémas enregistrés au Québec en reprenant des éléments d'échantillonnage de Belmati (2011). Deuxièmement, nous utilisons des éléments de la base de données assemblée par Bonnet (2009) qui compile l'information concernant les acteurs et les réalisateurs des films sélectionnés. Troisièmement, nous créons une base de données sur mesure qui contient la variable de lecture de journaux moyenne de la population dans une subdivision de recensement, le code postal de chaque cinéma, sa latitude, et sa longitude. Ces trois éléments sont réunis afin de pouvoir estimer les paramètres de plusieurs modèles linéaires mixtes.

4.3.1 Base de données CINÉAC

Description

La base de données de l'entreprise CINÉAC est une source de données précise et fiable car elle rassemble des mesures objectives des recettes des salles de cinéma au Québec. Nous y retrouvons les mesures concernant le box-office (en dollars et par rapport au nombre d'entrées), l'origine du film, sa durée, son genre, le type de visa, etc (figures reffig :accesscine et 4.2). La version disponible nous donne accès à des films sortis entre 2001 et 2009. Nous allons dans un premier temps présenter la structure de cette base de données avant de décrire la compilation et la transformation de toutes les variables pour assembler la première grande base de données.

Base de données d'acteurs et réalisateurs

Les caractéristiques qui concernent le *star power* sont reprises de la collecte d'information de Bonnet (2009) et transformées afin de garder un

FIGURE 4.2 – Base de données ACCESS pour les films

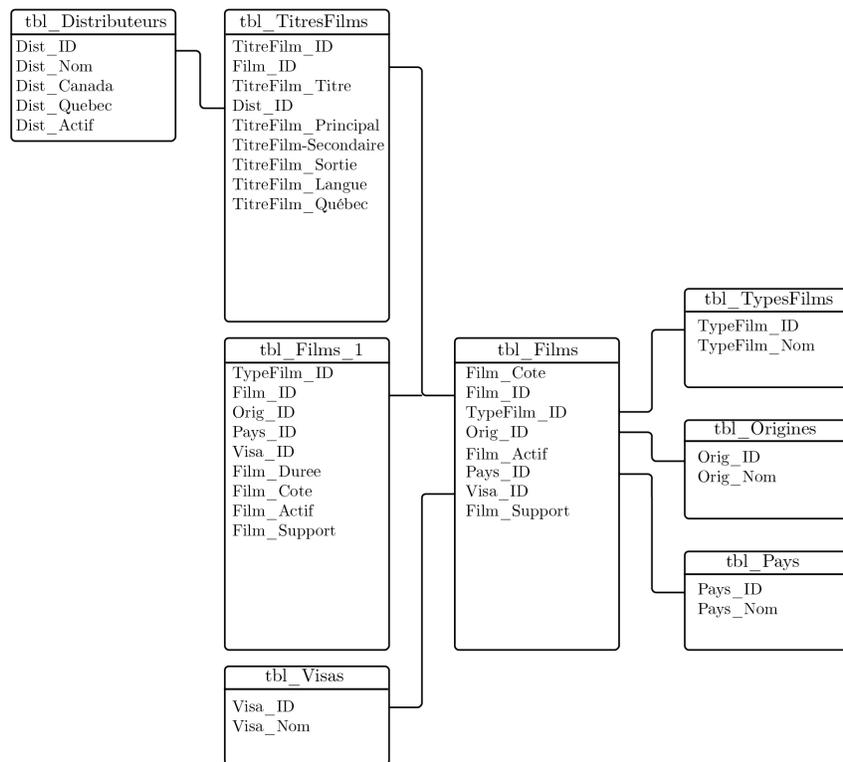
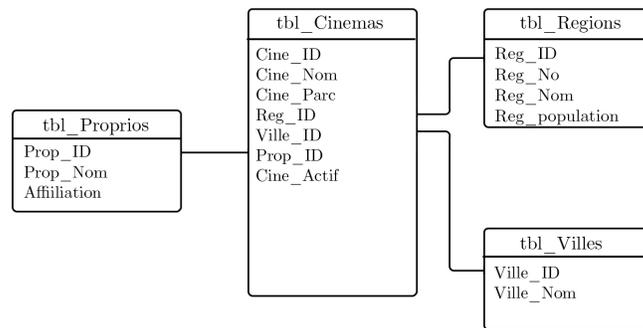


FIGURE 4.3 – Base de données ACCESS pour les cinémas



modèle parcimonieux. À la fin de notre manipulation, nous avons constitué une base de données rectangulaire, sans valeur manquante où chaque ligne représente un film identifié par la variable *Filmid*. Les prochaines étapes décrivent la manière dont nous avons procédé pour arriver à ce résultat.

Filtres et restrictions

La base de données telle que proposée par CINÉAC comporte 3274 films, 218 cinémas et 350 804 lignes de donnés. Pour mieux répondre à notre question de recherche, nous voulons avoir une population composée de cinémas ayant des caractéristiques reflétant un cinéma commercial qui s'adresse à un public relativement large. Nous utilisons donc plusieurs filtres afin de pouvoir retirer les cinémas qui ne sont pas actifs, les ciné-parcs et enfin les cinémas ne présentant pas au moins sept bobines au minimum par semaine. Nous retirons aussi manuellement certains cinémas ayant des critères pouvant altérer l'analyse.

En effet, nous remarquons que de nombreux petits cinémas ne présentent que peu de films. Cette caractéristique nous indique que ces cinémas sont plus à même de ne présenter qu'un certain type de films (souvent des films d'auteur) à un certain type de public (la plupart du temps des connaisseurs et habitués). La spécificité de ces cinémas les borne à avoir un box-office plutôt faible appuyant ainsi notre logique d'exclusion. Les ciné-parcs sont retirés quant à eux pour leur offre trop spécifique, trop faible et pour leur mise en conditions de consommation différente.

Agrégation des écrans par film et agrégation temporelle

La force de la base de données CINÉAC est de recenser les informations de chaque écran dans chaque cinéma par semaine. Nous reprenons

la méthodologie de Belmati (2011) pour agréger les données en regroupant tous les écrans pour un même film dans un même cinéma. En effet, étant donné le niveau d'agrégation des variables reliées à notre problématique, nous simplifions ainsi l'analyse sans perdre en qualité et en véracité. Après avoir créé nos lignes correspondant à un film donné dans un cinéma donné par semaine, nous construisons une base de données en sommant le box-office par semaine pour obtenir le box-office total généré par un film donné dans un cinéma donné. Nous nommons la base de données Box-Office Total (BOT). Nous construisons par la suite la base de données spécifique à la première semaine en supprimant les lignes où la semaine est supérieure à un. Nous nommons cette base de données Box-Office de la Première Semaine (BOPS). Nous construisons aussi la base de données des semaines subséquentes à la première semaine en éliminant respectivement la première semaine et en sommant les semaines restantes. Nous nommons cette base de données Box-Office des Semaines Subséquentes 1 (BOSS1). Nous regroupons ensuite les observations temporelles en deux grandes périodes dans la même base de données, les observations présentées la première semaine et les observations présentées les semaines subséquentes tel que $t = 1, 2$. Nous nommons la base de données Box-Office Composé (BOC) dans le reste de ce mémoire.

Valeurs extrêmes du box-office et vraie première semaine

Une variable qui nous donne de l'information sur le nombre de semaines totales qu'un film a passé en salle dans un cinéma est créée afin de pouvoir l'intégrer en tant que variable de contrôle dans l'analyse. De plus, nous retirons les films ayant un box-office total de plus de 400 000\$ qui sont considérées comme des valeurs extrêmes.

Nous redéfinissons la semaine d'ouverture dans les bases de données précédemment créées (BOT, BOPS, BOSS1, BOC) en identifiant les se-

maines d'avant-première pour tous les films à la manière de Belmati (2011). Les semaines d'avant-premières sont en effet des observations trompeuses car elles correspondent à des présentations limitées tant au niveau du nombre de salles qu'au niveau géographique. De plus, le nombre de billets est soumis au quota de la taille des salles. Il n'est pas faux de penser que les consommateurs assistant à ce genre d'événement puissent être des experts et que le box-office réagisse d'une manière différente à une sortie ordinaire. Les comportements des consommateurs pouvant être différents, nous préférons donc redéfinir la « vraie » première semaine de vie du film. Le critère du nombre de salles est retenu pour identifier la semaine de sortie d'un film. Dans la base de données, deux cas sont présents. Certains films peuvent avoir une semaine d'avant-première ou deux semaines d'avant-première. Nous définissons qu'une semaine d'avant-première satisfait aux conditions suivantes : Le nombre de salles à la semaine 1 est inférieur à 10 salles et le nombre de salles à la semaine 2 moins le nombre de salles à la semaine 1 est supérieur à 20. ou : le nombre de salles à la semaine 1 et à la semaine 2 est égal à 1 et le nombre de salles à la semaine 3 moins le nombre de salles à la semaine 2 est supérieur à 20. Dans les deux cas, c'est donc la semaine suivant la semaine d'avant-première qui constitue la semaine de sortie du film. Quand un film a une semaine d'avant-première, sa semaine de sortie est la deuxième semaine. Quand un film a deux semaines d'avant-première, sa semaine de sortie est la troisième semaine.

TABLE 4.1 – Semaine d'avant première - Première condition

Avant première	Fréquence	Pourcentage	Fréquence cumulative	% Cumulé
0	340018	99.99	340018	99.99
1	31	0.01	340049	100.00

Après avoir identifié les semaines d'avant-première, nous supprimons les 50 lignes correspondantes dans les bases de données.

TABLE 4.2 – Semaine d’avant première - Deuxième condition

Avant première	Fréquence	Pourcentage	Fréquence cumulative	% Cumulé
0	340030	99.99	340030	99.99
1	19	0.01	340049	100.00

Suppression des films avec données incomplètes

Enfin, notre analyse requiert d’avoir des données complètes concernant le box-office pour pouvoir calculer le box-office total, le box-office de la première semaine de chaque film et le box-office des semaines subséquentes, nous supprimons donc toutes les observations de l’année 2001. En effet, certains films sortis en fin 2000 ont continué d’être sur les écrans en 2001 ce qui fausserait le calcul du box-office total. Notre base de données finale comporte 1065 films et 87 cinémas.

4.3.2 Enquête Sociale Générale Cycle 19 : Emploi du temps

Description

Nous utilisons plusieurs sources de données qui seront retravaillées avant d’être combinées pour créer la base de données qui contient la variable clé de la recherche. Premièrement, l’enquête sociale générale (ESG) cycle 19, emploi du temps (2005), de la division de la statistique sociale et autochtone a été sélectionnée, car elle nous donne de l’information précise sur la lecture de journaux des consommateurs québécois. Cette enquête vise à collecter des données au niveau individuel pour étudier l’évolution des conditions de vie et du bien-être des Canadiens dans le temps et enregistre les réponses à des questions portant sur leur emploi du temps. La population des personnes de 15 ans et plus du Canada a été divisée en strates géographiques puis une composition aléatoire de numéros

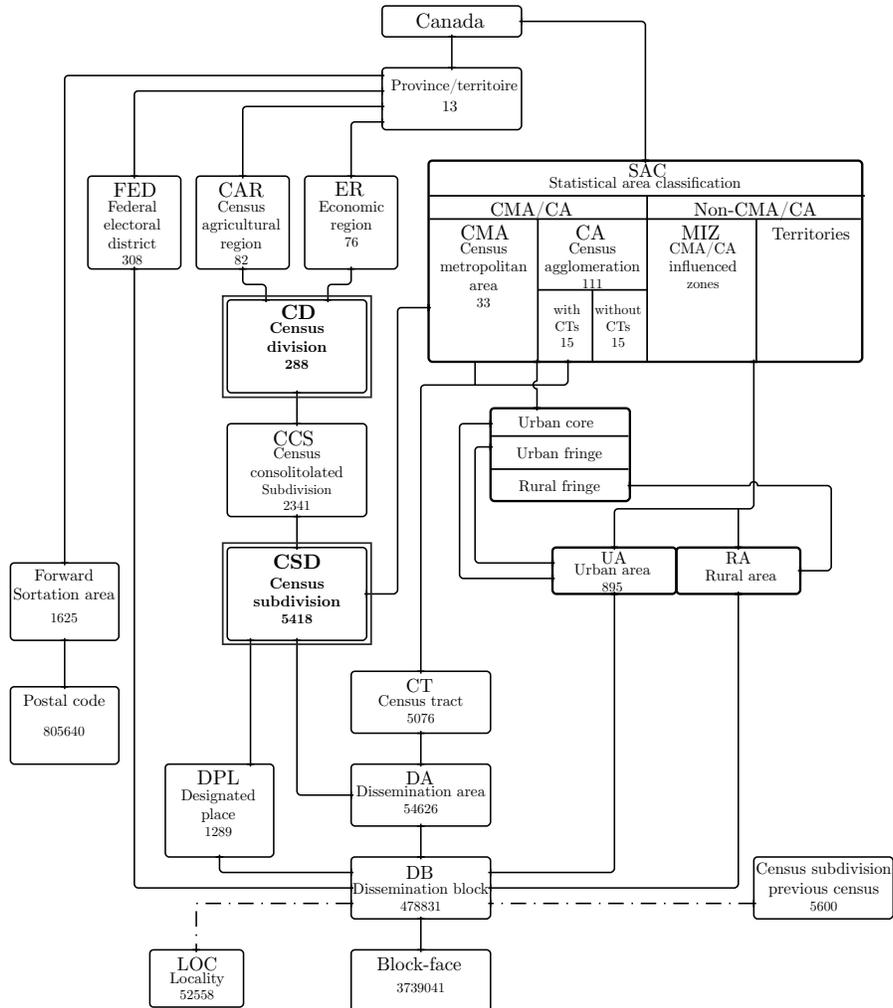
de téléphone a été effectuée au sein de chaque strate pour sélectionner les répondants. Une entrevue téléphonique assistée par ordinateur a été employée pour enregistrer les réponses. Cette enquête, administrée à 11 échantillons mensuels de janvier à novembre 2005, compte 19 597 répondants au Canada et 3 800 au Québec. Nous avons sélectionné la section 2 : Journal de l'emploi du temps et la variable DUR940 et nous avons commandé une base de données sur mesure à la division de la statistique sociale et autochtone à Statistique Canada. Nous détaillons dans la prochaine section cette base de données.

Variable d'intérêt

La variable d'intérêt DUR940 est définie comme étant la durée totale (en minutes) de lecture de journaux. Dans la base de données, 13,6% des répondants lisaient les journaux au moins 3 minutes sur une période de 24 heures. Nous détaillons la distribution de la variable extraite pour le Québec dans les sections suivantes. Étant donné la nature géographique des observations de la base de données CINÉAC au niveau des cinémas, nous avons pu exploiter cette dimension dans le cas de l'ESG cycle 19. Nous utilisons les niveaux géographiques définis par Statistique Canada qui proviennent de la codification disponible sur la figure 4.4.

Nous retenons la plus petite unité géographique disponible dans le cas de l'ESG, les subdivisions de recensement (SDR), pour avoir accès au nombre de minutes moyen de lecture de journaux par SDR. Étant donné l'échantillonnage et le nombre de répondants, il a aussi été nécessaire de calculer cette moyenne pour les divisions de recensement (DR), qui sont à un niveau d'agrégation géographique plus haut que les SDR. En effet, afin de ne pas perdre d'observations où les données concernant les SDR seraient manquantes, nous choisissons d'utiliser les niveaux agrégés à la hauteur de la DR. Les niveaux retenus sont représentés en gras et

FIGURE 4.4 – Structure des données disponibles



encadrés en gris sur le diagramme 4.4.

Liaison des données aux codes postaux

Afin de pouvoir lier les données de l'ESG aux cinémas, nous avons utilisé un outil nous permettant de créer une base de données intermédiaire pour assurer le lien entre les deux bases de données par codes postaux. Nous avons utilisé une base de données particulière CANSIM (*the Canadian Socio-Economic Information Management System*) at CHASS (*Computing in the Humanities and Social Sciences*) disponible sur le site de l'université de Toronto. Cette base de données associe le nom de DR et de SDR (en caractères) à chaque code postal, à une latitude et à une longitude.

4.4 Assemblage des bases de données

4.4.1 Base de données sur mesure et Cansim

La base de données CANSIM ne pouvant être exportée que par SDR dans différents fichiers séparés, nous avons dû fusionner chacune des cent bases de données (équivalent au nombre de DR) en une seule. Étant donné l'association par nom et le codage des caractères, nous avons aussi dû reformater les noms des SDR et des DR afin de pouvoir assembler les bases de données. En effet, certains noms comportaient des majuscules et de la ponctuation dans les bases de données tirées de Cansim à l'opposé de la base de données sur mesure. Comme des données de certaines régions n'étaient pas disponibles dans les bases de données de Cansim, nous avons dû retrouver les latitudes, longitudes, divisions de recensement et subdivisions de recensement de certains codes postaux en effectuant une double vérification à l'aide des cartes fournies sur le site

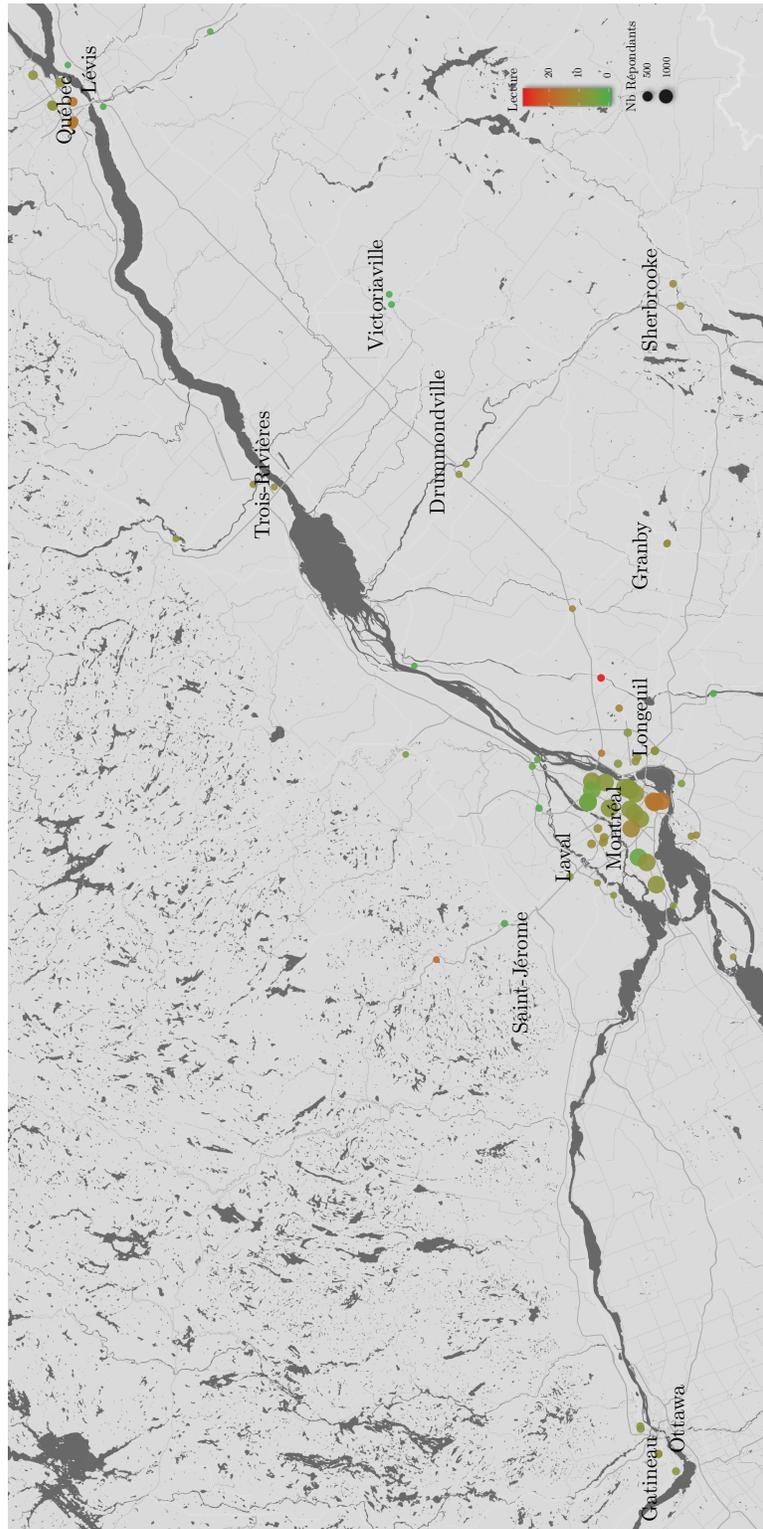
de Statistique Canada¹ pour les divisions et subdivisions de recensement et de l'outil Bing maps pour la latitude et la longitude. Par exemple la région de Longueuil n'était pas disponible et nous avons reconstruit l'information à partir des cartes et de Bing maps pour les cinémas de la base de données finale. Finalement, nous avons accès à l'association des codes postaux aux SDR, DR, latitudes et longitudes.

