CATÓLICA LISBON
SCHOOL OF BUSINESS & ECONOMICS

UNIVERSIDADE CATÓLICA PORTUGUESA

The Film Industry: sequel production and release lags

David Polido

September 2015

Advisor: Professor Pedro Raposo

Dissertation submitted in partial fulfillment of the requirements for the degree of MSc in Economics at Católica-Lisbon School of Business and Economics

Abstract[1]

The theatrical film industry can structurally be divided in three stages: production, distribution and exhibition. In each stage companies face strategic decisions that help determine the success of a given film. This thesis aims to optimize these strategic decisions on two issues: the sequel-making process, concerning producers, and the timing of release in international markets, decided by distributors.

Following from these issues, two research questions are established. First, if it's (still) worth it for a producer to make a sequel film. Second, if the time lag between releases in North America and Portugal affects a film's revenues in the latter market. To pursue these objectives, the approach established by Basuroy & Chatterjee (2008) was followed: a Generalized Estimating Equations (GEE) method, under a Gaussian family, with an identity link and exchangeable correlation structure.

For the producer's problem, data was obtained for 208 films released in 2011 in North America. The results show that sequel films consistently outperform non-sequels at the box office. Furthermore, this effect is moderated by longer time gaps between parent and target and producers should expect faster decaying weekly revenues from sequels.

For the distributor's problem, Portuguese data consists of 139 of those films covered in the producer's problem, which, due to release lags, were released in Portugal between 2009 and 2013. The results show that longer release lags have a significant and negative effect on revenues and that this effect is largely concentrated in the first few weeks of a films' release.

[1] My gratitude goes to Professor Pedro Raposo for guidance, encouragement and understanding, as well as his sincere interest in the subject. To my family, thank you for providing this and all opportunities and having the patience to stand by me. I am forever grateful to Rita Gomes for the love, friendship and adventures.

Contents

1. Introduction

The film industry is frequently associated with stars, explosions and popcorn. Among many other things, the industry is perhaps mostly associated with money. It's easy to understand why: in 2013 alone the 134.588 cinema screens worldwide reported film revenues in excess of $35.9 billion (MPAA 2013). Despite this penchant for large numbers, not every film results profitable. At least six out of any given ten major films is considered unprofitable with one of the remaining four only breaking even (Vogel 2011). Navigating through tough competition and facing a risky nature, managers in the industry must make sound decisions to avoid losses. The purpose of this thesis is to help managers optimize their strategic decisions by answering two specific questions. Is it (still) worth it for a producer to make a sequel film? And, does the time interval between a film's releases in North America and Portugal impact the revenues obtained in the latter market? Answering these two questions faced by producers and distributors, respectively, can provide insights that positively influence their decision-making and improve the likelihood of profits.

Ever since *The Fall of a Nation* (1916) was made, sequel films have been part of the film industry's landscape. Sequels have been observed to be a less risky avenue for which managers in the industry can opt as they try to manage their catalog (Ravid & Basuroy 2004; Ravid 1999; Hennig-Thurau et al. 2009). In an industry facing large investment needs on the short run – production and marketing – while expecting the potential returns only over the long run (Vogel 2011), risk reduction can be very valuable. However, the adoption of numerous sequel projects might be noticeably limited. As Ravid and Basuroy (2004) point out, producing a sequel is dependent on multiple factors such as owning the rights to a successful film, the return of key personnel (cast and even part of the crew) or just having the ability to do so story-wise. Despite these limitations the popularity of sequels at the box office[2] invites close inspection. From 2011 to 2013, the majority of the top ten highest-grossing films worldwide were sequels: six in each of the 2012 and 2013 top ten[3], with the nine of 2011 almost keeping non-sequels out entirely. Interestingly, this dominance was not so pronounced twenty or even ten years ago: the 1991-1993 period saw one, three and zero sequels in the top ten, respectively; and from 2001 to 2003 the same ranking showed three, five and six sequel films, respectively.[4] In line with these are indications that until recently – specifically, 2008 – the number of sequels produced had not substantially increased (Dhar et al. 2011). However, there are some signs that managers might be attempting to change this situation: 105 sequel films were reported as being in production in 2012, with 2014 seeing this number ascend to 129 (Brew 2012; Brew 2014). Taken together, these signals suggest a greater awareness of the potential of producing sequels and a

[2] Industry term for film exhibition in cinemas.

[3] Of the remaining combined eight films, only three are unquestionably not part of a series. For a brief explanation of some of the industry's concepts and terminology, see Appendix 1.

[4] All rankings come from BoxOfficeMojo.com, an IMDb.com company.

consequential propensity to increase the number of these types of production. While the literature shows that sequel films tend to outperform contemporaneous non-sequels, the majority of these studies do not cover recent years. As a somewhat extreme example, one of the most cited papers focused on this issue, Basuroy and Chatterjee (2008), covers a sample of films from 1991 to 1993.[5] Even more recent studies such as Hennig-Thurau, Houston and Heitjans (2009) and Dhar, Sun and Weinberg (2011) use samples ending in 2006 and 2008, respectively. As consequence of this gap, managers are making a strategic decision to invest in a sequel based on insights that reflect a different time period. In a field where consumers' tastes and responses are seen to shift often (Vogel 2011), this gap can be very consequential. Evidently, academic analysis can never be truly contemporaneous with decision-making, but reducing the gap between the two as well as discerning possible changes in the industry remains worthwhile.

To address this issue, this thesis' main objective concerning producers is to learn if sequel films outperform contemporaneous non-sequels for a more recent period, using a 2011 sample of the North American market. This producer's problem is lent further strength by using Basuroy and Chatterjee's (2008) approach to characterize sequels in four additional dimensions specific to these releases. The revenues obtained by the parent film[6], the time gap between said parent and the target sequel[7], the number of intervening sequels[8] and the behavior of sequel films during release are all issues that can either help shape the characteristics of a sequel release towards success, or serve as clues to determine if the existing conditions are unfavorable for such a release.

International distribution has become a subject of increased interest lately. While North America[9] remains the biggest market by revenues – as well as the one with the most global presence – its position of dominance has been slowly eroding throughout recent years. If in 2000 this market represented over 48% of the global box office, after a steady decline this percentage fell in 2013, for the first time, below 31% (MPAA 2005; 2013). Although this market grew considerably during this period – increasing its size by 45% – the growth of all other markets combined was even more remarkable: from $8.4 billion to $25 billion, almost tripling their combined size. Both the growth of international markets and the coinciding decrease in weight of the American market demands attention from managers. American managers in particular, as the leading exporters in the industry, seem to be aware of this evolution and willing to explore it (Kolesnikov-Jessop 2011). Recognition of strategies less conducive to global success, such as films targeted at narrow audiences and small theatrical releases with almost immediate release on

[5] This specific sample was created by Ravid (1999) and has been used in multiple studies over the years (Basuroy et al. 2003; Ravid & Basuroy 2004; Basuroy et al. 2006).

[6] First in a series.

[7] The film of a series present in the sample.

[8] Films between the parent and the target sequel.

[9] United States and Canada combined market. Another synonym for this joint market are U.S./Canada.

ancillary channels (such as video-on-demand), point to an increased interest in the international markets' performance. With estimates indicating that about 70% of the total revenues obtained by North American films originate from foreign markets, this increased attention seems justified (Brook 2013).

One of the most important aspects of an international release, shown to influence revenues obtained by films, is timing of release (Einav 2007; Elberse & Eliashberg 2003). For this thesis, the focus lies on examining the time gap (lag) between releases in two different countries. To accomplish this goal the model of the producer's problem is extended and applied to this distributor's problem for the Portuguese market. Studying the exposure between North American and Portuguese markets, through understanding how these lags might affect the box office in the latter market, can provide distributors with insights on large-small market dynamics as well as help optimize the sequence/timing of releases for these experiential products.

This thesis is divided in seven chapters. After this introduction, chapter 2 contextualizes the discussion by characterizing the industry structure, with the theoretical basis being covered in chapter 3. Beginning with a literature review, this chapter extends into hypothesis construction and formulation. Following, chapters 4 and 5 describe and explore the data and the methodology used, respectively, with chapter 6 presenting a discussion of the empirical results obtained. In a pursuit for clarity, chapters 3, 4 and 6 divide their contents between the two main research questions: the producer's problem regarding sequels, and the distributor's problem of release lags. The thesis ends with chapter 7, which offers the concluding remarks by bridging the empirical findings with managerial decision-making.

2. Industry Structure

The theatrical film industry[10] structure can be divided into three key stages: production, distribution and exhibition.[11] Production typically starts with the acquisition of rights to a story idea or script, which is then developed until the final draft is ready to be presented to financers and investors.[12] Throughout, the role of the producer attached to the project includes finding the desirable personnel (director, cast and crew), scouting locations, forecasting costs and projecting sets, among other tasks. Depending on the project, the director and/or specific cast members can fulfill multiple roles and be involved with previous steps. For example, Tom Cruise producing and starring in the *Mission: Impossible* series. After securing financing the film goes through planning, filming and post-production until one copy (or print) is finished. Numerous strategic decisions about the film are made in this first stage, as production defines the majority of the film characteristics which can later impact the decision-making of a filmgoer and, consequently, the film's revenues. As one example, the budget[13] – generally considered a significant determinant of revenues – is defined at this stage.

After the first single copy of the film is produced, the distribution stage begins. Here, the film progresses from single to multiple copies, which are then delivered to theaters. As in production, and since the bulk of the marketing effort is made at this stage, this is a decisive period for films. Strategic decisions concerning release dates, advertising and release strategies, as well as international distribution – along with the resulting negotiations with partners – are all made at this juncture. It's worth noting the close ties between production and distribution with regards to strategic decisions. It's not coincidental that the biggest films – both in investment and performance – are exceedingly more common when both stages are covered by the same companies, as most North American studios[14] do. The increased control over the whole range of production-distribution decisions are viewed by these companies as risk-reducers. From the three main types of risks faced in the whole process – financing, completion and performance – two become drastically reduced and the third becomes more manageable (Vogel 2011).

The final stage is exhibition, where the product (film) finally reaches the consumer (filmgoer). For a single film, aside from being part of the negotiation for the number of screens/theaters, the extent of

---

[10] Films produced to be released in cinemas. For a discussion of ancillary markets/channels see Appendix 2.

[11] Eliashberg, Elberse and Leenders (2006) and Vogel (2011) serve as sources for the industry structure.

[12] This description is intended as a guide to the most usual sequence of events. Financing only an idea, certain personnel being attached to a project from the start or even not securing the full financing before the production start are all alternative possibilities.

[13] Industry term synonymous with production costs. These do not include advertising and other marketing costs which are bundled with printing costs under P&A (Prints & Advertising) costs.

[14] The type of company, operating in both production and distribution stages, most associated with Hollywood. While typically producing films in-house (with smaller co-producers), they can also function solely as financers or/and distributors.

the critical decision making for an exhibitor is in large part limited to scheduling the film's showings (Eliashberg et al. 2008). As a consequence, fewer variables from this stage tend to be included in the analysis of the determinants of film revenues, when compared to production and distribution. Instead, exhibition lends itself to other types of studies, especially those concerned with their key decision of deciding theater locations.[15]

In light of this thesis distributor's problem, it's worth noting that international releases result in additional distributors and exhibitors entering the value chain. The additional distributors can be local subsidiaries/affiliates of the original distributor or new companies negotiating the distribution rights. Deals can be made for single films or the whole catalog for a specific period of time. In terms of control, an exporting distributor[16] from a strong exporting country (e.g. U.S.) usually holds more sway than the importing distributor.[17] However, since both companies are interested in maximizing revenues, the benefits from optimizing decisions concerning the release dates are mutual. Exhibitors are almost always local companies that enter in agreements with the importing distributors.

Different markets have differences in structure that are a reflection of their specific contexts.[18] For this thesis it's worth observing that while the North American market shows an oligopolistic structure in both production/distribution[19] and exhibition, the importing nature of the Portuguese market results in production being comprised by a greater number of smaller companies. Portuguese distribution and exhibition stages follow an oligopoly structure, as North America, but production is closer to perfect competition. Tied with this arrangement is the fact that Portuguese production of films, unlike that of North America, is significantly dependent on grants and governmental support, this being specifically part of the mission of *Instituto do Cinema e Audiovisuais* (ICA).

[15] Other studies can focus, for example, on market structures, prices, cannibalization, or entry.

[16] The holder of international distribution rights. These companies are usually from the producing country and function either as a primary distributor or a co-distributor responsible for international distribution.

[17] The buyer of the local distribution rights.

[18] For brief overview of the worldwide, North American and Portuguese market contexts see Appendices 3, 4 and 5, respectively.

[19] Despite the largest North American companies, the studios, operating in both production and distribution, tradition dictates that the scrutiny is to lay mostly on distribution. Partly, this is tied with studios' practice of establishing each film they produce as a single, independent corporation, nullifying any statistic on a producing corporation basis. This setup allows the management of expenditures in order to better control royalties and profit-sharing agreements (USA Today 2007; Thompson 2011; Epstein 2012). Most of the time this results in reported losses, even in very strong performing, moderately budgeted films. This is a wildly reported phenomenon sometimes known as "Hollywood Accounting".

3. Theory

3.1. Literature review

The film industry has over a hundred years of history and, for most of it, has been a fully-fledged business with considerable economic influence. However, throughout its greatest expansion phase, a period that is considered its golden age, the industry operated with little to no research (Jowett 1985). As a consequence, managers adopted the widespread use of "common sense" beliefs, resulting in an industry that is in many ways reliant on tradition, rules of thumb and conventional wisdom (Eliashberg et al. 2006). Despite this somewhat late development of the literature, the body of literature has considerably grown in recent times: if from 1977 to 1995 the two most prolific years saw the publication of four papers, between 1996 and 2006 the number of yearly papers published never dropped below five (Hadida 2009). Consequently, only the specific issues addressed in this thesis are surveyed here.

The literature is commonly divided in two approaches (Eliashberg et al. 2006; Chang & Ki 2005). The first, the older psychological approach, tends to look towards understanding the motivations and choices of consumers, often using surveys of individuals. Drawing heavily from the psychology field it covers topics such as differences in movie enjoyment (Möller & Karppinen 1983; D'Astous & Touil 1999), effect of prior information (Burzynski & Bayer 1977), segmentation of audiences (Cuadrado & Frasquet 1999) and consumer's evaluation and choice of films (Austin 1986; Gazley et al. 2011). The second, the economic approach, looks to explain the financial performance of films using aggregate movie-level variables such as budgets, release dates and ratings (Litman 1983; Prag & Casavant 1994; Sochay 1994). It's reasonable to say that while the psychological approach is focused on moviegoers and their characteristics, the economical approach looks at films and how their features can impact financial results. Partly, the delay in this type of research might be due to the confidentiality to which some data, such as production cost, are bound (Litman 1983; Chang & Ki 2005). Although lessened by the internet's arrival and circumvented by industry experts' approximations, this limitation in the access to precise values still occurs.

The first specific issue concerns critical evaluation. The opinion of critics are mostly used as explanatory factors in the context of financial performance.[20] One of the first, and most influential, studies on critical evaluations was done by Eliashberg and Shugan (1997), where two possible roles are attributed to film critics: influencers, opinion leaders that persuade consumers to watch a film; or predictors, indicators of film quality and expected performance. Finding strong correlation of critics' evaluations with later weeks' box office, but not with early weeks, they reason that critics serve better as predictors than

[20] Simonton (2009) and Hadida (2009) propose critical evaluations as a success criteria, i.e. as measure of film performance.

as influencers. Contrastingly, critics have been said to fill both roles as positive and negative reviews relate to box office results over a period of eight weeks (Basuroy et al. 2003). Further studies have looked for interactions of critics' evaluations with other variables, finding that: the effects of star power[21] and critic reviews is diminished for genres familiar to the audience (Desai & Basuroy 2005), positive reviews are particularly impactful on the demand of certain genres and when competition between films abounds (Reinstein & Snyder 2005) and valence[22] is more important than volume (Chintagunta et al. 2010). On the whole, while the specific role of film critics' evaluations remains inconclusive, evaluations and the quality of a film seem to have a strong positive relation to box office performance (Litman & Kohl 1989; Ravid 1999; Collins et al. 2002; Zuckerman & Kim 2003).

Another popular subject within the film is the role stars play in the success of films. This has been studied using a large variety of methods: using awards[23] (Basuroy & Chatterjee 2008), high-profile film participation[24] (Litman 1983) or even using personal industry knowledge to decide on superstardom status (Prag & Casavant 1994). Ravid (1999) suggests that stars can possibly influence the success of films in two ways: as a signal of project quality or by capturing their economic rent.[25] While finding results consistent with the economic rent perspective, the author also notes that higher production costs may signal higher revenues regardless of the expense's source. Essentially, because stars expected to positively impact results also command higher salaries, their influence cannot be fully separated from the positive effect of an increased budget. Elberse (2007), using casting announcements on simulated and real stock market settings, found more concrete evidence of a star's presence positively influencing expected revenues. Interestingly, announcing the recruitment of a new star for an already strong cast has a greater impact than when the cast is unknown or less proven.

Budgets have, in general, been found to significantly and positively affect revenues (Litman & Kohl 1989; Prag & Casavant 1994; Ravid 1999; Elberse & Eliashberg 2003; Chang & Ki 2005; Liu 2006). Still, one point worth noting is that the inclusion of Prints & Advertising spending has been observed to deflate the impact of budget on box office performance (Prag & Casavant 1994), while being a positive predictor itself (Basuroy et al. 2006; Hennig-Thurau et al. 2006). Parental guidance ratings[26] is another subject, regularly found to impact revenues (Sochay 1994; Basuroy et al. 2003), addressed in this thesis. With most of these studies finding that G and PG ratings tend to help performance the most, there was an observable trend

---

[21] See next paragraph.

[22] The attributed rating or score.

[23] Commonly Academy Awards, as Oscar nominations – and sometimes Oscar wins – in the main categories generate significant box office gains (Nelson et al. 2001; Deuchert et al. 2005). Basuroy, Chatterjee and Ravid (2003), for example, created an index of Oscar nominations and wins.

[24] Like a Top10 film in previous years.

[25] Successful films increase stars' values which might then result in higher revenues for their next film.

[26] See Appendices 6 and 7 for examples.

in the industry to produce many R-rated films, a situation known as the "R-rating puzzle" (Ravid & Basuroy 2004). However, studies on this subject have shown somewhat mixed results. While De Vany and Walls (2002) found that producers should shift their investment from R to G and PG-rated films, Ravid and Basuroy (2004) note that R-rated films enjoy increased revenues and tend to lose money less often. Recently, the R-rating prevalence seems to have subsided, partly due to the current blockbuster-oriented path of the industry (Ravid & Basuroy 2004; Vogel 2011). Screen coverage is possibly the most universally agreed-upon variable, with the overwhelming majority of the literature finding more screens to have a significant positive effect on revenues (Sochay 1994; Neelamegham & Chintagunta 1999; Zuckerman & Kim 2003; Ainslie et al. 2005). Curiously, Chang and Ki (2005) find both positive and negative effects: the first on revenues and the second on the length of run.[27]

3.2. The producer's problem – A sequel's worth

One early example of the inclusion of sequels as a characteristic of interest is Litman and Kohl (1989), where a binary variable was found to have a significant and positive effect on rentals.[28] Prag and Casavant (1994) extended the approach by differentiating between the first and further sequels, confirming previous results for the former and finding the latter not significantly different from zero. Following these first findings, the presence of sequel-related variables was rather irregular. Still, when included, sequels were predominantly found to positively impact the different success criteria used. Namely, sequels were found to improve the probability of a hit (De Vany & Walls 1999) and garner more revenues than non-sequels (Ravid 1999; Simonoff & Sparrow 2000). Aside from this direct impact, sequel films have been noted to have stronger openings[29] and breaking even[30] much more frequently (Ravid & Basuroy 2004; Chang & Ki 2005). However, this seems to occur at the cost of a comparatively higher initial investment or budget (Basuroy et al. 2006). Partly explaining these effects are findings that these types of release generate higher online buzz as a result of being increasingly anticipated by audiences (Karniouchina 2011).

Studies specifically focused on sequels have been developed along two perspectives. The first views sequel films as brand extensions of an experiential good (Sood & Drèze 2006; Basuroy & Chatterjee 2008): successful films create a brand which is then extended by producers with the introduction of new products (sequels). This constitutes a different, less risky, strategy of product introduction. Sood and Drèze (2006) first examined this perspective and found that, as in other types of extensions, perceived similarity between two products (original and extension) crucially influences the extension's evaluation. However,

[27] Time spent in theatrical release.

[28] Distributor's share of revenues.

[29] Higher first weekend revenues.

[30] Recovering budgets.

conflicting with other types of extensions, consumers are found to value dissimilarity over similarity in sequels, suggesting the presence of a satiation process. The second perspective proposes sequels as quality signals. As Basuroy, Desai and Talukdar (2006) explain, the observable increased cost of sequel-production can function as a signal to consumers that the new film is of sufficient quality to warrant the investment. Signal credibility is attributed to the fact that a sequel of inferior quality is expected to reduce both the potential subsequent revenues (e.g. further sequels) and the value of existing films in ensuing deals, incentivizing a commitment to quality.

Following this discussion and previous research, the producer's problem hypothesis is

H_1: A sequel film's box office revenues will be positively related to the contemporaneous non-sequel films.

Alongside other control variables, this thesis also addresses four further sequel-specific factors with the purpose of strengthening this hypothesis. Following Basuroy and Chatterjee (2008), the revenues of the parent film, the time elapsed between the releases of the parent film and the target sequel[31], the number of intervening sequels and a measure of the revenue decay of sequels films at the box office are included. All these factors can either help shape the characteristics of a sequel release towards success, or serve as clues to determine if the existing conditions are unfavorable for such a release, in either case providing strength to the producer's problem hypothesis.

