Title
Essays on Industrial Organization

Author
Caoui, El Hadi

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays on Industrial Organization

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Economics

by

El Hadi Caoui

2019

ABSTRACT OF THE DISSERTATION

Essays on Industrial Organization

by

El Hadi Caoui

Doctor of Philosophy in Economics

University of California, Los Angeles, 2019

Professor John William Asker, Chair

This dissertation consists of two essays in the area of Industrial Organization.

In Chapter 1, I investigate the role of network effects in explaining the within-firm rate of technology adoption. I study the conversion of movie distribution and exhibition from 35mm film to digital technology. These industries constitute a hardware-software system with indirect network effects. I specify and estimate a dynamic oligopoly game of digital hardware adoption by movie theaters and digital movies (software) supply by movie distributors. Crucially, theaters' technology-adoption decisions are made at the screen level so diffusion occurs both within and across firms. Counterfactual simulations establish that: (1) at the industry level, diffusion occurs mainly within rather than across firms; (2) differences in technology adoption across firms, which are commonly attributed to scale economies and strategic incentives, are in part due to larger firms' ability to initially adopt the technology at a smaller scale. Therefore, explicitly accounting for intra-firm adoption dynamics is important to better explain aggregate diffusion and firm heterogeneity in technology adoption.

In Chapter 2, I study how non-cartel firms adjust their pricing to the supra-competitive level sustained by a cartel, and in doing so, may harm consumers via so-called "umbrella" damages. Such damages arise, in particular, when contracts are awarded through first-price procurement auctions. This chapter examines the bidding behavior of non-cartel firms bidding against the Texas school milk cartel between 1980 and 1992. Evidence is found that the largest non-cartel firm bid significantly higher when facing the cartel. Structural estimation of damages and inefficiencies due to the cartel agreement reveals that per contract: (1) damages from non-cartel firms overbidding are at least 47% of damages caused by the cartel, (2) when the outcome of the auction is inefficient, damages due to misallocation amount to 64% of cartel damages. Finally, inefficiencies raise the winner's cost by 3.7%. These results shed light on the potential importance of umbrella damages from a civil liability perspective.

The dissertation of El Hadi Caoui is approved.

Simon Adrian Board

Martin B Hackmann

Brett William Hollenbeck

John William Asker, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am deeply grateful to my thesis advisors—John Asker, Simon Board, Martin Hackmann, and Brett Hollenbeck—for their exceptional guidance, patience, and support through my years as a graduate student. I have learned tremendously from our interactions, and aspire to be as good of an advisor to my future students as they have been to me.

I am grateful for helpful discussions and guidance from Hugo Hopenhayn, Volker Nocke, and John Riley. Comments from seminar participants at the UCLA IO proseminar, USC Economics, UCLA Anderson helped shape this dissertation as well.

I would also like to acknowledge the hard work of Chiara Paz and Juliana Smith in administering the Economics graduate program at UCLA. Their support made navigating graduate life much smoother for all students.

I thank Renato Giroldo, Keyoung Lee, Rustin Partow, Ksenia Shakhgildyan, Mengshan Cui, Jonathan Gu, Vladimir Pecheu, Jesper Riis-Vestergaard Soerensen, Bruno Pellegrino, Marco Testoni, Maria Lucia Yanguas and the rest of my classmates for their continued friendship since we arrived at UCLA in 2014. I thank my parents for their constant encouragement. Finally, I owe heartfelt thanks to my partner, Ludovica, for her support and for sharing the ups and downs of this adventure of graduate school. I look forward to all of the new adventures awaiting us.

VITA

Education

2015	M.A. in Economics, University of California, Los Angeles
2014	M.A. in Economics, Paris School of Economics
2014	M.Sc. in Statistics. ENSAE ParisTech
2012	B.Sc. in Economics and Mathematics. Ecole Polytechnique

Fellowships

2018–2019	Dissertation Year Fellowship. UCLA Graduate Division
2018	Board of Visitors Fellowship. UCLA Economics
2017–2018	Jacqueline and George Mefferd Fellowship for best original work presented at the UCLA Industrial Organization proseminar
2016–2017	Jack Hirshleifer Endowed Fund recipient for Best Teaching Assistant
2014–2018	Graduate Student Fellowship. UCLA Economics

Teaching

2015–2018	Teaching Assistant. Introduction to Econometrics II (Ph.D. level), Microeconomic Theory, Introduction to Statistics

INTRODUCTION

This dissertation consists of two papers relating to the field of Industrial Organization. The first examines the role of network effects in technology adoption, focusing on the within-firm margin of adoption. As an illustration, I study the conversion of the movie industry from 35mm to digital. The second studies "umbrella damages," a type of damages caused to consumers when non-cartel firms adjust their pricing to the collusive level sustained by a cartel. The common thread throughout this work is the focus on classic questions of the Industrial Organization literature, such as the relationship between technology adoption and market structure (i.e, firm size distribution and concentration levels), and collusion between competitors in product markets.

The first chapter is motivated by the empirical observation that technological innovations diffuse over time, not only across firms, but also within firms. I contribute to the understanding of within-firm technology adoption by quantifying the role of network effects. Network effects arise when the benefit from adopting an innovation depends on the number of other firms adopting it. Focusing on the movie industry's conversion from 35mm film to digital technology between 2005 and 2014, I collect panel data on movie theaters' decisions to adopt digital technology. I specify a dynamic model of technology adoption by movie theaters, and calibrate it to the data. This allows the simulation of counterfactuals. The analysis show that technological diffusion within the firm is an important factor affecting the speed of industry-wide diffusion. Additionally, the ability to gradually adopt the innovation favors larger firms, and explains a significant fraction of the adoption time lag between small and large firms.

Chapter 2 aims attention at a potential collateral damage of cartel agreements. The latter are agreements between rival firms to cooperate and raise prices rather than compete. Such agreements harm consumers and cause “damages” to efficient production allocation. This chapter argues that firms which are *not part* of a cartel agreement can also adjust their pricing to the level sustained by a cartel, harming in the process consumers who do not buy directly from the cartel. Such “umbrella damages” broaden the scope of the overall damages caused by the cartel. The Texas school milk cartel in the 1990s is used as a case study. Estimates of umbrella damages indicate that they constitute a significant fraction of overall cartel damages.

CHAPTER 1

Intra-Firm Technology Adoption under Network Effects: Evidence from the Movie Industry

1.1 Introduction

Innovations contribute to economic growth only insofar as they are broadly adopted by firms. Understanding the factors affecting firms' adoption decisions is therefore essential to devise effective policies encouraging the spread of new technologies. One such factor are network effects, which have been the object of an extensive empirical literature. This literature focuses mainly on the inter-firm adoption margin by assuming firms make 0–1 adoption decisions or by studying firms' first adoption. By contrast, this paper investigates how network effects drive the intra-firm margin of technology adoption, that is, the rate at which a new technology replaces the old technology *within* a given firm. The objective is to evaluate how the latter margin contributes to industry-wide diffusion and shapes the relationship between market structure and technology adoption.

The importance of analyzing this margin appears at two separate levels. First, if intra-firm diffusion constitutes the main driver of industry diffusion, policy aiming at accelerating adoption should explicitly target inefficiently slow diffusion within firms. Second, differences in adoption across firms that are attributed to scale economies and strategic interactions may, in fact, be better explained when accounting for intra-firm adoption dynamics: In an environment with network effects, capital indivisibilities amplify the positive link between firm size and early adoption.

The study focuses on the conversion of movie distribution and exhibition from 35mm film to digital cinema in France, between 2005 and 2013. Digital cinema consists of distributing motion pictures to theaters over a digital support (internet or hard drives) as opposed to the historical use of 35mm film reels. To screen digital movies, theaters must equip their screens with digital video projectors instead of film projectors: The two technologies (digital and film) are incompatible.