4.4.2 Assemblage final

Nous avons premièrement recherché et créé une base de données avec les codes postaux et la variable Cine ID pour chaque cinéma présent dans notre base de données. Par la suite, nous avons inséré la variable postcode dans la base de données CINEAC. Nous avons enfin assemblé les deux bases de données par la variable postcode. À partir de cette base de données par semaine, nous créons nos trois bases de données finales selon nos agrégations temporelles nécessaires à la vérification de nos hypothèses. Nous nommons la base de données comprenant le box-office total agrégé BOT, la base de données comprenant le box-office de la première semaine BOPS, la base de données comprenant le box-office agrégé des semaines subséquentes à la première BOSS1 et enfin notre base de données principale BOC. La figure 4.5 représente les cinémas retenus et les minutes que les consommateurs les plus proches des cinémas passent à lire les journaux quotidiens. Cette figure est centrée sur les plus gros centre d'activité et nous omettons de représenter les cinémas éloignés de la base de données.

1. Les cartes sont référencées dans la bibliographie [Statistique Canada, 2006]

FIGURE 4.5 – Cinémas retenus et minutes de lecture de journaux



4.5 Variables utilisées

Afin de mieux comprendre les relations entre les facteurs dans notre modèle, nous détaillons la forme que prennent les variables que nous utilisons

4.5.1 Variables dépendantes

Nous nous intéressons à estimer un modèle pré-lancement afin d'analyser l'effet de différents facteurs et leur interaction sur le box-office des films. Pour y parvenir nous utilisons nos trois façons distinctes d'agrèger le box-office pour pouvoir identifier les différents effets que nous étudions. Tandis que le logarithme du box-office de la première semaine nous permettra de pouvoir étudier l'effet d'influence, le logarithme du box-office des semaines subséquentes nous permettra d'étudier l'effet de prédiction des critiques. Le logarithme du box-office total sera quant à lui utilisé pour comparer les résultats de nos modèles.

4.5.2 Variables d'intérêt

Cote

Alors que la variable *cote* originale comporte 7 modalités, nous regroupons les catégories ayant trop peu d'observations. Ci-dessous, les modalités d'origine

$$cote = \begin{cases} 1 & \text{Chef-d'oeuvre} \\ 2 & \text{Remarquable} \\ 3 & \text{Très bon} \\ 4 & \text{Bon} \\ 5 & \text{Moyen} \\ 6 & \text{Pauvre} \\ 7 & \text{Minable} \end{cases}$$

Nous recodons donc la variable concernant la cote des critiques cinématographique sur une échelle de 1 à 4.

$$cote = \begin{cases} 1 & \text{si la cote du film est de 6 ou plus} \\ 2 & \text{si la cote du film est de 5} \\ 3 & \text{si la cote du film est de 4} \\ 4 & \text{si la cote du film est de 3 ou moins} \end{cases}$$

Lecture de journaux

Pour créer la variable concernant la lecture de journaux pour chaque cinéma, nous utilisons le nombre de minutes de lecture de journaux au niveau de la subdivision de recensement. Si cette information n'est pas disponible, nous utilisons un niveau supérieur d'agrégation, les divisions de recensement.

Indicatif de la première semaine

Avec notre base de données dans la configuration temporelle en deux parties (BOC), c'est à dire une ligne pour chaque film observé dans chaque cinéma à chaque période temporelle, nous pouvons estimer les paramètres d'intérêt avec une seule équation. Nous créons donc une variable binaire

définie comme suit

$$s1 = \begin{cases} 1 & \text{si l'observation a été observée la première semaine} \\ 0 & \text{sinon} \end{cases}$$

Nous l'utilisons pour indiquer quels sont les coefficients qui sont propres à la première semaine et quels sont les coefficients propres aux semaines subséquentes. Par la suite, nous incluons les covariables relatives à la première semaine et les covariables relatives aux semaines subséquentes pour que les estimateurs des paramètres concernant nos facteurs d'intérêt soient non-biaisés. Nous utilisons une interaction entre les variables et les covariables pour représenter la première semaine. L'interaction entre nos deux facteurs continus pour chaque semaine nous permet d'obtenir les deux estimateurs d'intérêt.

Interactions

Afin de tester notre hypothèse de recherche, nous incluons une interaction entre les deux facteurs à l'étude. Nous pouvons ainsi capturer des différences de l'effet de la cote des critiques selon le niveau de lecture des consommateurs (variable *minlecf*) et donc tester s'il existe un effet de modération des minutes de lecture de journaux sur l'effet de la cote.

4.5.3 Covariables

Visa

Le visa d'exploitation fait partie de nos variables de contrôle étant donné sa prise en compte justifiée dans la littérature. Dans nos modèles la va-

riable concernant le visa est définie ainsi

$$visa = \begin{cases} 1 & \text{si le film a un visa général} \\ 2 & \text{si le film a un visa 13 et plus} \\ 3 & \text{si le film a un visa 16 et plus ou un visa 18 et plus} \end{cases}$$

Genre

Dans leur analyse, Chang et Ki (2005) avancent qu'une recherche plus précise doit être effectuée sur l'effet du genre directement sur le box-office des films. Nous introduisons le genre comme covariable dans les équations de nos différents modèles. De fait, nous contrôlons l'effet moyen de chaque genre lorsque nous estimons les effets des deux facteurs étudiés sur le box-office. La variable genre est définie comme suit

$$genre = \begin{cases} 1 & \text{si le film est un drame} \\ 2 & \text{si le film est une comédie} \\ 3 & \text{si le film est un film de suspense} \\ 4 & \text{si le film est un film d'action et aventure} \\ 5 & \text{si le film est un film pour enfant ou pour la famille ou un documentaire} \\ 6 & \text{si le film est un film de science-fiction} \\ 7 & \text{si le film est un film d'horreur} \end{cases}$$

Durée

La durée du film en nombre de minutes est inclus dans chacun de nos modèles. Nous la définissons comme suit

$$duree = \begin{cases} 1 & \text{si le film a une durée comprise dans l'intervalle } [40 \text{ min}; 91 \text{ min}[\\ 2 & \text{si le film a une durée comprise dans l'intervalle } [91 \text{ min}; 100 \text{ min}[\\ 3 & \text{si le film a une durée comprise dans l'intervalle } [100 \text{ min}; 110 \text{ min}[\\ 4 & \text{si le film a une durée comprise dans l'intervalle } [110 \text{ min}; 124 \text{ min}[\\ 5 & \text{si le film a une durée comprise dans l'intervalle } [124 \text{ min}; 222 \text{ min}[\end{cases}$$

Mois

Nous incluons onze variables dichotomiques (une pour chaque mois moins la catégorie de base) qui nous permettent de contrôler la saisonnalité dans nos modèles.

Nombre d'écran la première semaine

Le budget publicitaire et le budget de production ont tous deux été démontrés comme étant reliés au nombre d'écrans au premier weekend (Elberse et Eliashberg, 2003 ; Hennig-Thurau et al., 2007). Beuscart et Mellet (2012) trouvent que le nombre de bobines distribuées est fortement corrélé au budget publicitaire (0,81). Nous utiliserons donc cette relation mise au jour dans la littérature pour prendre indirectement en compte l'investissement dans les éléments de promotion d'un film. Le nombre d'écrans la première semaine est utilisé comme un proxy des dépenses en marketing dans nos modèles.

Distributeur

Nous considérons neuf distributeurs dans ce mémoire. La variable distributeur est définie comme suit

$$\text{distributeur} = \begin{cases} 1 & \text{si le distributeur est Viva film} \\ 2 & \text{si le distributeur est Disney} \\ 3 & \text{si le distributeur est Sony} \\ 4 & \text{si le distributeur est Dreamworks} \\ 5 & \text{si le distributeur est Fox} \\ 6 & \text{si le distributeur est Paramount} \\ 7 & \text{si le distributeur est Universal} \\ 8 & \text{si le distributeur est Warner} \\ 9 & \text{si le distributeur est indépendant ou autre} \end{cases}$$

Star power et éléments de marque

La méthodologie de création des variables concernant le star power est directement reprise de la recherche de Bonnet (2009). Les deux principaux objets considérés ici sont les acteurs et les réalisateurs. Au niveau temporel, nous considérons que l'effet star est actif durant les trois années qui suivent la sortie du film (Bonnet, 2009). Après avoir établi un palmarès des 20 plus grands succès commerciaux entre 1998 et 2010, il a été possible d'enregistrer le nom de l'acteur principal, de l'acteur de soutien et du réalisateur. Ces informations ont été recueillies en ligne sur la base de données IMDb (the Internet Movie Database).

La filmographie de chaque acteur et de chaque réalisateur a ensuite été réalisée pour les trois années précédant chaque film avant d'être mise en opposition avec les films les plus populaires au Québec. En tout, neuf variables binaires sont créées afin de pouvoir dériver la mesure objective

du star power. Nous définissons premièrement l'ensemble A_i dans lequel on dit qu'un film $a_i \in A_{i-n}$ si et seulement s'il fait partie du top 20 créé ou i est l'année de sortie du film et $i - n$ l'année n avant la sortie du film. Trois variables sont créées pour l'acteur principal, l'acteur de soutien et le réalisateur.

Pour l'acteur principal nous avons

$$Star1.1 = \begin{cases} 1 & \text{si l'acteur principal était à l'affiche d'un film } a_i \in A_{i-1} \\ 0 & \text{sinon} \end{cases}$$

$$Star1.2 = \begin{cases} 1 & \text{si l'acteur principal était à l'affiche d'un film } a_i \in A_{i-2} \\ 0 & \text{sinon} \end{cases}$$

$$Star1.3 = \begin{cases} 1 & \text{si l'acteur principal était à l'affiche d'un film } a_i \in A_{i-3} \\ 0 & \text{sinon} \end{cases}$$

Pour l'acteur de soutien nous avons

$$Star2.1 = \begin{cases} 1 & \text{si l'acteur de soutien était à l'affiche d'un film } a_i \in A_{i-1} \\ 0 & \text{sinon} \end{cases}$$

$$Star2.2 = \begin{cases} 1 & \text{si l'acteur de soutien était à l'affiche d'un film } a_i \in A_{i-2} \\ 0 & \text{sinon} \end{cases}$$

$$Star2.3 = \begin{cases} 1 & \text{si l'acteur de soutien était à l'affiche d'un film } a_i \in A_{i-3} \\ 0 & \text{sinon} \end{cases}$$

Pour le réalisateur nous avons

$$director1 = \begin{cases} 1 & \text{si le réalisateur était à l'affiche d'un film } a_i \in A_{i-1} \\ 0 & \text{sinon} \end{cases}$$

$$director2 = \begin{cases} 1 & \text{si le réalisateur était à l'affiche d'un film } a_i \in A_{i-2} \\ 0 & \text{sinon} \end{cases}$$

$$director3 = \begin{cases} 1 & \text{si le réalisateur était à l'affiche d'un film } a_i \in A_{i-3} \\ 0 & \text{sinon} \end{cases}$$

Nous reprenons la méthode d'agrégation des neuf variables créée par Bel-mati (2011) pour avoir un modèle plus parcimonieux. Nous redéfinissons donc les trois variables suivantes

Pour le réalisateur nous avons

$$Star1 = \begin{cases} 1 & \text{si Star1.1} = 1 \text{ ou si Star1.2} = 1 \text{ ou si Star1.3} = 1 \\ 0 & \text{sinon} \end{cases}$$

$$Star2 = \begin{cases} 1 & \text{si Star2.1} = 1 \text{ ou si Star2.2} = 1 \text{ ou si Star2.3} = 1 \\ 0 & \text{sinon} \end{cases}$$

$$director = \begin{cases} 1 & \text{si director1} = 1 \text{ ou si director2} = 1 \text{ ou si director3} = 1 \\ 0 & \text{sinon} \end{cases}$$

Notre base de données finale comporte donc ces trois dernières variables qui représentent l'effet du *star power*.

4.6 Modèles

Dans ce mémoire, nous utiliserons des techniques développées en premier lieu par Fisher (1919), leurs avancées et les outils permettant la sélection de modèle et leur diagnostic (Henderson, 1953; Hartley and Rao, 1967; Diggle, Liang and Zeger, 1994; Bozdogan, 1987) pour modéliser la structure des relations entre les revenus du box-office, les caractéristiques des films sélectionnés et les caractéristiques des cinémas sélectionnés. Les premières applications des modèles linéaires mixtes sont apparues en biométrie et sont fréquemment utilisées en bio statistiques, psychométrie et statistiques appliquées à l'agriculture. Leur structure, plus complète que les modèles de régression linéaire de base, permet de relaxer l'hypothèse d'indépendance (les observations ne sont plus indépendamment et identiquement distribuées) entre les observations et de prendre en compte les différents niveaux de corrélation entre les éléments étudiés.

Pour exprimer les effets fixes du modèle nous utilisons une interaction entre notre effet de traitement et l'index de critiques cinématographiques. Selon Angrist et Krueger (1998) ce type d'effet de traitement peut être interprété d'une façon non-linéaire. Il est donc possible de parler de l'effet croisé du traitement et de la cote des critiques cinématographiques lorsque le traitement passe de $t=1$ à $t=2$ (par exemple) à un niveau de cote fixé. De même, nous pouvons estimer l'effet croisé du traitement de la cote à un niveau de lecture fixé. Cette manière de représenter les facteurs influençant le box-office permet de réduire le biais des estimateurs des moindres carrés ordinaires mis au jour par Reinstein et Snyder (2005). Selon eux, l'index des critiques est endogène au modèle étant donné qu'il existe des variables non-observables que nous ne prenons pas en compte. En effet, nous ne pouvons pas contrôler pour la qualité du jeu d'acteur, la qualité des effets spéciaux ou encore certains efforts marketing qui sont corrélés avec l'index des critiques. Il peut alors être démontré que deux

sources de biais peuvent être présentes et que nous ne pouvons pas savoir précisément de quel signe est l'effet net obtenu par moindres carrés ordinaires. Reinstein et Snyder (2005) utilisent une méthode de différences en différences qui permet de traiter ce problème sous certaines conditions.

Notre méthodologie est appuyée par une des remarques des deux auteurs. En effet, ils recommandent d'enrichir l'analyse avec des variables proxys observables des facteurs de qualité non observables des films et d'inclure toute variable qui pourrait être contenue dans le terme d'erreur et corrélée avec une des variables indépendante. Cela permet ainsi de réduire le biais présent dans les estimateurs des facteurs étudiés. Par rapport aux études précédentes, notre analyse fait ressortir l'effet des critiques cinématographiques d'une nouvelle manière tout en apportant un niveau de détail plus avancé. Nous explorons ainsi les effets d'un traitement continu de l'exposition aux critiques grâce à la variable concernant la lecture de journaux quotidiens. Dans notre base de données, nous identifions de plus différents niveaux de regroupement que nous prenons en compte pour garantir une inférence valide.

4.6.1 Structure des données

Premièrement, nous introduisons quelques notations nous permettant de représenter la structure de la base de données avant de dériver le modèle complet pour mieux dissocier les différents effets que nous allons estimer. Le logarithme du box-office du film f présenté dans le cinéma c lors de la semaine t est noté Y_{cft} . Chaque observation a un vecteur de caractéristiques observables et de proxys qui varient selon le film $x_{1f}, x_{2f}, \dots, x_{kf}$ et un vecteur de caractéristiques observables et de proxys qui varient selon le cinéma $q_{1c}, q_{2c}, \dots, q_{mc}$. Ou $k = 1, \dots, K$, K étant le nombre total de variables indépendantes propres aux films, et $m = 1, \dots, M$, M étant le nombre total de variables indépendantes propres au cinéma.

Plusieurs régressions seront utilisées pour analyser les données afin de répondre à nos questions de recherche. Nous utiliserons les trois configurations finales de la base de données pour estimer les effets des facteurs sur le logarithme du box-office total des films, sur le logarithme du box-office de la première semaine des films, et sur le logarithme du box-office total des films des semaines strictement supérieures à 1. La dernière configuration que nous créons à partir de la base de données de la première semaine et la base de données des semaines subséquentes nous permet de modéliser conjointement les effets des facteurs pour ces deux périodes importantes dans la commercialisation des films. Cette base de données est donc composée de deux périodes temporelles distinctes ; $logbototal_{cf1}$ correspond à la première période tandis que $\sum_{t=2}^T logbototal_{cft}$ correspond à la deuxième période.

Étant donné la différence au niveau de la dimension temporelle, les équations estimées utiliseront des facteurs différents. La base de données peut être représentée sous une forme rectangulaire avec un nombre de ligne égal à la somme pour tous les films du nombre de période (t) où chaque film retenu (f) est présenté dans chaque cinéma retenu (c).

4.6.2 Modélisation des effets fixes et des effets aléatoires

Premièrement nous présentons en quoi les effets aléatoires nous permettent de construire un modèle qui prend en compte les spécificités des données. Nous illustrons cette configuration dans un graphique récapitulatif qui présentera les relations entre les observations d'une manière intuitive.