3.3. The distributor's problem – Timing an international release

The timing of a film's release has long been discussed in the literature. One way to approach the issue is through the use of binary variables that signal a release within peak periods of attendance. These studies tend to find diverging results. While Litman (1983) found positive significance for the Christmas/New Year period, Litman and Kohl (1989) and Sochay (1994) found summer releases to be the only ones to perform significantly better.[32] Meanwhile, similar results have been found for both Christmas and Easter peak periods (Chang & Ki 2005), as well as a combination of summer and Christmas releases (Hennig-Thurau et al. 2006). Another approach is through the use of a coefficient that varies for each film depending on the specific week of release: using Vogel's (2011) graphic of normalized weekly attendance for North America (1969-1984)[33], authors extract a coefficient between 0 and 1[34] displaying the attractiveness of a specific week depending on observed seasonality in attendance. The results, however,

[31] The one present in the sample.

[32] Out of three peak periods: Christmas-New Year, Easter and summer.

[33] Vogel (2011) shows *Variety* as the source with the copyright 1984 by A. D. Murphy.

[34] Or 0 and 100 for Elberse and Eliashberg (2003).

have not been very promising as it was found not significant in almost all instances where it has been used (Basuroy et al. 2003; Ravid & Basuroy 2004; Basuroy et al. 2006; Basuroy & Chatterjee 2008).

Multiple studies focused specifically on the subject have been published. Krider and Weinberg (1998) examined the competition for audience between two films, finding the existence of a trade-off between high season release benefits and the consequential increased competition. Avoiding competition against films with similar target audiences and minding a film's potential length of run can help improve the delaying/competing decision and mitigate the trade-off. Similarly, Radas and Shugan (1998) create a model where time moves faster during peak periods to find that timing decisions are linked to a film's life cycle, reckoning that peak season might not increase but accelerate revenues. Finally, Einav (2007) finds that observed seasonality is a result of two underlying patterns: the seasonality in the number and quality of films and the seasonality in underlying demand. As prices are mostly stable throughout the year the timing decisions made by distributors – number and quality of films released – amplify the patterns of underlying demand. As consequence, improvements in revenues can potentially be made by shifting some films from these high sales seasons to low sales season.

One particular way in which these timing decisions can be extended is towards international releases. The most common path has distributors releasing films in international markets after the domestic release.[35] Maintaining or extending these time gaps (lags) is supported by two main reasons.[36] First, specific for North American films, is the historical abundance of film theaters in North America and their relative scarcity abroad. The North American release/run can serve as a filtering mechanism that signals successful films and allows scarcer foreign theater owners to, potentially, better allocate screens. Second, lags allow the initial domestic performance to be observed and affords distributors the opportunity to adjust the marketing strategy for foreign markets. Although additional reasons have been asserted, their relevance has been somewhat lost recently. Due to industry evolution common concerns such as print costs[37] or subtitle/dubbing time[38] are largely being rendered inconsequential. Also, the use of key personnel in marketing efforts, an oft-used reasoning, is unfitting for the Portuguese market.[39] Conversely, a couple of reasons point towards the reduction of the release lags. First, using simultaneous

[35] As opposed to earlier international releases or simultaneous (day-and-date) releases.

[36] Throughout this paragraph Hadida (2009), Danaher and Waldfogel (2012) and Eliashberg, Elberse and Leenders (2006) serve as sources.

[37] Cost of producing physical copies. Because physical prints have been increasingly replaced with digital prints, the argument that reusing prints from domestic distribution in foreign venues greatly reduces costs has lost weight. The remarkable growth of digital cinemas – 16.383 in 2009 to 111.809 digital screens worldwide in 2013 – and its universal adoption – 80% of all screens – enable this notion (MPAA 2013).

[38] While not refutable altogether most studios include sufficient time in post-production phase to allow for these tasks to be accomplished, hence the existence of simultaneous releases in the first place.

[39] Although the use of a film's stars or/and director for promotion – e.g. appearing in premieres around the world – is common and requires release staggering, in countries where these occurrences are rare or almost non-existent (e.g. Portugal), this reason doesn't hold up.

or quasi-simultaneous releases allows economies of scale in advertisement, a point clearly tied with globalization effects. Second, the minimization or prevention of cannibalization of revenues by piracy, particularly in a big market/small market dynamic that facilitates the availability of pirating tools.

Academically, two different studies (partly) address this issue. Eliashberg and Elberse (2003) develop a model to understand the consumer-exhibitor relation and its drivers. Included is the use of time lags between releases, which the authors find to be a moderator of the relationship between North American and major European markets' revenues, an effect that flows through theater allocation decisions of exhibitors. The buzz or momentum generated for a film in the U.S./Canada market quickly wears out and fails to influence other markets' decision-makers if too much time passes. Danaher and Waldfogel (2012) frame their study around the emergence of BitTorrent service in 2003. By looking at the impact of time lags before and after the introduction of this popular piracy tool, they find considerable differences originated by its introduction. Lending substance piracy concerns, longer lags are found to be related to decreases in box office revenues in both periods. With the intention of ascertaining the effect of the relationship between North America and Portugal in terms of film revenues, the producer's problem model is adapted to the Portuguese market and this interaction. On the whole, the expectation stemming from these results is that, akin to previous results,

H_2: a film's box office revenues obtained in Portugal will be negatively related to the time lag between its releases in the North American and Portuguese markets.

Three unique aspects of this thesis are worth mentioning. First, instead of looking at the time interval as a moderator, this thesis is focused on its' role as a determinant of revenues. Second, the specific North America/Portugal relationship tackled here is not included in neither of the two aforementioned studies. Third, while both of those studies consider only North American films, this thesis extends the analysis to films produced in other countries. Consequences of this last difference are covered under sensitivity analysis in chapter 6.

4. Data

4.1. The producer's problem

The data sample is composed of films released in 2011 in the North American market. Using this specific year, instead of a more recent period, allows the concerns of cutting off films with long time lags[40] to be reduced. Two differences from previous literature are worth mentioning. First, the time period covered is substantially more recent than that of most other studies[41], a strong motivation for this thesis. Second, Basuroy and Chatterjee's (2008) sample, in particular, covers films between late 1991 and early 1993, roughly 1.5 years, while here only a single year is used, meaning that the number of films in specific periods of the year might be greater.[42] However, this should not significantly affect results due to seasonality effects being controlled for and because the global number of films in the sample is increased. Furthermore, reducing this interval to one year might help mitigate potential effects of price changes on revenues, improving the comparability of films within the sample.

The producer's problem depended variable is the log of weekly box office revenues obtained by a film, in the North American box office. For each film, weekly revenues are collected from the first week of release up to the fifteenth. If length of run exceeds fifteen weeks, only the first fifteen are included. If a film runs fifteen weeks or less, only those when revenues were obtained are covered. The log transformation, following the example of Basuroy and Chatterjee (2008), is intended to pull extreme box office revenues closer to the center and thus thrust this variable's distribution towards normal.

The website BoxOfficeMojo.com reports that 602 films were released during 2011 in the North American market. Since the most confidential data has historically been budget information, the availability of this variable is used as a decider on film inclusion. Although not strictly a random variable, descriptive statistics suggest that this doesn't constitute an obvious sample selection. The weekly revenues collected, for example, range from $56 to $226 million with a mean of $3.9 million and a standard deviation of $11.7 million. Extracting the budget information, with BoxOfficeMojo.com as primary and The-Numbers.com as secondary sources, leads to a sample of 212 films.[43] Four of these films are eliminated because of missing data: not reaching a full week of revenues or no records on parent film. Additionally, one last film is dropped due to having a unique Motion Picture Association of America (MPAA) rating.[44] The result is a data sample of 207 films, both North American and foreign, released in the U.S./Canada market in 2011. Table 1 below presents relevant descriptive statistics for the sample. Specific descriptions

[40] See distributor's problem.

[41] For Basuroy and Chatterjee (2008), the basis of the approach to sequels, this is a 20 year update.

[42] Namely, films from the early and late parts of year, as they can "appear" twice.

[43] BoxOfficeMojo.com is an IMDb company. The-Numbers.com is property of Nash Information Services.

[44] See Appendix 6.

of the hypothesis and remaining fifteen variables, as well as the sources used, are made in Appendix 6. Preliminary analysis is discussed below.

Table 1: Producer's problem descriptive statistics

Variable	Mean	Std. Deviation	Minimum sample value	Maximum sample value
Sequel	0.12	0.32	0	1
Revenue of Parent [a]	22.29	72.01	0	403.36
Sequel Time Gap	1.07	3.92	0	32
Number of Intervening Sequels	0.18	0.75	0	6
G	0.03	0.17	0	1
PG	0.13	0.34	0	1
PG-13	0.41	0.49	0	1
R	0.39	0.49	0	1
Unrated	0.04	0.2	0	1
Award	0.22	0.41	0	1
Budget [a]	39.7	45.77	0.14	250
Total Number	23.67	10.38	0	41
Positive Ratio	0.45	0.3	0	1
Seasonality	0.63	0.13	0.36	1
Number of Theaters in the first week	1854.29	1500.83	1	4375
Weekly revenue for the first week [a]	19.32	30.17	0.001	226.12

[a] in millions of dollars

The hypothesis variable, **Sequel**, directly measures the validity of H_1 by serving as comparison of the box office revenues obtained by sequel and non-sequel films. There are 24 sequel films in the sample, representing approximately 11.6%.[45] This sample's value is consistent with recent percentages, around 13% (Dhar et al. 2011). Descriptive statistics indicate that there is a large difference in the average total revenues between sequels and non-sequels: $135.5 million vs. $33.7 million. Using a crude measure of return on budget (total revenues to budget ratio) still indicates sequels are better performers on average: a ratio of 2.24 against the 1.56 of non-sequel. This presents a promising sign for the confirmation of the first hypothesis.

There are four strength variables. **Revenue of Parent** tends to be highly and positively correlated with the total revenues obtained (0.70).[46] However, sequels' total revenues compare negatively to those

[45] Larger than the 11 out of 167 films (6.58%) in Basuroy and Chatterjee's (2008) sample.

[46] See Appendix 8.

of the parent film.[47] Since the correlation indicates a positive coefficient but a negative comparison suggests the opposite, predicting the coefficient sign remains difficult.[48] **Sequel Time Gap**, on the other hand, is clearly expected to have a negative coefficient: there is a higher concentration of large revenue values for films with smaller time gaps. In the sample, excluding non-sequels, 61.5% of the films below the average gap of 9.21 years made at least $100 million in revenues, compared to only 40% of those above the average. The **Number of Intervening Sequels** appears to negatively impact the revenues obtained by a film as higher total revenues are more prevalent for films with fewer intervening sequels. This might be offset by *Harry Potter and the Deathly Hallows Part 2*¸ the highest earner in the sample, as it has the large value of six intervening sequels. The remaining strength variable, **Sequel x Week**, is an interaction term not suitable for preliminary analysis.

The remaining eleven variables are used to control for some commonly used factors. The first four variables (**G**, **PG**, **PG-13** and **R**) cover the most common film ratings issued by the MPAA. The overwhelming majority of films, and all major films by major studios, are rated by the MPAA.[49] While theoretically each successive rating implies a decreased audience size, since unrated films tend to differ significantly from the rated films in features and even distribution methods, all four coefficients are expected to be positive.[50] Neither the star power variable, **Award**, nor the seasonality effects variable, **Seasonality**, offer particularly strong preliminary indications observable. From the discussion made in the previous chapter, however, these variables should have a positive relation with revenues. Graphical analysis shows that high **Budget** values are associated with high total revenue values, a point reinforced by a fairly high correlation value (0.75), suggesting a strong positive coefficient.[51] Another common category concerns the evaluation – i.e. reviews – of film critics. Here, while the **Total Number** of reviews appears to positively influence the box office revenues of a film[52], the **Positive Ratio** doesn't offer any clear indication. The remaining two variables, **Number of Theaters** and its square, are related to film availability. Two particular points stand out when looking at the first week values. First, the clear exponential progression: as the number of theaters grows, the total revenues for a film tend to grow in an increasing manner. Second, the greatest divergence from this trajectory occurs in films with very small number of theaters. Examining an auxiliary

[47] Only 4 of the 24 (16.7%) sequels in the sample had revenues that exceed their CPI-adjusted parent film's revenues.

[48] See chapter 6, Results, for further reasoning.

[49] Only 9 out of 207 (4.3%) in the sample were unrated. For comparison Basuroy and Chatterjee (2008) had a percentage of around 3.5%.

[50] Unrated is the default when all these variables assume the value 0.

[51] Basuroy and Chatterjee (2008) found considerably lower mean and standard deviation values for the budget. This can be both a consequence of the difference in the time period analyzed as well as a representation of the recent industry trend of inflated production costs (Vogel 2011 p.97; Einav 2007).

[52] Of the 7 films that grossed $200 million or more, 5 received at least 35 reviews (the remaining received 27 and 29). Furthermore, while only 5 of the 108 films (4.8%) below the median number of reviews (25) reached $100 million, 24 films (24.2%) of the remaining reached that level.

variable of the maximum number of theaters achieved by a film in any of the weeks reported[53] shows the disappearance of this deviation. This suggests that there are two types of low availability films: the small release films (low availability throughout) and the platform release films (low initial availability with growth).[54] Since the square of the number of theaters aims to capture possible diminishing returns associated with greater releases, its coefficient is expected to be negative.

4.2. The distributor's problem

The data sample for the distributor's problem comes from the Portuguese market. Since the focus is on the interaction between the North American and the Portuguese markets in terms of release date lags, the data sample is an extension of the one used for the producer's problem. From the final 207 films of the producer's sample, 147 were released in Portugal.[55] In a way, this reduction already serves as an unintentional but unavoidable sample selection. Clearly, not all films released in a given market are released in another. From a managerial standpoint it stands to reason that the films that are released in an international market are expected by distributors to perform to a certain level, motivating the international release in the first place.[56] Yet, the actual performance of a film in the importing market is still to be decided. In this sense, improving the strategic decisions available at this point remains important for all parties involved.

Of the 147 films released in Portugal, only 139 are present in the final sample. As most variables are directly extended from the previous sample, the only variable whose availability conditions a film's presence is Revenue of Parent. A database supplied by *Instituto do Cinema e Audiovisuais* (ICA) serves as one source for this variable, with BoxOfficeMojo.com complementarily used[57] in order to minimize the exclusions.[58] Still, eight films are removed from the sample due to neither source bring able to supply the parent's revenue information. Consequently, the final sample is comprised of 139 films, with the dependent variable being the log of weekly revenues – up to the fifteenth week – obtained in the Portuguese market.

---

[53] Essentially, the highest value in the Number of Theaters variable for each film.

[54] The industry recognizes two main terms for releases: wide and limited. BoxOfficeMojo.com defines wide release as a film opening in 600 or more theaters (Rentrak and the MPAA use an 800 value). Consequently, a limited release occurs when films open below this threshold. Thus, a platform release is a limited release (first weeks) that grows towards wider release as the film gains momentum (e.g. through word of mouth or marketing efforts).

[55] Between November 2010 and May 2013, due to the release lags. CPI-adjustments were made to films not released in 2011.

[56] A Heckman 2-step procedure revealed no selection problems from the extension (i.e. using, in the first step, variables from the producer's problem, including initial results, as possible determinants of a Portuguese release).

[57] With the necessary adaptations: dollar to euro/escudo conversions and the CPI adjustment (from *Instituto Nacional de Estatística*'s website: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ipc).

[58] The ICA database extends only from 2007 to 2013, not covering a significant number of parent films.

The explanatory variables are divided between the single hypothesis variable, **Release Lag**, and the fifteen control variables. Appendix 7 describes the variables and their sources, as well as clarifications on the adaptations needed to extend the producer's problem.[59] Table 2 below covers some descriptive statistics of the sample.

Table 2: Distributor's problem descriptive statistics

Variable	Mean	Std. Deviation	Minimum sample value	Maximum sample value
Release Lag [a]	8.78	12.34	-10	94
Sequel	0.1	0.29	0	1
Revenue of Parent [b]	120.74	413.71	0	2,354.83
Sequel Time Gap	0.63	2.27	0	15
Number of Intervening Sequels	0.12	0.51	0	3
M/4	0.02	0.15	0	1
M/6	0.13	0.34	0	1
M/12	0.58	0.5	0	1
M/16	0.27	0.45	0	1
Award	0.29	0.46	0	1
Budget [c]	48.58	45.95	0.8	200
Total Number	26.89	8.95	0	41
Positive Ratio	0.48	0.31	0	1
Seasonality	0.62	0.13	0.36	1
Number of Showings in the first week	1001.77	711.11	38	3409
Weekly revenue in the first week [b]	162.94	213.06	1.93	1138.7

[a] Separating between sequels and non-sequels the mean values are 2.36 and 9.50, respectively.
[b] in thousands of euros
[c] in millions of dollars

The hypothesis variable, **Release Lag**, directly measures the validity of H_2 by measuring the time between releases in North America and Portugal. Graphical analysis suggests that a longer span of time between release dates is negatively related to the total revenues obtained by a film, with the highest revenues being obtained by films with lags between 0 and +5.[60] Furthermore, out of the thirteen films that surpass EUR 1 million in revenues, only three (23.1%) fall outside the 0 to 5 weeks range: two with negative lags (-2 and -8) and the last with 12 weeks between releases. Constructing a measure of revenues per showing[61] for the first week, accounting for availability, paints a similar picture with the highest values

[59] Particularly the film ratings category which saw a reduction in the number of variables, from four to three.

[60] All five highest earners – the only ones above the EUR 2 million threshold – fall within this interval.

[61] See Appendix 7's description of Number of Showings variable.

being concentrated around very small release lags. Together, these indications suggest a negative relation to revenues, supporting H_3.

Out of the fifteen control variables, the first five are directly related to sequels. The fourteen (10.1%) films present in this sample defined as **Sequels** are expected to perform better than non-sequels: the average total revenue of sequels is over four times higher than that of non-sequels. Even accounting for budget effects, again using the return on budget measure, sequels can be expected to outperform non-sequels (1.6 times higher returns). The **Revenue of Parent** variable presents mixed signals: while the positive correlation with total revenues is not as strong as in the producer's problem (0.58)[62], the number of films that exceed their CPI-adjusted parent's revenues is greater here.[63] The **Sequel Time Gap** lost some of its variability due to its maximum being considerably diminished (from 32 to 15), a difference that might lead to the variable being less clear-cut. Still, because total revenues seem to follow a downward trend as the gap increases, this variable is expected to be negatively related to revenues. Preliminary analysis on the **Number of Intervening Sequels** indicates that, excluding first sequels[64], the average total revenues increases with the number of intervening sequels. Separating first sequels from non-sequels reinforces that the revenues of first sequels are considerably higher than those of non-sequels.[65] Overall, the sample appears to support a growing trend between the values of 1 and 3 with the inclusion of first sequels in the 0 category potentially counteracting this expected positive coefficient. As in the producer's problem, the last variable pertaining sequels is the interaction term **Sequel x Week**.

The remaining ten control variables largely follow those of the producer's problem. As measured by **Award**, only four films with Oscar winners were dropped from the producer's sample, mildly suggesting that these films are more prone to follow through to international markets. Although films with Oscar winning personnel tend to perform slightly better[66], the differences in preliminary analysis do not appear substantial enough for this variable to positively and significantly impact revenues. Looking at the **Budget** shows that only four out of the fourteen (28.6%) films that surpassed EUR 1 million in revenues have a budget below $100 million.[67] Reusing the first week revenues per showing metric confirms that this positive relation, while slightly diluted, is still present after controlling for availability. The **Seasonality** and critical evaluation variables, **Total Number** and **Positive Ratio**, do not present any distinct indications from

[62] See Appendix 9.

[63] 6 out of the 14 films (42.9%).

[64] That, similar to non-sequels, are valued at 0.

[65] A difference which, in a preliminary analysis, cannot be separated from the sequel effect itself.

[66] In terms of first week and total revenues as well as stronger returns on budget.

[67] And the lowest of these, at $45 million, is still above the median and close to the average budget ($35 million and $48.6 million, respectively).

a preliminary analysis standpoint, with the seasonality variable even having no correlation with total revenues (0.0). The availability, measured by the **Number of Showings** and its square, shows a similar exponential behavior (first week's showings against total revenues) to that of the producer's problem. Furthermore, only five of the fourteen films (35.7%) that reached at least EUR 1 million have less than 2000 showings in the first week. The remaining three variables pertain the IGAC ratings (**M/6**, **M/12** and **M/16)**. Aside from being the biggest difference from the producer's problem, the default (M/4) classification is only surpassed by M/6 in average total box office revenues, signaling a positive coefficient for M/6 and negative ones for M/12 and M/16.

5. Methodology

The method of estimation used here is Generalized Estimating Equations (GEE), a procedure first introduced by Liang and Zeger (1986) and Zeger and Liang (1986). This follows directly from the approach made in Basuroy and Chatterjee (2008), as it functions as basis for the producer's problem addressed here. Furthermore, since the distributor's problem is developed as an extension of the producer's problem, with the focus shifting towards different variables (consequence of a different hypothesis), GEE will be used throughout. As most variables are either equal or similar for both problems, this option should also help in improving comparability of the results.

The GEE method is usually presented as an extension of standard Generalized Linear Models (GLMs), that uses quasi-likelihood estimation to work on correlated longitudinal and clustered data. Instead of depending on the joint distribution of the outcome's observations being fully specified, GEE uses the likelihood for the marginal distributions of the outcome variable and a working correlation matrix for the vector of repeated observations from each subject. The dependence between responses on the same subject – i.e. weekly revenues of the same film – are treated as a nuisance when estimating the marginal mean structure (Rabe-Hesketh & Everitt 2006). In addition, this method assumes heterogeneity across subjects, which means that weekly box office revenues should be independent across films. This, admittedly, can be a limitation of the model as films commonly compete for audience.