Digital cinema is well suited for analyzing the role of network effects in intra-firm technology adoption for two reasons. First, the movie distribution-exhibition industries constitute a hardware-software system with indirect network effects (Katz and Shapiro (1985)). Adoption of digital projectors—the hardware—by theaters is contingent on the availability of digital movies—the software—supplied by distributors. Conversely, software availability depends on the hardware installed base.

Second, indirect network effects lead to intra-firm technology diffusion (i.e.,*within* theaters). Indeed, the benefit of replacing a film projector with a digital projector can initially be small because of the limited availability of digital movies. As a consequence, it is only optimal for a given theater to initially convert a small fraction of its capital stock of screens to digital projection. As the industry-wide share of screens equipped with digital projectors grows over time, so does the availability of digital movies. The latter in turn further increases theaters' marginal benefit from adoption, and leads to the process of technological diffusion *within* theaters. According to industry professionals, network effects were a major factor affecting adoption.

This mechanism applies more generally to other industries. A recent and still developing example is the trucking industry's adoption of electric vehicles (hardware), which depends on the availability of charging infrastructures (software).[1] At a given point in time, a trucking firm's rate of adoption of electric vehicles will reflect the overlap between its distribution routes and the network of available charging stations.

The paper leverages three novel datasets: (1) a panel recording adoption of digital projectors at the theater-screen level, as well as information on local market conditions and theater characteristics, (2) a time series of hardware prices, and (3) a time series reporting the share of movies distributed in digital.

To quantify the contribution of the intra-firm margin to aggregate diffusion and the cross-sectional heterogeneity in adoption, I specify a structural model and simulate counterfactuals shutting down this margin. Theaters' technology-adoption choices are modeled as a dynamic oligopoly game, allowing for rich theater and market heterogeneity. Every period, theaters choose the number of screens to equip with the digital projection hardware, given their competitors' adoption decisions, the adoption cost, and the availability of digital movies. In turn, the availability of digital movies depends on the number of digitally equipped screens in the industry. Because network effects are at the industry level, with a few hundred theaters adopting, this framework generates a particularly high-dimensional state space. To alleviate the computational burden, the paper assumes firms condition their adoption decisions on moments summarizing the industry state, rather than all possible realizations of it: the equilibrium concept employed in this paper follows the *moment-based equilibrium* defined in Ifrach and Weintraub (2017).[2]

Theaters' single-period profits are estimated using the two-step estimator of Bajari, Benkard, and Levin (2007) (hereafter BBL (2007)). By recovering the equilibrium actually played in the data, this approach allows me to deal with equilibrium multiplicity, a prevalent issue in games with network effects. The estimation approach

[1]See "How Tesla's first truck charging stations will be built" - Reuters 02/01/2018

[2]This framework has been used in recent empirical work by Jeon (2017) and Gerarden (2017).

exploits differences in adoption behavior across theaters (e.g., differences in adoption times, units of new technology acquired, and adoption costs) to estimate how theaters and market characteristics affect single-period profits.

Using the estimated model, the paper first evaluates the extent to which industry diffusion is driven by intra-firm diffusion. The equilibrium industry diffusion (i.e., variance in adoption times across all screens) is decomposed into an intra-firm and inter-firm margins. To separate the two margins, a counterfactual diffusion path is simulated, restricting every theater' adoption-strategy space to a binary 0–1 adoption decision. In this sense, theaters are restricted to converting their entire capital stock of screens at once, conditional on adoption, thus shutting down the intra-firm margin. Importantly, theaters take the aggregate share of digital movies over time as given in the equilibrium played in the data: this approach, therefore, computes a counterfactual *best-response*.[3] The analysis shows aggregate diffusion (i.e., the dispersion in adoption times across capital units) is mainly explained (69%) by the diffusion within rather than across theaters.

Second, the analysis moves from the industry to the local market level. The estimated model is used to evaluate the role of the intra-firm margin in explaining the observed heterogeneity in adoption rates across firms. Such differences have been historically attributed to two important factors: firm size (economies of scale) and market concentration (strategic incentives).[4] The objective is to isolate the role of the intra-firm margin from the latter two factors. The intra-firm margin plays a role because large theaters are able to initially convert a smaller fraction of their stock of screens than are small theaters. It is optimal to do so due to the presence

[3]A counterfactual best-response is sufficient to decompose the equilibrium aggregate diffusion rate into an intra-firm and inter-firm margins. While interesting in itself for welfare analysis, computation of counterfactual equilibria is complicated by multiplicity.

[4]Differences can also arise due to heterogeneity in firm characteristics (type of programming, integration etc.) which the model allows and controls for.

of indirect network effects: the benefit from adopting depends on the availability of digital movies, and initially, only a small fraction of movies is released in digital. As a result, larger theaters and more concentrated local markets (i.e., markets with fewer theaters, keeping the total stock of screens fixed) introduce the technology faster—all else equal.

The introduction lag, or difference in expected time to first adoption between a small and large theater (resp. between a competitive and concentrated market), is simulated under (1) the equilibrium adoption strategy played in the data and (2) the counterfactual best response outlined above (intra-firm margin shut down)—in a given local market, fixing theaters and market characteristics. By comparing the introduction lags simulated under adoption strategies (1) and (2), one can isolate the role of the intra-firm margin from other factors (scale economies, strategic interactions). For the average urban local market with more than $100,000$ inhabitants, a significant fraction of the introduction lag (30% to 42% for small/large theaters, and 43% to 69% for less/more concentrated market) is due to the intra-firm adoption margin. Therefore, in addition to economies of scale and strategic incentives, intra-firm adoption dynamics are an important factor explaining differences in adoption behavior across firms and shape the relationship between market structure and technology adoption.

The rest of the paper is organized as follows. The next section reviews the literature and highlights the main points of departure from it. Section 1.3 presents the movie distribution and movie exhibition industries, describes the technology and highlights the specificities of the French market. Section 1.4 describes the data and gives preliminary descriptive statistics. Section 1.5 quantifies the magnitude of indirect network effects via reduced-form analysis. Section 1.6 develops the dynamic model of technology adoption. Section 1.7 estimates the industry model. Finally, section 1.8 presents the counterfactual analysis.

1.2 Related Literature

Previous empirical work estimating the extent of network effects in technology adoption has focused mainly on the inter-firm (or extensive) margin, modelling adoption as a 0–1 decision. This approach is appropriate when adopters are end-consumers with a unit demand for the technological good or when, in the case of firm adoption, the technology is not embodied in capital. Recent examples in the case of consumer goods include video games platforms (Clements and Ohashi (2005), Corts and Lederman (2009), Dubé, Hitsch, and Chintagunta (2010), Lee (2013)), DVD players (Karaca-Mandic (2003)), and home computers (Goolsbee and Klenow (1999)). More recently, Ryan and Tucker (2012) studies the diffusion of videocalling within a multinational firm. In their model, the within-firm rate of adoption is measured by the number of employees using the technology, with each employee making 0–1 adoption decisions; network effects arise within the firm.[5]

In the case of firm adoption, examples include the US Fax market (Economides and Himmelberg (1995)), Automatic Teller Machines (Saloner and Shepard (1995)), electronic switching in the US telecommunication industry (Majumdar and Venkataraman (1998)), and the Automated clearinghouse payment system (Gowrisankaran and Stavins (2004), and Ackerberg and Gowrisankaran (2006)). In contrast to this literature, this paper analyzes the extent to which network effects, at the industry level, determine not only whether the technology is adopted, but also how the intra-firm adoption rate is affected.