Hétérogénéité inter-observations

L'utilisation des modèles linéaires mixtes nous permet d'attribuer un effet propre à chaque film et à chaque cinéma. Notre modèle admet ainsi qu'il existe des caractéristiques non-observables qui sont liées à la performance d'un film en particulier. Par exemple, la présence de certains effets spéciaux ou le lien du film avec un sujet d'actualité pourrait affecter ses performances au box-office. De ce fait, pour un film en particulier, toutes les observations y étant reliées dans notre base de données auront une performance avec un élément en commun. Cette performance commune pour le même film qui peut être présenté dans des cinémas différents vient introduire de la corrélation entre les observations pour ce film. Pour compléter la structure de corrélation, nous différencions la composante commune à chaque film selon la période à laquelle il a été présenté.

De même, un effet spécifique à chaque cinéma est considéré, ce qui nous permet de prendre en compte les caractéristiques communes de l'environnement expérientiel et la localisation de chaque centre de distribution. Nous permettons au box-office des films d'un même cinéma d'être influencé par des caractéristiques non-observables propres au cinéma. Par exemple, les services connexes comme la restauration au sein de l'établissement ou encore la proximité avec un centre d'achat peuvent influencer la performance des films présentés. Encore une fois, l'élément commun contenu dans la performance au box-office des films présentés dans le même cinéma introduira de la corrélation entre ces observations.

Hétérogénéité intra-observations

Notre nouvelle façon de modéliser le box-office prend en compte sa dimension temporelle. En effet, les revenus liés à la diffusion des films après

la première semaine sont corrélés fortement aux revenus de la première semaine. Nous ajustons donc un effet spécifique à chaque film présenté dans chaque cinéma. La structure de corrélation admet qu'il existe un élément commun entre les revenus d'un film particulier présenté dans un cinéma particulier à travers nos deux périodes considérées. Nous restreignons cette corrélation qui est présente uniquement entre deux observations d'un même film présenté dans le même cinéma. Nous détaillerons plus loin la structure complète de variance covariance et chacune des corrélations estimées.

Représentation graphique

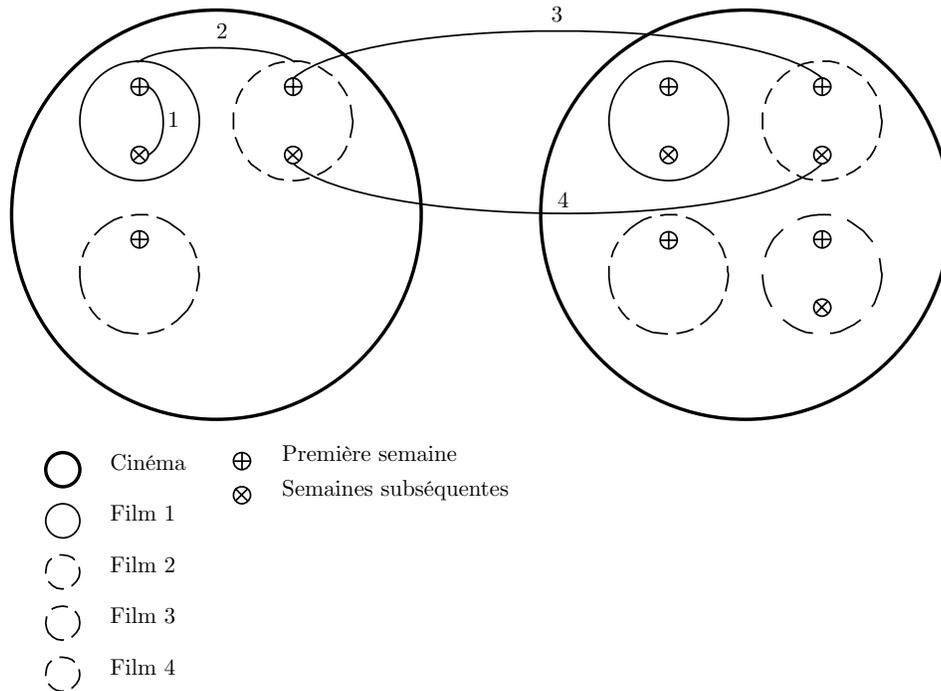
Afin de mieux comprendre la structure théorique des effets que nous allons estimer nous la représentons graphiquement à la figure 4.6 où nous définissons :

1. la corrélation pour le même film présenté dans un cinéma à travers les deux périodes
2. la corrélation entre deux films différents présentés dans le même cinéma
3. la corrélation entre deux observations du même film présenté dans deux cinémas différents à la première période
4. la corrélation entre deux observations du même film présenté dans deux cinémas différents à la deuxième période

Formulation mathématique du modèle proposé

Pour mieux comprendre quelles sont les propriétés des effets que nous allons estimer, nous exposons le modèle et les relations que nous supposons entre les variables et les observations. Nous présentons la version finale

FIGURE 4.6 – Structure de corrélation



utilisée pour répondre à notre question de recherche. L'équation 4.1 représente la forme générale du modèle estimé tandis que les énoncés 4.2 sont les hypothèses que nous retenons pour la modélisation.

$$\begin{aligned}
 Y_{cft} = & \mathbf{X}_f \boldsymbol{\beta} + \mathbf{Q}_c \boldsymbol{\alpha} + s1(\mathbf{X}_f \boldsymbol{\beta} + \mathbf{Q}_c \boldsymbol{\alpha}) & (4.1) \\
 & + \lambda_{ft} + \xi_c + \omega_g + \varepsilon_{cft}
 \end{aligned}$$

avec

$$\begin{aligned}\mathbf{X}_f \boldsymbol{\beta} &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_{kf} \\ \mathbf{Q}_c \boldsymbol{\alpha} &= \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_{kc}\end{aligned}$$

où le vecteur \mathbf{X}_f représente le vecteur des caractéristiques des films tandis que le vecteur \mathbf{Q}_c représente le vecteur des caractéristiques des cinémas. Nous définissons les effets aléatoires et le terme d'erreur comme suit

$$\lambda_{ft} \sim N(0, \sigma_{\lambda_t}^2), \xi_c \sim N(0, \sigma_{\xi}^2), \omega_g \sim N(0, \sigma_{\omega}^2), \epsilon_{cft} \sim N(0, \sigma_{\epsilon}^2) \quad (4.2)$$

où λ_{ft} correspond à l'intercept aléatoire concernant les films à la période t , ξ_c correspond à l'intercept aléatoire concernant les films, ω_g correspond à l'intercept aléatoire au niveau d'un film dans un cinéma particulier et ϵ_{cft} correspond à l'erreur. Toutes ces composantes aléatoires sont supposées indépendantes les unes des autres. Notons que $t=1,2$ où 1 représente la première semaine tandis que 2 représente les semaines subséquentes. Sous cette forme, le modèle admet une structure de covariance complexe introduite par les effets aléatoires. La corrélation entre deux films différents présentés dans le même cinéma la première semaine s'exprime

$$\frac{\sigma_{\xi}^2}{\sigma_{\lambda_1}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2}$$

La corrélation entre deux films différents présentés dans le même cinéma les semaines subséquentes s'exprime

$$\frac{\sigma_{\xi}^2}{\sigma_{\lambda_2}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2}$$

La corrélation pour le même film présenté dans deux cinémas différents

la première semaine s'exprime

$$\frac{\sigma_{\lambda_1}^2}{\sigma_{\lambda_1}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2}$$

La corrélation pour le même film présenté dans deux cinémas différents les semaines subséquentes s'exprime

$$\frac{\sigma_{\lambda_2}^2}{\sigma_{\lambda_2}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2}$$

La corrélation pour le même film présenté dans le même cinéma entre les deux périodes s'exprime

$$\frac{\sigma_{\omega}^2}{\sqrt{\sigma_{\lambda_1}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2} * \sqrt{\sigma_{\lambda_2}^2 + \sigma_{\xi}^2 + \sigma_{\omega}^2 + \sigma_{\epsilon}^2}}$$

Codage des variables catégorielles pour l'estimation

Afin d'estimer les effets des covariables et des facteurs catégoriels nous avons choisi les catégories de bases suivantes

- **genre** : 1, Drame
- **durée** : 1, le film dure entre 40 et 90 minutes inclus
- **mois** : 1, janvier
- **distributeur** : 1, Viva film

Pour avoir une estimation des paramètres pour chaque niveau ayant une interprétation par rapport à la moyenne nous effectuons aussi l'estimation après avoir utilisé le *simple coding*. Nous transformons les matrices de variables binaires utilisées pour chaque variable catégorielle de façon à ce que l'ordonnée à l'origine corresponde à la moyenne de toutes les catégories. Par exemple, pour la variable visa nous avons la matrice originale

pour les variables binaires

$$binaires = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Nous créons ensuite cette matrice en soustrayant $1/k$ où k est le nombre de classes.

$$\begin{aligned} M_{\text{simple coding Visa}} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.3333333 & 0.3333333 \\ 0.3333333 & 0.3333333 \\ 0.3333333 & 0.3333333 \end{bmatrix} \\ &= \begin{bmatrix} -0.3333333 & -0.3333333 \\ 0.6666667 & -0.3333333 \\ -0.3333333 & 0.6666667 \end{bmatrix} \end{aligned}$$

Ces nouvelles matrices nous permettront d'interpréter l'effet de chaque modalité de chaque variable catégorielle comme étant l'effet marginal par rapport à la moyenne, *ceteris paribus*.

4.6.3 Estimation avec le logiciel R

À des fins d'estimation, nous créons la variable binaire complémentaire de $s1$, $ns1$ définie ci-dessous

$$ns1 = \begin{cases} 1 & \text{si l'observation a été observée les semaines subséquentes} \\ 0 & \text{sinon} \end{cases}$$

Cette astuce nous permet de contourner les restrictions de l'outil d'estimation, le package *lme4* utilisé avec le logiciel R. Nous pouvons donc

réécrire notre équation

$$Y_{cft} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + s1(\mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha}) \quad (4.3)$$

$$+ s1\lambda_{f1} + ns1\lambda_{f2} + \xi_c + \omega_g + \varepsilon_{cft}$$

On voit alors que cette configuration nous permet de faire apparaître λ_{f1} dans le cas où l'observation correspond à une présentation du film à la période 1. Si l'observation correspond à une présentation du film à la période 2, c'est alors λ_{f2} qui apparaît dans l'équation. Nous pouvons ainsi prendre en compte l'hétérogénéité inter-film.

Formulation mathématique des modèles existants

Afin de pouvoir comparer notre modèle, nous sélectionnons plusieurs manières communes de représenter la dépendance entre les facteurs et le box-office dans la littérature. Premièrement nous séparons entièrement le problème pour traiter indépendamment l'analyse de l'influence des facteurs sur le box-office et l'analyse des facteurs qui prédisent le box-office. Nous utilisons les données de la première semaine seulement, pour effectuer une première régression afin d'explorer l'effet d'influence en introduisant un intercept aléatoire pour le film et un intercept aléatoire pour le cinéma

$$Y_{cf1} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + \lambda_{f1} + \xi_c + \varepsilon_{cf} \quad (4.4)$$

Nous estimons ensuite un modèle utilisant le box-office total comme variable dépendante pour explorer l'effet de prédiction avec un intercept aléatoire pour le film et un intercept aléatoire pour le cinéma

$$Y_{cf} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + \lambda_f + \xi_c + \varepsilon_{cf} \quad (4.5)$$

Nous estimons enfin un modèle utilisant le box-office des semaines stric-

tement supérieures à 1 comme variable dépendante pour explorer l'effet de prédiction avec un intercept aléatoire pour le film et un intercept aléatoire pour le cinéma

$$Y_{cft>1} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + \lambda_f + \xi_c + \varepsilon_{cf} \quad (4.6)$$

Ces trois modèles ne dissocient pas l'effet d'influence et l'effet de prédiction d'une façon distincte car le box-office total inclut le box-office de la première semaine et la structure de corrélation temporelle n'est pas prise en compte. Il est donc délicat de pouvoir donner une interprétation claire sur l'effet qui prévaut.

4.6.4 Algorithmes et techniques d'estimation utilisées

Afin de pouvoir correctement estimer les effets, nous utilisons une estimation par maximum de vraisemblance restreint (*REML*) qui nous permet d'avoir des estimateurs non-biaisés des composants de la variance si le modèle est dans sa forme correcte. L'algorithme PIRLS (*penalized iteratively reweighted least squares*) est utilisé pour trouver les estimateurs des effets aléatoires et des effets fixes en minimisant le critère de maximum de vraisemblance restreint (Bates, 2013).

4.6.5 Vecteurs de variables de contrôle

Nous détaillons ici la structure des vecteurs de variable de contrôle liés aux films et aux cinémas. Nous composons donc le vecteur de caractéristiques pour les films avec les variables suivantes

$$\mathbf{X}_f = \left\{ \begin{array}{ll} s1 & : \text{ une variable indicatrice de la première semaine} \\ cote & : \text{ la note que les critiques ont attribué au film} \\ genre & : \text{ le genre du film} \\ visa & : \text{ le visa associé au film} \\ duree & : \text{ la durée du film} \\ mois & : \text{ le mois de sortie du film} \\ necranpremssem & : \text{ le nombre d'écrans où le film a été présenté lors de sa sortie} \\ distributeur & : \text{ le distributeur du film} \\ star1 & : \text{ la présence d'une star pour le premier rôle} \\ star2 & : \text{ la présence d'une star pour un rôle secondaire} \\ director1 & : \text{ la présence d'un réalisateur reconnu} \end{array} \right.$$

Le vecteur des caractéristiques pour les cinémas est composé d'un seul élément qui nous sert ici à étoffer l'effet des critiques cinématographiques

$$\mathbf{Q}_c = \left\{ \begin{array}{ll} minlecfm & : \text{ le nombre de minutes de lecture de journaux quotidiens} \end{array} \right.$$

La forme matricielle est ici préférée afin de pouvoir considérer des extensions du modèle qui pourraient inclure de nouveaux facteurs et variables de contrôle. Nous développerons cet aspect du modèle dans la partie discussion de cette étude. Nous utilisons aussi des interactions entre les variables *minlecfm*, *s1* et *cote* qui nous permettent de répondre à notre problématique de recherche.

Chapitre 5

Analyses et résultats

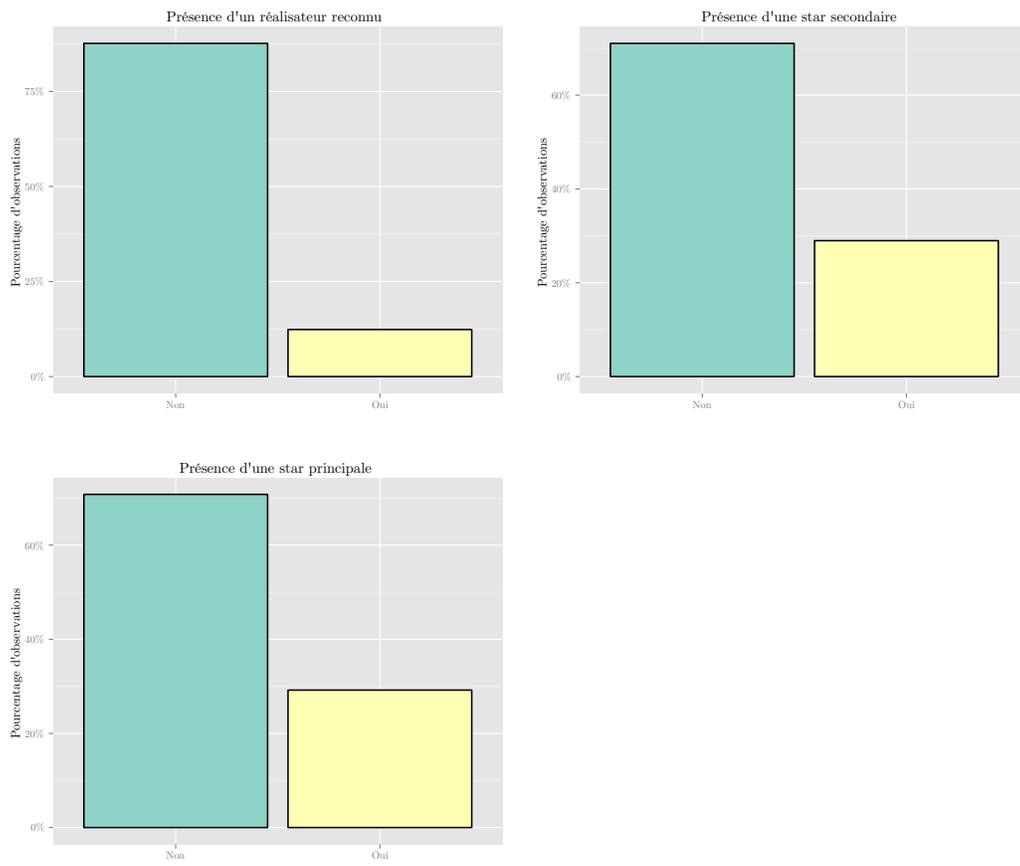
Dans cette partie, nous débutons l'analyse de notre modèle en exposant les spécificités des bases de données d'une manière descriptive. Par la suite, nous étoffons l'architecture des effets aléatoires présentés schématiquement à la figure 4.6. Puis, nous décrivons et nous interprétons les résultats de l'estimation des effets fixes qui constituent l'élément central de notre étude. Afin de mieux exposer les avantages de notre modèle, nous comparerons nos résultats avec ceux de la littérature d'un point de vue technique et d'un point de vue intuitif.

5.1 Statistiques descriptives

Nous explorons dans un premier temps les bases de données que nous avons utilisées pour notre analyse¹ afin de pouvoir faire ressortir les éléments descriptifs importants pour la suite de notre analyse. Nous comparons globalement les statistiques descriptives tout en nous appuyant sur une représentation graphique des données.

1. Nous utilisons la base de données du box-office total pour réaliser les graphiques et les tableaux qui n'exploitent pas la dimension temporelle des données.

FIGURE 5.1 – Éléments de marque



Dans les variables représentant l'élément de marque d'un film (table 5.1 et figure 5.1) nous observons que 29,20% des films comptent un acteur considéré comme une star pour le premier rôle par notre classification. Nous voyons aussi que 28,97% des films comptent une star dans un rôle secondaire et que 12,37% des films comptent un réalisateur qui a été reconnu.

TABLE 5.1 – Distribution des modalités des variables concernant les éléments de marque

		%
Présence d'une star pour le premier rôle	Non	70.80
	Oui	29.20
Présence d'une star dans un rôle secondaire	Non	71.03
	Oui	28.97
Présence d'un réalisateur connu	Non	87.63
	Oui	12.37

Plus de la majorité des films de notre échantillon (58,36%) dispose d'un visa général tandis que 36,72% disposent d'un Visa 13 et plus et 4,91% disposent ou d'un visa 16 et plus ou d'un visa 18 et plus (table 5.2 et figure 5.2).

TABLE 5.2 – Distribution des modalités de la variable visa

	%
Visa Visa général	58.36
Visa 13 et plus	36.72
Visa 16 et plus ou Visa 18 et plus	4.91

Les comédies sont le genre de film le plus représenté avec 39,45% des observations tandis que les films les moins représentés sont les films d'action et aventure avec 7,22% des observations (table 5.3 et figure 5.2).