Aside from the difference in the fact that dependence between responses are treated differently, as Rabe-Hesketh and Everitt (2006) point out, there are three further differences between the GEE and random effects (RE) modelling. First, contrarily to RE, GEE is an estimation method that is not based on a statistical model. Second, the coefficients in RE modelling are subject-specific in the sense that the source of heterogeneity is being explicitly modeled and coefficients are interpreted on a subject (here, film) basis; GEE, on the other hand, returns marginal outcomes averaged over the population, meaning the coefficients represent the response averaged over the population. Third and perhaps most important, the parameter estimates of GEE are still consistent even for a misspecified correlation structure (if the mean structure is correct), unlike RE modelling. It's worth noting that under the specifications used here (Gaussian family with the identity link) the estimation of GEE would yield the same results as a RE estimation for balance panels, which is not the case here.[68]

From a practical estimation standpoint, two specific choices must be made in any analysis: the distribution family and the working correlation matrix structure (Horton & Lipsitz 1999). Additional, some estimation packages might allow for the specification of the link function, outside of the default canonical

[68] In any case, the differences between the estimated coefficients are minimal. See Appendices 10 and 14 for the Stata output of RE estimations.

link in each distribution. Since this thesis uses the Stata 13 software package, it's relevant to overview the options available. In terms of distribution family, Stata offer ranges between Bernoulli/binomial, gamma, Gaussian, inverse Gaussian, negative binomial and Poisson. The most common working correlation structures, such as independent, exchangeable, unstructured or auto-regressive, are all included as options. As touched upon above, a Gaussian family is chosen here, along with the corresponding identity link, with an exchangeable correlation structure, where the correlation across different weeks for each film is constant, is assumed.

Moving towards the procedure description, in the GEE the marginal regression model is assumed as

$$g(\mu_{ij}) = x'_{ij}\beta$$

where $\mu_{ij} = E(y_{ij})$, y_{ij} are the box office revenues of a film *i* at week *j*, x_{ij} is the *p* x 1 vector of study variables for the *i*th film at the *j*th week, β is a vector of the *p* parameters of interest and $g(\cdot)$ is the link function. Describing variance of y_{ij} as a function of the mean:

$$V(y_{ij}) = v(\mu_{ij})\emptyset ,$$

where $\emptyset$ is the scale parameter and $v(\cdot)$ is a known variance function. The link function mentioned above can be any monotonic differentiable function, and varies with the data. In this specific case, since it is expected – following the log transformation – that box office revenues are normally distributed, there is an "identity link" and the link function becomes

$$g(\mu_{ij}) = \mu_{ij} \text{ , leading to}$$

$$v(\mu_{ij}) = 1 \text{ and } V(y_{ij}) = \emptyset$$

Additionally, GEE models a working variance-covariance matrix for y_i as

$$V(\alpha) = \emptyset A_i^{1/2} R_i(\alpha) A_i^{1/2} ,$$

where A_i is the *n* x *n* diagonal matrix with $V\left(\mu_{ij}\right)$ as the *j*th diagonal element, and $R_i(\alpha)$ is the *n* x *n* working correlation matrix. Simplifying for normally distributed outcomes yields

$$V(\alpha) = \emptyset R_i(\alpha)$$

Finally, the GEE estimator of β is the solution of

$$\sum_{i=1}^{N} D'_i [V(\hat{\alpha})]^{-1} (y_i - \mu_i) = 0,$$

where $\hat{\alpha}$ is a consistent estimate of α and $D_i = {}^{\partial \mu_i}/_{\partial \beta}$.

Using the link function and the consequent simplifications, $\mu_i = X_i\beta$, $D_i = X_i$ and $V(\hat{\alpha}) = \widehat{\emptyset} R_i(\hat{\alpha})$ results in $\sum_{i=1}^{N} X_i'[R_i(\hat{\alpha})]^{-1}(y_i - X_i\beta) = 0$, with the GEE estimate of β becoming

$$\hat{\beta} = \left[\sum_{i=1}^{N} X_i'[R_i(\hat{\alpha})]^{-1}X_i\right]^{-1}\left[\sum_{i=1}^{N} X_i'[R_i(\hat{\alpha})]^{-1}y_i\right]$$

In general terms, the framework for this procedure starts with a guess of the working correlation matrix $R_i(\alpha)$ and the scale parameter $\emptyset$ which allows the estimation of β using iteratively reweighted least squares (GLIM is offered as an example by Zeger and Liang (1986)). Given these estimates, and using the standardized residuals, consistent estimates of α and $\emptyset$ can then be obtained. These steps are then iterated until convergence is achieved.

6. Results

6.1. The producer's problem

The main results concerning the producer's problem can be found in table 3, below.

Table 3: Producer's problem results – GEE estimation

Dependent variable: Log (Weekly Revenues) Wald χ2 = 12393.74 , p<0.00			
Variable	Coefficient (robust std. error)	z-value	
Sequel	**+ 1.207 (0.199)**	**6.05**	***
Revenue of Parent	– 0.001 (0.001)	– 1.22	
Sequel Time Gap	– 0.024 (0.006)	– 3.98	***
Number of Intervening Sequels	– 0.047 (0.061)	– 0.78	
Sequel * Week	– 0.056 (0.011)	– 4.86	***
G	+ 1.639 (0.394)	4.16	***
PG	+ 1.451 (0.370)	3.92	***
PG-13	+ 1.571 (0.374)	4.2	***
R	+ 0.970 (0.371)	2.61	**
Award	+ 0.197 (0.133)	1.48	
Budget	+ 0.004 (0.001)	3.02	**
Positive Ratio	+ 1.164 (0.229)	5.08	***
Total Number	+ 0.022 (0.008)	2.9	**
Seasonality	+ 0.180 (0.458)	0.39	
Number of Theaters	+ 0.003 (0.000)	29.02	***
$(\text{Number of Theaters})^2$	– 0.000 (0.000)	– 15.45	***
Constant	– 5.342 (0.462)	– 11.57	***

*** Significant at p<0.001 | ** Significant at p<0.01 | * Significant at p<0.05

First and foremost these results strongly support the producer's hypothesis. The significant positive coefficient shows that, as H_1 predicted, sequel films clearly outperform contemporaneous non-sequels in terms of box office revenues. This is particularly important in line of the discussion of the changes that the industry has been going through, which could have led to a diminished worth of producing sequels. The strength variables, however, offer mixed results. While the time gap between sequels and the sequel-week interaction have found strong significance, the revenues of the parent film and the number of intervening sequels are not significant. The revenue of the parent film can be seen as either a parent-target sequel revenue comparison or an indication of brand strength. The negative coefficient, although not significant, is more in line with the view that sequels tend to perform worse than the parent film in terms of box office revenues. Timing appears to be a very important moderator of the advantage of producing a sequel. Preserving strong associations in memory and taking advantage of brand equity in a timely manner favor a reduced gap between films. The number of intervening sequels can be seen in two ways: a quality signal for each subsequent sequel that is produced[69] or an indication of satiation effects. The negative coefficient appears to substantiate the latter in that too many extensions of the same brand might diminish the revenues of subsequent sequels. Additionally, managers should expect from sequels a different behavior at the box office: whether through satiation effects or higher anticipation from consumers, a stronger concentration of revenues on the first weeks can be expected.

The control variables results show that the only factors that don't significantly affect box office performance are the presence of an Oscar winner (cast or director), and the release date as measured by the seasonality coefficient. The effects of these two variables might be diluted in other variables – for example the effect of the cast through budget[70] – or altogether needed to be modeled in a different way. The most surprising finding, in comparison to Basuroy and Chatterjee (2008), is that the ratio of positive reviews is highly significant here, a result that might stem from the extension of the timing, the enlargement of their geographical origin[71], or both. Aside from small variations in significance levels, it's worth emphasizing that the use of number of theaters in substitution of the number of screens doesn't appear problematic as the two variables, as in their paper, found strong significance as determinants of box office revenues.

6.1.1. Sensitivity Analysis

Under the Gaussian family, identity link and exchangeable correlation structure specifications, GEE assumes the dependent variable to be normally distributed. However, the log of weekly revenues' histograms (Figure 1 below) show a sort of bimodality, particularly for the first two weeks. As mentioned previously, plotting the first week's number of theaters shows a concentration of films on very small values. The incredible difference in the scope of film availability, with the highest number of theaters (4375) being extremely distant from those low numbers, suggests that grouping films with such strong differences might be a cause of the bimodality problem. In order to avoid a direct revenue-based cut or trimming, a separation of the sample into two groups is made using the aforementioned maximum number of theaters variable. The specific threshold value is established at 10% of this variable's maximum (437.5), with Group A containing films below this value and Group B containing the remaining films. In other words, the sample is divided between films that, at any point in their run, were available in 10% of the market[72]

[69] A rational studio would not produce multiple sequels out of negatively received previous films. Seeing that a film is the fourth, fifth or sixth film in a series, for example, can push the consumer towards perceiving the series and the target sequel more favorably (Basuroy & Chatterjee 2008).

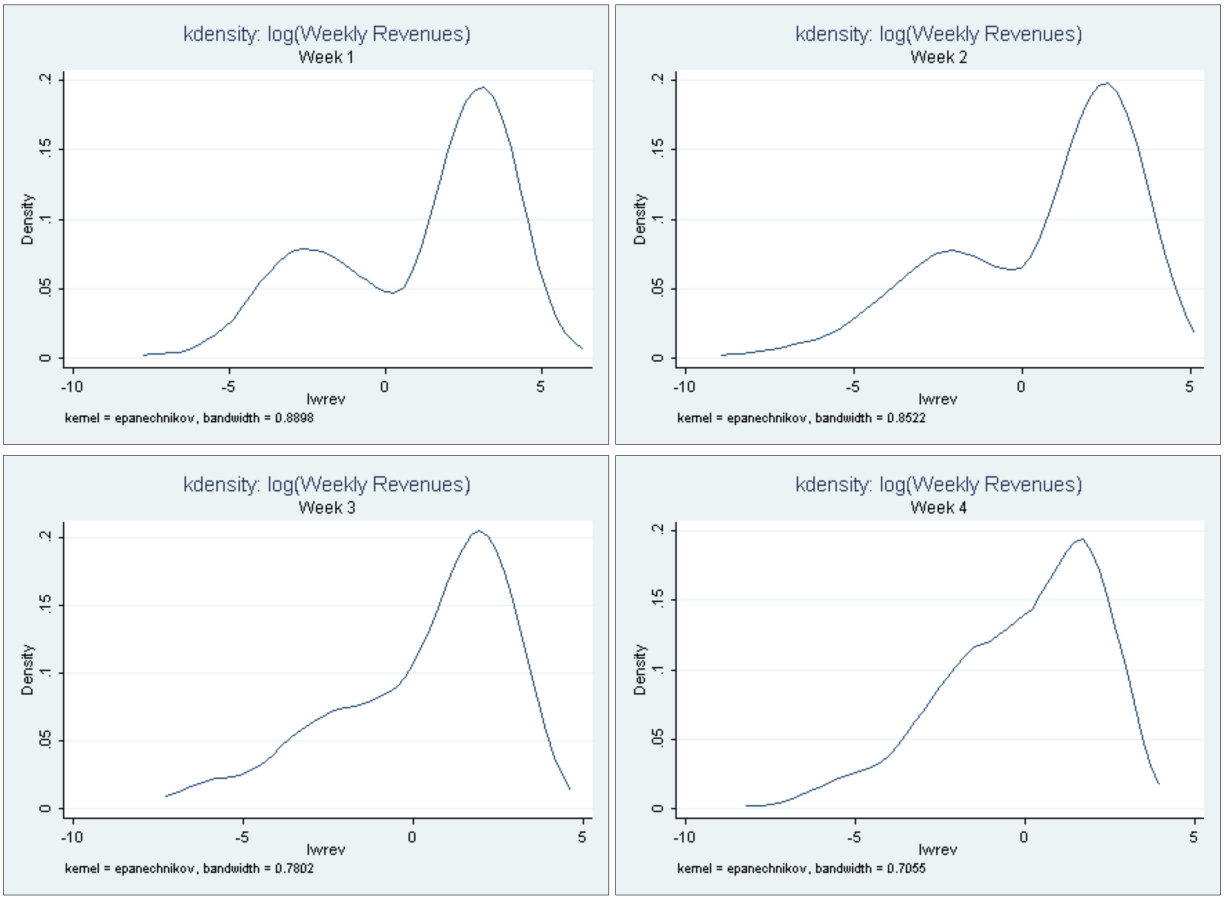
[70] Oscar winners and more renowned personnel can be expected to command higher salaries, which increase production costs.

[71] Reviews published in the major cities vs. published on national publications or online.

[72] As measured by the maximum number of theaters.

and those that failed to reach that value.[73] Two points are worth noting. First, this method allows platform releases to be present in both groups.[74] Second, since it's not a direct revenue-based cut there is some leeway towards the presence of successful/unsuccessful films in each of the groups. Most importantly, the bimodality of the dependent variable is eliminated with both groups presenting a distribution very close to the normal (Appendix 11).

Figure 1: kDensity graphs of the dependent variable, log of weekly revenues.



Descriptive statistics of the two groups (Appendix 12) present interesting nuances. Group A is composed of 69 films and contains neither sequels nor any film rated G. While the absence of sequels is to be expected, increased budgets and anticipation encouraging availability, the lack of G-rated films is somewhat surprising. Group B on the other hand, is composed of 138 films. No unrated film are present in this group, as these films have a clear tendency of being small-scoped. The upshot is that a model change is needed with regards to ratings. With the intent of keeping the coefficients positive, the rating with the lowest average total revenues, R, is the new default rating for this group. Comparing Group A's maximum values with Group B's minimum values yields curious insights. It's interesting to think that a film in which $36 million were invested into the budget can fail to be available in a significant way, suggesting that even

[73] Initially the division was based on reaching a wide release status. However, since different sources point towards different values for this wide release threshold (400, 600 or 800 or even more) the choice fell on the round 10%.

[74] Even Group A has the "smaller platform releases", i.e. under 20 initial theaters to about 400.

after hefty investments the distributor-exhibitor negotiations still filter expected bad performers. Looking at the total number of reviews indicates that scope decisions are not always in tune with the degree of critical attention, with some films of Group B receiving no critical reviews and multiple Group A films largely exceeding Group B's average number of reviews.

For completeness sake, GEE is applied to both groups with the results being presented in Appendix 13. Since the interest lies in analyzing the possible impacts of the bimodality and because several variables (six) must be omitted from Group A, the focus is on Group B. In this group, the R rating variable is omitted from the estimated as it functions as default. While the majority of the results are similar to those of the full sample, three surprising results can be found in Group B. First, there is a positive and significant relation between parent and target sequel's revenues, which suggests that the partial effect of a stronger brand (leading to stronger extensions) offsets the negative comparison between revenues. Second, Award is now a significant predictor, validating, at least for widely available films, studios' constant search for star power and bankable stars. Third, the total number of reviews variable loses its significance, an indication that the effect of this variable might be more powerful for films with fewer sources of information, making critics and their buzz a deciding factor in filmgoers' decision-making process. Since only one film released by the six major studios[75] is present in Group A and over 70% of the films of Group B come from these studios, these findings are especially interesting for the major studios. Overall, this prompts an interesting discussion on how to approach the study of this industry. Perhaps trying to identify the characteristics that influence the financial performance of all films, regardless of their dimension, can be misleading. A better choice, for optimizing specific decisions[76], might be to focus on Group B's results. However, since the process is dynamic and availability is also a strategic decision, the argument persists. It's not a coincidence that a large part of the research opted to address these issues by performing sample selection from the start, for example including only top 50 films.

6.2. The distributor's problem

Performing the GEE estimation for the Portuguese market yields the results presented in Table 4, below. The table of results supports the second hypothesis of this thesis. With a p-value of 0.012, the time lag between North American and Portuguese markets' releases is found to negatively influence the box office revenues in the latter market. This means that distributors in the Portuguese market should look to decrease, as much as possible, the lag between the releases in these markets in order to maximize the potential revenues. Auxiliary regressions show that releasing first in the Portuguese market yields neither particular gains nor losses.

---

[75] Which hold consistently over 80% of the production/distribution market.

[76] In this case, for bigger companies with larger catalogues.

Table 4: Distributor's problem results – GEE estimation

Dependent variable: Log (Weekly Revenues) Wald χ2 = 4349.78 , p<0.00		
Variable	Coefficient (robust std. error)	z-value
Release Lag	**– 0.012 (0.005)**	**– 2.50 ***
Sequel	+ 1.337 (0.407)	3.29 ***
Revenue of Parent	– 0.000 (0.000)	– 0.01
Sequel Time Gap	– 0.103 (0.036)	– 2.89 **
Number of Intervening Sequels	+ 0.127 (0.112)	1.13
Sequel * Week	– 0.025 (0.043)	– 0.60
M/6	+ 0.497 (0.379)	1.31
M/12	+ 0.128 (0.369)	0.35
M/16	+ 0.106 (0.372)	0.29
Award	+ 0.096 (0.085)	1.13
Budget	– 0.004 (0.001)	– 3.62 ***
Positive Ratio	+ 0.771 (0.155)	4.98 ***
Total Number	– 0.014 (0.007)	– 1.86
Seasonality	+ 0.112 (0.366)	0.31
Number of Showings	+ 0.008 (0.000)	19.92 ***
(Number of Showings)^2	– 0.000 (0.000)	– 8.51 ***
Constant	– 0.561 (0.437)	– 1.28

*** Significant at p<0.001 | ** Significant at p<0.01 | * Significant at p<0.05

The most surprising aspect of these findings is the significance of a negative coefficient for budget. Portuguese audiences appear to not only be unmoved by higher production values – many times likened to an "input" measure of quality – but indeed shy away from these blockbuster types in favor of other films, a characteristic of interest for specific marketing efforts in the Portuguese market. Assuming the audience doesn't outright prefer worse production values (such as costumes, sets etc.) and since budget values are not widely publicized or even available, this variable might not be influential on the basis of its specific value but rather on its perceivable clues. In other words, shying away from big-budgeted films may not be a consequence of a specific budget value but of the film characteristics into which it was translated. This indication signals that it might be of interest to managers to adjust the marketing effort to highlight other, more valued, characteristics. This result is made more surprising by both the sequel and the positive ratio significant results. Due to sequels' tendency for higher budgets, their positive influence on box office revenues would be expected to parallel a positive budget coefficient. In the case of the positive ratio, the surprise comes from both this variable and the budget being commonly seen as complementary quality assessments (Dhar et al. 2011).

Concerning ratings, the results echo those of North America[77] in that the rated-unrated contrast seems the most impactful part of the rating system: in the absence of unrated films – here, a result of the

[77] By Group B's results.

market's own inner workings – the differences between specific ratings appear insufficient to promote significant coefficients. Thus, the rating system remains mostly informative and not determinant of performance. As for critical evaluations, the positive ratio again finds strong support for a positive impact on revenues. Total number of reviews, however, ends up with a negative coefficient that fails to reach the highest threshold of 5% by a small margin (p-value of 0.062). It's important to keep in mind that both variables come directly from North America. At face value these findings indicate that film quality is important for Portuguese audiences but higher critical attention from North American critics tends to have a negative impact. This last result seems peculiar until the specifics of the variable are taken into account: the total number contains reviews published during the week previous to the release date and the opening weekend. This timing is appropriate for U.S./Canada market because the anticipation of a film is peaking in this period. Even conceding that the exposure between countries might lead to some anticipation from this critical attention to spillover to Portugal, expecting it to remain relevant when numerous weeks pass until a film's release – due to release lags – is not entirely realistic. Its inclusion is solely based on an attempt to keep the model as close as possible to that of the producer's problem. Still, its removal from the estimation has only minor impacts on coefficients with all significant levels remaining similar (Appendix 15).

The variables associated with sequels display mixed findings. While sequel and time gap between sequels remain strong determinants of revenues, revenue of parent and number of sequels failed to find significance. In contrast to North American findings, sequels in the Portuguese market do not behave in a consistently different manner along their run, with no significant faster revenue decay to be found. Additionally, the measure of availability, weekly number of showings (and its square) significantly and positively (negatively) impacts the revenues.[78] One last point worth noting is that, on the whole, the estimation finds considerably fewer significant determinants of box office revenues. Since no other studies for this market are available, it's difficult to understand if this is due to the characteristics of the market itself or a signal of further need for adaptations.