I also contribute to the literature analyzing the drivers of intra-firm technology adoption. The development of this literature has been limited by the lack of detailed data at the unit of capital level. In order to observe the intra-firm spread of an

[5]Note that in all the aforementioned papers, except in the cases of home computers and videocalling, network effects are indirect. This type of network effects have been formally modelled by Chou and Shy (1990), and further developed by Church and Gandal (1992). Recent theoretical contributions include Markovich (2008), and Markovich and Moenius (2013).

innovation, each firm has to be observed over long time periods since the date of initial adoption. Intra-firm adoption was first analyzed in Mansfield (1963), who studies railroads' conversion from steam to diesel powered locomotives. In this model, a new technology diffuses within the firm as the risk attached to the payoff from this technology is reduced over time through inter-firm and intra-firm learning. Other empirical studies based on the learning approach include Nabseth and Ray (1974), Romeo (1975), Levin, Levin, and Meisel (1992), and Fuentelsaz, Gomez, and Polo (2003).[6]

Recently, alternatives to the learning model were proposed. In particular, Battisti and Stoneman (2005) propose stock effects as an important driver of the time intensity of intra-firm adoption, using the case of Computer Numerically Controlled Machine tools within firms in the UK engineering and metalworking sectors.[7] Other studies following this approach include Hollenstein (2004) and Hollenstein and Woerter (2008). This strand of the literature relies on cross-sectional survey data and is constrained to estimating the effect of firm- or environment-specific factors on inter- and intra-firm adoption at a given point in time. By contrast, this paper relies on a panel of adoption decisions for all firms active in a given industry, and explicitly models the dynamics of stock effects.

This paper makes a contribution to the literature on market structure and innovation (including technology adoption). The recent research on this topic builds dynamic structural models and simulates the effect of competition on innovation (e.g.,

---

[6]A literature on the "depth" of adoption considers the intensity of use of a given stock of new technology. One example is Astebrot's (2004) study of the effect of learning sunk costs. The depth of adoption differs from the intra-firm margin of adoption, in that the stock of new technology adopted is fixed while the extent to which this stock's technological capabilities are exploited varies.

[7]The stock-effect approach argues intra-firm diffusion will not be instantaneous, because the marginal profit gain from increased use of new technology decreases with use. In their case, the decreasing marginal returns from further adoption are implied by a Cobb-Douglas specification of the production function (implying decreasing marginal productivity of capital embodying the new technology).

Goettler and Gordon (2011), Igami (2017), Igami and Uetake (2017)) or technology adoption (Schmidt-Dengler (2006)). This paper follows this methodological approach and analyzes how market structure, and in particular firm size, can impact the timing of technology adoption specifically via the intra-firm margin.

Finally, this paper contributes to the empirical literature studying the movie industry. This literature has considered many facets of the industry: the effect of vertical integration (Gil (2008)), seasonality (Einav (2007)), strategic entry and exit, and spatial retail competition (Davis (2006a),Davis (2006b), Takahashi (2015), Gil, Houde, and Takahashi (2015)), and the quality-variety trade-off in screening brought about by digital projection (Rao and Hartmann (2015)). In the strand focusing on technology adoption, the closest paper is Gil and Lampe (2014), who analyze Hollywood's conversion to color in the 1940s-1950s. Another related paper in this strand is Waldfogel (2016). The latter paper studies the effect of digital movie production, alternative distribution channels (streaming), and online film criticism on new product releases. The present paper focuses on the digitalization of the movie distribution and exhibition sectors, with theater releases as the main channel of distribution.

1.3 Industry Background

This section describes the movie-distribution and movie-exhibition industries before and after the advent of digital technology. It presents costs and benefits of digital cinema from the perspective of distributors and exhibitors, and discusses the effect of digital cinema on movie ticket prices and quality. Finally, this section highlights specificities of the French distribution and exhibition markets and important stylized facts.

1.3.1 From 35mm Film to Digital

For most of the 20th century, movies reached viewers after going through a series of specified steps in a vertically structured industry. After the movie is shot and produced, distributors print the movie onto 35mm film reels and ship the reels to movie theaters. At the theater, a projectionist inspects the print, attaches the reels together, and positions them so they can be fed to the screening platter of a film projector. When the movie's run is over, the print is broken back down into shipping reels and either sent to the next theater venue or returned to the distributor.

On January 19, 2000, the Society of Motion Picture and Television Engineers, in the US, initiated the first standards group dedicated to developing digital cinema. The technology would entail (1) movie distribution on a digital support (via the internet or hard drives), instead of the historical uses of film reels and (2) movie projection via digital projection hardware instead of the film-projection technology.

To screen a digital movie, theaters must equip their screens with digital projectors. Four manufacturers supply digital cinema projectors worldwide: Sony, Barco, Christie, and NEC. The average list price of a digital projector (in 2010 euros) was €88,000 in 2005, €50,000 in 2010, and €40,000 by 2012. In addition to the digital projector, a digital cinema requires a powerful computer known as a "server." A digital movie is supplied to the theater as a digital file called a Digital Cinema Package (DCP). The DCP is copied onto the internal hard drives of the server, usually via a USB port.

Digital projection automates all the technical tasks that were previously performed by the projectionist. Unskilled staff can control the playback of the content (movie featured, trailers, ads), the projector, sound system, auditorium lighting, and tab curtains through automation cues in the server.

1.3.1.1 Distributors' supply of digital movies

Digital distribution of movies drastically cuts printing and shipping costs for movie distributors. To print an 80-minute feature film can cost US$1,500 to $2,500 per print. By contrast, a feature-length movie can be stored on an off-the-shelf 300GB hard drive for $50.[8] In addition, hard drives can be returned to distributors for reuse. With several hundred movies distributed every year, the US distribution industry saves over $1 billion annually.

1.3.1.2 Exhibitors' adoption of digital projectors

Digital projection allows exhibitors to cut down on operating costs. Screening film prints is a technical task, requiring mechanical skills that are growing rare. Film projectionists are commonly represented by powerful unions and are therefore expensive.[9] By contrast, operating a digital projector is a simple task: untrained staff can easily compose a playlist and launch a projection as on a regular computer. One consequence is that uncertainty about the benefits of the technology and learning, which are usually thought of as important determinants of intra-firm diffusion, are not central to adoption decisions. Digital projection also opens up the possibility of using theaters for âalternative contentâ such as pop concerts, opera broadcasts, and sports events.

---

[8]The latter figure of $50 does not include the price of encryption-key generation, transportation, and storage, which add approximately $200–$300. For French/European movies, these figures are around €950 for film print distribution and around €350 for digital print distribution, including shipping, encryption-key generation, and storing.

[9] See interview in l'Obs 07/14/2010 (in French) "Frederic, projectionniste chez UGC pour 1800 euros par mois" The collective bargaining agreements set the minimum monthly salary to €1,500 over the period of interest.

1.3.1.3 Multi-homing by movie distributors and theaters

Multi-homing in movie distribution consists of the distribution of a given movie on both film and digital supports. It was initially very common: most movies released in digital format were also distributed on film. As the technology diffused, the share of multi-homed movies decreased toward 0.

Multi-homing in movie exhibition refers to equipping a given screen with both a digital and film projector. This type of multi-homing was rare for practical reasons (limited space in screening booth, heavy and sensitive projection equipment), and because theaters laid off their projectionists following the adoption of digital projection.

1.3.1.4 The Virtual Print Fee system

A large fraction of the cost savings from digital cinema is realized by distributors. For this reason, theaters have been reluctant to switch without a cost-sharing arrangement with movie distributors. An agreement was reached with the Virtual Print Fee (VPF) system. The VPF system was born in the US market and was rapidly adopted in the rest of the world. Under this system, the distributor pays a fee per digital movie to help finance the digital hardware acquired by the theater. The VPF contract would typically cover 50% of the hardware adoption cost; the rest has to be paid for by the exhibitor.