Concernant la durée de film, les observations sont réparties d'une façon symétrique avec un mode correspondant aux films ayant une durée comprise entre 100 et 109 minutes (table 5.4 et figure 5.2).

FIGURE 5.2 – Caractéristiques des films

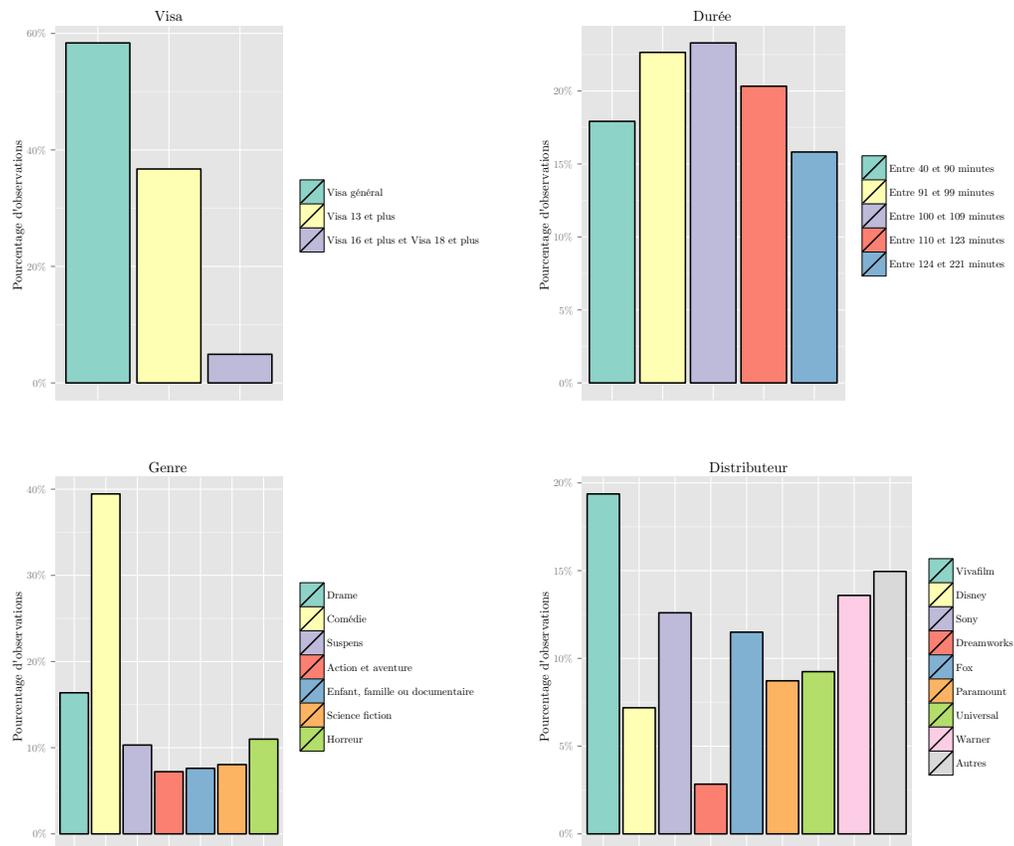


TABLE 5.3 – Distribution des modalités de la variable genre

	%
Genre Drame	16.38
Comédie	39.45
Suspens	10.31
Action et aventure	7.22
Enfant, famille ou documentaire	7.61
Science fiction	8.04
Horreur	10.99

TABLE 5.4 – Distribution des modalités de la variable durée

	%
Durée	
Entre 40 et 90 minutes	17.92
Entre 91 et 99 minutes	22.64
Entre 100 et 109 minutes	23.29
Entre 110 et 123 minutes	20.33
Entre 124 et 221 minutes	15.82

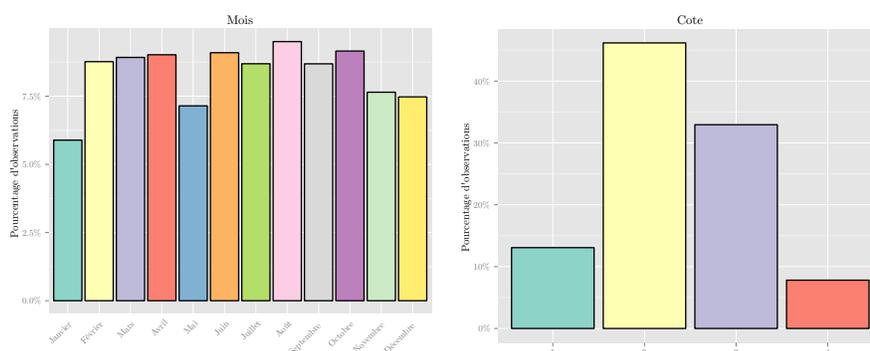
Vivafilm est le distributeur qui affiche le plus grand nombre de films à l'écran. La catégorie qui regroupe les producteurs les moins présents individuellement arrive en deuxième position avec 14,96% d'observations et Warner arrive en troisième position avec 13,59% d'observations. Des distributeurs les plus connus, Dreamworks place le moins grand nombre de films en salles dans notre échantillon (table 5.5 et figure 5.2). Nous pouvons constater qu'il est possible que les distributeurs qui placent individuellement le plus de films soit dotés d'un certain pouvoir de marché. En effet, leur capacité à pouvoir produire de nombreux films peut les avantager dans la négociation de la parution de leurs produits dans les différents cinémas

TABLE 5.5 – Distribution des modalités de la variable distributeur

	%
Distributeur	
Vivafilm	19.37
Disney	7.19
Sony	12.60
Dreamworks	2.83
Fox	11.50
Paramount	8.72
Universal	9.24
Warner	13.59
Autres	14.96

Si nous regardons la fréquence de sortie des films par mois, nous remarquons que les mois d'août, octobre et juin sont les mois de sortie les plus représentés avec respectivement 9,50%, 9,15% et 9,10% des observations).

FIGURE 5.3 – Saisonnalité et cote



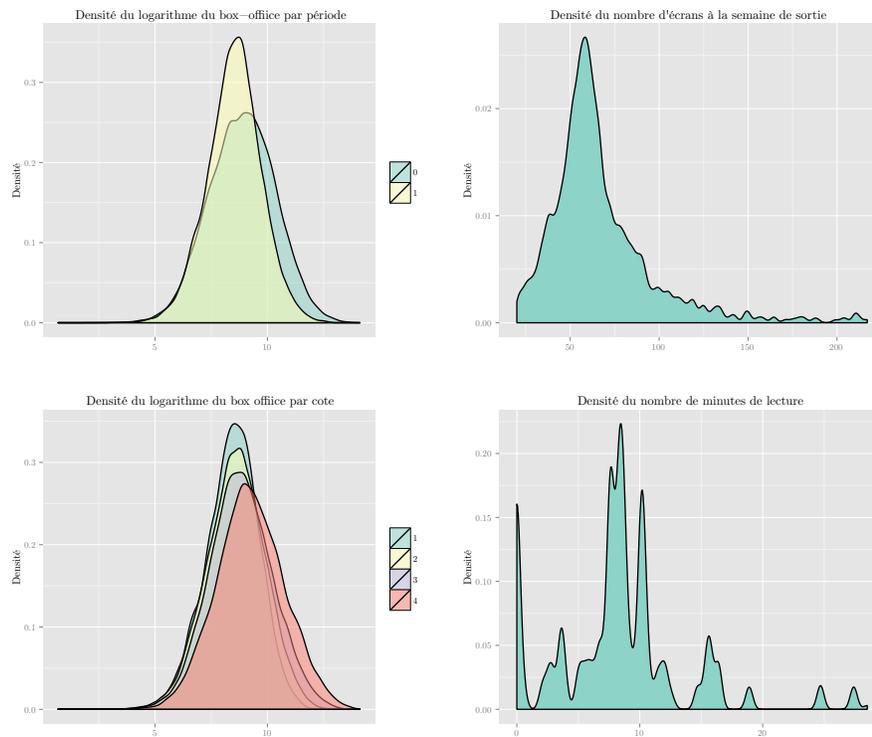
Les mois de janvier, mai et décembre sont les mois de sortie les moins représentés avec respectivement 5,89%, 7,15% et 7,48% des observations (table 5.6 et figure 5.3).

TABLE 5.6 – Distribution des modalités de la variable mois

		%
Mois	Janvier	5.89
	Février	8.77
	Mars	8.92
	Avril	9.02
	Mai	7.15
	Juin	9.10
	Juillet	8.69
	Août	9.50
	Septembre	8.69
	Octobre	9.16
	Novembre	7.65
	Décembre	7.48

Globalement, les films ayant une cote de 2 sont les plus représentés dans l'échantillon. Nous pouvons voir que moins de 10% des films sont cotés 4 (figure 5.2). Nous voyons donc qu'il existe une dominance claire du nombre de visas généraux, du nombre de comédies et du nombre de films distribués par Vivafilm.

FIGURE 5.4 – Variables continues



Pour les variables continues de notre base de données (figure 5.4), nous remarquons que la distribution de la variable *minlecfm* est gonflée à 0 avec une queue plus longue à droite. La distribution de la variable *necranpresem* est asymétrique à droite tandis que la distribution du box-office par période et par cote reste symétrique et semble être très proche d'une normale. Nous observons globalement que la variance du logarithme du box-office est plus grande pour la somme des semaines subséquentes que la variance du logarithme du box-office de la première semaine. La variance du logarithme du box-office semble aussi augmenter légèrement quand la cote des films augmente.

5.2 Paramètres de variance-covariance

5.2.1 Sélection de la structure

Lors de la modélisation nous avons estimé plusieurs structures de corrélation afin d'utiliser le modèle le plus adéquat pour représenter le modèle théorique sous-jacent. Nous expliquons ici comment nous avons sélectionné l'équation 4.1. Premièrement, nous insérons un effet aléatoire qui introduit de la corrélation entre le même film présenté dans un même cinéma. Cette corrélation représente en réalité la nature temporelle des données qui doit être prise en compte. Par construction de la base de données, nous devons tenir compte de la dépendance temporelle et nous estimons l'équation suivante

$$Y_{cft} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + s1(\mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha}) + \omega_g + \varepsilon_{cft} \quad (5.1)$$

Par la suite, nous insérons un intercept aléatoire sur le film, ce qui introduit de la corrélation pour le même film présenté dans deux cinémas

différents quelle que soit la période

$$Y_{cft} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + s1(\mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha}) + \lambda_f + \omega_g + \varepsilon_{cft} \quad (5.2)$$

Nous supposons ensuite que cet intercepte aléatoire sur le film ne devrait pas être le même à chaque période, ce qui introduit de la corrélation pour le même film présenté dans deux cinémas différents à la même période, le paramètre estimé λ varie donc selon un indice t

$$Y_{cft} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + s1(\mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha}) + \lambda_{ft} + \omega_g + \varepsilon_{cft} \quad (5.3)$$

Enfin nous ajoutons un intercepte aléatoire sur les cinémas, ce qui introduit de la corrélation pour deux films différents présentés dans le même cinéma quelle que soit la période

$$Y_{cft} = \mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha} + s1(\mathbf{X}_f\boldsymbol{\beta} + \mathbf{Q}_c\boldsymbol{\alpha}) + \lambda_{ft} + \xi_c + \omega_g + \varepsilon_{cft} \quad (5.4)$$

Pour choisir adéquatement la structure de corrélation que nous utilisons, nous estimons les modèles par maximum de vraisemblance afin de pouvoir utiliser le *Akaike Information Criterion* (AIC) et le *Bayesian Information Criterion* (BIC). Selon ces critères d'information, nous choisissons le modèle 5.1 car les critères d'information sont minimums.

TABLE 5.7 – Choix de la structure de variance-covariance- Critères d'information

	AIC	BIC
Mod. 5.1	323769.21	324540.30
Mod. 5.2	312777.13	313557.86
Mod. 5.3	272489.61	273279.98
Mod. 5.4	204415.76	205215.77

Les critères d'information nous dirigent bien vers une structure de corré-

lation complexe quand nous modélisons le box-office avec le cadre conceptuel que nous utilisons. Sa prise en compte est cruciale afin que l'inférence de nos résultats soit valide lors de l'interprétation de la significativité des facteurs que nous voulons tester. De plus, l'ajout d'intercepts aléatoires vient apporter une précision accrue aux prévisions du modèle.

5.3 Effets fixes

Dans un premier temps, nous regardons quels facteurs sont significatifs en effectuant un test des effets de type III avec une approximation pour les degrés de liberté de Satterthwaite afin de tenir compte de l'estimation des paramètres de variance covariance (Satterthwaite, 1946). Par la suite, nous explorons plus en détails la significativité des facteurs et nous analysons leur valence dans le cadre de l'industrie cinématographique. Nous décomposons cette dernière partie en une analyse de la significativité de chaque paramètre estimé et en une analyse à la marge des effets des variables en nous aidant de graphiques. C'est sur cette dernière partie que nous confrontons nos hypothèses de recherche à nos résultats empiriques.

En observant les tests des effets de type III dans la table 5.8, nous constatons que l'interaction double ($s1 * minlecfm$), que le mois de sortie et le nombre d'écrans à la première semaine sont significatifs au seuil de 1% pour la première semaine.

L'interaction triple ($s1 * minlecfm * cote$) est significative au seuil de 5%. Nous voyons que le visa (5%), la durée (5%), le mois de sortie (1%) et le nombre d'écrans la première semaine (1%) sont significatifs pour les semaines subséquentes.

Globalement, nous confirmons qu'il existe un effet de la cote différent selon le nombre de minutes de lecture de journaux et que cet effet est

TABLE 5.8 – BOC - Test des effets de type III (Satterthwaite)

	Df	Som. Carrés	Carrés Moy.	Valeur F	Dénom	Pr(>F)
s1	1	7.37	7.37	24.17	1903.35	0.0000
s1 :minlecfm	1	18.29	18.29	27.54	53751.72	0.0000
s1 :cote	1	9.20	9.20	11.64	1928.86	0.0007
s1 :genre	6	5.17	0.86	0.92	1878.91	0.4786
s1 :visa	2	1.70	0.85	2.09	1880.09	0.1238
s1 :duree	4	2.83	0.71	0.39	1879.23	0.8173
s1 :mois	11	8.12	0.74	4.50	1925.40	0.0000
s1 :necranpremssem	1	10.12	10.12	53.33	1880.43	0.0000
s1 :distributeur	8	2.24	0.28	1.58	1880.88	0.1257
s1 :star1	1	0.19	0.19	1.10	1877.68	0.2941
s1 :star2	1	0.04	0.04	0.29	1877.11	0.5924
s1 :director1	1	0.03	0.03	0.19	1876.74	0.6613
s1 :minlecfm :cote	1	0.80	0.80	4.46	53772.16	0.0347
minlecfm	1	0.08	0.08	0.15	86.22	0.6953
cote	1	29.30	29.30	12.19	1069.41	0.0005
genre	6	25.19	4.20	1.87	1023.66	0.0837
visa	2	0.46	0.23	3.22	1024.57	0.0402
duree	4	22.77	5.69	2.43	1024.12	0.0461
mois	11	15.57	1.42	5.04	1041.17	0.0000
necranpremssem	1	234.45	234.45	717.65	1024.56	0.0000
distributeur	8	18.54	2.32	8.70	1023.02	0.0000
star1	1	0.03	0.03	0.19	1022.59	0.6617
star2	1	0.05	0.05	0.02	1022.26	0.8759
director1	1	0.26	0.26	1.18	1021.64	0.2779
minlecfm :cote	1	22.56	22.56	114.08	91463.79	0.0000

différent à travers les périodes que nous avons considérées.

TABLE 5.9 – BOC - Effets fixes

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	6.73409	0.23828	28.261	< 0.0001	***
s1	0.74002	0.20039	3.693	0.000228	***
minlecfm	-0.00674	0.01714	-0.393	0.695285	
cote	0.12352	0.03537	3.492	0.000499	***
minlecfm :cote	0.00662	0.00062	10.681	< 0.0001	***
s1 :minlecfm	0.00790	0.00151	5.248	1.55e-07	***
s1 :cote	-0.15114	0.04431	-3.411	0.000660	***
s1 :minlecfm :cote	-0.00127	0.00060	-2.111	0.034748	*

Signif. : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Avant de nous pencher sur l'interprétation à la marge de l'effet des facteurs, nous observons plus préciment le sens des effets dans la table 5.9 où figurent les estimations pour les variables sur lesquelles nous avons posé nos hypothèses de recherche. Le paramètre estimé pour l'interaction double entre les variables *minlecfm* et *cote* est significativement différent de 0 en permettant une erreur de première espèce de 1%. Le paramètre estimé pour l'interaction triple entre les variables *s1*, *minlecfm* et *cote* est lui aussi significativement différent de 0 en considérant une erreur de première espèce de 5%. Considérés seuls, les facteurs *s1* (1%) et *cote* (1%) et les interactions *s1 * minlecfm* (1%) et *s1 * cote* (1%) ont un effet significativement différent de 0 en considérant le seuil indiqué entre parenthèses. Nous voyons clairement que l'inclusion du facteur *minlecfm* et des interactions est nécessaire pour expliquer la variation du logarithme du box-office.

5.3.1 box-office composé et interprétation graphique

Selon les résultats de nos estimations, nous pouvons constater que l'interaction significative avec un paramètre estimé positif pour les semaines subséquentes signifie que la lecture de journaux quotidiens accentue l'effet de la cote des films émis par les critiques. L'effet marginal des minutes de lectures pour les semaines subséquentes dépend du niveau de la cote et il peut s'exprimer ainsi

$$\frac{\partial Y_{cf2}}{\partial minlecfin} = 0.0066172 * cote - 0.0067360$$

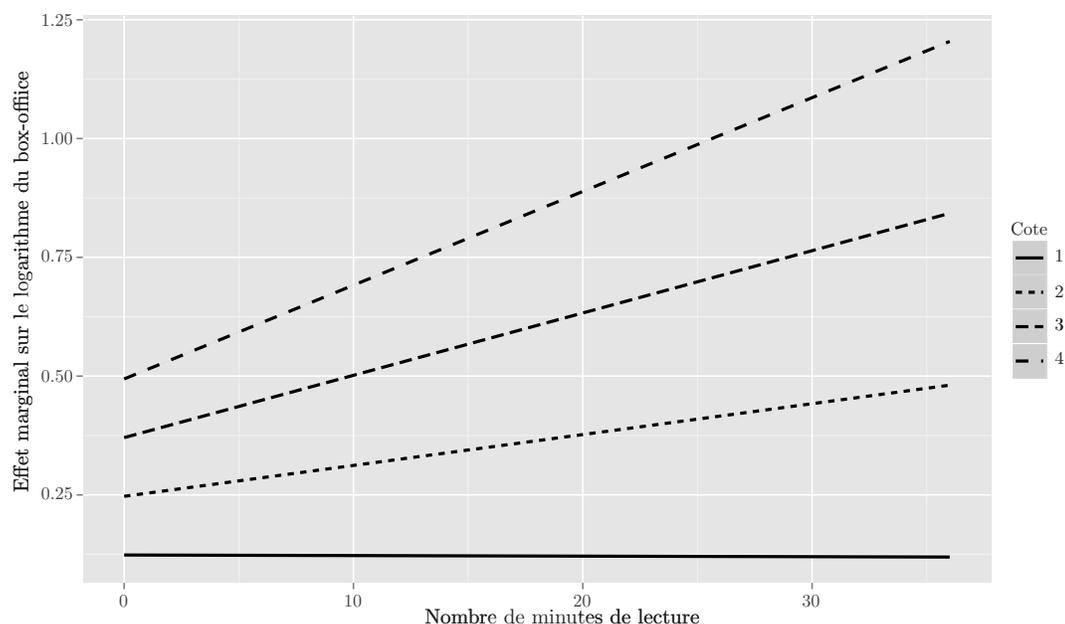
Pour la première semaine, l'interaction significative avec un paramètre estimé négatif (-0.0012676) et inférieur en valeur absolue à l'effet des semaines subséquentes (0.0066172) signifie que l'effet est différent la première semaine. L'effet net de l'interaction est donc positif (0,0053496) et le nombre de minutes de lecture de journaux accentue l'effet de la cote des films émis par les critiques. L'effet marginal des minutes de lectures pour la première semaine dépend lui aussi du niveau de la cote et il peut s'exprimer ainsi

$$\begin{aligned} \frac{\partial Y_{cf1}}{\partial minlecfin} &= (0.0066172 - 0.0012676) * cote + (0.0078984 - 0.0067360) \\ &= (0.0053496) * cote + 0.0011624 \end{aligned}$$

Ainsi, plus les consommateurs sont exposés aux informations, plus ils seront influencés par la valence des critiques et l'effet est plus important pour la première semaine. En effet, nous constatons que les pentes des droites pour les semaines subséquentes sur la figure 5.5 sont moins accentuées et révèlent donc un effet moins marqué de l'information par rapport aux pentes des droites de la figure 5.6.