6.2.1. Sensitivity Analysis

6.2.1.1. North American productions or co-productions

The intention behind extending the producer's problem towards the distribution decision of increasing or reducing the release lag is to understand the exposure between the North American and Portuguese markets. Hence the inclusion of all films from the producer's problem with complete data in this sample. Some of these films, however, are not North American productions or co-productions. This

[78] For a better comparison against the producer's problem a separate estimation for North America is made, using only the 139 films released in Portugal (Appendix 18). This reduced sample shows a similar loss of significance of the total number of reviews and, as in Group B's results, shows that the measure of star power becomes an influential factor.

means that for the majority of these films three relevant releases exist: producing country, U.S./Canada and Portugal. Although the insights obtained concern the interaction of two of these markets (while disregarding the film's source), understanding if the dynamics shift or persist for North American productions/co-productions remains pertinent. Managerial implications might even be more fruitful as negotiations to correctly coordinate three releases is bound to be more troublesome.[79] Table 5 below presents the results of the estimation only for the 120 North American productions/co-productions.[80]

Table 5: Distributor's problem results, NA prod./co-prod. only – GEE estimation

Dependent variable: Log (Weekly Revenues) Wald χ2 = 4495.78 , p<0.00			
Variable	Coefficient (robust std. error)	z-value	
Release Lag	**+ 0.002 (0.008)**	**0.34**	
Sequel	+ 1.117 (0.358)	3.12	**
Revenue of Parent	+ 0.000 (0.000)	0.08	
Sequel Time Gap	– 0.087 (0.032)	– 2.69	**
Number of Intervening Sequels	+ 0.132 (0.102)	1.3	
Sequel * Week	– 0.011 (0.045)	– 0.25	
M/6	+ 0.302 (0.287)	1.05	
M/12	– 0.166 (0.272)	– 0.61	
M/16	– 0.094 (0.273)	– 0.34	
Award	+ 0.079 (0.095)	0.83	
Budget	– 0.003 (0.001)	– 2.77	**
Positive Ratio	+ 0.483 (0.163)	2.97	**
Total Number	– 0.002 (0.008)	– 0.23	
Seasonality	+ 0.135 (0.389)	0.35	
Number of Showings	+ 0.008 (0.000)	19.7	***
(Number of Showings)^2	– 0.000 (0.000)	– 8.42	***
Constant	– 0.710 (0.351)	– 2.02	*

*** Significant at p<0.001 | ** Significant at p<0.01 | * Significant at p<0.05

As can be seen immediately in the table the variable release lag lost all significance. Looking at the nineteen non-North American films that were dropped offers some clarification. Seventeen out of the nineteen films (89.5%) have positive release lags, with ten (52.6%) having at least 10 weeks between releases. Furthermore, out of the twelve films in the whole distributor's sample that had release lags above 20 weeks, half were foreign productions, including the outlier of +94 weeks. The variable's average dropped from 8.8 to 7.1 and the standard deviation from 12.3 to 8.1 weeks, following these changes in the sample. Since the majority of these films tend to have worse performances than North American productions/co-productions – EUR 124.4 thousand against EUR 484.3 thousand in average total revenues

[79] Besides, the producing country/North America relation has never been specifically studied for most countries, leaving the insights garnered here interesting but only part of a difficult to apply puzzle.

[80] As above, the removal of the Total Number variable does not significantly alter the results. For consistency, the variable is kept.

– the sample became much more concentrated. The combination of the loss of many of the more deviated observations with these being among the ones with lowest revenues leads to this table's result.

6.2.1.2. Number of weeks

Elberse and Eliashberg (2003) indicate that the strong relationship of performances in North America and foreign markets occurs for the first week's revenues and that release lags moderate this relationship. The extension of the model from the producer's problem resulted in the preservation of the fifteen week period. However, if the effect of a lag is concentrated on the first or first few weeks of release in the international market, it's possible that a length of fifteen weeks would dilute this effect. In other words, even if these initial weeks were strongly influenced by the release lag, the remaining weeks might lead the overall result towards non-significance. To understand how a different number of weeks of study might influence the results, the model is estimated for different study lengths. While Appendix 16 shows the Stata outputs for the whole sample, capping successively from weeks 4 to 11, Appendix 17 presents similar Stata outputs for North American productions/co-productions only.

As expected, the results for the whole sample show that the release lag remains negatively significant throughout. However, focusing on North American films shows a very strong effect for the first few weeks that is then lost as the length is increased. Until a week 7 capping, it has a negative significant influence on revenues, in agreement with H_2. This comes as confirmation of Elberse and Eliashberg (2003) findings that this inter-market effect is concentrated in the initial box office results, which seems entirely reasonable. If a film that took three weeks to arrive in Portugal is released in the same week as another film that took only one, for example, it seems unlikely that the effects of the extra two weeks are still (comparatively) penalizing the first film's revenues five, six, or seven further weeks down the length of run. Filmgoers who were exposed to the buzz generated in the North American release or first results are more likely to act on the added anticipation on the first opportunity available than long after the film has been released. For this reason, using the weekly capped estimations to counterbalance the results of the full fifteen weeks results seems to be a way to account for all factors.

It's interesting to note the different ways that some factors interact with box office results and evolve over the weeks. Looking at the estimations for the whole sample, sequel is clearly a significant positive influence on the box office throughout the entire run. As are both variables of the number of showings and the positive ratio. Yet, the M/6 rating, the award variable and the sequel-week interaction term start out by strongly impacting revenues in the first few weeks and then slowly lose their importance as weeks, in which they aren't particularly relevant, start to be added. The time gap between sequels, on the other hand appears to have no importance for shorter estimations but ends up finding significance in the later weeks. North American productions/co-productions have similar progressions with most differences happening in the limited 4 weeks' estimation.

The relevance of studying the first weeks, and a possible explanation for these behaviors, might lie with the fact that the large bulk of the whole revenues is made within the first few weeks. In fact, for this sample's films, the first five weeks' revenues (a third of the total length) represents on average 94.0% of the total revenues for all films and 95.2% for North American productions/co-productions. The release lag being significant in explaining the first six weeks, for example, is akin to saying that it is a significant factor in explaining 97.5% of all revenues obtained by North American films. For most purposes that would be a significant amount to have control of.

6.2.1.3. Number of Showings allocation

The last issue, also tied with Elberse and Eliashberg's (2003) findings, concerns the path through which the release lags influence revenues. In their paper, the authors suggest that the relationship between the North American and the foreign country's box office is indeed moderated by release lags, but mostly in an indirect manner: by influencing the decisions of the distributor-exhibitor duo in theater allocation. Longer release lags would lead this decision towards fewer theaters as distributors and exhibitors acknowledge the perishability of the buzz. Table 6 below shows the results from estimating the number of showings[81] on the remaining variables for the whole sample.[82] [83]

The result found by Elberse and Eliashberg (2003) is supported here: the time lag between releases has a significant negative influence in the availability decisions. This result shows that alongside the observed direct effect, the release lag has another, indirect, path to affect box office revenues. It's difficult, within this construction, to understand how simultaneous the two effects are. The coexistence, with significance, on the previous sections suggests that neither the number of showings contains the whole weight of the influence, nor does the release gap presence removes the impact of availability.

The positive significance of the parent film's revenue can be framed as an expectation of the repetition of success on the part of the decision-makers. The brand extension that results from a better performer is reasoned to have increased chances of success. Interestingly, due to the contrast against previous results, the budget positively affects availability decisions. This reflects the expected outcome, in most of the literature, of the perception of quality associated with this variable. Taken together, it can be said that the budget affects revenues both positively (albeit indirectly) and negatively. Lastly, the behavior

[81] Which, as mentioned functions as a substitute for number of theaters.

[82] The results for North American productions/co-productions are virtually identical, with the only difference lying on the total number of reviews narrowly failing to find significance for these productions.

[83] In addition, GEE regressions with week capping were also applied to this section. The results were very close to those of this full-run table. The sole exception being that the sequel time gap was found to significantly and negatively affect distributor-exhibitor decisions for earlier weeks, remaining significant at the 5% level until the 10 week capping.

of sequels is clearly expected by decision-makers to be different from that of non-sequels as both the dummy sequel and its interaction with the week are significant.

Table 6: Number of Showings estimation results – GEE estimation

Dependent variable: Number of Showings Wald χ2 = 246.01 , p<0.00			
Variable	Coef. (robust std. error)	z-value	
Release Lag	**– 4.68 (1.66)**	**– 2.82**	**
Sequel	1024.12 (219.02)	4.68	***
Revenue of Parent	0.37 (0.10)	3.67	***
Sequel Time Gap	– 32.10 (18.83)	– 1.70	
Number of Intervening Sequels	61.02 (63.23)	0.97	
Sequel * Week	– 153.45 (13.81)	– 11.11	***
M/6	42.13 (127.53)	– 0.33	
M/12	– 76.64 (118.05)	– 0.65	
M/16	– 100.64 (120.59)	– 0.83	
Award	20.32 (31.97)	0.64	
Budget	1.56 (0.47)	3.3	***
Positive Ratio	– 66.55 (46.45)	– 1.43	
Total Number	66.55 (1.50)	2.11	*
Seasonality	76.05 (115.09)	0.66	
Constant	226.1 (141.31)	1.6	

*** Significant at p<0.001 | ** Significant at p<0.01 | * Significant at p<0.05

7. Conclusion

One of the more prominent issues currently influencing film production is centered on sequels. Explicitly, the producer must determine if it's (still) worth it to produce a sequel. The answer is a resounding yes. All estimates throughout this thesis, even those not focused on this issue, reveal that sequel films consistently and significantly outperform contemporaneous non-sequels at the box office. The increased interest and production of this type of film is, for now, warranted. Still, managers in the field should acknowledge some limitations. As Ravid and Basuroy (2004) point out, the characteristics of a sequel are, by definition, a restrictive factor. When producers reach a point where the production of sequels is being forced, be it story-wise or in any other capacity, they might find consistent, strong financial performance hard to maintain. Moreover, the problem of cost is important. A balance between how far to invest and how much revenues to reasonably expect is important, even if sequels improve the latter.

Exploring strength variables concerning specific characteristics of a sequel release in the North American market presented two factors as worthy of attention by producers. First, producers should be concerned with the time elapsed between the parent and the target sequel, as it moderates the performance of the sequel at the box office. Going forward, with the evolution of the industry's sophistication with regards to brands and series – e.g. shared universes – this will tend to become a critical issue. Even if closer gaps are currently stimulating the performance of films, repeatedly recurring to small intervals between films in a series can lead to satiation effects. Second, managers should expect strong openings with faster decay of revenues from sequel films. Making the best of the short-run by finding less competitive release dates or fine-tuning the marketing effort for pre-release buzz can help them maximize initial revenues. An interesting challenge for managers is taking advantage of existing technology to translate word of mouth from inflated first week admissions into an extended plateau of strong weekly revenues.

Two other issues, although less sizable, are worth being acknowledged. The revenue of the parent films failing to find significance suggests that producers cannot be confident if a film will outperform or underperform its parent film, contrasting previous findings that sequels consistently underperform. For this difference to occur, either the sequels included here did not perform as poorly – in comparison to the parent – as in earlier studies, or distributors are finding better ways to leverage the stronger brand strength of better performing parent films. Likewise, the number of sequels in a series seems to be a mostly inconsequential characteristic, a result that might be tied to the permeation of the landscape with multiple types of sequels and series. The latest *Mission Impossible* or *The Fast and the Furious* films send very different signals, and elicit a different involvement from audiences, from those of a literary series such as *The Hunger Games*, for example. The investment in the story made by the audience differs between sequels of different series, meaning that the anticipation of additional sequels has varying intensities and is handled differently. For example, the fourth film out of a four film series is bound to be subjected to different expectations from the fourth film of an open-ended series. With occurrences like these, films can push both ways and a less well-defined effect is the combined result.

The second question of this thesis, the distributor's problem, is concerned with addressing the time gap between the releases of two markets, North American and Portuguese, and understanding how it affects the box office performance in the latter market. The results present evidence that postponing a release and expanding the release lag decreases the prospects of strong box office results. This effect is at least partly driven by the availability of a film, as decided by the distributors and exhibitors. This thesis differs from previous findings by studying the time interval itself as a determinant of box office performance, instead of its role as a moderator of the relationship between two countries' revenues. The effect of these lags is found to be limited in timing (particularly for North American productions/co-productions), with these being a much stronger driver for the initial weeks of a film's run. In order to take advantage of these findings distributors can follow two approaches: either push for the smallest gap possible, especially if the film is an expected fast-earner, such as sequels or most blockbusters; or, if lag reduction is not possible, adjust the marketing effort by negotiating longer (if smaller) availability in theaters and pushing the publicity and visibility promotion as deep into the film's run as possible.

References

Adams, R., 2014. The Shawshank Residuals. *The Wall Street Journal*. Available at: http://www.wsj.com/articles/SB10001424052702304536104579560021265554240 [Accessed April 22, 2015].

Ainslie, A., Drèze, X. & Zufryden, F., 2005. Modeling Movie Life Cycles and Market Share. *Marketing Science*, 24(3), pp.508–517.

American Film Institute, Rush (2013). Available at: http://www.afi.com/members/catalog/DetailView.aspx?s=&Movie=69693 [Accessed May 15, 2015].

Austin, B.A., 1986. Motivations for movie attendance. *Communication Quarterly*, 34(2), pp.115–126.

Basuroy, S. & Chatterjee, S., 2008. Fast and frequent: Investigating box office revenues of motion picture sequels. *Journal of Business Research*, 61(7), pp.798–803.

Basuroy, S., Chatterjee, S. & Ravid, S.A., 2003. How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, 67(October), pp.103–118.

Basuroy, S., Desai, K.K. & Talukdar, D., 2006. An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research*, 43(2), pp.287–295.

BoxOfficeCine.fr, Classement des meilleurs films au box office pour l'année 2013. Available at: http://www.boxofficecine.fr/top-2013.html [Accessed April 27, 2015].

BoxOfficeMojo.com, 2009. China Yearly Box Office. Available at: http://www.boxofficemojo.com/intl/china/yearly/?yr=2009&p=.htm [Accessed April 29, 2015].

Brew, S., 2012. 105 Movie Sequels Currently in the Works. *Den of Geek*. Available at: http://www.denofgeek.com/movies/18891/105-movie-sequels-currently-in-the-works [Accessed April 4, 2015].

Brew, S., 2014. 129 Movie Sequels Currently in the Works. *Den of Geek*. Available at: http://www.denofgeek.com/movies/sequels/30022/129-movie-sequels-currently-in-the-works [Accessed April 4, 2015].

Brook, T., 2013. How the Global Box Office is Changing Hollywood. *BBC.com*. Available at: http://www.bbc.com/culture/story/20130620-is-china-hollywoods-future [Accessed May 1, 2015].

Burzynski, M.H. & Bayer, D.J., 1977. The effect of positive and negative prior information on motion picture appreciation. *The Journal of Social Psychology*, 101(2), pp.215–218.

Chang, B.H. & 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), pp.247–269.

Chintagunta, P.K., Gopinath, S. & Venkataraman, S., 2010. The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science*, 29(5), pp.944–957.

Collins, A., Hand, C. & Snell, M.C., 2002. What makes a blockbuster? Economic analysis of film success in the United Kingdom. *Managerial and Decision Economics*, 23(6), pp.343–354.

Cuadrado, M. & Frasquet, M., 1999. Segmentation of Cinema Audiences : An Exploratory Study Applied to Young Consumers. *Journal of Cultural Economics*, 23, pp.257–267.

D'Astous, A. & Touil, N., 1999. Consumer Evaluations of Movies on the Basis of Critics' Judgments. *Psychology & Marketing*, 16(8), pp.677–694.

Danaher, B. & Waldfogel, J., 2012. Reel Piracy: The Effect of Online Film Piracy on International Box Office Sales. *Available at SSRN 1986299*.

Daniels, B., Leedy, D. & Sills, S.D., 1998. *Movie Money: Understanding Hollywood's (Creative) Accounting Practices. Silman-James Press, Los Angeles, CA.No Title* 1st ed., Los Angeles, CA: Silman-James Press.

De Vany, A. & 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), pp.425–451.

De Vany, A. & 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), pp.285–318.

Desai, K.K. & Basuroy, S., 2005. 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), pp.203–223.

Deuchert, E., Adjamah, K. & Pauly, F., 2005. For Oscar Glory Or Oscar Money? *Journal of Cultural Economics*, 29(3), pp.159–176.

Dhar, T., Sun, G. & Weinberg, C.B., 2011. The long-term box office performance of sequel movies. *Marketing Letters*, 23(1), pp.13–29.

Eaton, A., 2013. Why being British is no longer a handicap when it comes to films. *The Guardian2*. Available at: http://www.theguardian.com/film/2013/dec/16/british-film-industry-andrew-eaton [Accessed May 5, 2015].

Einav, L., 2007. Seasonality in the US motion picture industry. *The Rand Journal of Economics*, 38(1), pp.127–145.

Elberse, A., 2007. The power of stars: Do star actors drive the success of movies? *Journal of Marketing*, 71(4), pp.102–120.

Elberse, A. & Eliashberg, J., 2003. Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, 22(3), pp.329–354.

Eliashberg, J., Elberse, A. & Leenders, M.A.A.M., 2006. The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Science*, 25(6), pp.638–661.

Eliashberg, J. & Shugan, S.M., 1997. Film critics: Influencers or predictors? *The Journal of Marketing*, 61(2), pp.68–78.

Eliashberg, J., Weinberg, C.B. & Hui, S.K., 2008. Decision Models for the Movie Industry. In C. B. Weinberg, ed. *Handbook of Marketing Decision Models*. pp. 437–468.

Epstein, E.J., 2012. *The Hollywood Economist 2.0: The Hidden Financial Reality Behind the Movies*, Melville House.

Fischer, R., 2014. Is "Pirates 5" Conceived as a Soft Franchise Reboot? */Film*. Available at: http://www.slashfilm.com/pirates-of-the-caribbean-reboot/ [Accessed April 23, 2015].

Gada, K., 2013. Netflix and the Culture of Creation. *Forbes*. Available at: http://www.forbes.com/sites/onmarketing/2013/04/24/netflix-and-the-culture-of-creation/ [Accessed April 26, 2015].

Gazley, A., Clark, G. & Sinha, A., 2011. Understanding preferences for motion pictures. *Journal of Business Research*, 64(8), pp.854–861.

Guiness World Records, First Feature Film. Available at: http://www.guinnessworldrecords.com/records-1/first-feature-film/ [Accessed April 18, 2015].

Hadida, A.L., 2009. Motion picture performance: A review and research agenda. *International Journal of Management Reviews*, 11, pp.297–335.

Hennig-Thurau, T., Houston, M.B. & Heitjans, T., 2009. Conceptualizing and Measuring the Monetary Value of Brand Extensions : The Case of Motion Pictures. *Journal of Marketing*, 73(November), pp.167–183.

Hennig-Thurau, T., Houston, M.B. & Walsh, G., 2006. The differing role of success drivers across sequential channels: an application to the motion picture industry. *Journal of the Academy of Marketing Science*, 34(4), pp.559–575.

Horton, N.J. & Lipsitz, S.R., 1999. Review of software to fit generalized estimating equation regression models. *The American Statistician*, 53(May), pp.160–169.

Ingram, M., 2015. Reed Hastings on the future of non-linear television. *Fortune*. Available at: http://fortune.com/2015/05/11/reed-hastings-on-the-future-of-non-linear-television/ [Accessed May 14, 2015].

Instituto do Cinema e do Audiovisual (ICA), 2009. Anuário Estatístico. , pp.1–68.

Instituto do Cinema e do Audiovisual (ICA), 2010. Anuário Estatístico. , pp.1–84.

Instituto do Cinema e do Audiovisual (ICA), 2011. Anuário Estatístico. , pp.1–72.

Instituto do Cinema e do Audiovisual (ICA), 2012. Anuário Estatístico. , pp.1–73.

Instituto do Cinema e do Audiovisual (ICA), 2013. Anuário Estatístico. , pp.1–132.

Instituto do Cinema e do Audiovisual (ICA), 2008. Cinema - Estatísticas 2008. , pp.1–53.

Jowett, G.S., 1985. Giving them what they want: Movie audience research before 1950. *Current Research in Film: Audiences, Economics, and Law*, 1, pp.19–36.

Karniouchina, E. V., 2011. Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing*, 28(1), pp.62–74.

Kolesnikov-Jessop, S., 2011. Hollywood Presses Its Global Agenda. *The New York Times*. Available at: http://www.nytimes.com/2011/05/23/business/media/23film-screensingapore.html?_r=0 [Accessed May 3, 2015].

Krider, R.E. & Weinberg, C.B., 1998. Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game. *Journal of Marketing Research*, 35(1), pp.1–15.

Lafrance, A., 2015. How Technicolor Changed Storytelling. *The Atlantic*. Available at: http://www.theatlantic.com/technology/archive/2015/02/technicolor-at-100/385039/ [Accessed February 17, 2015].

Levine, R., 2011. How the internet has all but destroyed the market for films, music and newspapers. *The Guardian*. Available at: http://www.theguardian.com/media/2011/aug/14/robert-levine-digital-free-ride [Accessed May 1, 2015].

Liang, K.Y. & Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), pp.13–22.

Litman, B.R., 1983. Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, 16(4), pp.159–175.

Litman, B.R. & Kohl, L.S., 1989. Predicting financial success of motion pictures: The ’80s experience. *Journal of Media Economics*, 2(2), pp.35–50.

Liu, Y., 2006. Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, 2(July), pp.74–89.

Lumiere, Rush (2013). Available at: http://lumiere.obs.coe.int/web/film_info/?id=43115 [Accessed May 12, 2015].

Möller, K.E. & Karppinen, P., 1983. Role of motives and attributes in consumer motion picture choice. *Journal of Economic Psychology*, 4(3), pp.239–262.

Motion Picture Association of America (MPAA), 2014. *The Economic Contribution of the Motion Picture & Television Industry to the United States Fact Sheet*, [Motion Picture Association of America].

Motion Picture Association of America (MPAA), 2011. Theatrical Market Statistics. , pp.1–21.

Motion Picture Association of America (MPAA), 2012. Theatrical Market Statistics. , pp.1–25.

Motion Picture Association of America (MPAA), 2013. Theatrical Market Statistics, pp.1–31.

Motion Picture Association of America (MPAA), 2005. U. S. Theatrical Market: 2005 Statistics. , pp.1–26.

National Association of Theatre Owners (NATO), Theatrical release window. Available at: http://natoonline.org/data/windows/ [Accessed May 2, 2015].

Neelamegham, R. & Chintagunta, P.K., 1999. A Bayesian model to forecast new product performance in domestic and international markets. *Marketing Science*, 18(2), pp.115–136.