1.3.1.5 Impact on ticket prices and movie quality; and the role of 3D

Excluding 3D movies, the film-digital quality differential was small enough not to warrant any impact on ticket prices.[10] Although 3D movies, and in particular *Avatar*

[10] As noted by Davis (2006a), theaters' ability to set ticket prices is constrained by distributors' incentives. Because the conversion to digital distribution affected mainly the cost of a movie print,

(released in the winter 2009, grossing $2.7 billion worldwide), were initially a major selling point for digital projection, exhibitors quickly realized it was not expanding the audience as promised.[11] The vast majority of movies released over the diffusion period were in 2D.

1.3.1.6 Welfare implications

Although the paper does not discuss welfare, digital cinema is expected to increase consumer surplus by reducing the cost of making movies. A first consequence of such reduction in costs are wider releases with increased access for theaters located in small and rural markets. Second, cost reductions in movie making will lead to new product entry.[12]

1.3.2 The French Distribution and Exhibition Markets

1.3.2.1 The French Exhibition Industry

The French exhibition industry is fragmented, with a large fraction of small theaters. Figure 1.1 represents the distribution of theaters by size, defined as the number of screens owned by the theater. Half of the theaters are mono-screen. An additional 15% are two-screen theaters. The largest theater chains by share of total screens in 2014 (end of the diffusion of digital cinema) are Gaumont-Pathe (13.6% of screens), CGR (7.8%), and UGC (7.5%). These three chains make up 50.1% of total box

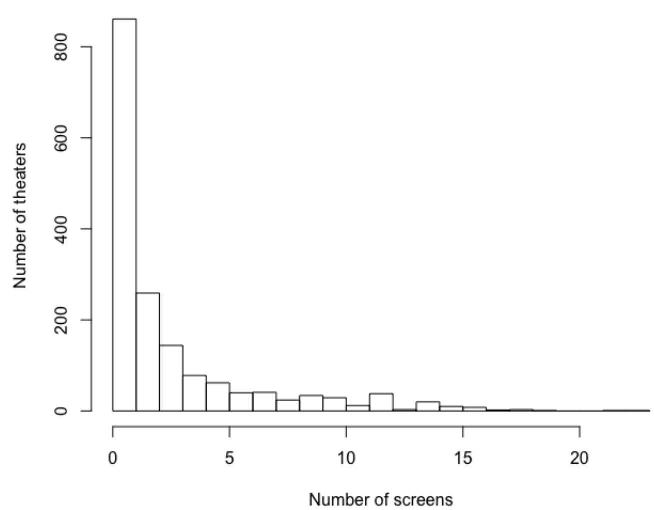
which is a *fixed* cost from the point of view of movie tickets, it did not significantly impact ticket pricing.

[11]See Bordwell (2013).

[12]Aguiar and Waldfogel (2018) show that if product quality is unpredictable at the time of investment (as is typically the case with cultural products such as movies), new product entry can have large welfare benefits.

office revenue.[13] In the early phase of the diffusion period (in 2007), these shares were: Gaumont-Pathe (12.1% of screens), CGR (7.1%), and UGC (7.0%). Market shares were relatively stable over the diffusion period. The French exhibition industry experienced small entry and exit rates over the diffusion period (around 1.5% per year). As a result, the majority of digital projectors acquired were replacing old film projectors, enabling the analysis of intra-firm adoption decisions.

Figure 1.1: Distribution of theaters by size (number of screens)



1.3.2.2 The French Distribution Industry

The French distribution industry is less concentrated than its US counterpart. In 2014, for example, the four-firm concentration ratio was 35.2% in France and 57.4% in the US. Over the diffusion period $2005 - 2014$, US movies had an average 47% market share (of total box-office revenue), French movies had a 39% market share,

[13]See Kopp (2016).

and European and other nationalities made up 14% of the box office.[14] An important point to note is that US studios distribute their movies via national subsidiaries (e.g., Universal France or Warner Bros. France). Subsidiaries tailor their advertising and distribution campaigns to the national market they operate in. Therefore, the support—film or digital—over which US movies are distributed in France depends primarily on the installed bases of film and digital projectors in France.

1.3.2.3 The VPF and Government Subsidies

In the US, the VPF system was the result of bilateral negotiations between distributors and exhibitors. This was initially the case in France as well, until a law was passed on September 2010 making VPF contributions mandatory: any movie distributor willing to distribute digital copies of its movie must pay a fixed fee to the theater booking the digital copy. As in the US, the VPF would go toward covering 50% of the digital projector cost, the rest being paid by the exhibitor.

Government and regional subsidies to small theaters were another important feature of the hardware-acquisition process in France. Many small "continuation" theaters, which receive movies only two or three weeks after their national release, did not generate enough VPF to be able to acquire the digital-projection hardware. The government, along with the regions, stepped in to help these theaters finance their digital conversion. These aids were allocated to theaters that owned less than three screens and were not part of a chain controlling 50 screens or more.

1.3.2.4 Art House Theaters

French theaters can acquire the "art house" label if they screen a minimum share of independent and art house movies. This share depends on the theater location

[14] Based on CNC 2014 annual report.

(market size). The label, awarded every year, entitles the theater to government financial support (in the form of a lump-sum subsidy). A priori, operating profits may differ for art house theaters compared to non-art house theaters. Therefore, art house theaters may differ in their adoption behavior.

1.4 Data and Descriptive Statistics

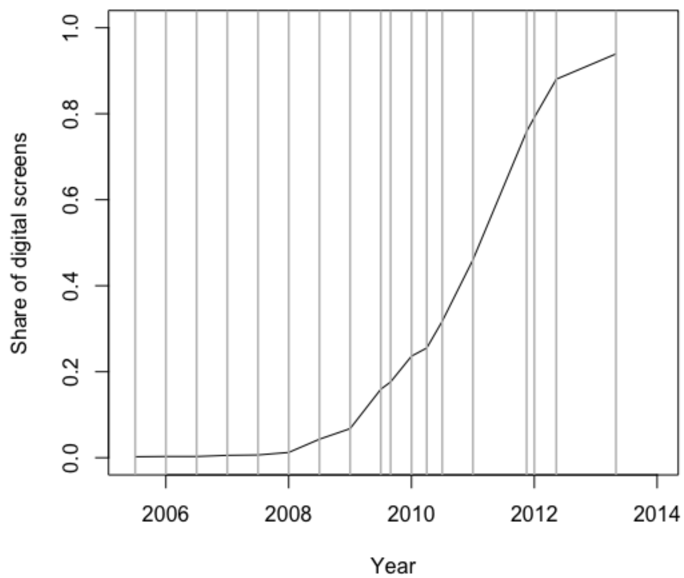
This section describes the data and presents descriptive statistics. The data contain information on theaters' digital-adoption decisions, theater characteristics, adoption costs, and availability of digital movies over time. This information is used to study the role of indirect network effects in driving within-theater adoption of digital.

The main dataset is a panel describing digital adoptions by theaters. This dataset was collected from two sources: the European Cinema Yearbooks published by Media Salles, and an online database maintained by Cinego, a private digital platform.[15] Both sources are public and provide snapshots of the digital-exhibition industry at different time periods spanning June 2005 through March 2013, in France. Thirteen dates are obtained from the Cinema Yearbooks, and 5 dates from the Cinego database. At each of the 18 observation dates, the number of digital projectors acquired is known for every active theater. The observation dates and source are detailed in Appendix A.1. Figure 1.2 represents the 18 observation dates along the industry share of screens equipped with digital projection. As seen in this figure, the panel is aperiodic (starting in 2008) and stops before the diffusion is complete in 2014. Five periods are dropped to ensure a relative periodicity in the sample (6 months). Details about this procedure can be found in Appendix A.1.