Nous pouvons observer graphiquement les effets avec une analyse séparée par période afin de mieux mettre en exergue les différences qu'il existe.

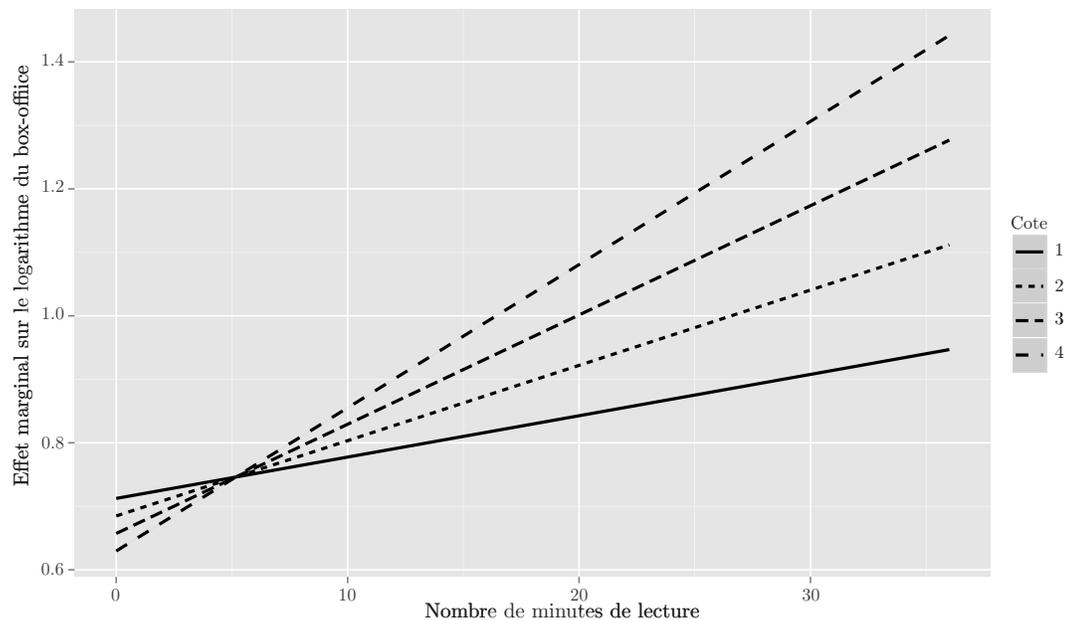
FIGURE 5.5 – Effets marginaux du nombre de minutes de lecture pour les semaines subséquentes



Plus particulièrement, pour les semaines subséquentes, plus les consommateurs sont exposés aux critiques et plus l'effet de la cote est accentué. Nous voyons clairement sur le graphique 5.5 que l'écart entre les droites s'accroît lorsque les films sont présentés dans des subdivisions de recensement où les consommateurs sont plus exposés à l'information. Nous pouvons comprendre qu'il existe une congruence entre les critiques cinématographiques et les préférences des lecteurs et que cette congruence est plus forte si les lecteurs sont plus exposés à l'information. Les ordonnées à l'origine représentent les effets non contrôlés d'autres sources d'informations externes comme celui du bouche à oreille en interaction avec la critique.

Si nous nous intéressons à l'effet marginal des variables *cote*, *minlecf*, *s1* et de leur interaction pour la première semaine, nous constatons que la réaction des consommateurs est différente. En effet, les consommateurs

FIGURE 5.6 – Effets marginaux du nombre de minutes de lecture pour la première semaine



qui ont une exposition inférieure à 5,17 minutes (le point d'intersection entre les quatre droites) ont un comportement flou en ce qui à trait à la cote. Notre modèle estime un effet qui semble être à l'inverse de la réponse des consommateurs observées les semaines subséquentes ; une cote plus forte entraîne des revenus plus bas sur cette première portion. Ce comportement reflète le besoin pour consommateur d'accumuler de l'information.

La deuxième partie des effets réagit d'une manière intuitive car plus la cote augmente plus les revenus sont élevés. De plus, nous voyons que plus les consommateurs sont exposés à l'information et plus l'effet de la valence de la cote est important avec des pentes plus marquées que pour les semaines subséquentes.

Ces résultats soutiennent directement nos hypothèses de recherche **H1**

et sont intuitifs à interpréter. La première semaine, l'asymétrie d'information est forte pour un bien expérientiel. Nous voyons que l'effet de la cote est flou lorsque l'exposition à l'information est nulle ou très faible. Il faut alors qu'un certain niveau d'information soit accumulé pour que les consommateurs puissent considérer le niveau de la cote.

Pour les semaines subséquentes, nous pouvons voir que même si les consommateurs ne sont pas exposés aux critiques, il existe une différence selon le niveau de la cote. Cet effet peut traduire le fait que l'information sur le produit est plus accessible étant donné que d'autres consommateurs en ont eu l'expérience et ont pu la partager. De même, de nouvelles sources d'information auront pu faire leur apparition étant donné l'exclusivité que peuvent obtenir certains médias avant la sortie des films.

Lorsque le contexte informationnel est restreint pour la première semaine, nos effets estimés suggèrent que les critiques sont des sources influentes après l'accumulation d'un certain niveau d'information. Quand le contexte informationnel est plus large, les consommateurs ont déjà acquis assez d'information grâce aux autres sources pour que la valence des critiques ait un effet clair à partir d'un niveau d'exposition à cette source supérieur à 0 minutes de lecture.

Le degré d'exposition à des sources d'information externes est une composante non-négligeable lorsque l'on cherche à expliquer la performance des films au box-office. De plus, notre modélisation conjointe des effets pour les deux périodes considérées dans ce contexte est très pertinente pour faire ressortir la présence d'un effet d'influence qui diffère à travers le temps. Nous avons montré qu'il existe un effet temporel intéressant lié à l'accumulation d'information qui s'explique par la difficulté qu'un consommateur éprouve pour évaluer les caractéristiques intangibles d'un bien expérientiel. Nous pouvons noter que le mécanisme de réduction de l'asymétrie de l'information joue un rôle important dans le processus

de mise en marché de ce type de produits.

5.3.2 Analyse comparative des modèles estimés

Nous nous intéressons ici aux différences entre la modélisation conjointe des deux semaines et la modélisation séparée.

TABLE 5.10 – Comparatif des effets fixes

		Est. conjointe		Est. séparée	
		Estimés	Pr(> t)	Estimés	Pr(> t)
Prem. Sem.	(s1) minlecfm	0.00110	< 0.0001	0.00042	0.979406
	(s1) cote	-0.02762	0.000660	-0.02701	0.325023
	(s1) minlecfm :cote	0.00535	0.034748	0.00523	< 0.0001
	s1	0.74002	< 0.0001		
Sem. Subséqu.	minlecfm	-0.00674	0.695285	-0.00570	0.750738
	cote	0.12352	0.000499	0.12919	0.000305
	minlecfm :cote	0.00662	< 0.0001	0.00681	< 0.0001
Total	minlecfm			-0.00196	0.907547
	cote			0.07714	0.003841
	minlecfm :cote			0.00570	< 0.0001

La table 5.10 présente les résultats des estimations concernant les facteurs considérés dans nos hypothèses de recherche. Nous pouvons observer que dans tous les cas, l'interaction entre les minutes de lectures de journaux quotidiens et la cote des films est significative. Les estimations de notre modèle sont cohérentes avec les estimations séparées et nous constatons qu'il existe une certaine différence entre les paramètres estimés. L'estimation séparée estime un effet plus faible conjoint de la cote et des minutes de lecture pour la première semaine. Concernant les semaines subséquentes, les effets sont très similaires avec une légère surestimation des paramètres lorsqu'ils sont estimés séparément. Nous constatons que les cinq modèles donnent une interaction significative que l'on peut interpréter comme une différence de cote selon le niveau de lecture ce qui soutient notre hypothèse de recherche principales **H1**. Selon le cadre posé

par nos hypothèses et la littérature sur l'effet d'influence et l'effet de prédiction, nous confirmons l'hypothèse **H4** et nous infirmons les hypothèses **H2** et **H3**. En effet, nous ne pouvons pas clairement distinguer l'un ou l'autre des effets car ils sont tous les deux présents. Nous considérons ici la significativité des paramètres estimés pour notre modélisation conjointe et notre modélisation du box-office total. Ces résultats montrent que les deux effets existent et que l'effet de prédiction est plus marqué que l'effet d'influence de part la valeur de l'interaction entre la variable *cote* et la variable *minlecfin*.

La taille de notre échantillon nous permet d'appuyer avec confiance la présence d'un effet d'influence et celle d'un effet de prédiction. Nous avons un total de 1065 films contre 175 dans l'étude de Basuroy, Chatterjee et Ravid (2003) et 56 dans celle de Eliashberg et Shugan (1997). De plus, nous avons accès à des données sur une période plus longue (2002 à 2009) que pour les échantillons sélectionnés dans les articles parus. La note des critiques utilisée est différente au niveau de l'échelle et de la source ce qui peut aussi expliquer que nos effets sont significatifs pour la première semaine contrairement aux résultats de Eliashberg et Shugan (1997) et à ceux de Basuroy, Chatterjee et Ravid (2003) qui trouvent un effet des mauvaises critiques significatif et ayant un impact plus marqué que les bonnes critiques.

Cette congruence des résultats avec différentes façons de modéliser les relations font preuve de leur robustesse. Nous pouvons donc comparer sereinement nos résultats avec les résultats publiés dans la littérature.

5.3.3 Comparaison avec la littérature

Nous allons dans un deuxième temps comparer nos résultats avec les résultats développés par la littérature. Nous commençons l'analyse des

variables que nous étudions et nous continuons avec l'analyse des covariables qui ont été incluses dans le modèle. La table B.1 dans l'annexe B donnent accès aux estimations des paramètres des modèles estimés avec un groupe de base correspondant à la moyenne d'une modalité. La table B.2 dans l'annexe B donne accès aux estimations des paramètres estimés avec un groupe de base correspondant à la moyenne de toutes les modalités des variables (*simple coding*). Nous utilisons les deux méthodes pour interpréter les résultats que nous avons estimés.

Cote

Pour l'effet d'influence, nous mesurons un effet marginal qui est positif sur les revenus du box-office à partir de 5,17 minutes de lectures de journaux. La significativité du coefficient devant les trois facteurs nous assure qu'il existe un effet conjoint de la cote, des minutes de lectures de journaux de la période et de l'interaction triple. Notre résultat va à l'inverse de celui de Eliashberg et Shugan (1997) mis au jour sur le marché des États-Unis et qui stipule que les critiques cinématographiques n'ont qu'un effet de prédiction et celui de Ravid (1999) qui découvre que les critiques positives n'ont pas d'impact sur les revenus du box-office la première semaine. D'un autre côté, nous confirmons les résultats de Basuroy, Chatterjee et Ravid (2003) qui trouvent un effet significatif des critiques la première semaine. Nos estimations vont aussi dans le sens des résultats de Reinstein et Snyder (2000) qui trouvent aussi un effet significatif positif des critiques sur les revenus du box-office la première semaine.

Nous estimons un effet marginal de la cote, significatif et positif sur les revenus du box-office pour les semaines subséquentes ($0.1235159 + 0.0066172 * minlecfin$) comme la variable *minlecfin* est strictement positive. Ce résultat coïncide avec les estimations de nombreux auteurs qui de leur côté considèrent les revenus du box-office cumulés (Litman 1982,

Litman et Kohl 1989, Prag et Casavant 1994, Sawhney et Eliashberg 1996, Eliashberg et Shugan 1997, Litman et Ahn 1998, Jedidi et al. 1998, Ravid 1999, Zufryden 2000, Brewer Kelley Jozefowicz).

Genre

Pour les premières semaines, il n'existe aucun effet du genre du film sur les revenus du box-office (table 5.8). Pour la deuxième semaine, nous observons une différence en permettant une erreur de deuxième espèce de 10% (table 5.8). Les revenus du box-office sont significativement moins élevés pour les comédies, les films à suspense et les films d'horreur par rapport aux drames (table B.1) lorsque nous contrôlons pour les effets des autres variables. Par rapport à la moyenne de toutes les modalités un film de suspense a un revenu moins élevé de 21,3% par rapport à la moyenne et un film d'horreur un revenu 26,9% moins élevé par rapport à la moyenne (table B.2). Ces résultats diffèrent des observations de Simonoff et Sparrow (2000), de Terry, Butler et de Armond (2011) et de Terry, Cooley et Zachary (2010). Ces études admettent que les films d'action ou pour enfant (Terry, Cooley et Zachary 2008, Terry, Butler et de Armond 2011) et les films d'horreur ou de science fiction (Simonoff et Sparrow, 2000) génèrent des revenus plus élevés que la moyenne du reste des catégories. Il est à noter que la classification des films dans certains genre est différente entre les études.

Visa

Dans notre modèle, le visa n'a pas un effet significatif sur les revenus du box-office comme mis au jour par Brewer, Kelley et Jozefowicz (2009) pour la première semaine (table 5.8). L'effet restrictif des films interdits au moins de 16 ans ou au moins de 18 ans sur le logarithme du

box-office des semaines subséquentes (table 5.8) est par contre confirmé, comme dans les études de Leenders et Eliashberg (2011), De Vany et Walls (2002) , Ravid (1999) , Ravid et Basuroy (2004), Walls (2005) et Dhar, Sun et Weinberg (2012) . Un film ayant un visa 16 et plus ou 18 et plus a un revenu 30,3% moins élevé que la moyenne des deux autres catégories (table B.2. En effet, un film qui s'adresse à un groupe restreint de consommateurs à moins de chance d'avoir un nombre d'entrées élevé qu'un film tout public. Étant donné le prix relativement fixe des billets de cinéma, les revenus des films à public restreint est donc moins élevé.

Durée

Alors que pour la première semaine, la durée n'a pas un impact significatif sur le logarithme des revenus du box-office, nous pouvons constater que pour les semaines subséquentes, il existe une différence de revenu positif entre les films qui ont une durée comprise entre 40 et 90 minutes inclus et les films qui ont une durée comprise entre 110 minutes et 124 minutes inclus (table 5.8 et table B.1). Nous estimons qu'il existe une différence positive de 27,42% en faveur des films qui ont une durée comprise entre 110 et 124 minutes inclus et la moyenne de tous les films (table B.2). Par rapport aux résultats de Bonnet (2009) qui mesure une relation significative positive entre la durée d'un film et ses revenus au box-office, nous constatons que la durée du film n'a pas d'influence sur les revenus du box-office de la première semaine. Nous voyons que les films plus long performant mieux que la moyenne et nous pouvons attribuer cet effet d'une part au prix fixe d'un billet de cinéma et d'autre part au choix d'allocation en capital mental (Holbrook et Hirschman,1982) et d'allocation du temps (Becker, 1965)). En effet, le consommateur doit choisir d'allouer un certain capital mental aux activités cognitives et optimise son utilisé sous contrainte de temps. Nous nous attendons à ce

qu'il existe un niveau pour lequel il puisse substituer les minutes supplémentaires à regarder un film à une autre activité qui lui apportera une utilité marginale plus grande.

Mois

Pour les revenus de la première semaine, et des semaines subséquentes nous observons un effet significatif de la variable *mois* qui vient contrôler pour l'effet de saisonnalité (table 5.8). En particulier, le mois de décembre a un effet négatif sur le box-office par rapport à janvier la première semaine. Pour les semaines subséquentes, avril, mai, juin, septembre et octobre sont des mois de sortie qui ont un effet négatif sur les revenus du box-office par rapport à janvier. Nous observons des revenus moins élevés de 27,3% en avril, des revenus moins élevés de 61,1% en juillet, et des revenus moins élevés de 36,4% en août comparés à la moyenne de notre échantillon (table B.2). Ces résultats confirment qu'il existe une compétition marquée en saison (correspondant à la période d'été) qui touche particulièrement le box-office des semaines subséquentes. En effet, étant donné le prix relativement fixe des billets de cinéma, le nombre de films sur le marché reflète le niveau de compétition. Nous avons vu dans la section descriptive que la plupart des films sortaient entre avril et octobre, entraînant ainsi une compétition plus élevée. Nos résultats coïncident avec les observations de Cartier et Liarte (2012) et Einav (2007) qui estiment qu'il existe des périodes propices à la sortie de films qui leur garantissent un plus grand succès commercial. Einav(2007) découvrent que cet effet est amplifié par la décision des distributeurs de lancer leur films durant les mêmes périodes. Cartier et Liarte (2012) ont mis au jour les éléments de saisonnalité qui entraînent une agglomération des sorties de film en saison. Ces décisions de lancement sont révélées comme sous optimales dans nos résultats et sont expliquées par Cartier et Liarte

(2012) comme une surestimation de la demande de films pour certaines périodes.

Distributeur

Pour la première semaine, il n'existe pas d'effet spécifique à la variable distributeur dans nos estimations (table 5.8). Par contre, l'effet est présent pour les semaines subséquentes et nous pouvons constater que les films distribués par Disney engendrent un revenu 40,7% supérieur à la moyenne (table 5.8 et table B.2). De l'autre côté, les films distribués par Warner engendrent un revenu 39,9% inférieur à la moyenne et les films distribués par les distributeurs indépendants et autres engendrent un revenu 20,4% inférieur à la moyenne (table B.2). L'effet positif concernant les films Disney peut être considéré sous l'angle du public ciblé. En effet, la plupart des films distribués par Disney sont tout public et s'adressent donc à une population plus large ce qui peut expliquer leur succès. De plus, Disney dispose d'un élément de marque plus important que les autres distributeurs. Pour les petits distributeurs de la catégorie autre, c'est la position de force des grands distributeurs et leur faible pouvoir de négociation qui peut expliquer la performance plus faible de leurs films au box-office.