Nelson, R.A. et al., 2001. What’s an Oscar worth? *Economic Inquiry*, 39(1), pp.1–16.

Nelson, R.A. & Glotfelty, R., 2012. Movie stars and box office revenues: an empirical analysis. *Journal of Cultural Economics*, 36(2), pp.141–166.

Prag, J. & Casavant, J., 1994. An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *Journal of Cultural Economics*, 18(3), pp.217–235.

PwC, Global entertainment and media outlook: 2014-2018. Available at: http://www.pwc.com/gx/en/global-entertainment-media-outlook [Accessed April 29, 2015].

PwC, 2013. *Portugal Entertainment and Media Outlook: 2013-2017*,

Rabe-Hesketh, S. & Everitt, B.S., 2006. *A handbook of statistical analyses using Stata* 4th ed., New York: CRC Press.

Radas, S. & Shugan, S.M., 1998. Seasonal marketing and timing new product introductions. *Journal of Marketing Research*, XXXV(August), pp.296–315.

Ravid, S.A., 1999. Information, Blockbusters, and Stars: A Study of the Film Industry*. *The Journal of Business*, 72(4), pp.463–492.

Ravid, S.A. & Basuroy, S., 2004. Managerial Objectives, the R- Rating Puzzle, and the Production of Violent Films. *The Journal of Business*, 77(S2), pp.S155–S192.

Reinstein, D.A. & Snyder, C.M., 2005. The influence of expert reviews on consumer demand for experience goods: A case study of movie critics*. *The Journal of Industrial Economics*, 53(1), pp.27–51.

Schedeen, J., 2010. The History of 3D Movie Tech. *IGN*.

Shao, H., 2013. China's Box Office to Surpass the U.S. by 2019, Says IMAX CEO. *Forbes*. Available at: http://www.forbes.com/sites/hengshao/2013/10/17/chinas-box-office-to-surpass-the-u-s-by-2018-says-imax-ceo/ [Accessed May 8, 2015].

Simonoff, J.S. & Sparrow, I.R., 2000. Predicting movie grosses: Winners and losers, blockbusters and sleepers. *Chance*, 13(3), pp.15–24.

Simonton, D.K., 2009. Cinematic success criteria and their predictors: The art and business of the film industry. *Psychology & Marketing*, 26(5), pp.400–420.

Sochay, S., 1994. Predicting the performance of motion pictures. *Journal of Media Economics*, 7(4), pp.1–20.

Sood, S. & Drèze, X., 2006. Brand extensions of experiential goods: Movie sequel evaluations. *Journal of Consumer Research*, 33(December), pp.352–361.

Sun, S., 2011. Top 10 Box Office Films of 2010. Available at: https://shaoyis.wordpress.com/2011/01/28/top-10-box-office-films-of-2010/ [Accessed April 28, 2015].

Sun, S., 2012. Top 10 Box Office Films of 2011 in China. Available at: https://shaoyis.wordpress.com/2012/01/13/top-10-box-office-films-of-2011-in-china/ [Accessed April 28, 2015].

Sun, S., 2013. Top 10 Box Office Films of 2012 in China. Available at: https://shaoyis.wordpress.com/2013/01/06/top-10-box-office-films-of-2012-in-china/ [Accessed April 29, 2015].

TCM.com, Rush (2013). Available at: http://www.tcm.com/tcmdb/title/871444/Rush/ [Accessed May 1, 2015].

Thompson, D., 2011. How Hollywood Accounting can Make a $450 Million Movie "Unprofitable." *The Atlantic*. Available at: http://www.theatlantic.com/business/archive/2011/09/how-hollywood-accounting-can-make-a-450-million-movie-unprofitable/245134/ [Accessed May 10, 2015].

Thompson, D., 2013. The Big Business of Big Hits: How Blockbusters Conquered Movies, TV, and Music. *The Atlantic2*. Available at: http://www.theatlantic.com/business/archive/2013/10/the-big-business-of-big-hits-how-blockbusters-conquered-movies-tv-and-music/280298/ [Accessed February 2, 2015].

Tsui, C., 2013. China Box Office: 10 Highest-Grossing Movies of All Time. *The Hollywood Reporter2*. Available at: http://www.hollywoodreporter.com/gallery/china-box-office-10-highest-425083/1-avatar-james-cameron-2010-2219-million [Accessed May 2, 2015].

UNESCO, 2013. Emerging Markets and the Digitalization of the Film Industry. , (14), p.39.

USA Today, 2007. 15 actors sue New Line Cinema over “Lord of the Rings” profits. Available at: http://usatoday30.usatoday.com/life/movies/news/2007-06-06-rings-suit_N.htm [Accessed May 11, 2015].

Vogel, H.L., 2011. *Entertainment Industry Economics: a Guide for Financial Analysis* 8th ed., Cambridge University Press.

Zeger, S.L. & Liang, K.Y., 1986. Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics*, 42(1), pp.121–130.

Zuckerman, E.W. & Kim, T.Y., 2003. The critical trade-off: identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change*, 12(1), pp.27–67.

Appendices

Appendix 1 – Film characterization terminology

The basic terminology used by the industry to characterize films, with regards to its standing within a series, can be reduced to sequels, prequels, remakes, reboots, spin-offs and shared universes.

Sequel: a film that is released after the film that originates the series, portraying events that happen after those of the film that precede it. Usually shares characters, settings and more with the other films in the series. This is the most common option, with most series being composed of more than one sequel.

Prequel: a film that, despite being produced and released after the film that initiates the series (or another film in the series), conveys a story that takes place earlier in the timeline. Usually shares characters, settings and more with the other films in the series. Perhaps the most famous example is the *Star Wars* prequel trilogy (Episodes I, II and III) which arrived in theaters sixteen years after the last film in the original trilogy, *Return of the Jedi* (1983) (also known as Episode VI).

Remake: a film that presents a story previously portrayed by another film (usually not a recently one). With the intention of updating it for a new audience, it typically approaches the story in a similar but not identical way. Thus, the core plot points tend to match and the film works as a new version of the previous one. As an example, *Arthur* (2011) is a remake of the 1981 film with the same name.

Reboot: a film that restarts an established series by ignoring the story portrayed by the preceding films. Normally occurs when a series is not progressing in the desired manner. Taking advantage of characters or settings known to the audience it cuts all story ties (and personnel) with the previous entries in the series. One example is the 2012 film *The Amazing Spider-Man*, which serves as a reboot to the *Spider-Man* trilogy released between 2002 and 2007.

Spin-off: a film that is focused on a particular feature (commonly a secondary character) of the film or series from which it derives. The film *Get Him to the Greek* (2010) is centered on the character Aldous Snow which first appeared in *Forgetting Sarah Marshall* (2008).

Shared Universe: a fictional universe where multiple films are set. Sharing characters, settings and other story elements, having an overarching story or the existence of continuity between films are all possibilities that vary from universe to universe. The most current example of a share universe in film is the *Marvel Cinematic Universe* where characters like Iron Man, Thor, The Hulk or Captain America exist within same universe, allowing them to crossover and appear in *Marvel's The Avengers* (2012).

To complicate matters even further, not only aren't these categories mutually exclusive but also further variations are created regularly. *Puss in Boots* (2011) is considered both a prequel and a spin-off, while *Pirates of the Caribbean: Dead Men Tell No Tales* (exp. 2017), the fifth film of the successful series – currently under development – is rumored to be a soft reboot, which in this particular case amounts to ignoring *only* the fourth film (Fischer 2014).

Appendix 2 – Ancillary markets/channels

The network of companies that operate in the three stages has become, over the years, a complex system of cross-ownerships and joint-venture arrangements (Vogel 2011). One reason for this situation is the need to face the uncertainties and risks involved in this industry, which led major companies to diversification and no longer being solely devoted to the theatrical film industry. Particularly, the growth of ancillary markets and channels – such as home entertainment, television (both pay and network), streaming and other ancillary rights (merchandising, video games and more) – has made them very desirable for companies. In this sense it's important to recognize that while a theatrical release is an extremely significant channel, financial success can be found despite weak box office results and even without a theatrical release, as exemplified by cult films and television/direct-to-video films. *Shawshank Redemption* (1994), for example, made only $28 million in North America and $30 million internationally on a $25 million budget, which by most standards is considered a failure when advertising costs are factored in. However, since its release in 1994 it went on to make about $80 million on the video rental market, has to date aired on 15 unique basic cable networks, filled 151 hours of basic cable airtime in 2013 alone and is still, almost 20 years later, generating substantial residuals to at least one of its actors (Adams 2014). With new and evolving platforms such as streaming, ancillary channels will become increasingly diverse and potentially rewarding for companies that are able to adapt and meet different types of consumers (PwC). A tendency towards a diminished weight of the theatrical channel can be seen by looking at the theatrical release window:[84] information from National Association of Theatre Owners (NATO) for the North American market shows that this window has been steadily declining, from close to 170 days in 2000 to an industry average below 120 days in 2013. Interestingly, while the visible part of the discussion is centered on financial success being decided by box office revenues themselves, these might be more significant by serving as catalysts. In other words, theatrical performance might not be where these companies profit from their investment, but where the success (and profits) of subsequent release windows is being critically shaped (Eliashberg et al. 2006).[85] Ultimately, two points should be clear: ancillary channels should be recognized as important sources of revenues to be explored by managers and a theatrical release is still relevant, even if for different reasons.

[84] Temporal window where a film is exclusively in theatrical release, that is, before release on ancillary channels.

[85] "Theatrical exhibition is the major factor in persuading the public what they want to see, even if that public never sets foot inside a motion picture theater. And how well and how long a picture plays in theaters has everything to do with its value in other markets" (Daniels et al. 1998).

Appendix 3 – Worldwide market context

Ever since the first feature film, *The Story of the Kelly Gang*, was made in Melbourne, Australia in 1906 for £450 (Guiness World Records), the industry has evolved tremendously. Nowadays, budgets above $100 million are common. In the North American market alone, the 2010-2013 period saw the release of 105 films with budgets in excess of $100 million, with 23 of these costing more than $200 million. Yet, these substantial numbers show only part of the considerable reach of the industry. According to the Motion Picture Association of America (MPAA), 1.9 million jobs are supported in the U.S. domestic market alone, of which 293.000 and 360.000 were direct and related distribution jobs, respectively. Along with the $46 billion generated in wages, $38 billion were made in payments by the industry to the network of 330.000 local businesses around the country, in 2012 (MPAA 2014).

Although these values underscore the weight of the Hollywood film industry, it's worth mentioning that this is not the largest market by all measures. While the North American market was, in 2013, the largest in terms of box office with $10.9 billion revenues – followed by China ($3.6 billion) and Japan ($2.4 billion) – changing measures alters the standings. In number of admissions, for 2011, India is the undisputed leader with over 2.9 billion, followed by the US/Canada with close to 1.3 billion and China with 370 million (UNESCO 2013). Considering the number of films produced again reveals India as the largest market. The average of 1178 films produced yearly between 2005 and 2009 kept it ahead of Nigeria (1094) and North America (555) (UNESCO 2013). Interestingly, although India clearly tops admissions, it appears only as sixth in the box office ranking with $1.47 billion.[86] Despite more than doubling the United States' number of admissions, noticeably dominating the quantity, it reached only about 14.4% of its revenues, suggesting a huge discrepancy in the pricing, along with exchange rate effects.

The recent evolution of the industry paints an attractive picture for managers: the global box office has been growing significantly. From the $15.9 billion of 2000 – the first year reported by the MPAA – it grew to the $35.9 billion of 2013. Although reservations on the accuracy of this data are reasonable, since this is a worldwide measure dependent on many primary and secondary sources, the magnitude of the growth – approximately $20 billion (close to 126%) – doesn't allow dispute over the global trend. However, despite this remarkable aggregate growth, looking at the top countries in the last three years makes it clear that not all markets tread the same path (MPAA 2011, 2012, 2013). China, following the global trend, exhibits remarkable growth – 80% from 2011 to 2013, over 33% yearly – and has become, indisputably, the second biggest market. However, countries like Japan, Russia, India and Brazil have followed more gradual growth trajectories with increases between $0.1 and $0.2 billion (4.5% to 16.7%) over the same period. On different paths are most large European markets, along with Australia. While the U.K. and Germany showed perfect stability over the three years, the Australian market displays some volatility by growing and declining the exact same $0.1 billion

[86] For the same year, 2011

in consecutive years. Falling on the negative range of the spectrum are Italy and France with $0.1 billion (-11.1%) and $0.4 billion (-20%) decreases, respectively.

Looking towards the future, the global growth is expected to continue. Estimates put global box office values at $45.9 billion in 2018, with 8.89 billion admissions (PwC). These are centered on the expectation that, even with the predictable price increases being confirmed, watching a film in theaters will continue to be a financially accessible option for most consumers. Moreover, the slower growth expected of the North American market is predicted to be offset by the opportunities in developing markets such as China and Brazil, where a deficit of venues should, when addressed, stimulate these markets into considerable expansion (PwC 2013). So remarkable are these prospective growths that some sources predict the Chinese market to surpass its North American counterpart as soon as 2018, and to double its value in 2025 (Shao 2013). Alongside these, and possibly constraining, is the notion that “filmed entertainment” distribution will continue to undergo changes in order to adapt to new players and platforms. One example that has already affected the television segment is the incursion of Netflix, an on-demand internet streaming provider, into content creation (Gada 2013; Ingram 2015). As in this case, new options for consumers will call for adaptations from the industry and one direction in which these challenges seem to be pushing is towards a greater focus on the industry’s most unique product: the big budget production (Thompson 2013).[87] Embracing this course would lead (even) larger investments on a reduced number of films: global brands with strong international influence (PwC 2013). This perspective underscores the value that sequels and series are expected to have moving forward. With continued reliable performances at the box office, it’s expected that these types of production remain a prevalent feature of the market in the coming years (PwC). On the whole, significant opportunities will clearly be available for companies with the ability to adapt and develop through persistent challenges.

[87] From Vogel (2011 p.96): “One of the most noticeable tendencies, for instance, has been the virtual dichotomization of the theatrical market into a relative handful of “hits” and a mass of also-rans. This can be seen from several recent peak-season box-office experiences, in which four out of perhaps a dozen major releases have generated as much as 80% of total revenues.”

Appendix 4 – North American market context

Two main characteristics of the North American market are worth highlighting. The first concerns this market's longevity, paralleled by continuous shifts in dynamics. The long and often turbulent history of Hollywood is permeated with many examples. One, rather abrupt, example is the 1948 decree that dictated the end of the "studio system" (Vogel 2011). Before this decree industry operators were fully vertically integrated throughout the value chain (i.e. controlling production, distribution and exhibition), the landscape was composed of many single screened "palatial size" theaters and the majority of the talent (stars, directors, writers etc.) was under exclusive contracts with the major studios. In just a few years the landscape of the market had been completely transformed. Although other changes are bound to be more subtle, their influence cannot be denied. Technological advances, in particular, can be incredibly powerful (Vogel 2011). From Technicolor a century ago, through the advent of computers and digital filmmaking, the resurgence of 3D to the current effects of the internet, the shifts have been plenty and impactful (Lafrance 2015; Schedeen 2010; Levine 2011).

The second characteristic is this market's influence, as seen by the disparity between the penetration of Hollywood films in other markets and the penetration of foreign films in this market. North American films – i.e. produced or co-produced by the U.S. – unfailingly dominate the North American charts: the latest completely foreign-produced film to reach a yearly top 10 of highest-grossing films in North America was *Crocodile Dundee* (1982, Australia); more recently, only *The Crying Game* (1992, Japan and U.K.) and *Taken* (2009, France) have manage to appear in the top 20[88] (BoxOfficeMojo.com). At the same time, Hollywood films have kept a strong international presence. In France, for example, only four, seven and three films of the 2011, 2012 and 2013 yearly top 20s were non-Hollywood productions, respectively (BoxOfficeCine.fr). Even in the second biggest market in the world by revenues, China, the penetration is remarkable: four of the ten highest-grossing films before 2013 were Hollywood blockbusters (Tsui 2013), as were over half of the yearly top ten between 2009 and 2012 (BoxOfficeMojo.com 2009; Sun 2011, 2012, 2013). This second characteristic is, however, compounded by two factors. First, the global presence and partnerships of North American studios result in the majority of North American releases being considered, at least, U.S. co-productions. This reduces the number of completely foreign films and, importantly, this reduction occurs for many English-language films, which traditionally perform better in the U.S./Canada market than those in other languages (Vogel 2011). Second, a lack of consistency in the labeling of films results in an increased difficulty in determining a film's origin country. Due to the vague nature of the public information pertaining the deals between companies, different sources can state different details for the same film. *Rush* (2013) serves as an example: depending on the source it is considered a British film (Eaton 2013), a German film (TCM.com), a British-German co-production (Eaton 2013), a British-American co-production (American Film Institute) or even a three way co-production between the U.K., Germany and the United States (Lumiere).

[88] Both appearing in the twentieth position.

Appendix 5 – Portuguese market context

The Portuguese market reaches far more discreet values than many of its international counterparts. Films such as *Fast & Furious 6*, *Frozen*, *Despicable Me 2* and *The Croods*, all among the top 10 performers of the year in Portugal, obtained between 0.21% and 0.52% of their worldwide revenues in this market (ICA 2013). In fact, 41 *individual films* achieved, in North America alone, higher revenues than those obtained by *all films* released in the entire Portuguese market, in 2013. Still, these circumstances can, conceivably, provide some flexibility for Portuguese companies. Due to this market being considered less critical by larger companies of the global landscape, the definition of the relevant strategic decisions – such as deciding a more favorable release date or implementing specific marketing efforts – could be achieved beneficially.

One aspect that differs from the North American market is the irregularity in the aggregate yearly box office revenues, with both fast growth and sharp declines occurring. Following a decrease (6.8%) between 2004 and 2005, the market entered a growth period until 2010: first steadily (around and below 3%) and then explosively (5.6% and 11.4% jumps). Since 2010, the best year in record, the market has been falling increasingly – 2.8%, 7.5% and 11.4% - towards the € 65.5 million of 2013, the lowest value recorded (ICA 2008, 2009, 2010, 2011, 2012, 2013).[89] Furthermore, recently released provisional figures from ICA indicate that 2014 continued this downward trend with a € 62.7 million box office and 12.1 million in admissions, both establishing new lows for the respective categories.

Parallel to this, North American productions have dominated this market by generating over 90% of revenues (and admissions) on a 59% market share of number of releases between 2009 and 2012.[90] Although 2013 saw a decrease in these numbers (79.5% in revenues, 78.6% in admissions and 55.2% in number of releases), provisional figures for 2014 show a recovery in both revenues and admissions (86.2% and 85.9%, respectively) with the market share still declining (51.8%).Portuguese productions, on the other hand, reach significantly smaller values with the best recent year being 2012 where these films aggregately reached 4.9% of the total revenues of all films. Looking at Portugal's top performing domestic productions paints the same picture. Not one of the top 40 most seen films in the period between 2004 and 2014 was produced or even co-produced in Portugal.[91]

---

[89] It's worth noting that the first consistently collected data is from 2004 (ICA 2009).

[90] Values from the previously mentioned ICA reports.

[91] During this period ICA reports a grand total of 230 feature-length (which excludes short-films and serials) films produced in Portugal.

Appendix 6 – Variables for the producer's problem

Hypothesis variable

Sequel. Dummy variable that assumes a value of 1 for sequel films and 0 for non-sequels. Both IMDb.com and BoxOfficeMojo.com served as sources to distinguish the films that are part of a series.[92] For this thesis' purposes, every film that, on those sources, is established as being a part of a series as a sequel, prequel, reboot or spin-off is included here as a sequel film. Remakes are not included in this variable since the ties between a remake and the film it's remaking are much looser than those in the other categories. Shared universes remain a rather novel concept in film, with few examples and study, so the option fell to exclusion – therefore entering the sample as non-sequels.

Strength variables

Revenue of Parent. The revenues obtained by the parent film – i.e. the first in the series– in the North American market, in millions of dollars. While all non-sequel films in the sample have a value of 0 for this variable[93], sequel films can theoretically have any positive value. In practice, for sequels the interval observed in the sample falls between 21.73 and 403.36. The source used for this variable is BoxOfficeMojo.com with CPI adjustments made using the U.S. Bureau of Labor Statistics' CPI inflation calculator.[94] This CPI-adjustment to 2011 from the respective years is due to the existence of film series consisting of multiple films and the possibility that any sequel might be made several years after the parent film or a previous sequel.

Sequel Time Gap. The time gap (in years) between the releases of the parent film in the series and the target sequel, in the North American market. As is the case for the revenue of the parent film, this variable will assume the value zero for all the non-sequel films in the sample. Although it could theoretically range between 0 and 105,[95] in the sample this variable assumes values between 1 and 32. However, all but two sequels in the sample have a time gap of 15 years or less. All release dates were obtained from BoxOfficeMojo.com, with the subsequent calculation being made in Excel.

Number of Intervening Sequels. The number of intervening films in the series between the parent film and the target sequel. As with the other variables pertaining sequel's characteristics, non-sequel films were given a value of zero. For sequels it's worth noting that by counting the number of intervening films instead of the number of *previous* films in the series, the value attributed to the first sequel film – i.e. the second of the

---

92 Even though BoxOfficeMojo.com is an IMDb.com company, there are content differences.

93 Since they aren't part of a series, they do not have a parent film. In accordance with Basuroy and Chatterjee (2008).

94 http://www.bls.gov/data/inflation_calculator.htm

95 If the first and any additional films in a franchise were released in the same year or if *The Story of the Kelly Gang* (1906) had a 2011 sequel, respectively. While the second definitively didn't occur, the first possibility is still highly unlikely as producers tend to wait for at least some preliminary results or reaction by public and critics before starting production of a sequel. Excluding the parent film it is possible, and has occurred more than once recently (e.g. *The Matrix* second and third films, *Harry Potter* series final two films), that two films in a series have been release within a year.

series – is also zero. This is the case for eight sequels in this sample. The remaining sixteen sequels have between one and six previous sequels in their series, i.e. between three and eight films in the whole brand.[96] The film series' compositions needed for the construction of this variable were obtained from BoxOfficeMojo.com and IMDb.com.