Two auxilliary datasets complement the main adoption panel dataset. The first

[15] Raw data available at: `http://www.mediasalles.it/yearbook.htm` and `https://cinego.net/basedessalles` (via the Internet Archive)

Figure 1.2: Observation times (vertical lines) for the diffusion of digital projectors



is obtained from the French National Center of Cinematography (CNC hereafter). The CNC dataset provides a rich set of information on theaters' characteristics, local market demand, and the share of movies available in digital, between 2005 and 2015. More precisely, this annual dataset contains: (1) lists of all active theaters, (2) the number of screens, the number of seats, the address, the owner's identity (theater chain, individual), and art house status for each active theater, (3) market population (categorical) at the urban/rural unit level (defined below), (4) the share of movies released in digital (distributed partially or entirely in digital), and (5) in 2015 only, a categorical variable for ticket sales (e.g., "150,000 to 200,000 tickets sold"), number of movies screened, total number of screenings, shares of movies screened that year by type (art house vs. non-art house), and by nationality (US, France, Europe, and other) for each active theater.

The second auxiliary dataset, obtained from the European Audiovisual Observatory, provides time-series information on digital-projector acquisition costs.[16] Namely, the time-series for the hardware adoption cost is constructed by adding (1) the price of a digital projector (net of VPF contributions) to (2) ancillary costs. The time-series for digital-projector prices is based on a survey of projector manufacturers. Actual prices paid by specific theaters are not public due to nondisclosure agreements between theaters and manufacturers. This time-series is taken as representative of the "list" price (or manufacturer's suggested retail price MSRP) of digital projectors. The analysis accounts for the VPF subsidies, which cover 50% of the projector price.[17] Ancillary costs include the price of other equipment (the server and the digital sound processor), Theater Management software, and labor costs (installation). Estimates of ancillary costs were collected by the European Audiovisual Observatory, but are only available for 2010. In the analysis, these ancillary costs are assumed to have stayed constant over the sample period. This assumption seems reasonable for labor costs. According to the Observatory, price declines for the server and digital sound processor are more limited than for the digital projector. The hardware-adoption cost is adjusted to 2010 constant euros.[18]. The hardware adoption cost is interpolated to obtain estimates at the 13 observation dates. Figure 1.3 shows the time series for this variable.

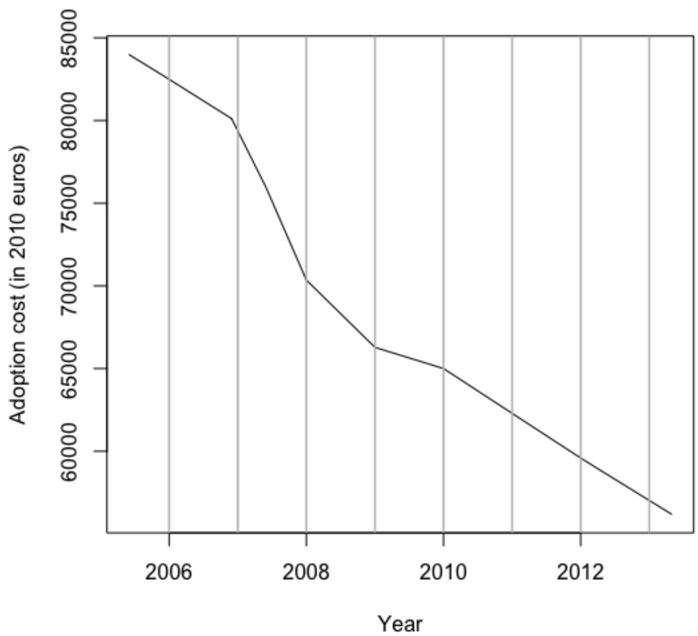
The analysis is conducted on the data after the following preparation. Itinerant theaters, which account for 5% of active theaters, are dropped. Because the focus is on firms' decision to convert existing capital from film to digital, theaters that enter during the diffusion period already equipped with digital projectors are excluded

[16]See "The European Digital Cinema Report - Understanding digital cinema roll-out" (Council of Europe, 2012)

[17]Although the law mandating VPF subsidies was only enacted in September 2010, it was retroactive. Moreover, anecdotal evidence indicates that pre-2010, the majority of projectors were purchased under VPF agreements. The model will assume digital projector-purchases before 2010 benefited from VPF subsidies.

[18]The GDP Implicit Price Deflator for France is used.

Figure 1.3: Hardware adoption cost



from the model. Their contribution to the overall installed base of digital screens is, however, accounted for and taken as exogenous. Firms exiting before conversion to digital are also excluded.[19] Rates of entry and exit are, however, low (around 1.5% of firms enter or exit every year). Theaters in French overseas territories are excluded. The final sample includes 1,671 theaters, located in 1,169 markets (urban or rural units, defined below), and observed over 13 dates between June 2005 and April 2012. The sample covers 87% of all non-itinerant theaters located in Metropolitan France, which were active in 2005 or entered before 2008 equipped with the old technology. A description of variables used in the analysis is shown in Table 1.1.

[19]The inability to convert is, however, not a significant cause for exit, because the CNC and regional governments subsidized digital adoption for the smallest and less financially sound theaters.

Table 1.1: List of variables

Type	Variable	Description
Movie theater	digital screens: s_{it}	number of screens converted to digital by time t
	screens : S_i	total number of screens
	seats	average number of seats per screen
	art house	equals 1 if art house theater
	chain	identifier of movie theater chain
	competitors d-screens	competitors' digital screens by time t
	competitors f-screens	competitors' film screens by time t
	box-office*	number of tickets sold (annual - categorical)
	art house movies*	share of art house movies
	screenings*	total number of screenings
Market demand	region	identifier for the 22 administrative regions
	market size	identifier for: Paris, Paris inner suburbs ("petite couronne")
		Paris outer suburbs ("grande couronne"),
		urban unit with more than 100 thousands inhabitants,
		urban unit with 20 to 100 thousands inhabitants,
		urban unit with less than 20 thousands inhabitants and rural
Digital projector	adoption cost: p_t	list prices for 2K digital projectors
		(including VPF subsidies and ancillary costs) in 2010 euros
Movie distribution	digital movies h_t	share of movies released in digital format

*Note: The first three categories in market size (Paris and suburbs) are colinear with the regional dummy for "Ile-de-France", the latter is therefore excluded. * variables only available for* 2015

Local Market Definition and Competitors

Local market demand and competition is defined with respect to the urban or rural unit in which the theater is located. An urban unit is defined by the INSEE, the French National Statistics Office, for the measurement of contiguously built-up areas. It is a "commune" alone or a grouping of communes that form a single unbroken spread of urban development, with no distance between habitations greater than 200 meters, and have a total population greater than 2,000 inhabitants. Communes not belonging to an urban unit are considered rural.[20] In 2010, Metropolitan France contained 2,243 urban units and about 33,700 rural units.

For the largest cities (Paris, Lyon, Marseille), the urban unit division is not appropriate, because the resulting local markets are too large. In these cases, the relevant market within each city is the *"arrondissement"* (equivalent to zipcode in the US).[21] In the rest of the paper, a theater's competition is measured using the number of competing screens in the same local market.

Descriptive Statistics

The analysis focuses on theaters with at least four screens, due to the prevalence of government and regional subsidies for small theaters (fewer than three screens).[22] Table 1.2 and 1.3 report cross-sectional summary statistics, and highlight the market and firm heterogeneity captured by the data.

Table 1.2 shows summary statistics for the 399 theaters with at least four screens. A significant fraction of these theaters, 33%, are art house theaters. The average theater has eight screens, and 1,538 seats. Thirty-five percent of theaters are part of the three largest theaters chains: Gaumont-PathÃ©, CGR, and UGC. In total, 53.4% of theaters are miniplexes (4-7 screens) and 46.6% are multiplexes/megaplexes (8 screens or more).

Table 1.3 reports summary statistics by market type. Paris and its suburbs are controlled for separately because attendance rates are significantly higher in the cap-

[20]Communes correspond to civil townships and incorporated municipalities in the US.