Effet star

Que ce soit pour la première semaine ou pour les semaines subséquentes, les variables *star1*, *star2* et *director1* n'ont pas un effet significatif sur le box-office (table 5.8). Ces résultats soutiennent les études de Litman (1983), Austin (1989), DeVany et Walls (1999), Litman et Ahn (1998), Ravid (1999) et vont à l'opposé des résultats de Levin et al. (1997), Ainslie, Drèze et Zufryden (2005), Albert (1998), Basuroy, Chartterjee

and Ravid (2003), Litman et Kohl (1989), Desay et Basuroy (2005) et Elberse (2007).

5.3.4 Nombre d'écrans à la première semaine

Pour la première semaine comme pour les semaines subséquentes, la variable *necranpremsem*, qui reflète l'effort marketing investi par les distributeurs, a un effet très significatif sur les revenus du box-office (table 5.8). *Ceteris paribus*, l'ajout d'un écran supplémentaire la semaine de sortie d'un film bonifie les revenus de 2,78%. Autrement dit, plus le budget marketing pour la sortie d'un film est élevé et plus son succès commercial est assuré. Nous soutenons donc les résultats de Prag et Casavant (1994),

5.3.5 Synthèse de l'analyse

Nous avons confirmé de nombreux résultats déjà parus dans la littérature, notamment pour l'effet de la cote qui est une composante de l'effet d'interaction que nous avons estimé significatif. En plus de pouvoir interpréter séparément l'interaction de nos trois variables d'intérêt entre les deux périodes, nous pouvons confirmer que l'effet évolue à travers le temps. Cette dynamique est importante à l'heure où l'information est accessible par la quasi-totalité des consommateurs sur le marché et où la stratégie média peut être rapidement adaptée. D'autre part, la significativité des effets concernant le nombre d'écran à la première semaine, qui représente le budget marketing, est l'un des facteurs influençant le plus les revenus. Ceci confirme que la diffusion de l'information, que ce soit par des sources externes ou directement par le distributeur, permet au film de gagner en popularité et d'être mieux cerné par les consommateurs. Préciser la structure et la magnitude des effets des variables qui gravitent

autour du processus d'acquisition d'information de la part du consommateur est donc crucial afin de comprendre pourquoi et comment les films performant. La dynamique de cette acquisition d'information doit être prise en compte par les distributeurs et les propriétaires de salles afin de pouvoir s'adapter aux réactions du marché à chaque semaine. Ces acteurs de l'industrie peuvent pousser la communication envers leurs films dans les régions les plus à même de générer un profit élevé en identifiant les régions géographiques où les consommateurs sont exposés plus de 5,17² minutes aux critiques cinématographiques la première semaine. Par la suite, ils se doivent d'ajuster leur stratégie de diffusion et prendre en compte la cote d'une manière plus restrictive étant donné son effet qui peut être négatif pour les films répertoriés comme faibles. Nous remarquons donc que la stratégie de diffusion d'information concernant un film et la décision d'un propriétaire de salle de garder le film à l'écran sont étroitement liées.

2. Il est à noter que notre seuil correspond spécifiquement à notre échantillon et qu'il serait plus prudent d'estimer un intervalle de confiance afin de rendre notre estimation relative à la dispersion des données.

Chapitre 6

Conclusion

6.1 Objectifs de recherche

Dans ce travail de recherche, nous avons modélisé l'effet des critiques cinématographiques sur la performance des films au box-office en utilisant un niveau d'observation précis, que ce soit sur la dimension temporelle ou sur la dimension géographique. Nous nous sommes servi de ce niveau de précision afin de construire des mesures permettant d'estimer d'une nouvelle manière les relations de dépendance liant les caractéristiques des produits et les informations disponibles les concernant à leur succès commercial. Nous avons effectué une analyse qui nous a amené à confirmer qu'il existe un effet différent de la valence d'une source d'information (variable *cote*) selon le niveau auquel le consommateur y est exposé (variable *minlecfm*) sur le succès commercial d'un produit. Nous avons pris en compte la nature temporelle de cet effet et avons relevé une différence entre la première semaine et les semaines subséquentes. Ces résultats nous permettent de confirmer l'hypothèse **H1** et donc l'importance de considérer une nouvelle dimension à prendre en compte à travers le temps lors du processus de mise en marché des biens expérientiels. La confirmation

de l'hypothèse **H4** nous indique la présence d'un effet d'influence et d'un effet de prédiction

6.2 Apports du modèle

Notre modèle amène un nouvel axe de développement dans l'analyse des déterminants du succès commercial et dans l'analyse des facteurs qui influencent l'achat des produits sur des marchés pour lesquels l'asymétrie d'information entre les caractéristiques des produits et les consommateurs est forte. Naturellement, nous pouvons étendre nos résultats à des industries similaires, où les produits ont un court cycle de vie et les prix restent stables, comme le marché des livres, des Blu-rays et de la vente de musique en ligne. Nous nous sommes intéressés à un niveau de prédiction précis qui prend en compte des effets de corrélation croisés avec des effets propres aux films, différents selon la période, des effets propres aux cinémas et des effets propres à chaque film présenté dans chaque cinéma. Les composantes aléatoires ainsi estimées permettent de prédire précisément les moyennes conditionnelles du box-office des films à l'aide des variables retenues pour l'analyse. Nous avons ainsi estimé un modèle en utilisant deux dimensions qui nous permettent de mieux situer les effets estimés et de prendre en compte des informations temporelles et des informations spatiales.

D'un point de vue managérial, le degré d'exposition aux sources d'information externe peut être quantifié et utilisé afin de diriger les efforts de communication aux bons endroits et au bon moment. La diffusion de l'information concernant les films au niveau géographique est en effet une étape cruciale qui permet de réduire l'asymétrie d'information et de faciliter le choix de consommateurs. Nous avons vu dans nos résultats que ce niveau d'information influence d'une manière importante le choix de

consommation des biens expérientiels quand ils bénéficient d'une réaction positive de sources reconnues comme les critiques cinématographiques.

Il est clair que certains distributeur ont un pouvoir de négociation plus important et que les petits distributeurs ne peuvent pas se permettre de diffuser le film en salles d'une manière extensive la première semaine. C'est pour cette raison que le découpage de l'effet d'exposition à l'information d'une manière géographique grâce à la variable *minlecf* se révèle importante. La capacité de pouvoir mesurer plus précisément le retour sur investissement et d'établir une stratégie de communication dans les médias sur une base géographique permet au petits distributeurs et aux propriétaires de salles de cibler les régions plus à même de générer du profit. Après avoir estimé précisément les effets du degré d'exposition aux critiques, il est possible de pouvoir considérer cette nouvelle variable afin d'établir une stratégie de diffusion de l'information visant à optimiser le profit des films distribués en salles.

6.3 Limites et avenues de recherche

Les modèles estimés pour nos analyses considèrent un proxy (*minlecf*) de l'exposition des consommateurs aux critiques cinématographiques. Cette méthode est concluante et intuitive lors de l'interprétation des effets. Cependant, notre travail est directement effectué sur la variable proxy et son interaction entre le box-office et la variable temporelle. Il serait donc préférable d'accéder à une variable plus précise concernant l'exposition aux sources d'information externe et des variables concernant le bouche-à-oreille.

En effet, les consommateurs peuvent accéder à de nombreuses sources d'informations sur Internet (Gazley, Gemma et Sinha 2011) telle que des critiques ou encore les notes et commentaires des internautes et il est

possible, grâce à la construction d'outils sur mesure et l'exploitation des capacités de certains sites Internet, de pouvoir télécharger de nombreuses informations sur la fréquentation de telles sources d'information. Étant donné la disponibilité des données, il serait intéressant de pouvoir préciser la variable concernant la cote en utilisant des bases de données que l'on peut exploiter en ligne, tel que Rotten Tomatoes ou IMDb. Il serait aussi possible de trouver des variables qui représentent correctement le bouche à oreille, que nous n'avons pas inclus dans nos estimations.

Grâce à des variables plus précises et comportant une plus grande variation, il serait aussi possible d'estimer plus précisément les effets sur la dimension temporelle. En effet, alors que nous avons agrégé le box-office pour les semaines subséquentes, un modèle qui prend en compte chaque semaine individuellement peut être estimé. Il faudrait alors prendre en compte l'élément de sélection des films une longévité plus grande. Nous avons utilisé une composante géographique de notre variable proxy et il est possible d'estimer des modèles économétriques spatiaux qui peuvent faire ressortir les effets d'une nouvelle manière grâce aux longitudes et latitudes de chaque cinéma. Il serait aussi possible d'inclure des variables sociodémographiques telles que le revenu, le niveau d'études ou encore la diversité socioculturelle qui peuvent avoir une influence sur le box-office et pourraient elles aussi avoir un effet de modération sur la variable concernant la cote. Globalement, le travail méthodologique effectué dans cette étude ouvre de nombreuses voies de recherches qu'il serait possible d'emprunter pour enrichir la compréhension du processus de choix du consommateur et de l'industrie cinématographique.

Annexe A

Modèles estimés

Effets aléatoires - Modèle 1

$$\begin{aligned} Y_{cft} = & \beta_1 \text{minlecfinc} + \beta_2 \text{cote}_f + \beta_3 \text{minlecfinc} * \text{cote}_f + \beta_4 \text{genre}_f + \beta_5 \text{visa}_f \\ & + \beta_6 \text{duree}_f + \beta_7 \text{mois}_f + \beta_8 \text{necranpresem}_f + \beta_9 \text{distributeur}_f + \beta_{10} \text{star1}_f \\ & + \beta_{11} \text{star2}_f + \beta_{12} \text{director1}_f + \beta_{13} \text{s1}_f + \beta_{14} \text{s1}_f * \text{minlecfinc} \\ & + \beta_{15} \text{s1}_f * \text{cote}_f + \beta_{16} \text{s1}_f * \text{minlecfinc} * \text{cote}_f + \beta_{17} \text{s1}_f * \text{cote}_f \\ & + \beta_{18} \text{s1}_f * \text{genre}_f + \beta_{19} \text{s1}_f * \text{visa}_f + \beta_{20} \text{s1}_f * \text{duree}_f \\ & + \beta_{21} \text{s1}_f * \text{mois}_f + \beta_{22} \text{s1}_f * \text{necranpresem}_f + \beta_{23} \text{s1}_f * \text{distributeur}_f \\ & + \beta_{24} \text{s1}_f * \text{star1}_f + \beta_{25} \text{s1}_f * \text{star2}_f + \beta_{26} \text{s1}_f * \text{director1}_f \\ & + \omega_g + \varepsilon_{cft} \end{aligned} \tag{A.1}$$

Effets aléatoires - Modèle 2

$$\begin{aligned}
Y_{cft} = & \beta_1 \text{minlecfinc} + \beta_2 \text{cote}_f + \beta_3 \text{minlecfinc} * \text{cote}_f + \beta_4 \text{genre}_f + \beta_5 \text{visa}_f \\
& + \beta_6 \text{duree}_f + \beta_7 \text{mois}_f + \beta_8 \text{necranpremssem}_f + \beta_9 \text{distributeur}_f + \beta_{10} \text{star1}_f \\
& + \beta_{11} \text{star2}_f + \beta_{12} \text{director1}_f + \beta_{13} \text{s1}_f + \beta_{14} \text{s1}_f * \text{minlecfinc} \\
& + \beta_{15} \text{s1}_f * \text{cote}_f + \beta_{16} \text{s1}_f * \text{minlecfinc} * \text{cote}_f + \beta_{17} \text{s1}_f * \text{cote}_f \\
& + \beta_{18} \text{s1}_f * \text{genre}_f + \beta_{19} \text{s1}_f * \text{visa}_f + \beta_{20} \text{s1}_f * \text{duree}_f \\
& + \beta_{21} \text{s1}_f * \text{mois}_f + \beta_{22} \text{s1}_f * \text{necranpremssem}_f + \beta_{23} \text{s1}_f * \text{distributeur}_f \\
& + \beta_{24} \text{s1}_f * \text{star1}_f + \beta_{25} \text{s1}_f * \text{star2}_f + \beta_{26} \text{s1}_f * \text{director1}_f \\
& + \lambda_f + \omega_g + \varepsilon_{cft}
\end{aligned} \tag{A.2}$$

Effets aléatoires - Modèle 3

$$\begin{aligned}
Y_{cft} = & \beta_1 \text{minlecfinc} + \beta_2 \text{cote}_f + \beta_3 \text{minlecfinc} * \text{cote}_f + \beta_4 \text{genre}_f + \beta_5 \text{visa}_f \\
& + \beta_6 \text{duree}_f + \beta_7 \text{mois}_f + \beta_8 \text{necranpremssem}_f + \beta_9 \text{distributeur}_f + \beta_{10} \text{star1}_f \\
& + \beta_{11} \text{star2}_f + \beta_{12} \text{director1}_f + \beta_{13} \text{s1}_f + \beta_{14} \text{s1}_f * \text{minlecfinc} \\
& + \beta_{15} \text{s1}_f * \text{cote}_f + \beta_{16} \text{s1}_f * \text{minlecfinc} * \text{cote}_f + \beta_{17} \text{s1}_f * \text{cote}_f \\
& + \beta_{18} \text{s1}_f * \text{genre}_f + \beta_{19} \text{s1}_f * \text{visa}_f + \beta_{20} \text{s1}_f * \text{duree}_f \\
& + \beta_{21} \text{s1}_f * \text{mois}_f + \beta_{22} \text{s1}_f * \text{necranpremssem}_f + \beta_{23} \text{s1}_f * \text{distributeur}_f \\
& + \beta_{24} \text{s1}_f * \text{star1}_f + \beta_{25} \text{s1}_f * \text{star2}_f + \beta_{26} \text{s1}_f * \text{director1}_f \\
& + \lambda_{ft} + \omega_g + \varepsilon_{cft}
\end{aligned} \tag{A.3}$$

Effets aléatoires - Modèle 4

$$\begin{aligned}
Y_{cft} = & \beta_1 \text{minlecfinc} + \beta_2 \text{cote}_f + \beta_3 \text{minlecfinc} * \text{cote}_f + \beta_4 \text{genre}_f + \beta_5 \text{visa}_f \\
& + \beta_6 \text{duree}_f + \beta_7 \text{mois}_f + \beta_8 \text{necranpremssem}_f + \beta_9 \text{distributeur}_f + \beta_{10} \text{star1}_f \\
& + \beta_{11} \text{star2}_f + \beta_{12} \text{director1}_f + \beta_{13} s1_f + \beta_{14} s1_f * \text{minlecfinc} \\
& + \beta_{15} s1_f * \text{cote}_f + \beta_{16} s1_f * \text{minlecfinc} * \text{cote}_f + \beta_{17} s1_f * \text{cote}_f \\
& + \beta_{18} s1_f * \text{genre}_f + \beta_{19} s1_f * \text{visa}_f + \beta_{20} s1_f * \text{duree}_f \\
& + \beta_{21} s1_f * \text{mois}_f + \beta_{22} s1_f * \text{necranpremssem}_f + \beta_{23} s1_f * \text{distributeur}_f \\
& + \beta_{24} s1_f * \text{star1}_f + \beta_{25} s1_f * \text{star2}_f + \beta_{26} s1_f * \text{director1}_f \\
& + \lambda_{ft} + \xi_c + \omega_g + \varepsilon_{cft}
\end{aligned} \tag{A.4}$$

Annexe B

Tableaux des paramètres estimés

TABLE B.1 – BOC - Effets fixes

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	6.7340918	0.2382789	28.261	< 0.0001	***
s1	0.7400235	0.2003871	3.693	0.000228	***
minlecfm	-0.0067360	0.0171397	-0.393	0.695285	
cote	0.1235159	0.0353710	3.492	0.000499	***
genre2	-0.1535470	0.0801015	-1.917	0.055528	.
genre3	-0.2130591	0.0986610	-2.160	0.031043	*
genre4	-0.1032632	0.1115628	-0.926	0.354869	
genre5	-0.1769287	0.1296589	-1.365	0.172688	
genre6	0.0117318	0.1115921	0.105	0.916292	
genre7	-0.2692550	0.1022587	-2.633	0.008589	**
visa2	-0.0928589	0.0607967	-1.527	0.126978	
visa3	-0.3026523	0.1246624	-2.428	0.015363	*
duree2	0.1479190	0.0786386	1.881	0.060256	.
duree3	0.1406555	0.0809405	1.738	0.082552	.
duree4	0.2742291	0.0884321	3.101	0.001981	**

Continue sur la prochaine page. . .

TABLE B.1 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
duree5	0.1946875	0.1065839	1.827	0.068049	.
mois2	-0.1168566	0.1256045	-0.930	0.352407	
mois3	-0.0611074	0.1267685	-0.482	0.629881	
mois4	-0.2725811	0.1263769	-2.157	0.031246	*
mois5	-0.6106966	0.1379294	-4.428	1.06e-05	***
mois6	-0.3635029	0.1299828	-2.797	0.005262	**
mois7	-0.0898996	0.1275186	-0.705	0.480970	
mois8	-0.2349729	0.1253447	-1.875	0.061126	.
mois9	-0.2625285	0.1263076	-2.078	0.037909	*
mois10	-0.2882859	0.1244431	-2.317	0.020721	*
mois11	-0.1702739	0.1356061	-1.256	0.209529	
mois12	0.1989831	0.1322419	1.505	0.132711	
necranpremsem	0.0278417	0.0010393	26.789	< 0.0001	***
distributeur2	0.4064860	0.1118941	3.633	0.000294	***
distributeur3	0.0155237	0.0906684	0.171	0.864089	
distributeur4	-0.2910660	0.1595841	-1.824	0.068459	.
distributeur5	-0.1316766	0.0913415	-1.442	0.149725	
distributeur6	0.0477573	0.0995701	0.480	0.631589	
distributeur7	0.1382946	0.0981269	1.409	0.159037	
distributeur8	-0.3985861	0.0886336	-4.497	7.68e-06	***
distributeur9	-0.2036772	0.0799617	-2.547	0.011004	*
star11	-0.0258931	0.0591646	-0.438	0.661735	
star21	0.0089598	0.0573380	0.156	0.875857	
director11	-0.0910187	0.0838340	-1.086	0.277867	
minlecfm :cote	0.0066172	0.0006195	10.681	< 0.0001	***
s1 :minlecfm	0.0078984	0.0015051	5.248	1.55e-07	***
s1 :cote	-0.1511423	0.0443071	-3.411	0.000660	***
s1 :genre2	0.0328708	0.1007666	0.326	0.744304	
s1 :genre3	0.2165963	0.1241270	1.745	0.081156	.

Continue sur la prochaine page...