Sequel x Week. Interaction term between the dummy variable for a sequel film and the number of each week since release. This variable is included to find if the revenues of sequel films have an increased decay when compared to non-sequel films, i.e. present a different behavior in their revenue generation. Should they behave in this manner, this variable's coefficient will be negative. The construction of this variable is simply made by interacting the preexisting sequel variable with the week variable created with the shaping of the sample into panel form.

Control variables

MPAA Ratings: *G, PG, PG-13 & R*. Four binary variables based on the ratings attributed by the MPAA. These ratings are grounded upon the content and how it's portrayed in a film – in particular, factors such as violence, sex, language and drug use – with the purpose of indicating its level of content and suitability for younger audiences. The four mutually exclusive variables divide films between G (all ages admitted), PG (some material not suited for children), PG-13 (some material may be inappropriate for children under 13) and R (under 17 requires accompanying parent or adult guardian), with unrated films as default (i.e. valued at zero in all four variables). The MPAA attributes a fifth, exceptionally rare, additional rating[97], NC-17 (no one 17 or under admitted). As in Basuroy and Chatterjee (2008), this variable is not included and the single NC-17 film, *Shame*, is the final elimination in the sample. The MPAA rating comes from the individual film pages in BoxOfficeMojo.com and is then transformed into its dummy state.

Award. Dummy variable representing films with Academy Award (Oscar) winning directors and/or actors.[98] A list of eight actors and the director or directors' names from each film was collected from IMDb.com and compared to a listing of all Oscar recipients. If a film contains at least one winner, the variable is valued at 1. Otherwise, it assumed the value 0. This variable is intended to provide a measure of a film's attractiveness in terms of personnel. Additionally, since Oscar winners were able to produce work of enough quality to warrant this recognition, it can be seen as a measure of ex-ante quality of the film. This argument is tied both

---

96 Although Basuroy and Chatterjee do not expand on the reasons for this specific construct, it's possible that their intention for this variable was to discern only the effect of further sequels allowing the dummy variable sequel to absorb part of the impact of the first sequel. This is in agreement with their expectations that the variables sequel and number of intervening sequels should have a negative and a negative coefficient, respectively.

97 Between 2002 and 2011, only 10 NC-17 films were released. For any given year in this period, the combined grosses of all films with this rating never exceed 1.1% of the yearly total box office. In six out of those ten year not a single NC-17 film was released.

98 Although Basuroy and Chatterjee (2008) don't clarify the specific categories considered, here the option is to include only Best Actor, Best Actress and Best Director.

to critics' reviews and budget. Of the 208 films in the sample, 45 contained at least one Oscar winner, approximately 21.6%.

Budget. Production cost of the film in millions. These values were primarily retrieved from the individual film pages in BoxOfficeMojo.com with further info obtained from Nash Information Services (TheNumbers.com). In-sample values range from $0.135 million to $250 million, with an average value of $39.7 million. The comparison between this mean and the median value of $25 million, displays the larger concentration on smaller budget values that can be found in the sample. This variable can be seen as proxy of production quality. However, inflated salaries or mismanagement of the budget can lead to a less accurate proxy.

Critics' Reviews: *Total Number* & *Positive Ratio.* Two variables intend to assess the critical opinion of a film. Using Metacritic.com, the number of critics' reviews, divided between positive, mixed and negative, were collected. This website aggregates in its "Critic Reviews" section a body of reviews that were published on major North American publications (e.g. *Empire*, *Variety* or *The New Yorker)* and recognized well-respected critics (e.g. Roger Ebert). The sum of these three values led to the variable Total Number which functions as a measure of popularity and availability of the film within critics' circles. Then, using this variable to compose a ratio between the number of positive reviews and the total number of reviews, the variable Positive Ratio was created. The total number of reviews varies from 0, not a single high-profile review published, to 41 reviews.

One point of difference from this variable to its counterpart in Basuroy and Chatterjee (2008) is that the collected reviews were published from the start of the previous week to the end of the opening weekend – i.e. from the Friday of the week before to the Sunday of the release week, both inclusive – while in their paper the timing is not entirely clear.[99] Even despite this potential extension of the timing window, both the average and the standard deviation of the number of total reviews were higher in their paper. If this increased number was due to the inclusion of less influential critics it's possible that some distortion might appear.

Seasonality. Coefficient that measures the attractiveness of the specific week when a film is released. From Vogel's (2011) graph of normalized weekly attendance, the week of release was matched with the corresponding coefficient value. This coefficient varies between 0.36 for early December and 1 for the maximum peak period to be found on Christmas. As discussed before, the inclusion of this variable is intended to account for variations in seasonality throughout the year, improving the comparability between films by extracting this specific factor.

Number of Theaters. Weekly varying variable of the number of theaters where a film is being exhibited. This is the only explanatory variable that varies across films and across weeks. These values come from BoxOfficeMojo.com where, alongside the revenue numbers collected as dependent variable, the

99 They mention only that reviews were collected from Variety.

numbers theaters in which a given film was played are reported for the entire length of run. This variable differs from Basuroy and Chatterjee (2008). There, the variable used to indicate the dimension of release/availability of a film is the number of screens. Since these values are not reported by BoxOfficeMojo.com, the number of theaters is included here as a substitute. It's possible that the number of screens might give a closer measure of availability of a specific film. Still, the numbers obtained in the sample for theaters do not differ greatly from multiple examples in the literature concerning the number of screens. Nelson and Glotfelty (2012), for example, use a 1999-2005 sample in which the number of screens varies from 999 to 4223, which is close to the values in this sample.[100] Likewise, Hennig-Thurau, Houston and Walsh (2006) report that their 1999-2001 sample has, for the number of screens in the first week, an average of 1527 and a standard deviation of 1134. Seeing these studies present similar values to the ones reported here gives credibility to the notion that this difference should not significantly impact the estimation results.

Number of Theaters Squared. An additional variable associated with this availability of films directly constructed by squaring the Number of Theaters variable. The reasoning behind the inclusion of this variable is that there might be diminishing returns associated with greater releases.

On the timing of decisions

All variables related to sequels (disregarding the interaction term)[101] are defined in production. Furthermore, the producer is also responsible for deciding the cast and the value he is willing invest (or raise from external investors) into a film, which means that the award and budget variables are largely under his supervision. The rating, in its fundamentals, is also decided in production as the film is shot and then edited. However, there have been instances where a studio is adamant about having a different rating (usually transforming an R into a PG-13) and so a reediting occurs in order for this variable to be defined. The border between producer and distributor in a major studio is very blurry, and this reediting for rating is rather rare from other types of companies. Still, these variables can be defined in both stages. While the seasonality coefficient is clearly a consequence of the date of release that is decided by the distributor, the variables related to critics' evaluations are much less obvious. At their core they are a product of the critics themselves, however, the film under evaluation is made by the producer and the decision to have showings for critics is made by distributors. Both of these can affect these variables' values. Finally, the number of theaters is typically a product of negotiation between distributors and exhibitors.

---

[100] They study only the top50 for each year, which explains the rather high minimum value (900).

[101] It's worth noting that despite the revenue of the parent not being decided by anyone but the audiences, the decision to produce a sequel from that specific parent is still made by producers.

Appendix 7 – Variables for the distributor's problem

Hypothesis variable

Release Lag. The time, in weeks, between the North American and Portuguese releases of a film. While theoretically this variable could vary between large negative and large positive numbers, there is a tendency for this number to be positive.[102] In practice, this variable ranges from -10 to +94 weeks with a large concentration of films between 0 and 5 week lags, with a mode of 2 and a median of 5. The variable is a result of the difference between the release dates, with the values for the Portuguese market coming from ICA's database.

Control variables

Sequel. Equal to producer's problem variable. Fourteen sequel films are present in this sample.

***Revenue of Parent*. ** Equivalent of the producer's problem variable. The revenues obtained by the parent film are now from the Portuguese market, and appear in thousands of euros. The CPI adjustment is made for the Portuguese economy by using the calculator available on the website of *Instituto Nacional de Estatística* (INE).

***Sequel Time Gap*. ** Equivalent of the producer's problem variable. The only difference between this variable and its producer's problem counterpart is the fact that the release dates used here are from the Portuguese market, instead of the North American. The range of this variable is considerably diminished with the exclusion of part of the sequel films. From a maximum value of 32, the highest sequel time gap for the distributor's problem is less than half, 15.

***Number of Intervening Sequels*. ** Equal to producer's problem variable.

***Sequel x Week*. ** Equal to producer's problem variable.

IGAC Ratings: *M/6, M/12 & M/16.* The parental guidance ratings for the Portuguese market are issued by *Inspecção Geral das Actividades Culturais* (IGAC), which classifies films as M/4, M/6, M/12, M/16 and M/18.[103] The numbers in the names signify the minimum age required to watch a particular film. Allowing for one year differences, these classifications are close to the ones used by the MPAA.[104] The biggest difference from the producer's problem lies on the fact that films cannot be released in the Portuguese market without having been attributed an IGAC rating. In practical terms, this means that the model setup must be changed

---

102 As discussed in chapter 2, Theory. In the sample, only 6 out of the 139 films (4.3%) have negative lags.

103 In addition, all these ratings can have a quality seal attached, for example: "M/4 – Qualidade". The advantage in having the quality seal is that distributors are exempt of paying a EUR 150 fee on films that the IGAC deems deserving, which is usually reserved for time-tested classics' re-releases. None of the films present in the sample had a quality seal.

104 M/4 = G, M/6 = PG, M/12 = PG-13, M/16 = R, M/18 = NC-17.

as Unrated cannot remain the default rating. The upshot is that the number of rating variables is reduced to three – M/6, M/12 and M/16 – with M/4, the classification with the least number of films, becoming now the default.[105]

***Award*.** Equal to producer's problem variable. The sample contains 41 films (29.5%) that feature Oscar winning personnel.

***Budget*.** Equal to producer's problem variable. Since the production costs do not cover any marketing costs, this variable remains unaltered from the producer's problem.

Critics' Reviews*: Total Number* & *Positive Ratio*. Equal to producer's problem variable. The usage of the exact same variables from the North American market was motivated by a lack of consistent review production and aggregation in the Portuguese markets. While there are examples of publications with regular publication of reviews, their number is too limited to be of particular use here. The positive ratio, in particular, was found to be almost exclusively either 0 or 1, as the sole review found for many films decided by itself this variable's value. Since broadening the spectrum of the variables to include other review sources, for example blogs, would lead to further complications stemming from the reach (or lack thereof) of these same sources, the option turned to the inclusion of the exact same variables used in the producer's problem. As a consequence, the variable's meaning has to be adapted to this situation. In this sense it's interesting to look at the positive ratio as a measure of film quality – biased as it may be. The total number, on the other hand, is expected to be less straightforward as it's more of a measure of critical buzz or attentiveness. Adopting these perspectives, both variables are expected to have a positive impact on revenues.

Seasonality. Equivalent of the producer's problem variable. The new release dates were matched against the Vogel (2011) graph with normalized weekly revenues data and attributed the coefficient that effectively varies between 0.36 and 1. Despite the fact that the source graph was constructed for the North American market it's expected that it provides a reasonable approximation of the Portuguese market behavior, as ICA reports show the existence of similar peak periods throughout the year (Easter, summer and Christmas, most notably).

Number of Showings. The total number of showings[106] of a film in the given week, from the first up to the fifteenth. This variable comes from the ICA database and is used here as a close substitute of the number of theaters as a measure of availability. Due to the structure of the data collection methodology by ICA and the resulting database supplied, the number of screens or theaters are not reported on a film by film basis. In some ways, as a measure of availability, the number of showings is an improvement over the number of

---

[105] The M/18 was not used in the sample. The only film that was given this rating, *Shame*¸ was dropped from the sample in the producer's problem for this very reason.

[106] A showing is the individual exhibition of a film in a film theatre. Most theater venues have multiple showings per day per film.

screens or theaters. That is, it shows exactly how many times the film was available for viewing during a given week. In the same manner that most theaters can have more than one screen showing a film, so can a screen have more or less showings. For example, if a particular screen only has one showing per week its contribution to availability is below that of a screen with three daily showings. The use of the number of showings improves on this account. On the other hand, if the showings are largely concentrated on a small number of venues, a large number of showings might incorrectly indicate a film as widely available when in fact it is very geographically limited.[107] Overall it's expected that the resulting impact comes out a balance between the two possibilities.

Number of Showings Squared. Similar construct to the producer's problem squared variable. Presents a measure of diminishing returns in the scheduling of too many showings.

[107] The same argument can be used for the screens-theaters dichotomy, for that matter. It might, however, be less likely for the multiple screens on a single theatre to occur, than multiple showings per theatre.

Appendix 8 – Producer's problem correlation table

	Sequel	Revenue of Parent	Sequel Time Gap	Number of Int. Sequels	G	PG	PG-13	R	Unrated
Sequel	1.00								
Rev. of Parent	0.86	1.00							
Sequel Time Gap	0.76	0.60	1.00						
Number of Intervening Sequels	0.68	0.73	0.72	1.00					
G	0.12	0.15	0.02	0.00	1.00				
PG	0.22	0.14	0.17	0.10	- 0.07	1.00			
PG-13	- 0.06	0.03	- 0.05	0.06	- 0.14	- 0.32	1.00		
R	- 0.10	- 0.15	- 0.05	- 0.10	- 0.14	- 0.31	- 0.66	1.00	
Unrated	- 0.08	- 0.07	- 0.06	- 0.05	- 0.04	- 0.08	- 0.18	- 0.17	1.00
Award	- 0.08	- 0.01	- 0.10	- 0.08	0.12	- 0.03	0.06	- 0.03	- 0.11
Budget	0.42	0.56	0.26	0.33	0.12	0.25	0.18	- 0.32	- 0.17
Total Number	0.12	0.18	0.12	0.14	0.01	0.07	0.14	- 0.08	- 0.27
Positive Ratio	- 0.01	0.04	0.06	0.15	- 0.03	- 0.01	- 0.10	0.10	0.05
Seasonality	- 0.04	0.02	- 0.05	0.01	0.00	0.06	0.04	- 0.08	0.01
Number of Theaters – W1	0.39	0.39	0.28	0.30	0.15	0.27	0.17	- 0.30	- 0.26
Max. Theaters	0.38	0.39	0.28	0.30	0.14	0.27	0.19	- 0.31	- 0.28
Weekly Rev. – W1	0.53	0.69	0.28	0.55	0.13	0.05	0.21	- 0.24	- 0.14
Total Rev.	0.52	0.70	0.29	0.51	0.17	0.10	0.20	- 0.27	- 0.15

	Award	Budget	Total Number	Positive Ratio	Season.	Number of Theaters – W1	Max. Theaters	Weekly Rev. – W1	Total Rev.
Award	1.00								
Budget	0.15	1.00							
Total Number	0.20	0.46	1.00						
Positive Ratio	- 0.03	- 0.02	0.24	1.00					
Seasonality	- 0.01	0.08	0.02	0.02	1.00				
Number of Theaters – W1	0.09	0.66	0.64	- 0.16	0.01	1.00			
Max. Theaters	0.11	0.68	0.66	- 0.12	0.02	0.99	1.00		
Weekly Rev. – W1	0.07	0.71	0.47	0.04	0.13	0.65	0.66	1.00	
Total Rev.	0.12	0.75	0.51	0.11	0.12	0.68	0.69	0.95	1.00

Appendix 9 – Distributor's problem correlation table

	Release Lag	Sequel	Revenue of Parent	Sequel Time Gap	Number of Int. Sequels	M/4	M/6	M/12	M/16
Release Lag	1.00								
Sequel	- 0.17	1.00							
Rev. of Parent	- 0.17	0.88	1.00						
Sequel Time Gap	- 0.15	0.84	0.80	1.00					
Number of Intervening Sequels	- 0.13	0.67	0.53	0.82	1.00				
M/4	- 0.13	0.11	0.03	0.05	0.06	1.00			
M/6	- 0.14	0.30	0.36	0.23	0.04	- 0.06	1.00		
M/12	- 0.10	- 0.20	- 0.16	- 0.11	- 0.03	- 0.17	- 0.45	1.00	
M/16	0.26	- 0.04	- 0.10	- 0.06	- 0.01	- 0.09	- 0.24	- 0.71	1.00
Award	- 0.01	- 0.06	- 0.01	- 0.10	- 0.12	0.01	0.03	0.08	- 0.11
Budget	- 0.36	0.43	0.45	0.39	0.31	- 0.01	0.35	0.01	- 0.27
Total Number	- 0.31	0.03	0.03	0.01	0.03	- 0.09	- 0.01	0.21	- 0.19
Positive Ratio	0.14	- 0.09	- 0.02	0.01	0.03	- 0.01	- 0.09	0.11	- 0.06
Seasonality	0.06	- 0.15	- 0.14	- 0.17	- 0.15	- 0.09	- 0.08	0.21	- 0.14
Number of Showings – W1	- 0.49	0.40	0.39	0.30	0.30	0.04	0.33	0.01	- 0.27
Weekly Rev. – W1	- 0.34	0.49	0.47	0.36	0.36	0.04	0.40	- 0.07	- 0.24
Total Rev.	-0.33	0.54	0.58	0.44	0.37	0.10	0.44	- 0.10	- 0.25

	Award	Budget	Total Number	Positive Ratio	Seasonality	Number of Showings – W1	Weekly Rev. – W1	Total Rev.
Award	1.00							
Budget	0.17	1.00						
Total Number	0.15	0.37	1.00					
Positive Ratio	- 0.05	- 0.13	0.10	1.00				
Seasonality	0.03	0.03	0.02	- 0.06	1.00			
Number of Showings – Week 1	0.11	0.76	0.44	- 0.15	0.02	1.00		
Weekly Rev. – W1	0.12	0.70	0.28	- 0.04	0.04	0.85	1.00	
Total Rev.	0.12	0.62	0.20	- 0.03	0.00	0.76	0.94	1.00

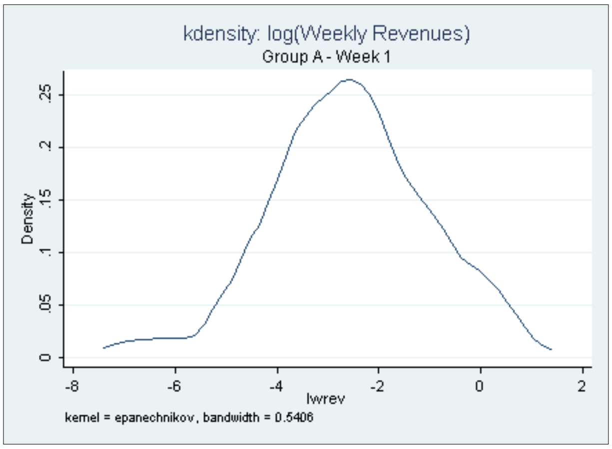
Appendix 10 – Random Effects estimation of producer’s problem (Stata)

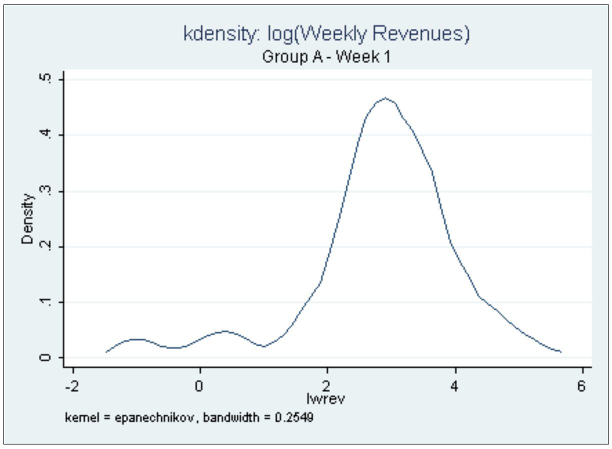
```
Random-effects GLS regression                   Number of obs      =      2329
Group variable: id                              Number of groups   =       207

R-sq:  within  = 0.7827                         Obs per group: min =         2
       between = 0.8320                                        avg =      11.3
       overall = 0.8012                                        max =        15

                                                Wald chi2(16)      =  12682.72
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                    (Std. Err. adjusted for 207 clusters in id)
```