[21]The subdivision by *arrondissement* is arbitrary, given that theaters are engaged in spatial competition. However, the use of a distance measure instead to define a firm's rivals (in which the relevant market is theater-specific) would make the dynamic model intractable.

[22]These subsidies covered part or all of a theater's adoption costs. This restriction allows the paper to avoid having to model firms' beliefs regarding the distribution of future government and regional subsidies (and subsequent increase in the number of digitalized screens). Although the model does not formally include subsidized firms, their adoption decisions will matter for the aggregate diffusion of digital projectors: their contribution to the network of digital screens is accounted for but assumed to be exogenous to the model.

Table 1.2: Summary statistics (theaters with at least 4 screens)

Variable	Minimum	Mean	Maximum	Std. Deviation
Theater characteristics				
Screens	4	8.118	23	3.736
Seats	112	1,538	7,408	961
Art House	0	0.336	1	0.473
# Competitors (theaters)	0	3.118	14	3.414
Theater size (indicators)				
Miniplexe (4-7 screens)	0	0.534	1	0.499
Multi/Megaplexe (8 screens or more)	0	0.466	1	0.499
Theater chains (indicators)				
UGC	0	0.083	1	0.276
Gaumont-Pathé	0	0.168	1	0.374
CGR	0	0.103	1	0.304
Per screen cost of digital conversion (in 2010 euros)	56,176	68,366	84,000	8,800

Table 1.3: Summary statistics by market size (theaters with at least 4 screens)

	Theaters	Markets	Theater size (mean)	Art house (share)	Screens per market (mean)	Screens per market (sd)
Urban unit - >100k inhab	174	101	9.167	0.213	15.792	8.431
Urban unit - 20 to 100k inhab	126	116	7.024	0.563	7.629	2.986
Urban unit - <20k inhab and rural	17	17	6.647	0.647	6.647	3.552
Paris	37	15	7.189	0.135	17.733	9.565
Paris - inner suburbs	18	18	9.222	0.278	9.222	4.977
Paris - outer suburbs	27	26	7.926	0.185	8.231	4.320
National	399	293	8.117	0.337	11.055	7.262

ital compared to national averages.[23] As expected, the stock of screens grows with the market size. A larger fraction of theaters are art house in rural areas, because the CNC's threshold requirements to qualify are lower for relatively less dense areas. Theater size increases on average with market size (except for Paris, where the scarcity of space limits theater size).

Preliminary analysis of the data shows that the intra-firm margin is quantitatively important at the aggregate level, and that there is substantial heterogeneity in adoption rates by firm size.

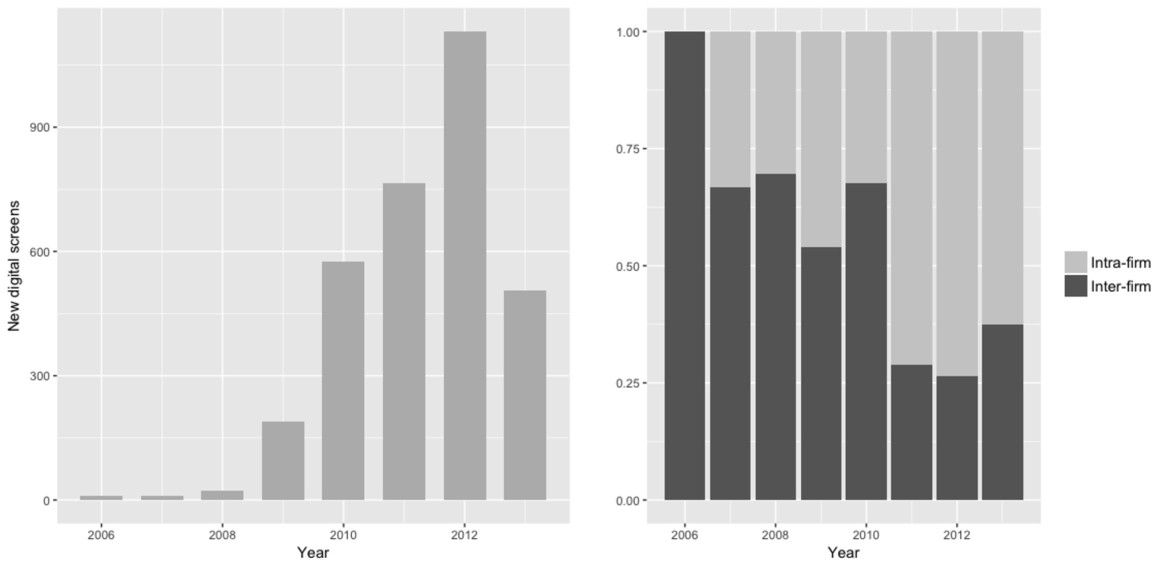
Figure 1.4a shows the number of new digital screens equipped per year. Figure 1.4b decomposes this number into: (1) screens installed by new adopters (theaters with no digital screens in $t-1$), and (2) screens installed by theaters with some digital screens by $t-1$. (1) is informative about the degree of inter-firm adoption, whereas (2) is informative about the degree of intra-firm adoption. Starting in 2008, a large fraction of screens converted to digital per year belong to theaters that have already adopted at least one digital screen in previous periods, highlighting the importance of the intra-firm margin. An alternative way of measuring the contribution of intra-firm adoption is to decompose the sample variance in adoption times across all screens in the industry into within-theater and between-theater variances. The former sample variance is 1.28 (corresponding to a standard deviation of 1.13 years), the latter is 1.02 (or a standard deviation of 1.01 years). Therefore, within-theater variance in adoption times explains about 56% of total variance across all screens in the industry.[24]

With respect to firm heterogeneity in adoption, figure 1.5 shows the share of digital screens over time for theaters grouped by size. Larger theaters adopt the technology faster (i.e., there is a first-order stochastic shift of the adoption path as firm size

[23] Moviegoers in Paris visit theaters on average 12 times a year, compared to a national average of $4-5$ times over the diffusion period.

[24] This likely underestimates the contribution of the intra-firm margin because the panel stops before the end of the diffusion.

Figure 1.4: Inter- and intra-firm decomposition

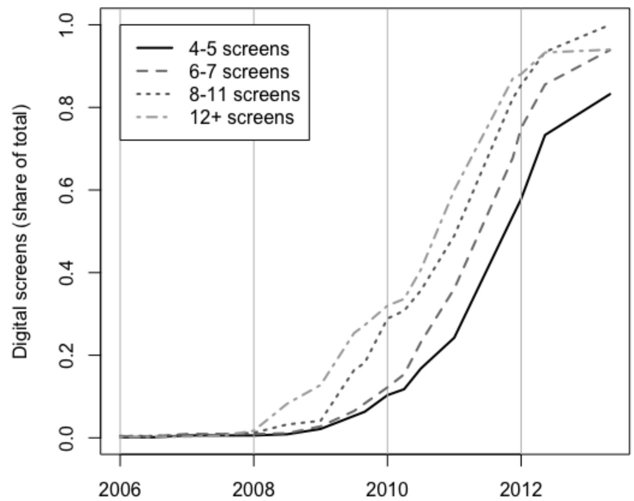


(a) *New digital screens* (b) *Decomposition: Inter/Intra-firm*

Note: "Inter-firm" corresponds to screens installed by new adopters (no digital screens by $t-1$). "Intra-firm" corresponds to screens installed by theaters with some digital screens by $t-1$. Subsidized theaters (3 screens or fewer) excluded.

decreases). On the one hand, this pattern can be due to size-related factors affecting theater's profits and adoption costs. In particular, theaters of varying size might differ in their characteristics (type of programming, ownership), in their competitive environment, and in profits per screen if economies of scale are important. On the other hand, capital indivisibilities at the screen level can also explain the initial lag in adoption of small theaters: if the share of digital movies is initially small, theaters have an incentive to initially convert a small share of screens to digital. Hence, only the largest theaters are able to adopt at the margin. To distinguish the contribution of capital indivisbilities from other potential factors explaining the delay, a structural model is needed. Before describing the model (section 1.6), the next section presents reduced-form evidence for the magnitude of network effects in this industry.