TABLE B.1 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
s1 :genre4	0.1785380	0.1403732	1.272	0.203573	
s1 :genre5	-0.0417667	0.1631309	-0.256	0.797954	
s1 :genre6	0.1370548	0.1404029	0.976	0.329114	
s1 :genre7	0.1343897	0.1286705	1.044	0.296412	
s1 :visa2	0.1358206	0.0764876	1.776	0.075941	.
s1 :visa3	0.2446169	0.1568314	1.560	0.118989	
s1 :duree2	-0.0887262	0.0989300	-0.897	0.369910	
s1 :duree3	-0.0525391	0.1018207	-0.516	0.605918	
s1 :duree4	-0.1236543	0.1112601	-1.111	0.266539	
s1 :duree5	-0.1188152	0.1341066	-0.886	0.375744	
s1 :mois2	-0.1358884	0.1573944	-0.863	0.388048	
s1 :mois3	0.0130635	0.1588318	0.082	0.934459	
s1 :mois4	0.0341628	0.1579120	0.216	0.828745	
s1 :mois5	0.2259296	0.1731979	1.304	0.192236	
s1 :mois6	0.0503126	0.1628321	0.309	0.757368	
s1 :mois7	-0.0449062	0.1595845	-0.281	0.778437	
s1 :mois8	-0.0200284	0.1571955	-0.127	0.898628	
s1 :mois9	0.0391540	0.1583837	0.247	0.804772	
s1 :mois10	0.0640223	0.1562095	0.410	0.681963	
s1 :mois11	-0.3134325	0.1702994	-1.840	0.065855	.
s1 :mois12	-0.6980638	0.1660044	-4.205	2.73e-05	***
s1 :necranpremsem	-0.0095472	0.0013074	-7.303	4.15e-13	***
s1 :distributeur2	-0.3866166	0.1407077	-2.748	0.006060	**
s1 :distributeur3	-0.0517219	0.1140206	-0.454	0.650156	
s1 :distributeur4	-0.0564566	0.2007990	-0.281	0.778619	
s1 :distributeur5	0.0023515	0.1146832	0.021	0.983643	
s1 :distributeur6	-0.1457599	0.1251875	-1.164	0.244437	
s1 :distributeur7	-0.0452644	0.1233772	-0.367	0.713751	
s1 :distributeur8	0.0545185	0.1114560	0.489	0.624794	

Continue sur la prochaine page...

TABLE B.1 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
s1 :distributeur9	0.0550463	0.1000162	0.550	0.582128	
s1 :star11	0.0781293	0.0744473	1.049	0.294103	
s1 :star21	-0.0386355	0.0721532	-0.535	0.592392	
s1 :director11	0.0462324	0.1055035	0.438	0.661287	
s1 :minlecfm :cote	-0.0012676	0.0006004	-2.111	0.034748	*

TABLE B.2 – BOC - Simple Coding

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	6.3889983	0.2119082	30.150	< 0.0001	***
s1	0.7531055	0.1466144	5.137	3.08e-07	***
minlecfm	-0.0067606	0.0171386	-0.394	0.694208	
cote	0.1235093	0.0353693	3.492	0.000499	***
genre.s2	-0.1535452	0.0800998	-1.917	0.055525	.
genre.s3	-0.2131246	0.0986559	-2.160	0.030982	*
genre.s4	-0.1032694	0.1115604	-0.926	0.354830	
genre.s5	-0.1769346	0.1296559	-1.365	0.172663	
genre.s6	0.0117401	0.1115897	0.105	0.916232	
genre.s7	-0.2692483	0.1022565	-2.633	0.008589	**
visa.s2	-0.0928661	0.0607951	-1.528	0.126939	
visa.s3	-0.3026112	0.1246596	-2.428	0.015375	*
duree.s2	0.1479195	0.0786369	1.881	0.060249	.
duree.s3	0.1406340	0.0809379	1.738	0.082590	.
duree.s4	0.2742243	0.0884301	3.101	0.001981	**
duree.s5	0.1946761	0.1065815	1.827	0.068059	.
mois.s2	-0.1168591	0.1256017	-0.930	0.352386	

Continue sur la prochaine page...

TABLE B.2 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
mois.s3	-0.0611162	0.1267657	-0.482	0.629824	
mois.s4	-0.2725720	0.1263741	-2.157	0.031248	*
mois.s5	-0.6107000	0.1379264	-4.428	1.05e-05	***
mois.s6	-0.3635142	0.1299799	-2.797	0.005260	**
mois.s7	-0.0899233	0.1275165	-0.705	0.480847	
mois.s8	-0.2349548	0.1253422	-1.875	0.061141	.
mois.s9	-0.2624888	0.1263051	-2.078	0.037934	*
mois.s10	-0.2882749	0.1244404	-2.317	0.020724	*
mois.s11	-0.1702823	0.1356031	-1.256	0.209497	
mois.s12	0.1989155	0.1322339	1.504	0.132819	
necranpremsem	0.0278418	0.0010393	26.790	< 0.0001	***
distributeur.s2	0.4065093	0.1118916	3.633	0.000294	***
distributeur.s3	0.0155691	0.0906665	0.172	0.863693	
distributeur.s4	-0.2910960	0.1595806	-1.824	0.068424	.
distributeur.s5	-0.1316553	0.0913395	-1.441	0.149782	
distributeur.s6	0.0477956	0.0995679	0.480	0.631308	
distributeur.s7	0.1383000	0.0981247	1.409	0.159012	
distributeur.s8	-0.3985620	0.0886316	-4.497	7.69e-06	***
distributeur.s9	-0.2036683	0.0799584	-2.547	0.011005	*
star1Oui	-0.0258775	0.0591631	-0.437	0.661919	
star2Oui	0.0089230	0.0573353	0.156	0.876356	
director1Oui	-0.0910207	0.0838321	-1.086	0.277846	
minlecfm :cote	0.0066210	0.0006195	10.688	< 0.0001	***
s1 :minlecfm	0.0079478	0.0015051	5.280	1.29e-07	***
s1 :cote	-0.1505670	0.0443186	-3.397	0.000694	***
s1 :genre.s2	0.0329272	0.1007943	0.327	0.743948	
s1 :genre.s3	0.2150927	0.1241590	1.732	0.083367	.
s1 :genre.s4	0.1786057	0.1404119	1.272	0.203526	
s1 :genre.s5	-0.0422380	0.1631757	-0.259	0.795780	

Continue sur la prochaine page...

TABLE B.2 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
s1 :genre.s6	0.1372052	0.1404415	0.977	0.328717	
s1 :genre.s7	0.1343314	0.1287059	1.044	0.296755	
s1 :visa.s2	0.1354091	0.0765084	1.770	0.076913	.
s1 :visa.s3	0.2449180	0.1568745	1.561	0.118637	
s1 :duree.s2	-0.0888178	0.0989573	-0.898	0.369547	
s1 :duree.s3	-0.0532474	0.1018481	-0.523	0.601167	
s1 :duree.s4	-0.1239068	0.1112907	-1.113	0.265696	
s1 :duree.s5	-0.1191069	0.1341435	-0.888	0.374705	
s1 :mois.s2	-0.1357615	0.1574365	-0.862	0.388618	
s1 :mois.s3	0.0130976	0.1588743	0.082	0.934306	
s1 :mois.s4	0.0343900	0.1579532	0.218	0.827669	
s1 :mois.s5	0.2257241	0.1732449	1.303	0.192761	
s1 :mois.s6	0.0500347	0.1628755	0.307	0.758728	
s1 :mois.s7	-0.0451272	0.1596277	-0.283	0.777434	
s1 :mois.s8	-0.0200467	0.1572383	-0.127	0.898564	
s1 :mois.s9	0.0394005	0.1584268	0.249	0.803620	
s1 :mois.s10	0.0641038	0.1562519	0.410	0.681662	
s1 :mois.s11	-0.3136235	0.1703457	-1.841	0.065764	.
s1 :mois.s12	-0.7003775	0.1660455	-4.218	2.58e-05	***
s1 :necranpremsem	-0.0095428	0.0013077	-7.297	4.31e-13	***
s1 :distributeur.s2	-0.3867962	0.1407463	-2.748	0.006050	**
s1 :distributeur.s3	-0.0516400	0.1140519	-0.453	0.650762	
s1 :distributeur.s4	-0.0567491	0.2008542	-0.283	0.777562	
s1 :distributeur.s5	0.0025780	0.1147143	0.022	0.982073	
s1 :distributeur.s6	-0.1456445	0.1252219	-1.163	0.244940	
s1 :distributeur.s7	-0.0452712	0.1234110	-0.367	0.713785	
s1 :distributeur.s8	0.0546400	0.1114866	0.490	0.624118	
s1 :distributeur.s9	0.0540697	0.1000415	0.540	0.588935	
s1 :star1Oui	0.0784027	0.0744677	1.053	0.292549	

Continue sur la prochaine page...

TABLE B.2 – Continué

	Estimés	Écart-type	Valeur-t	Pr(> t)	
s1 :star2Oui	-0.0394419	0.0721720	-0.546	0.584788	
s1 :director1Oui	0.0463040	0.1055326	0.439	0.660882	
s1 :minlecfm :cote	-0.0012850	0.0006004	-2.140	0.032344	*

TABLE B.3 – BOC - Estimateurs des variances

Effets aléatoires		
Nom du groupe	Variance	Écart-type.
idcinofilm (Intercept)	0.2023	0.4498
Film ID s1	0.3523	0.5935
Film ID ns1	0.5954	0.7716
Cine ID (Intercept)	0.7564	0.8697
Residual	0.1794	0.4235

Nombre d'obs : 113396, groupes : idcinofilm, 58541 ; Film ID, 1065 ; Cine ID, 87

TABLE B.4 – BOPS - Estimateurs des variances

Random effects		
Nom du groupe	Variance	Écart-type.
Film ID (Intercept)	0.3550	0.5958
Cine ID (Intercept)	0.6911	0.8313
Residual	0.3115	0.5581 .

Nombre d'obs : 58541, groupes : Film ID, 1065 ; Cine ID, 87

TABLE B.5 – BOPS - Tests des effets de type III

	DI	Som. Carrés	Carrés Moy.	Valeur F	Dénom	Pr(>F)
minlecfm	1	0.19	0.19	0.00	86.05	0.9794
cote	1	10.32	10.32	0.97	1085.94	0.3250
genre	6	23.75	3.96	3.09	1027.89	0.0053
visa	2	0.49	0.25	0.95	1029.48	0.3885
duree	4	13.50	3.37	1.27	1027.99	0.2796
mois	11	23.22	2.11	4.80	1061.36	0.0000
necranpremssem	1	177.43	177.43	507.04	1030.13	0.0000
distributeur	8	14.41	1.80	5.73	1037.39	0.0000
star1	1	0.30	0.30	1.31	1026.96	0.2527
star2	1	0.16	0.16	0.43	1026.58	0.5124
director1	1	0.13	0.13	0.43	1025.26	0.5106
minlecfm :cote	1	29.19	29.19	93.70	57387.49	0.0000

TABLE B.6 – BOPS - Effets fixes

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	7.4872843	0.2091467	35.80	< 0.0001	***
minlecfn	0.0004238	0.0163719	0.03	0.979406	
cote	-0.0270062	0.0274276	-0.98	0.325023	
genre2	-0.1185723	0.0619410	-1.91	0.055862	.
genre3	0.0043408	0.0762968	0.06	0.954641	
genre4	0.0727354	0.0862865	0.84	0.399452	
genre5	-0.2140925	0.1002911	-2.13	0.033020	*
genre6	0.1461818	0.0863044	1.69	0.090608	.
genre7	-0.1372426	0.0790754	-1.74	0.082936	.
visa2	0.0453155	0.0470194	0.96	0.335392	
visa3	-0.0572229	0.0964136	-0.59	0.552967	
duree2	0.0574937	0.0608155	0.95	0.344688	
duree3	0.0884653	0.0625993	1.41	0.157899	
duree4	0.1480945	0.0683920	2.17	0.030588	*
duree5	0.0783448	0.0824276	0.95	0.342098	
mois2	-0.2637821	0.0962293	-2.74	0.006224	**
mois3	-0.0441309	0.0970638	-0.45	0.649447	
mois4	-0.2544226	0.0961322	-2.65	0.008247	**
mois5	-0.3818932	0.1062126	-3.60	0.000339	***
mois6	-0.3123001	0.0994799	-3.14	0.001740	**
mois7	-0.1289411	0.0983546	-1.31	0.190142	
mois8	-0.2594994	0.0966233	-2.69	0.007351	**
mois9	-0.2278115	0.0973932	-2.34	0.019515	*
mois10	-0.2266718	0.0957516	-2.37	0.018099	*
mois11	-0.4861767	0.1043933	-4.66	3.62e-06	***
mois12	-0.4987766	0.1017249	-4.90	1.09e-06	***
necranpresem	0.0180969	0.0008037	22.52	< 0.0001	***
distributeur2	0.0217211	0.0864318	0.25	0.801625	
distributeur3	-0.0336167	0.0699934	-0.48	0.631127	
distributeur4	-0.3431927	0.1233735	-2.78	0.005506	**
distributeur5	-0.1271316	0.0702842	-1.81	0.070765	.
distributeur6	-0.0944826	0.0768874	-1.23	0.219409	
distributeur7	0.0921197	0.0757559	1.22	0.224259	
distributeur8	-0.3405866	0.0684313	-4.98	7.56e-07	***
distributeur9	-0.1385472	0.0610295	-2.27	0.023392	*
star11	0.0523583	0.0457523	1.14	0.252730	
star21	-0.0290591	0.0443452	-0.66	0.512426	
director11	-0.0426690	0.0648316	-0.66	0.510589	
minlecfn :cote	0.0052283	0.0005401	9.68	< 0.0001	***

Signif. : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TABLE B.7 – BOSS - Estimateurs des variances

Random effects		
Nom du groupe	Variance	Écart-type.
Film ID (Intercept)	0.6016	0.7756
Cine ID (Intercept)	0.8223	0.9068
Residual	0.4351	0.6596

Nombre d'obs : 54859, groupes : Film ID, 1065 ; Cine ID, 87

TABLE B.8 – BOSS - Tests des effets de type III

	Df	Som. Carrés	Carrés Moy.	Valeur F	Dénom	Pr(>F)
minlecfm	1	0.15	0.15	0.10	86.37	0.7507
cote	1	81.12	81.12	13.13	1075.55	0.0003
genre	6	39.38	6.56	1.88	1022.38	0.0814
visa	2	4.53	2.26	3.21	1023.40	0.0409
duree	4	43.28	10.82	2.66	1023.16	0.0315
mois	11	25.96	2.36	5.55	1023.49	0.0000
necranpremssem	1	340.98	340.98	712.29	1023.20	0.0000
distributeur	8	31.01	3.88	8.99	1021.83	0.0000
star1	1	0.13	0.13	0.21	1021.24	0.6472
star2	1	0.00	0.00	0.02	1020.87	0.8758
director1	1	0.45	0.45	1.04	1019.90	0.3088
minlecfm :cote	1	44.76	44.76	102.87	53705.12	0.0000

TABLE B.9 – BOSS - Effets fixes

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	6.7411668	0.2446615	27.553	< 0.0001	***
minlecfm	-0.0056974	0.0178780	-0.319	0.750738	
cote	0.1291877	0.0356588	3.623	0.000305	***
genre2	-0.1542013	0.0806277	-1.913	0.056089	.
genre3	-0.2112470	0.0992863	-2.128	0.033604	*
genre4	-0.1095844	0.1122649	-0.976	0.329234	
genre5	-0.1865672	0.1304849	-1.430	0.153080	
genre6	0.0289138	0.1123091	0.257	0.796884	
genre7	-0.2617786	0.1029069	-2.544	0.011111	*
visa2	-0.0949316	0.0611899	-1.551	0.121109	
visa3	-0.3024883	0.1254512	-2.411	0.016075	*
duree2	0.1512008	0.0791430	1.910	0.056351	.
duree3	0.1349859	0.0814651	1.657	0.097830	.
duree4	0.2865541	0.0890005	3.220	0.001324	**
duree5	0.1877825	0.1072610	1.751	0.080296	.
mois2	-0.1248984	0.1263927	-0.988	0.323299	
mois3	-0.0611151	0.1275618	-0.479	0.631968	
mois4	-0.2767075	0.1271672	-2.176	0.029788	*
mois5	-0.6276392	0.1388013	-4.522	6.85e-06	***
mois6	-0.3761562	0.1307939	-2.876	0.004112	**
mois7	-0.1014097	0.1295861	-0.783	0.434063	
mois8	-0.2468725	0.1266523	-1.949	0.051543	.
mois9	-0.2654422	0.1276825	-2.079	0.037873	*
mois10	-0.2981435	0.1252353	-2.381	0.017463	*
mois11	-0.1860720	0.1364484	-1.364	0.172969	
mois12	0.2400175	0.1331259	1.803	0.071691	.
necranpresem	0.0279178	0.0010461	26.689	< 0.0001	***
distributeur2	0.4003673	0.1125896	3.556	0.000394	***
distributeur3	0.0181219	0.0912334	0.199	0.842590	
distributeur4	-0.3013485	0.1605602	-1.877	0.060823	.
distributeur5	-0.1310941	0.0919190	-1.426	0.154119	
distributeur6	0.0525170	0.1002054	0.524	0.600327	
distributeur7	0.1575074	0.0987477	1.595	0.111010	
distributeur8	-0.4119031	0.0891782	-4.619	4.35e-06	***
distributeur9	-0.2063123	0.0804898	-2.563	0.010512	*
star11	-0.0272510	0.0595353	-0.458	0.647244	
star21	0.0090199	0.0576944	0.156	0.875797	
director11	-0.0858928	0.0843592	-1.018	0.308835	
minlecfm :cote	0.0068094	0.0006714	10.143	< 0.0001	***

Signif. : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TABLE B.10 – BOT - Estimateurs des variances

Random effects		
Nom du groupe	Variance	Écart-type.
Film ID (Intercept)	0.3345	0.5783
Cine ID (Intercept)	0.7322	0.8557
Residual	0.2601	0.5100

Nombre d'obs : 54848, groupes : Film ID, 1065 ; Cine ID, 87

TABLE B.11 – BOT - Tests des effets de type III

	Df	Som. Carrés	Carrés Moy.	Valeur F	Dénom	Pr(>F)
minlecfm	1	0.12	0.12	0.01	85.90	0.9075
cote	1	47.31	47.31	8.39	1079.92	0.0038
genre	6	29.80	4.97	2.19	1022.77	0.0420
visa	2	1.58	0.79	2.53	1023.86	0.0805
duree	4	28.76	7.19	3.03	1023.60	0.0169
mois	11	17.73	1.61	4.82	1023.95	0.0000
necranpremssem	1	251.90	251.90	878.27	1023.64	0.0000
distributeur	8	20.57	2.57	9.92	1022.18	0.0000
star1	1	0.01	0.01	0.01	1021.54	0.9272
star2	1	0.00	0.00	0.00	1021.14	0.9579
director1	1	0.15	0.15	0.56	1020.10	0.4556
minlecfm :cote	1	31.35	31.35	120.54	53705.83	0.0000