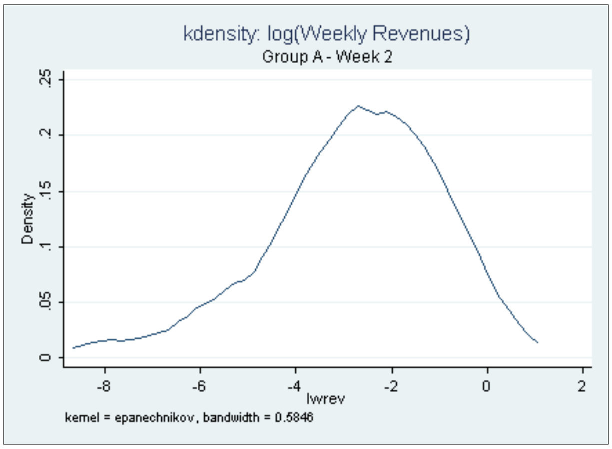
lwrev	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
sequel	1.20404	.1999536	6.02	0.000	.812138	1.595942
revparent	-.0011623	.0009631	-1.21	0.228	-.00305	.0007254
gapseq	-.024168	.0061058	-3.96	0.000	-.0361351	-.0122009
numseq	-.0473105	.0610167	-0.78	0.438	-.166901	.0722799
seqweek	-.0555337	.0114623	-4.84	0.000	-.0779994	-.033068
mpaa_g	1.635742	.3947721	4.14	0.000	.8620028	2.409481
mpaa_pg	1.448669	.3704416	3.91	0.000	.7226173	2.174722
mpaa_pg13	1.568031	.3741945	4.19	0.000	.8346237	2.301439
mpaa_r	.9684837	.3719342	2.60	0.009	.2395062	1.697461
award	.1976657	.1337491	1.48	0.139	-.0644777	.459809
budget	.0036154	.0012022	3.01	0.003	.0012591	.0059716
rmeta_pr	1.163538	.2294065	5.07	0.000	.7139094	1.613166
rmeta_tn	.0216631	.0075603	2.87	0.004	.0068451	.036481
season	.1808488	.4589008	0.39	0.694	-.7185802	1.080278
theaters	.0030421	.0001052	28.92	0.000	.002836	.0032483
theaterssq	-4.65e-07	3.02e-08	-15.41	0.000	-5.25e-07	-4.06e-07
_cons	-5.337559	.4627313	-11.53	0.000	-6.244496	-4.430622
sigma_u	.68865874					
sigma_e	.83190885					
rho	.40662045	(fraction of variance due to u_i)				

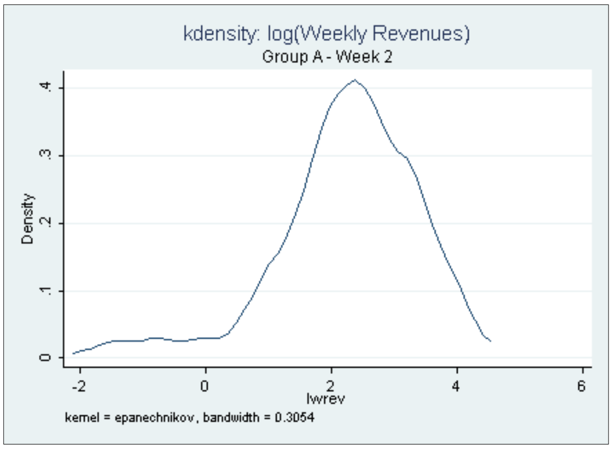
Appendix 11 – kDensity of dependent variable, by Group

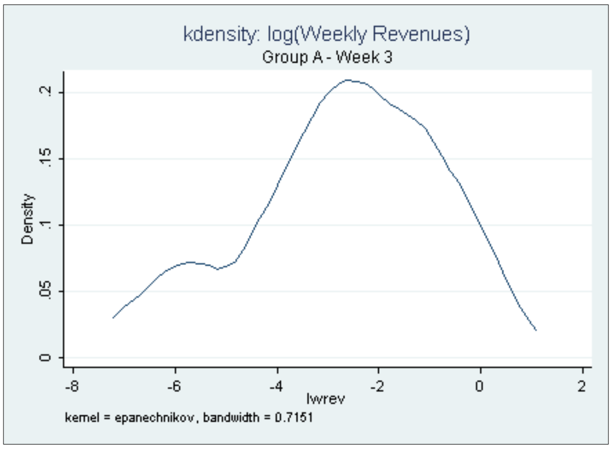
Group A	**Group B**

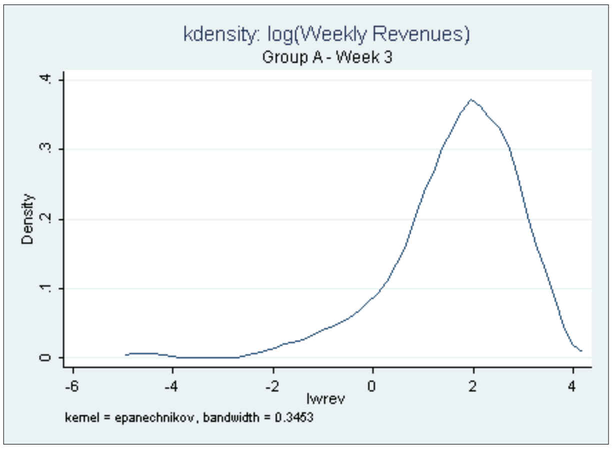


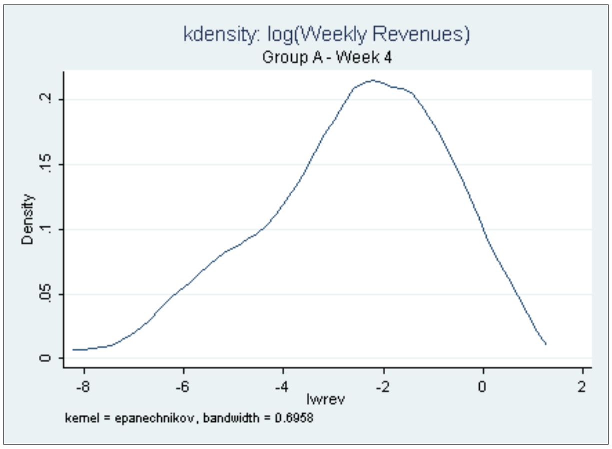


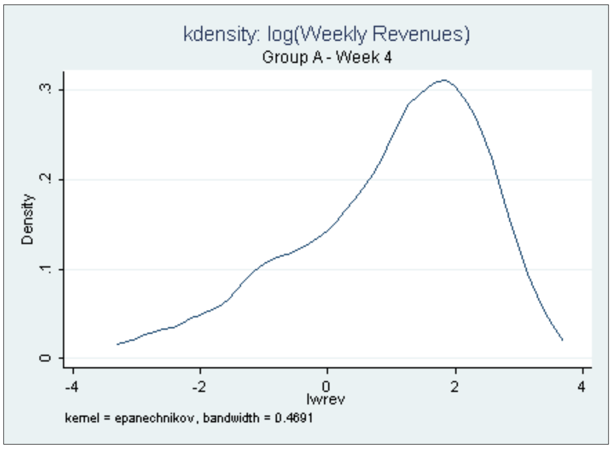












Appendix 12 – Descriptive statistics, by Group

	Group A (MaxTheaters < 438)				**Group B** (MaxTheaters ≥ 438)			
	Number of films: 69				**Number of films:** 138			
Variable	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Sequel	–	–	–	–	0.17	0.38	0	1
Revenue of Parent [a]	–	–	–	–	33.44	86.15	0	403.36
Sequel Time Gap	–	–	–	–	1.61	4.71	0	32
Number of Intervening Sequels	–	–	–	–	0.28	0.90	0	6
G	–	–	–	–	0.04	0.20	0	1
PG	0.01	0.12	0	1	0.19	0.39	0	1
PG-13	0.26	0.44	0	1	0.49	0.50	0	1
R	0.59	0.49	0	1	0.28	0.45	0	1
Unrated	0.13	0.34	0	1	–	–	–	–
Award	0.13	0.40	0	1	0.26	0.44	0	1
Budget [a]	7.75	8.38	0.14	36	55.68	48.41	1.5	250
Total Number	14.46	7.59	0	36	28.27	8.33	0	41
Positive Ratio	0.51	0.33	0	1	0.42	0.29	0	1
Seasonality	0.64	0.13	0.47	1	0.62	0.13	0.36	1
Number of Theaters in the first week	27.86	62.16	1	352	2767.51	929.98	4	4375
Weekly revenue in the first week [a]	0.22	0.38	0.001	2.41	28.87	33.06	0.29	226.12
Total Revenues	1.65	2.46	0.002	13.30	67.42	66.83	1.19	381.01
R.O.B. (Revenues/Budget)	1.04	2.78	0.001	13.58	1.93	3.80	0.06	36.01

[a] in millions of dollars

Appendix 13 – Estimation results (Stata), by Group

Group A:

```
GEE population-averaged model                   Number of obs      =       710
Group variable:                         id      Number of groups   =        69
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =      10.3
Correlation:                  exchangeable                     max =        15
                                                Wald chi2(10)      =    344.61
Scale parameter:                  1.348206      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
```

lwrev	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
mpaa_pg	.6458331	.2895232	2.23	0.026	.0783779	1.213288
mpaa_pg13	.3394466	.3494424	0.97	0.331	-.3454479	1.024341
mpaa_r	.2485769	.3078016	0.81	0.419	-.3547032	.851857
award	.009048	.1739846	0.05	0.959	-.3319555	.3500516
budget	.0174599	.0084744	2.06	0.039	.0008505	.0340694
rmeta_pr	1.547791	.367658	4.21	0.000	.8271941	2.268387
rmeta_tn	-.0111449	.016015	-0.70	0.486	-.0425337	.0202439
season	.7974575	.6705864	1.19	0.234	-.5168676	2.111783
theaters	.031094	.0037738	8.24	0.000	.0236974	.0384906
theaterssq	-.000066	.0000114	-5.79	0.000	-.0000884	-.0000437
_cons	-5.899336	.5970328	-9.88	0.000	-7.069499	-4.729173

Group B:

```
GEE population-averaged model                   Number of obs      =      1619
Group variable:                         id      Number of groups   =       138
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =      11.7
Correlation:                  exchangeable                     max =        15
                                                Wald chi2(15)      =  11328.99
Scale parameter:                    .55028      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
```

lwrev	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
sequel	.4942337	.1708996	2.89	0.004	.1592766	.8291909
revparent	.0016027	.0007187	2.23	0.026	.0001941	.0030113
gapseq	-.0174815	.0074846	-2.34	0.020	-.032151	-.002812
numseq	-.0519419	.0423986	-1.23	0.221	-.1350416	.0311578
seqweek	-.0564441	.0117951	-4.79	0.000	-.0795622	-.0333261
mpaa_g	.0727418	.1721155	0.42	0.673	-.2645984	.4100819
mpaa_pg	.0314985	.1186884	0.27	0.791	-.2011265	.2641235
mpaa_pg13	.1616268	.1007136	1.60	0.109	-.0357681	.3590218
award	.1896434	.0897517	2.11	0.035	.0137334	.3655534
budget	.0021802	.0009413	2.32	0.021	.0003352	.0040251
rmeta_pr	.9851781	.207414	4.75	0.000	.5786542	1.391702
rmeta_tn	-.0070938	.0075086	-0.94	0.345	-.0218103	.0076227
season	.222663	.3149356	0.71	0.480	-.3945994	.8399253
theaters	.0028508	.0000977	29.18	0.000	.0026593	.0030423
theaterssq	-4.13e-07	2.72e-08	-15.18	0.000	-4.67e-07	-3.60e-07
_cons	-2.862867	.2589722	-11.05	0.000	-3.370443	-2.355291

Appendix 14 – Random Effects estimation of distributor's problem (Stata)

```
Random-effects GLS regression                   Number of obs      =      1503
Group variable: id                              Number of groups   =       139

R-sq:  within  = 0.7515                         Obs per group: min =         2
       between = 0.7297                                        avg =      10.8
       overall = 0.7535                                        max =        15

                                                Wald chi2(16)      =   4442.74
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                 (Std. Err. adjusted for 139 clusters in id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0117105	.0046072	-2.54	0.011	-.0207405	-.0026805
sequel	1.337777	.4104733	3.26	0.001	.533264	2.14229
revparente	8.52e-06	.0002218	0.04	0.969	-.0004262	.0004433
intseq	-.1026946	.0360239	-2.85	0.004	-.1733002	-.0320891
numseq	.1272256	.1132766	1.12	0.261	-.0947925	.3492437
seqweek	-.0273956	.042817	-0.64	0.522	-.1113153	.0565241
igac_m6	.5023221	.3850842	1.30	0.192	-.252429	1.257073
igac_m12	.1298349	.3751125	0.35	0.729	-.6053721	.865042
igac_m16	.1087702	.3777983	0.29	0.773	-.6317008	.8492412
award	.094986	.08556	1.11	0.267	-.0727085	.2626806
orcamento	-.0037564	.0010446	-3.60	0.000	-.0058038	-.001709
rmeta_pr	.7648146	.1563772	4.89	0.000	.4583209	1.071308
rmeta_tn	-.0137904	.0075029	-1.84	0.066	-.0284958	.000915
season	.1209762	.371172	0.33	0.744	-.6065076	.84846
ss	.0079538	.0004024	19.77	0.000	.0071651	.0087424
ss2	-2.38e-06	2.81e-07	-8.46	0.000	-2.93e-06	-1.83e-06
_cons	-.564107	.4438684	-1.27	0.204	-1.434073	.3058591
sigma_u	.41212763					
sigma_e	1.2500895					
rho	.09803292	(fraction of variance due to u_i)				

Appendix 15 – Distributor's problem results (Stata), Total Number removed

GEE population-averaged model		Number of obs =	1503
Group variable:	id	Number of groups =	139
Link:	identity	Obs per group: min =	2
Family:	Gaussian	avg =	10.8
Correlation:	exchangeable	max =	15
		Wald chi2(15) =	4429.36
Scale parameter:	1.703095	Prob > chi2 =	0.0000

(Std. Err. adjusted for clustering on id)

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0096779	.0045576	-2.12	0.034	-.0186106	-.0007452
sequel	1.299214	.3977438	3.27	0.001	.5196504	2.078777
revparente	.0000327	.0002163	0.15	0.880	-.0003913	.0004567
intseq	-.0959515	.0342028	-2.81	0.005	-.1629877	-.0289154
numseq	.1325086	.1118688	1.18	0.236	-.0867502	.3517674
seqweek	-.0270596	.0425345	-0.64	0.525	-.1104257	.0563065
igac_m6	.4473136	.375572	1.19	0.234	-.2887941	1.183421
igac_m12	.015007	.3584371	0.04	0.967	-.6875169	.7175309
igac_m16	.0110926	.3637593	0.03	0.976	-.7018625	.7240478
award	.0806446	.0907431	0.89	0.374	-.0972086	.2584977
orcamento	-.0046939	.0010494	-4.47	0.000	-.0067507	-.002637
rmeta_pr	.7156306	.159147	4.50	0.000	.4037083	1.027553
season	.1186461	.3605859	0.33	0.742	-.5880893	.8253815
ss	.0079421	.0003989	19.91	0.000	.0071602	.0087239
ss2	-2.37e-06	2.79e-07	-8.51	0.000	-2.92e-06	-1.83e-06
_cons	-.7864434	.4094253	-1.92	0.055	-1.588902	.0160154

Appendix 16 – Distributor's problem results (Stata), all films with weekly capping (4 to 11)

```
GEE population-averaged model                   Number of obs      =       545
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       3.9
Correlation:                  exchangeable                     max =         4
                                                Wald chi2(16)      =   2443.10
Scale parameter:                  .4650758      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0243151	.0030264	-8.03	0.000	-.0302467	-.0183835
sequel	1.296598	.3061705	4.23	0.000	.696515	1.896681
revparente	-.0002406	.0002006	-1.20	0.230	-.0006338	.0001526
intseq	-.0006089	.0334344	-0.02	0.985	-.0661391	.0649213
numseq	-.0545744	.1023819	-0.53	0.594	-.2552394	.1460905
seqweek	-.2478365	.1079967	-2.29	0.022	-.4595062	-.0361669
igac_m6	.4501867	.1575693	2.86	0.004	.1413567	.7590168
igac_m12	.1829219	.1480031	1.24	0.216	-.1071589	.4730027
igac_m16	-.0034932	.1620908	-0.02	0.983	-.3211854	.3141989
award	.2282791	.0711553	3.21	0.001	.0888173	.3677409
orcamento	-.0031924	.0008506	-3.75	0.000	-.0048597	-.0015252
rmeta_pr	.4100215	.1255067	3.27	0.001	.1640328	.6560101
rmeta_tn	.0026273	.0047827	0.55	0.583	-.0067466	.0120012
season	.0148549	.3763903	0.04	0.969	-.7228565	.7525663
ss	.0042702	.0002751	15.52	0.000	.0037311	.0048093
ss2	-9.86e-07	1.42e-07	-6.93	0.000	-1.27e-06	-7.07e-07
_cons	1.02853	.2678227	3.84	0.000	.5036069	1.553453

```
GEE population-averaged model                   Number of obs      =       670
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       4.8
Correlation:                  exchangeable                     max =         5
                                                Wald chi2(16)      =   2694.40
Scale parameter:                   .590104      Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0238664	.0033139	-7.20	0.000	-.0303616	-.0173712
sequel	1.278744	.3256998	3.93	0.000	.6403844	1.917104
revparente	-.0003015	.0001883	-1.60	0.109	-.0006705	.0000675
intseq	-.002738	.0317097	-0.09	0.931	-.0648879	.0594119
numseq	-.0352994	.0924562	-0.38	0.703	-.2165102	.1459114
seqweek	-.1947201	.0978049	-1.99	0.046	-.3864143	-.003026
igac_m6	.547438	.1776642	3.08	0.002	.1992226	.8956534
igac_m12	.1735379	.1743866	1.00	0.320	-.1682535	.5153293
igac_m16	.0406343	.1859635	0.22	0.827	-.3238475	.4051161
award	.1652735	.0743983	2.22	0.026	.0194556	.3110915
orcamento	-.0032581	.0008392	-3.88	0.000	-.0049028	-.0016134
rmeta_pr	.5393273	.1294637	4.17	0.000	.2855832	.7930714
rmeta_tn	.0038699	.0051559	0.75	0.453	-.0062354	.0139752
season	-.0255602	.3852206	-0.07	0.947	-.7805787	.7294584
ss	.0047499	.0002878	16.51	0.000	.004186	.0053139
ss2	-1.17e-06	1.58e-07	-7.40	0.000	-1.47e-06	-8.57e-07
_cons	.7160317	.280659	2.55	0.011	.1659502	1.266113

```
GEE population-averaged model                   Number of obs      =       786
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       5.7
Correlation:                  exchangeable                     max =         6
                                                Wald chi2(16)      =   2770.67
Scale parameter:                  .6932389      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0245765	.0044248	-5.55	0.000	-.033249	-.015904
sequel	1.302708	.3356116	3.88	0.000	.6449215	1.960495
revparente	-.0003069	.0001924	-1.60	0.111	-.000684	.0000702
intseq	-.0117705	.0318907	-0.37	0.712	-.0742751	.0507341
numseq	-.0347602	.0937969	-0.37	0.711	-.2185987	.1490783
seqweek	-.168172	.0843323	-1.99	0.046	-.3334603	-.0028836
igac_m6	.517997	.2382259	2.17	0.030	.0510828	.9849113
igac_m12	.1314998	.2383686	0.55	0.581	-.3356941	.5986937
igac_m16	.0083259	.2454878	0.03	0.973	-.4728212	.4894731
award	.1581834	.0815722	1.94	0.052	-.0016951	.3180619
orcamento	-.0027819	.0008991	-3.09	0.002	-.004544	-.0010197
rmeta_pr	.6433396	.137911	4.66	0.000	.373039	.9136401
rmeta_tn	.0053317	.005751	0.93	0.354	-.00594	.0166034
season	-.1284087	.3936033	-0.33	0.744	-.899857	.6430396
ss	.0050841	.000295	17.24	0.000	.004506	.0056623
ss2	-1.30e-06	1.69e-07	-7.71	0.000	-1.63e-06	-9.70e-07
_cons	.5431523	.3261619	1.67	0.096	-.0961132	1.182418

```
GEE population-averaged model                   Number of obs      =       891
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       6.4
Correlation:                  exchangeable                     max =         7
                                                Wald chi2(16)      =   3588.95
Scale parameter:                   .791999      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0226119	.0042265	-5.35	0.000	-.0308957	-.0143282
sequel	1.284332	.3543158	3.62	0.000	.5898856	1.978778
revparente	-.000319	.0001857	-1.72	0.086	-.0006829	.0000449
intseq	-.0289071	.0305848	-0.95	0.345	-.0888521	.031038
numseq	-.0308826	.0903302	-0.34	0.732	-.2079265	.1461613
seqweek	-.1075311	.0725499	-1.48	0.138	-.2497263	.034664
igac_m6	.5907708	.2687186	2.20	0.028	.0640919	1.11745
igac_m12	.238305	.269889	0.88	0.377	-.2906678	.7672777
igac_m16	.0910959	.2760059	0.33	0.741	-.4498658	.6320576
award	.1437885	.0836649	1.72	0.086	-.0201917	.3077687
orcamento	-.0023354	.0009099	-2.57	0.010	-.0041188	-.000552
rmeta_pr	.7229387	.1280456	5.65	0.000	.4719739	.9739036
rmeta_tn	.0019248	.0056474	0.34	0.733	-.0091438	.0129934
season	-.1954649	.3387829	-0.58	0.564	-.8594671	.4685373
ss	.0054823	.0002999	18.28	0.000	.0048946	.00607
ss2	-1.45e-06	1.79e-07	-8.09	0.000	-1.80e-06	-1.10e-06
_cons	.3215031	.3371231	0.95	0.340	-.3392461	.9822523

```
GEE population-averaged model                   Number of obs      =       998
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       7.2
Correlation:                  exchangeable                     max =         8
                                                Wald chi2(16)      =   4246.70
Scale parameter:                  .9409054      Prob > chi2        =    0.0000