Figure 1.5: Aggregate share of digital screens by theater size



1.5 Reduced-Form Analysis

The anecdotal evidence garnered from industry professionals suggests indirect network effects are at play in the diffusion of digital cinema. This section provides estimates of the magnitude of these network effects. The analysis focuses primarily on hardware adoption, because of the richness of the data there. The results indicate that digital-movie availability is an important variable affecting theaters' digital hardware adoption. This variable will therefore be an important component in the structural model analyzed in the rest of the paper.

As noted in Gowrisankaran and Stavins (2004), network effects (whether direct or indirect) are difficult to identify using only time-series data because the adoption cost is decreasing over time while the network size is increasing over time. Disentangling both effects is challenging.

The effect of digital-movie availability on theaters' adoption of digital projectors is identified here by leveraging differences in movie programming between art house theaters and commercial theaters. The underlying assumption is that art house theaters' adoption of digital projectors would depend on the availability of digital art house movies, whereas commercial theaters' adoption would depend on the availability of digital commercial (non-art house) movies. At a given period t, both types of theaters will face the same hardware adoption cost but different digital-movie availabilities, giving cross-sectional variation in this latter variable.[25]

A theater is defined as an "art house" if at least 80% of the movies it screened in 2015 were art house movies. A theater is defined as a commercial theater if at most 20% of the movies it screened in 2015 were art house movies.[26] This definition is preferable to the use of the "art house" label awarded by the CNC every year (and observed in the data), because thresholds to qualify for the award can be as low as 30% in small urban and rural markets. The preferred definition allows for better identification of theater' type but has the disadvantage that data on the share of art house movies screened by the theater are only available for 2015. The share in 2015 is assumed to reflect the average share of art house movies screened by the theater over the diffusion period $2005-2014$. The validity of this assumption rests on two observations. First, the data show that no changes occured in theaters' CNC-defined art house labels over the sample period. Art house theaters tend to maintain their art house label, whereas commercial theaters do not turn into art house theaters.[27] Second, the share of art house movies released per year remained relatively constant

[25] Differences in digital-movie availability for art house and commercial movies stems from heterogeneity among distributors. Distributors of commercial movies (e.g., the majority of US studios) started distributing on digital earlier than smaller art house distributors.

[26] The quantitative results (documented below) are robust to both relaxing and restricting the thresholds used to define art house and commercial theaters.

[27] One reason for the stability in art house labels is that art house theaters receive state aid, which acts as an incentive to maintain their programming.

between 2005 and 2015.[28] In particular, the conversion to digital did not have an impact on it.

Because the share of art house (resp. commercial) movies released on digital are not directly observed in the data, the reduced-form analysis assumes that the latter variables can be proxied by the share of art house (resp. commercial) screens converted to digital across the industry. Details on how these two variables are constructed are presented in Appendix A.2. This assumption is valid if the availability of digital movies is driven by the installed base of digital projectors. In addition to the anecdotal evidence stating so, the assumption is buttressed by regressing the share of movies released in digital on the industry share of screens equipped with a digital projector, instrumenting the latter by the hardware-adoption cost, and controlling for year fixed effects. The results indicate the installed base of digital projector has a positive and significant effect on digital-movie availability.[29]

The reduced-form model relates the share of digital screens in theater i to: (1) the industry-wide installed base of (art house or commercial) screens, (2) the adoption cost, (3) firm characteristics (number of screens and seats), and (4) market characteristics (market size, competitors' screens). Denote by S_i the total number of screens in theater i, and by s_{it} the number of digital screens in theater i by period t. The dependent variable is the share of screens converted to digital s_{it}/S_i. The analysis focuses on non-adopters' incentive to adopt, so only the case of *first* adoption is considered; that is, the dependent variable is s_{it}/S_i conditional on $s_{it-1}/S_i = 0$.[30] Capital

---

[28]According to statistics compiled by the CNC, the share of art house movies released per year fluctuates between 55% and 62% between 2005 and 2015. The average share is 58.7% over the sample period. Over the diffusion period, on average, 25% of US movies released in France are art house, whereas this share for French movies is 68%. See Bilan 2015 at `http://www.cnc.fr/web/fr/bilans/-/ressources/9217573`

[29]A 1% increment in the industry share of digital screens (instrumented by digital projector prices) implies a 1.51% increment in the share of movies released in digital.

[30]Every period, only theaters that have not yet adopted are included in the regression.

indivisibilities imply that the dependent variable is discrete and can take values in $s_{it}/S_i \in \{0, \frac{1}{S_i}, \frac{2}{S_i}, ..., 1\}$.

Let $\mathbf{x}_{it}$ be the list of regressors including the aggregate share of digital art house or commercial screens, the adoption cost, theater i's number of screens S_i, number of seats, CNC-defined art house label, competitors' digital screens, and competitors' film screens.

An ordered probit model relating the discrete dependent variable s_{it}/S_i to $\mathbf{x}_{it}$ is estimated by maximum likelihood. The sample is restricted to art house and commercial miniplexes (4–7 screens). The periodic sample with 13 dates is used, and it contains 42 art house theaters and 111 commercial theaters.[31] Table 1.4 presents the estimates of the ordered probit model under three specifications: (1) is the baseline specification, (2) includes dummies for regions, market size, chain membership, box-office revenue (in 2015), and (3) includes theater-level random effects.[32] In all specifications, year fixed effects are included. An ordered logit specification, included in Appendix A.2, predicts effects of similar magnitude.

The results indicate that the share of digital art house (resp. commercial) screens in the industry has a significant and positive effect on art house (resp. commercial) theaters' own adoption of digital screens. The effect of a 10% increment in the industry share of art house (resp. commercial) digital screens on the probability of adoption is represented in Figure 1.6a as a function of the initial industry share of art house (resp. commercial) digital screens.[33] An increase in the share of digital screens from 50% to 60% increases the probability of adoption by approximately 10%.

[31]Multi and megaplexes (8 screens or more) are excluded because the majority are commercial theaters.

[32]Focusing on first adoption might lead to selection of theaters over time. The inclusion of theater random effects alleviates this concern.

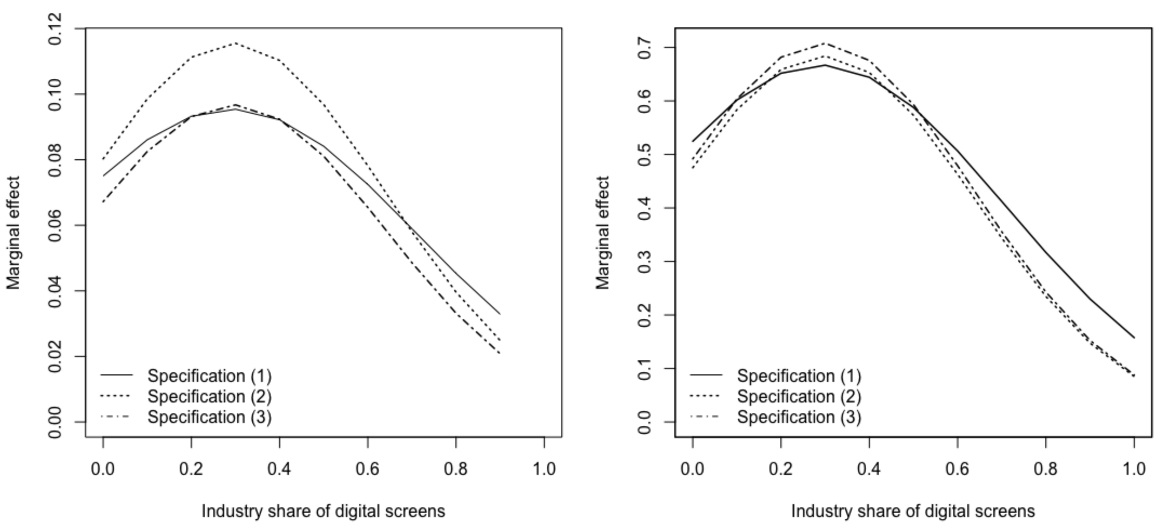
[33]Marginal effects are evaluated at the mean and are obtained by summing marginal effects on the probabilities of converting non-zero shares of digital screens $s_{it}/S_i > 0$.