TABLE B.12 – BOT - Effets fixes

	Estimés	Écart-type	Valeur-t	Pr(> t)	
(Intercept)	7.8567187	0.2113514	37.17	< 0.0001	***
minlecfn	-0.0019619	0.0168441	-0.12	0.907547	
cote	0.0771441	0.0266272	2.90	0.003841	**
genre2	-0.1278818	0.0601536	-2.13	0.033748	*
genre3	-0.1010685	0.0740717	-1.36	0.172720	
genre4	-0.0274467	0.0837529	-0.33	0.743198	
genre5	-0.1787550	0.0973476	-1.84	0.066610	.
genre6	0.0859196	0.0837873	1.03	0.305393	
genre7	-0.1635182	0.0767710	-2.13	0.033414	*
visa2	-0.0484731	0.0456502	-1.06	0.288560	
visa3	-0.2081797	0.0935927	-2.22	0.026345	*
duree2	0.1186708	0.0590453	2.01	0.044712	*
duree3	0.1167731	0.0607779	1.92	0.054970	.
duree4	0.2293580	0.0663986	3.45	0.000574	***
duree5	0.1452286	0.0800207	1.81	0.069834	.
mois2	-0.1565413	0.0942948	-1.66	0.097195	.
mois3	-0.0699394	0.0951657	-0.73	0.462555	
mois4	-0.2686362	0.0948716	-2.83	0.004723	**
mois5	-0.4989041	0.1035531	-4.82	1.67e-06	***
mois6	-0.3408012	0.0975760	-3.49	0.000499	***
mois7	-0.0903729	0.0966751	-0.93	0.350106	
mois8	-0.2328889	0.0944887	-2.46	0.013875	*
mois9	-0.2654505	0.0952571	-2.79	0.005424	**
mois10	-0.2813771	0.0934334	-3.01	0.002663	**
mois11	-0.2590232	0.1017933	-2.54	0.011087	*
mois12	0.0132163	0.0993254	0.13	0.894172	
necranpresem	0.0231276	0.0007804	29.64	< 0.0001	***
distributeur2	0.2293752	0.0839945	2.73	0.006426	**
distributeur3	-0.0268344	0.0680607	-0.39	0.693463	
distributeur4	-0.2819018	0.1197762	-2.35	0.018783	*
distributeur5	-0.1333697	0.0685741	-1.94	0.052061	.
distributeur6	-0.0247560	0.0747579	-0.33	0.740600	
distributeur7	0.1044494	0.0736693	1.42	0.156550	
distributeur8	-0.3880032	0.0665270	-5.83	7.33e-09	***
distributeur9	-0.1655006	0.0600543	-2.76	0.005957	**
star11	-0.0040602	0.0444143	-0.09	0.927180	
star21	-0.0022718	0.0430407	-0.05	0.957916	
director11	-0.0469720	0.0629319	-0.75	0.455601	
minlecfn :cote	0.0056987	0.0005191	10.98	< 0.0001	***

Signif. : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Bibliographie

- [Ainslie et al., 2005] Ainslie, A., Drèze, X., and Zufryden, F. (2005). Modeling movie life cycles and market share. *Marketing Science*, 24(3) :508–517.
- [Alba and Hutchinson, 1987] Alba, J. W. and Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of consumer research*, pages 411–454.
- [Albert, 1998] Albert, S. (1998). Movie stars and the distribution of financially successful films in the motion picture industry. *Journal of Cultural Economics*, 22(4) :249–270.
- [Angrist and Krueger, 1999] Angrist, J. D. and Krueger, A. B. (1999). Empirical strategies in labor economics. *Handbook of labor economics*, 3 :1277–1366.
- [Austin, 1984] Austin, B. A. (1984). Portrait of an art film audience. *Journal of Communication*.
- [Basuroy et al., 2003] Basuroy, S., Chatterjee, S., and Ravid, S. A. (2003). How critical are critical reviews? the box office effects of film critics, star power, and budgets. *Journal of Marketing*, pages 103–117.
- [Becker, 1965] Becker, G. S. (1965). A theory of the allocation of time. *The economic journal*, 75(299) :493–517.

- [Belmati, 2011] Belmati, S. (2011). La valeur de la marque au sein du réseau de distribution : le cas du cinéma québécois. *Mémoire de maîtrise - HEC Montréal*.
- [Bettman, 1979] Bettman, J. R. (1979). Memory factors in consumer choice : A review. *The Journal of Marketing*, pages 37–53.
- [Bonnet, 2009] Bonnet, A. (2009). Exploration d’outils statistiques de prédiction des recettes cinématographiques. *Mémoire de maîtrise - HEC Montréal*.
- [Boor, 1992] Boor, M. (1992). Relationships among ratings of motion pictures by viewers and six professional movie critics. *Psychological Reports*, 70(3c) :1011–1021.
- [Bozdogan, 1987] Bozdogan, H. (1987). Model selection and akaike’s information criterion (aic) : The general theory and its analytical extensions. *Psychometrika*, 52(3) :345–370.
- [Brewer et al., 2009] Brewer, S. M., Kelley, J. M., and Jozefowicz, J. J. (2009). A blueprint for success in the us film industry. *Applied Economics*, 41(5) :589–606.
- [Capon and Burke, 1977] Capon, N. and Burke, M. (1977). Information seeking behavior in consumer durable purchases. *Contemporary marketing thought*, pages 110–115.
- [Cartier and Liarte, 2012] Cartier, M. and Liarte, S. (2012). Impact of underlying demand and agglomeration of supply on seasonality : the case of the hollywood film industry.
- [Caves, 2000] Caves, R. E. (2000). *Creative industries : Contracts between art and commerce*. Number 20. Harvard University Press.
- [Caves, 2003] Caves, R. E. (2003). Contracts between art and commerce. *The Journal of Economic Perspectives*, 17(2) :73–84.
- [Chakravarty et al., 2010] Chakravarty, A., Liu, Y., and Mazumdar, T. (2010). The differential effects of online word of mouth and critics

- reviews on prerelease movie evaluation. *Journal of Interactive Marketing*, 24(3) :185–197.
- [Chang and Ki, 2005] Chang, B.-H. and Ki, E.-J. (2005). Devising a Practical Model for Predicting Theatrical Movie Success : Focusing on the Experience Good Property. *Journal of Media Economics*, 18(4) :247–269.
- [Chen et al., 2011] Chen, Y., Liu, Y., and Zhang, J. (2011). When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do? The Case of Media. 75(September) :116–134.
- [Cox, 1967] Cox, D. F. (1967). *Risk taking and information handling in consumer behavior*. Division of Research, Graduate School of Business Administration, Harvard University Boston.
- [Dahl, 2013] Dahl, D. B. (2013). *xtable : Export tables to LaTeX or HTML*. R package version 1.7-1.
- [d’Astous and Colbert, 2002] d’Astous, A. and Colbert, F. (2002). Moviegoers’ consultation of critical reviews : psychological antecedents and consequences. *International Journal of Arts Management*, pages 24–35.
- [D’Astous and Touil, 1999] D’Astous, A. and Touil, N. (1999). Consumer evaluations of movies on the basis of critics’ judgments. *Psychology and Marketing*, 16(8) :677–694.
- [De Vany and Walls, 1999] De Vany, A. and Walls, W. D. (1999). Uncertainty in the movie industry : Does star power reduce the terror of the box office? *Journal of Cultural Economics*, 23(4) :285–318.
- [De Vany and Walls, 2002] De Vany, A. and Walls, W. D. (2002). Does hollywood make too many r-rated movies? risk, stochastic dominance, and the illusion of expectation*. *The Journal of Business*, 75(3) :425–451.

- [Debenedetti, 2006] Debenedetti, S. (2006). L'impact de la critique de presse sur la consommation culturelle : un essai de synthèse dans le champ cinématographique. *Recherche et applications en marketing*, 21(2) :43–59.
- [Desai and Basuroy, 2005a] Desai, K. K. and Basuroy, S. (2005a). Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry : The case of motion pictures. *Psychology and Marketing*, 22(3) :203–223.
- [Desai and Basuroy, 2005b] Desai, K. K. and Basuroy, S. (2005b). Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry : The case of motion pictures. *Psychology and Marketing*, 22(3) :203–223.
- [Dhar et al., 2012] Dhar, T., Sun, G., and Weinberg, C. B. (2012). The long-term box office performance of sequel movies. *Marketing Letters*, 23(1) :13–29.
- [Diggle et al., 1994] Diggle, P., Liang, K.-Y., and Zeger, S. L. (1994). Longitudinal data analysis.
- [Einav, 2007] Einav, L. (2007). Seasonality in the U.S. motion picture industry. *The RAND Journal of Economics*, 38(1) :127–145.
- [Elberse, 2007] Elberse, A. (2007). The power of stars : Do star actors drive the success of movies? *Journal of Marketing*, pages 102–120.
- [Eliashberg and Sawhney, 1994] Eliashberg, J. and Sawhney, M. S. (1994). Modeling goes to hollywood : Predicting individual differences in movie enjoyment. *Management Science*, 40(9) :1151–1173.
- [Eliashberg and Shugan, 1997] Eliashberg, J. and Shugan, S. M. (1997). Film Critics : Influencers or Predictors? *Journal of Marketing*, 61(2) :68.

- [Elliott and Simmons, 2008] Elliott, C. and Simmons, R. (2008). Determinants of uk box office success : the impact of quality signals. *Review of Industrial Organization*, 33(2) :93–111.
- [Faber and O’Guinn, 1984] Faber, R. J. and O’Guinn, T. C. (1984). Effect of media advertising and other sources on movie selection. *Journalism Quarterly*, 61(2) :371–77.
- [Fellows and using the JMapView library by Jan Peter Stotz, 2013] Fellows, I. and using the JMapView library by Jan Peter Stotz (2013). *OpenStreetMap : Access to open street map raster images*. R package version 0.3.1.
- [Festinger, 1954] Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2) :117–140.
- [Fisher, 1919] Fisher, R. A. (1919). Xv the correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52(02) :399–433.
- [Gazley et al., 2011] Gazley, A., Clark, G., and Sinha, A. (2011). Understanding preferences for motion pictures. *Journal of Business Research*, 64(8) :854–861.
- [Gemser et al., 2006] Gemser, G., Oostrum, M., and Leenders, M. a. a. M. (2006). The impact of film reviews on the box office performance of art house versus mainstream motion pictures. *Journal of Cultural Economics*, 31(1) :43–63.
- [Hansen, 1972] Hansen, F. (1972). *Consumer choice behavior : A cognitive theory*, volume 10. Free Press New York.
- [Hartley and Rao, 1967] Hartley, H. O. and Rao, J. (1967). Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54(1-2) :93–108.

- [Hatchuel, 1995] Hatchuel, A. (1995). Les marchés à prescripteurs. crises de l'échange et genèse sociale. *L'inscription sociale du marché, Paris, L'Harmattan*, pages 205–225.
- [Henderson, 1953] Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics*, 9(2) :226–252.
- [Hennig-Thurau et al., 2007] Hennig-Thurau, T., Henning, V., and Sattler, H. (2007). Consumer file sharing of motion pictures. *Journal of Marketing*, pages 1–18.
- [Hennig-Thurau et al., 2010] Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., and Skiera, B. (2010). The impact of new media on customer relationships. *Journal of Service Research*, 13(3) :311–330.
- [Hirschman and Holbrook, 1982] Hirschman, E. C. and Holbrook, M. B. (1982). Hedonic consumption : emerging concepts, methods and propositions. *The Journal of Marketing*, pages 92–101.
- [Hirschman and Pieros, 1985] Hirschman, E. C. and Pieros, A. (1985). Relationships among indicators of success in Broadway plays and motion pictures. *Journal of Cultural Economics*, 9(1) :35–63.
- [Højsgaard, 2013] Højsgaard, U. H. S. (2013). *pbkrtest : Parametric bootstrap and Kenward Roger based methods for mixed model comparison*. R package version 0.3-7.
- [Holbrook and Hirschman, 1982] Holbrook, M. B. and Hirschman, E. C. (1982). The experiential aspects of consumption : consumer fantasies, feelings, and fun. *Journal of consumer research*, pages 132–140.
- [Hoyer, 2001] Hoyer, W. M. (2001). D. 2001, consumer behavior.
- [Izod, 1988] Izod, J. (1988). *Hollywood and the box office, 1895-1986*. Macmillan London.

- [Jacoby et al., 1977] Jacoby, J., Chestnut, R. W., and Silberman, W. (1977). Consumer use and comprehension of nutrition information. *Journal of Consumer Research*, pages 119–128.
- [Katz and Paul, 1955] Katz, E. and Paul, F. (1955). Lazarsfeld (1955), personal influence. *The Part Played by People in the Flow of Mass Communication*. New York.
- [Kuznetsova et al., 2013] Kuznetsova, A., Bruun Brockhoff, P., and Haubo Bojesen Christensen, R. (2013). *lmerTest : Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R package version 2.0-0.
- [Leigh and Rethans, 1984] Leigh, T. W. and Rethans, A. J. (1984). A script-theoretic analysis of industrial purchasing behavior. *The Journal of Marketing*, pages 22–32.
- [Levene, 1992] Levene, C. (1992). Marketing art-films to college students. *Working Paper*.
- [Levin et al., 1997] Levin, A. M., Levin, I. P., and Heath, C. E. (1997). Movie stars and authors as brand names : Measuring brand equity in experiential products. *Advances in consumer research*, 24 :175–181.
- [Litman and Kohl, 1989] Litman, B. and Kohl, L. (1989). Predicting financial success of motion pictures : The 80s experience. *Journal of Media Economics*, 2 :35–50.
- [Litman, 1983] Litman, B. R. (1983). Predicting success of theatrical movies : An empirical study. *The Journal of Popular Culture*, 16(4) :159–175.
- [Litman, 1998] Litman, B. R. (1998). *The motion picture mega-industry*. Allyn and Bacon Needham Heights, MA.
- [Liu, 2006] Liu, Y. (2006). Word of mouth for movies : Its dynamics and impact on box office revenue. *Journal of marketing*, pages 74–89.

- [Locander and Hermann, 1979] Locander, W. B. and Hermann, P. W. (1979). The effect of self-confidence and anxiety on information seeking in consumer risk reduction. *Journal of Marketing Research*, pages 268–274.
- [Lutz and Reilly, 1973] Lutz, R. J. and Reilly, P. J. (1973). An exploration of the effects of perceived social and performance risk on consumer information acquisition. *Advances in consumer research*, 1(4) :393–405.
- [Lynch Jr and Srull, 1982] Lynch Jr, J. G. and Srull, T. K. (1982). Memory and attentional factors in consumer choice : Concepts and research methods. *Journal of Consumer Research*, pages 18–37.
- [Murray, 1991] Murray, K. B. (1991). A test of services marketing theory : consumer information acquisition activities. *The Journal of Marketing*, pages 10–25.
- [Nelson, 1970] Nelson, P. (1970). Information and consumer behavior. *The Journal of Political Economy*, 78(2) :311–329.
- [R Core Team, 2013] R Core Team (2013). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [Ravid, 2003] Ravid, S. A. (2003). Are they all crazy or just risk averse ? some movie puzzles and possible solutions. *Contributions to economic analysis*, 260 :33–47.
- [Ravid and Basuroy, 2004] Ravid, S. A. and Basuroy, S. (2004). Managerial objectives, the r-rating puzzle, and the production of violent films*. *The Journal of Business*, 77(S2) :S155–S192.
- [Ravis, 1999] Ravis, S. A. (1999). Information, blockbusters and stars a study of the film industry : Fin-98-056. *Working Paper serie 1998*.
- [Reddy et al., 1998] Reddy, S. K., Swaminathan, V., and Motley, C. M. (1998). Exploring the determinants of Broadway show success. *Journal of Marketing Research*, pages 370–383.

- [Reinstein and Snyder, 2000] Reinstein, D. A. and Snyder, C. M. (2000). The influence of expert reviews on consumer demand for experience goods : A case study of movie critics. *Journal of Industrial Economics* (forthcoming).
- [Satterthwaite, 1946] Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin*, 2(6) :110–114.
- [Shrum, 1991] Shrum, W. (1991). Critics and publics : Cultural mediation in highbrow and popular performing arts. *American Journal of Sociology*, pages 347–375.
- [Simonoff and Sparrow, 2000] Simonoff, J. S. and Sparrow, I. R. (2000). Predicting movie grosses : Winners and losers, blockbusters and sleepers. *Chance*, 13(3) :15–24.
- [Sochay, 1994] Sochay, S. (1994). Predicting the performance of motion pictures. *Journal of Media Economics*, 7(4) :1–20.
- [Statistique Canada, 2006] Statistique Canada (2006). Classification géographique type (cgt) - num. 12-572. *Cartes de référence*, Volume II.
- [Statistiques Canada, 2012] Statistiques Canada (2012). Cinémas : tableaux de données 87f0009xwf. *Publications*.
- [Stigler, 1961] Stigler, G. J. (1961). The economics of information. *The journal of political economy*, 69(3) :213–225.
- [Terry and Butler,] Terry, N. and Butler, M. De armond, d. 2005. the determinants of domestic box office performance in the motion picture industry. *Southwestern Economic Review*, 32 :137–148.
- [Terry et al., 2011] Terry, N., Butler, M., DeArmond, D., et al. (2011). The determinants of domestic box office performance in the motion picture industry. *Southwestern Economic Review*, 32 :137–148.

- [Terry et al., 2010] Terry, N., Cooley, J. W., and Zachary, M. (2010). The determinants of foreign box office revenue for english language movies. *Journal of International Business and Cultural Studies*, 2(1) :12.
- [Van Raaij, 1977] Van Raaij, F. W. (1977). Consumer information processing for different information structures and formats. *Advances in consumer research*, 4 :176–184.
- [Varian and Shapiro, 1999] Varian, H. R. and Shapiro, C. (1999). Information rules : a strategic guide to the network economy. *Harvard Business School Press, Cambridge*.
- [Vézina, 1997] Vézina, R. (1997). One of the best movie of the year... : an experiment on the influence of critics recommendations on consumers. *Actes de la 26e Conférence EMAC, 3, 20-23 mai*, pages 1300–1312.
- [Wallace et al., 1993] Wallace, W. T., Seigerman, A., and Holbrook, M. B. (1993). The role of actors and actresses in the success of films : How much is a movie star worth? *Journal of cultural economics*, 17(1) :1–27.
- [Walls, 2005] Walls, W. D. (2005). Modeling movie success when 'nobody knows anything' : Conditional stable-distribution analysis of film returns. *Journal of Cultural Economics*, 29(3) :177–190.
- [Wickham, 2009] Wickham, H. (2009). *ggplot2 : elegant graphics for data analysis*. Springer New York.
- [Wyatt, 1991] Wyatt, J. (1991). High concept, product differentiation, and the contemporary us film industry. *Current research in film : Audiences, economics and law*, 5 :86–105.
- [Wyatt and Badger, 1984] Wyatt, R. O. and Badger, D. P. (1984). How Reviews Affect Interest in and Evaluation of Films. *Journalism & Mass Communication Quarterly*, 61(4) :874–878.
- [Zaichkowsky, 1985] Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of consumer research*, pages 341–352.

- [Zufryden, 1996] Zufryden, F. S. (1996). Linking advertising to box office performance of new film releases-a marketing planning model. *Journal of Advertising Research*, 36(4) :29-41.