                                       (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0198833	.0041384	-4.80	0.000	-.0279945	-.0117721
sequel	1.171342	.3662103	3.20	0.001	.4535825	1.889101
revparente	-.0003337	.0001884	-1.77	0.077	-.0007029	.0000356
intseq	-.0359708	.0309943	-1.16	0.246	-.0967186	.0247769
numseq	-.0310331	.0916879	-0.34	0.735	-.210738	.1486718
seqweek	-.0352183	.0640293	-0.55	0.582	-.1607134	.0902768
igac_m6	.6220779	.2735137	2.27	0.023	.0860008	1.158155
igac_m12	.2401785	.2742336	0.88	0.381	-.2973094	.7776664
igac_m16	.1112952	.2793688	0.40	0.690	-.4362576	.6588481
award	.1577406	.0822626	1.92	0.055	-.003491	.3189723
orcamento	-.0024316	.0009439	-2.58	0.010	-.0042817	-.0005816
rmeta_pr	.8121923	.1313299	6.18	0.000	.5547903	1.069594
rmeta_tn	.0001509	.0062067	0.02	0.981	-.012014	.0123158
season	-.0891068	.3102401	-0.29	0.774	-.6971663	.5189527
ss	.0059625	.0003144	18.97	0.000	.0053463	.0065786
ss2	-1.62e-06	1.94e-07	-8.34	0.000	-2.00e-06	-1.24e-06
_cons	-.0137383	.3459632	-0.04	0.968	-.6918137	.6643372

```
GEE population-averaged model                   Number of obs      =      1095
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       7.9
Correlation:                  exchangeable                     max =         9
                                                Wald chi2(16)      =   4444.95
Scale parameter:                  1.097867      Prob > chi2        =    0.0000

                                       (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0169862	.0041559	-4.09	0.000	-.0251315	-.0088409
sequel	1.125234	.3782617	2.97	0.003	.3838546	1.866613
revparente	-.000287	.000212	-1.35	0.176	-.0007026	.0001285
intseq	-.0377997	.0335053	-1.13	0.259	-.1034689	.0278696
numseq	-.0422159	.0988295	-0.43	0.669	-.2359183	.1514864
seqweek	-.0055733	.058818	-0.09	0.925	-.1208544	.1097079
igac_m6	.5511027	.3060077	1.80	0.072	-.0486615	1.150867
igac_m12	.2166816	.3049771	0.71	0.477	-.3810626	.8144258
igac_m16	.0875175	.3092104	0.28	0.777	-.5185237	.6935587
award	.1441999	.0840652	1.72	0.086	-.0205649	.3089647
orcamento	-.0028999	.0010638	-2.73	0.006	-.004985	-.0008148
rmeta_pr	.8967402	.1382625	6.49	0.000	.6257507	1.16773
rmeta_tn	-.0028721	.0068061	-0.42	0.673	-.0162118	.0104677
season	.0488526	.3071402	0.16	0.874	-.553131	.6508362
ss	.0064246	.0003323	19.33	0.000	.0057732	.0070759
ss2	-1.78e-06	2.11e-07	-8.43	0.000	-2.20e-06	-1.37e-06
_cons	-.2472613	.3742506	-0.66	0.509	-.980779	.4862563

```
GEE population-averaged model                   Number of obs      =      1181
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       8.5
Correlation:                  exchangeable                     max =        10
                                                Wald chi2(16)      =   3880.61
Scale parameter:                  1.218655      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
        lrec |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      rellag |  -.0160503   .0041389    -3.88   0.000    -.0241624   -.0079381
      sequel |   1.096015   .4007138     2.74   0.006     .3106306      1.8814
  revparente |  -.0001721   .0002181    -0.79   0.430    -.0005995    .0002553
      intseq |  -.0528826   .0343674    -1.54   0.124    -.1202415    .0144763
      numseq |  -.0249518   .0999056    -0.25   0.803    -.2207632    .1708596
     seqweek |   -.002476    .056222    -0.04   0.965    -.1126692    .1077172
     igac_m6 |   .4944587   .3340336     1.48   0.139    -.1602351    1.149152
    igac_m12 |   .1974191    .329654     0.60   0.549    -.4486909    .8435292
    igac_m16 |   .0872288   .3325972     0.26   0.793    -.5646496    .7391073
       award |   .1504643   .0853866     1.76   0.078    -.0168904     .317819
  orcamento  |  -.0028297    .001072    -2.64   0.008    -.0049307   -.0007286
    rmeta_pr |   .8944674   .1400969     6.38   0.000     .6198825    1.169052
    rmeta_tn |  -.0070927   .0068666    -1.03   0.302    -.0205509    .0063656
      season |    .083349    .328531     0.25   0.800      -.56056    .7272579
          ss |   .0067544    .000348    19.41   0.000     .0060723    .0074364
         ss2 |  -1.91e-06   2.27e-07    -8.44   0.000    -2.36e-06   -1.47e-06
       _cons |  -.2956987    .397177    -0.74   0.457    -1.074151     .482754
------------------------------------------------------------------------------
```

```
GEE population-averaged model                   Number of obs      =      1260
Group variable:                         id      Number of groups   =       139
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       9.1
Correlation:                  exchangeable                     max =        11
                                                Wald chi2(16)      =   4019.44
Scale parameter:                  1.368611      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |               Robust
        lrec |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      rellag |  -.0142551   .0043014    -3.31   0.001    -.0226856   -.0058245
      sequel |   .9741515   .3892395     2.50   0.012     .2112562    1.737047
  revparente |  -.0001424   .0002122    -0.67   0.502    -.0005584    .0002736
      intseq |  -.0571561   .0329161    -1.74   0.082    -.1216705    .0073583
      numseq |    .008229   .0893565     0.09   0.927    -.1669064    .1833645
     seqweek |    .023807   .0495476     0.48   0.631    -.0733045    .1209185
     igac_m6 |   .5726412   .3880755     1.48   0.140    -.1879729    1.333255
    igac_m12 |   .2571053   .3832194     0.67   0.502    -.4939911    1.008202
    igac_m16 |    .211574   .3848681     0.55   0.583    -.5427536    .9659017
       award |   .1497625   .0855121     1.75   0.080    -.0178382    .3173632
  orcamento  |   -.002823   .0010148    -2.78   0.005    -.0048119   -.0008341
    rmeta_pr |   .8383085   .1463232     5.73   0.000     .5515204    1.125097
    rmeta_tn |  -.0082002   .0069485    -1.18   0.238     -.021819    .0054187
      season |   .1550084   .3527619     0.44   0.660    -.5363922     .846409
          ss |   .0070625   .0003599    19.62   0.000     .0063571     .007768
         ss2 |  -2.02e-06   2.36e-07    -8.56   0.000    -2.48e-06   -1.56e-06
       _cons |  -.5245616   .4442767    -1.18   0.238    -1.395328    .3462048
------------------------------------------------------------------------------
```

Appendix 17 – Distributor's problem results (Stata), N.A. productions/co-productions with weekly capping (4 to 11)

```
GEE population-averaged model                   Number of obs      =       476
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       4.0
Correlation:                  exchangeable                     max =         4
                                                Wald chi2(16)      =   1844.36
Scale parameter:                  .3803393      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0122468	.0061269	-2.00	0.046	-.0242552	-.0002383
sequel	1.240592	.3069914	4.04	0.000	.6389002	1.842284
revparente	-.0002156	.0001855	-1.16	0.245	-.0005792	.000148
intseq	.0075261	.0304451	0.25	0.805	-.0521453	.0671974
numseq	-.0463798	.091246	-0.51	0.611	-.2252188	.1324592
seqweek	-.278378	.1049661	-2.65	0.008	-.4841078	-.0726482
igac_m6	.2942146	.121013	2.43	0.015	.0570334	.5313958
igac_m12	-.0207787	.1077347	-0.19	0.847	-.2319348	.1903775
igac_m16	-.09776	.1248929	-0.78	0.434	-.3425457	.1470256
award	.200709	.0713914	2.81	0.005	.0607845	.3406335
orcamento	-.0024562	.0008003	-3.07	0.002	-.0040248	-.0008877
rmeta_pr	.2284495	.131414	1.74	0.082	-.0291173	.4860163
rmeta_tn	.0086197	.0051937	1.66	0.097	-.0015597	.0187992
season	-.2041631	.3922962	-0.52	0.603	-.9730495	.5647233
ss	.0042108	.0002816	14.95	0.000	.0036589	.0047627
ss2	-9.50e-07	1.43e-07	-6.64	0.000	-1.23e-06	-6.70e-07
_cons	1.12906	.2404553	4.70	0.000	.6577759	1.600343

```
GEE population-averaged model                   Number of obs      =       588
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       4.9
Correlation:                  exchangeable                     max =         5
                                                Wald chi2(16)      =   2275.29
Scale parameter:                  .5363832      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0142805	.0049343	-2.89	0.004	-.0239515	-.0046094
sequel	1.215187	.3351045	3.63	0.000	.5583947	1.87198
revparente	-.0002786	.0001716	-1.62	0.105	-.000615	.0000578
intseq	.0060126	.0286456	0.21	0.834	-.0501317	.0621569
numseq	-.0301934	.0866647	-0.35	0.728	-.2000531	.1396663
seqweek	-.2091764	.1018834	-2.05	0.040	-.4088643	-.0094886
igac_m6	.4106297	.1272041	3.23	0.001	.1613143	.6599451
igac_m12	-.0230082	.1200252	-0.19	0.848	-.2582532	.2122367
igac_m16	-.0667869	.1355025	-0.49	0.622	-.3323668	.198793
award	.1439027	.075686	1.90	0.057	-.0044392	.2922446
orcamento	-.0026597	.0008533	-3.12	0.002	-.0043321	-.0009874
rmeta_pr	.3402017	.1278152	2.66	0.008	.0896885	.5907149
rmeta_tn	.0117433	.0054575	2.15	0.031	.0010469	.0224397
season	-.2079943	.4068339	-0.51	0.609	-1.005374	.5893855
ss	.0046965	.0002948	15.93	0.000	.0041186	.0052743
ss2	-1.14e-06	1.58e-07	-7.19	0.000	-1.45e-06	-8.26e-07
_cons	.7598028	.2504053	3.03	0.002	.2690174	1.250588

```
GEE population-averaged model                   Number of obs      =       690
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       5.8
Correlation:                  exchangeable                     max =         6
                                                Wald chi2(16)      =   2592.20
Scale parameter:                  .6240838      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0139258	.0048103	-2.89	0.004	-.0233538	-.0044978
sequel	1.207149	.3382074	3.57	0.000	.5442747	1.870023
revparente	-.000272	.0001742	-1.56	0.118	-.0006134	.0000694
intseq	-.0021961	.0284381	-0.08	0.938	-.0579338	.0535416
numseq	-.0282106	.0904121	-0.31	0.755	-.2054151	.1489939
seqweek	-.1758783	.0884676	-1.99	0.047	-.3492716	-.002485
igac_m6	.3669956	.1687043	2.18	0.030	.0363413	.6976499
igac_m12	-.0858158	.1663216	-0.52	0.606	-.4118002	.2401686
igac_m16	-.1061917	.178151	-0.60	0.551	-.4553612	.2429778
award	.1406449	.08107	1.73	0.083	-.0182493	.2995391
orcamento	-.0021513	.0008859	-2.43	0.015	-.0038877	-.0004149
rmeta_pr	.4138352	.1322273	3.13	0.002	.1546744	.672996
rmeta_tn	.0141475	.0061293	2.31	0.021	.0021344	.0261607
season	-.3120673	.4144362	-0.75	0.451	-1.124347	.5002127
ss	.0050063	.0003007	16.65	0.000	.0044169	.0055958
ss2	-1.26e-06	1.68e-07	-7.50	0.000	-1.59e-06	-9.32e-07
_cons	.5918796	.2794856	2.12	0.034	.0440979	1.139661

```
GEE population-averaged model                   Number of obs      =       783
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       6.5
Correlation:                  exchangeable                     max =         7
                                                Wald chi2(16)      =   3615.29
Scale parameter:                  .7238536      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0105787	.0048599	-2.18	0.029	-.0201039	-.0010535
sequel	1.170327	.3474795	3.37	0.001	.4892797	1.851374
revparente	-.0002817	.0001641	-1.72	0.086	-.0006035	.00004
intseq	-.0191463	.0266504	-0.72	0.472	-.07138	.0330875
numseq	-.0263538	.0844524	-0.31	0.755	-.1918774	.1391698
seqweek	-.1096291	.0766132	-1.43	0.152	-.2597883	.0405301
igac_m6	.4279713	.1878537	2.28	0.023	.0597848	.7961577
igac_m12	.0123628	.187431	0.07	0.947	-.3549951	.3797208
igac_m16	-.0356139	.197434	-0.18	0.857	-.4225774	.3513496
award	.1208803	.0842735	1.43	0.151	-.0442927	.2860533
orcamento	-.0016342	.0008946	-1.83	0.068	-.0033875	.0001192
rmeta_pr	.4845223	.1208757	4.01	0.000	.2476102	.7214343
rmeta_tn	.0107416	.0057789	1.86	0.063	-.0005849	.0220681
season	-.3684665	.3522564	-1.05	0.296	-1.058876	.3219433
ss	.0054001	.0003036	17.79	0.000	.0048052	.0059951
ss2	-1.41e-06	1.78e-07	-7.89	0.000	-1.76e-06	-1.06e-06
_cons	.366777	.2764101	1.33	0.185	-.1749769	.9085309

```
GEE population-averaged model                   Number of obs      =       879
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       7.3
Correlation:                  exchangeable                     max =         8
                                                Wald chi2(16)      =   4359.89
Scale parameter:                  .8965012      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0084042	.0054476	-1.54	0.123	-.0190812	.0022729
sequel	1.031083	.3599056	2.86	0.004	.3256813	1.736485
revparente	-.0002982	.0001682	-1.77	0.076	-.0006278	.0000315
intseq	-.0249541	.0274093	-0.91	0.363	-.0786754	.0287671
numseq	-.0194348	.0908297	-0.21	0.831	-.1974578	.1585881
seqweek	-.0328421	.0685377	-0.48	0.632	-.1671735	.1014893
igac_m6	.449474	.1895702	2.37	0.018	.0779233	.8210247
igac_m12	-.0070012	.1867741	-0.04	0.970	-.3730717	.3590693
igac_m16	-.0470986	.1953313	-0.24	0.809	-.429941	.3357438
award	.1431692	.0840448	1.70	0.088	-.0215557	.307894
orcamento	-.0017595	.0009411	-1.87	0.062	-.003604	.0000851
rmeta_pr	.5432471	.1234606	4.40	0.000	.3012687	.7852255
rmeta_tn	.0107855	.0063807	1.69	0.091	-.0017205	.0232914
season	-.191659	.3201539	-0.60	0.549	-.8191492	.4358312
ss	.0058959	.0003193	18.46	0.000	.0052701	.0065218
ss2	-1.58e-06	1.94e-07	-8.14	0.000	-1.96e-06	-1.20e-06
_cons	-.0445501	.282523	-0.16	0.875	-.5982851	.5091848

```
GEE population-averaged model                   Number of obs      =       967
Group variable:                         id      Number of groups   =       120
Link:                             identity      Obs per group: min =         2
Family:                           Gaussian                     avg =       8.1
Correlation:                  exchangeable                     max =         9
                                                Wald chi2(16)      =   4068.71
Scale parameter:                  1.077058      Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
```

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0045374	.006127	-0.74	0.459	-.016546	.0074713
sequel	.946591	.3615383	2.62	0.009	.2379889	1.655193
revparente	-.0002481	.000188	-1.32	0.187	-.0006166	.0001204
intseq	-.02719	.0292425	-0.93	0.352	-.0845043	.0301242
numseq	-.0238047	.0956899	-0.25	0.804	-.2113534	.163744
seqweek	.0019316	.0635137	0.03	0.976	-.122553	.1264162
igac_m6	.3612117	.2191385	1.65	0.099	-.0682919	.7907154
igac_m12	-.0524765	.2123266	-0.25	0.805	-.4686289	.3636759
igac_m16	-.0955507	.2184336	-0.44	0.662	-.5236728	.3325714
award	.1253671	.0868928	1.44	0.149	-.0449397	.295674
orcamento	-.0021706	.0010683	-2.03	0.042	-.0042643	-.0000769
rmeta_pr	.610043	.1336536	4.56	0.000	.3480869	.8719992
rmeta_tn	.008686	.0071892	1.21	0.227	-.0054047	.0227766
season	-.0440447	.3179879	-0.14	0.890	-.6672896	.5792001
ss	.0063722	.0003387	18.81	0.000	.0057083	.007036
ss2	-1.75e-06	2.12e-07	-8.23	0.000	-2.16e-06	-1.33e-06
_cons	-.3015291	.3022323	-1.00	0.318	-.8938936	.2908353

GEE population-averaged model		Number of obs	= 1044
Group variable:	id	Number of groups	= 120
Link:	identity	Obs per group: min	= 2
Family:	Gaussian	avg	= 8.7
Correlation:	exchangeable	max	= 10
		Wald chi2(16)	= 3528.01
Scale parameter:	1.208668	Prob > chi2	= 0.0000

(Std. Err. adjusted for clustering on id)

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0037524	.0064873	-0.58	0.563	-.0164672	.0089624
sequel	.8926528	.3836549	2.33	0.020	.140703	1.644603
revparente	-.0001497	.0001888	-0.79	0.428	-.0005199	.0002204
intseq	-.039539	.0297828	-1.33	0.184	-.0979122	.0188341
numseq	-.0191962	.0926653	-0.21	0.836	-.2008169	.1624244
seqweek	.0132033	.0603851	0.22	0.827	-.1051493	.1315559
igac_m6	.3103367	.2434835	1.27	0.202	-.1668822	.7875556
igac_m12	-.0688889	.2324452	-0.30	0.767	-.5244731	.3866953
igac_m16	-.0982768	.236513	-0.42	0.678	-.5618338	.3652802
award	.128189	.0887347	1.44	0.149	-.0457278	.3021059
orcamento	-.0020413	.0010954	-1.86	0.062	-.0041882	.0001057
rmeta_pr	.6016828	.1376303	4.37	0.000	.3319325	.8714332
rmeta_tn	.0044335	.0073203	0.61	0.545	-.009914	.018781
season	-.0017534	.3453067	-0.01	0.996	-.6785421	.6750352
ss	.0067041	.0003537	18.95	0.000	.0060108	.0073973
ss2	-1.87e-06	2.27e-07	-8.25	0.000	-2.32e-06	-1.43e-06
_cons	-.3597421	.3199553	-1.12	0.261	-.9868429	.2673587

GEE population-averaged model		Number of obs	= 1115
Group variable:	id	Number of groups	= 120
Link:	identity	Obs per group: min	= 2
Family:	Gaussian	avg	= 9.3
Correlation:	exchangeable	max	= 11
		Wald chi2(16)	= 3689.76
Scale parameter:	1.36003	Prob > chi2	= 0.0000

(Std. Err. adjusted for clustering on id)

lrec	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
rellag	-.0008734	.0068503	-0.13	0.899	-.0142998	.012553
sequel	.7388904	.3596715	2.05	0.040	.0339471	1.443834
revparente	-.0001174	.0001817	-0.65	0.518	-.0004735	.0002387
intseq	-.0427852	.0280643	-1.52	0.127	-.0977903	.0122198
numseq	.0185782	.0800672	0.23	0.817	-.1383507	.1755071
seqweek	.0417077	.0524756	0.79	0.427	-.0611427	.1445581
igac_m6	.3710183	.2845969	1.30	0.192	-.1867813	.9288179
igac_m12	-.0360859	.2714728	-0.13	0.894	-.5681629	.4959911
igac_m16	-.0002428	.273914	-0.00	0.999	-.5371043	.5366187
award	.1243602	.090642	1.37	0.170	-.0532949	.3020153
orcamento	-.0019577	.0010157	-1.93	0.054	-.0039486	.0000331
rmeta_pr	.5114152	.1455982	3.51	0.000	.226048	.7967824
rmeta_tn	.0046664	.0074274	0.63	0.530	-.0098911	.0192239
season	.0731307	.3704146	0.20	0.843	-.6528686	.7991299
ss	.0070116	.0003653	19.19	0.000	.0062956	.0077276
ss2	-1.98e-06	2.36e-07	-8.39	0.000	-2.44e-06	-1.52e-06
_cons	-.6058599	.3522809	-1.72	0.085	-1.296318	.0845979

Appendix 18 – Producer's problem results (Stata), reduced sample

GEE population-averaged model		Number of obs =	1667
Group variable:	id	Number of groups =	139
Link:	identity	Obs per group: min =	2
Family:	Gaussian	avg =	12.0
Correlation:	exchangeable	max =	15
		Wald chi2(16) =	10371.06
Scale parameter:	.8633743	Prob > chi2 =	0.0000

(Std. Err. adjusted for clustering on id)

lwrev	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
sequel	.9921053	.2395458	4.14	0.000	.5226042	1.461606
revparent	-.0004928	.0009128	-0.54	0.589	-.0022818	.0012962
gapseq	-.0625592	.0221777	-2.82	0.005	-.1060266	-.0190917
numseq	.0828005	.0818144	1.01	0.312	-.0775528	.2431539
seqweek	-.0423681	.014789	-2.86	0.004	-.0713541	-.0133821
mpaa_g	2.077954	.5639663	3.68	0.000	.9726004	3.183308
mpaa_pg	1.9262	.5677332	3.39	0.001	.8134638	3.038937
mpaa_pg13	1.953308	.5636728	3.47	0.001	.8485299	3.058087
mpaa_r	1.673384	.5549311	3.02	0.003	.5857391	2.761029
award	.2402575	.1153364	2.08	0.037	.0142023	.4663126
budget	.0037344	.0010503	3.56	0.000	.001676	.0057929
rmeta_pr	1.146112	.1950026	5.88	0.000	.7639137	1.52831
rmeta_tn	.0073664	.0068874	1.07	0.285	-.0061327	.0208656
season	.4693661	.423095	1.11	0.267	-.3598847	1.298617
theaters	.0029789	.0001102	27.04	0.000	.002763	.0031949
theaterssq	-4.45e-07	3.11e-08	-14.32	0.000	-5.06e-07	-3.84e-07
_cons	-5.539145	.6036758	-9.18	0.000	-6.722328	-4.355962