Table 1.4: Share of screens converted s_{it}/S_i conditional on $s_{i(t-1)}/S_i = 0$ (ordered probit)

	(1)		(2)		(3)	
	Estimate	s.e	Estimate	s.e	Estimate	s.e
Industry share of d-screens	2.392**	1.219	2.896**	1.263	2.425**	1.131
Adoption cost	-2.557**	1.002	-2.623**	1.043	-2.714**	1.018
Own screens	0.053	0.055	0.062	0.059		
Seats	-0.085	1.176	-0.288	1.302		
Art house	0.024	0.130	0.021	0.138		
Competitor d-screens	0.013	0.010	0.012	0.011	0.013	0.010
Competitor f-screens	-0.004	0.004	-0.005	0.004	-0.004	0.004
Year FE	Yes		Yes		Yes	
Region FE	No		Yes		No	
Market size FE	No		Yes		No	
Chain FE	No		Yes		No	
Box-office FE	No		Yes		No	
Theater RE	No		No		Yes	
Observations	1,563		1,563		1,562	
-log Likelihood	391.292		373.626		384.251	
AIC	816.584		805.251		798.502	

Note: $^{***}p < 0.01$;$^{**}p < 0.05$; $^{*}p < 0.1$. *d-screen = screen equipped with a digital projector. f-screen = screens equipped with a film projector. For market dummies, the omitted category is "urban unit with 20 to 100 thousands inhabitants." For the chain dummies, the omitted category is "single firm and small chains."*

Figure 1.6: Effects of the industry share of digital screens and adoption cost on the probability of adoption



(a) Effect of an increase in the share of digital screens

(b) Effect of a decrease in the adoption cost

Notes: Panel (a) shows the effect of a 10% *increment in the industry share of digital screen, on the probability of adoption, evaluated at the mean, as a function of the initial industry share of digital screens. Panel (b) shows the effect of a one standard deviation decrease in the adoption cost on the probability of adoption evaluated at the mean as a function of the industry share of digital screens.*

The hardware-adoption cost exerts a negative and significant effect on the probability that a theater adopts. Figure 1.6b shows that a one-standard-deviation decrease in the adoption cost (corresponding to €$6,500$) increases the probability of adoption by approximately 55%—when 50% of screens in the industry are digital.

In summary, these results indicate the magnitude of the network effect is significant. A given art house (resp. commercial) theater's likelihood of adoption responds to the share of art house (resp. commercial) movies released in digital, under the assumption that the latter variable can be proxied by the installed base of art house (resp. commercial) digital screens across the industry. The section relies on anecdotal evidence and a time-series regression to support this assumption.

1.6 Industry Model

This section presents the dynamic structural model. The model will be subsequently used to guide the estimation and recovery of theaters' operating profits under the film and digital technologies. These profits are required to study the role of the intra-firm margin, by simulating counterfactual adoption paths.

Theaters' technology adoption choices are modelled as a dynamic oligopoly game in the tradition of Ericson and Pakes (1995). The central part of the model specifies how theaters make their technology adoption decisions—both at the inter-firm and intra-firm margins—as a function of their type, the adoption cost, their rivals' adoption decisions, and the availability of technology-specific complementary goods (film or digital movies). Theaters adopt digital projectors for two reasons: (1) to be able to screen movies exclusively released on digital and (2) for cost-reduction purposes. For the distribution market, a reduced-form model is used. This part of the model is meant to capture, for a given movie to be released, a distributor's decision regarding on which support to distribute it (film and/or digital), given the technology-specific network size (number of screens equipped with film/digital projectors). Finally, an equilibrium of the distribution-exhibition industries is specified. In equilibrium, theaters convert their screens optimally to digital projection, given their information sets and beliefs about future states, and these beliefs are consistent with theaters' adoption decisions and distributors' optimal choices of distribution format.[34]

1.6.1 Adoption of Digital Projectors by Theaters

Time is discrete and infinite. A period corresponds to six months.

Firms: A firm is a movie theater. There are I firms indexed by $i \in \{1, ..., I\}$. This set is fixed throughout the game: no entry and exit occur.

Firm state space: Firm heterogeneity is reflected through firm states. In period t, the individual state of theater $i \in I$ is a vector denoted by $\mathbf{x}_{it} \in \mathscr{X}$. Firm state $\mathbf{x}_{it}$ is decomposed into $(\boldsymbol{\tau}_i, s_{it}, \mathbf{z}_{it})$:

- $\boldsymbol{\tau}_i$ is a vector representing theater i's type, which is fixed throughout the game. $\boldsymbol{\tau}_i$ includes firm size S_i (number of screens), local market characteristics (market size, denoted $market_i$, and number of competitors' screens, denoted S_{-i}), art house label $art_i \in \{0, 1\}$, and a chain identifier $chain_i \in \{0, 1, ..., C\}$ (with $chain_i = 0$ if i is not horizontally integrated).

- $s_{it} \in \{0, 1, ..., S_i\}$ represents the number of screens converted to digital by theater i, by the beginning of period t. The remaining $S_i - s_{it}$ screens operate using the film technology.

- $\mathbf{z}_{it}$ is a vector containing theater i's competitors' types and digital screens. This vector is at the *local market* level (urban or rural unit in which theater i is located), and differs from the *industry* state, which is at the national level and

[34]All vectors are denoted in bold.

defined below.

Let the industry state, $\mathbf{y}_t$, be a vector over individual firm states that specifies, for each firm state $\mathbf{x} \in \mathscr{X}$, the number of firms (across the industry) at $\mathbf{x}$ in period t. Focusing on symmetric and anonymous equilibrium strategies, this definition of the industry state is without loss of generality. Let $S = \sum_{i \in I} S_i$ denote the total number of screens in the industry, and let $s_t = \sum_{i \in I} s_{it}$ denote the total number of digital screens in the industry in period t.

Theaters that are part of the same chain are assumed to make their adoption decisions independently. This assumption is motivated by the fact that modelling adoption at the chain level is computationally burdensome: each chain's state should record firms' states for all theaters part of the chain. The resulting chain state vector is high-dimensional.[35] This modelling assumption is discussed in more length in section 1.6.3.

Transition dynamics: A theater can increase its number of digital screens, s_{it}, by paying an adoption cost. If firm i converts a_{it} screens to digital in period t, the firm transitions to a state s_{it+1} given by

$$s_{it+1} = s_{it} + a_{it} \quad \text{for} \;\; a_{it} \leq S_i - s_{it} \tag{1.1}$$

There is no uncertainty in state transition. A theater's state s_{it} is bounded above by its maximum capacity S_i.

Aggregate adoption cost: The aggregate adoption cost process, $\{p_t,\, t \geq 0\}$, includes the digital-projector price (net of VPF contributions) and ancillary costs, and is assumed to follow a finite Markov process, independent of all previously defined quantities:

$$p_t = p_{t-1} + \eta_{t-1} \tag{1.2}$$

[35] The chain adoption state vector for Gaumont-PathÃ© (70 theaters) has dimension 1,088,430—assuming all theaters have four screens and ignoring theaters' types and rivals states.

where η_{t-1} is a discrete random variable, i.i.d across periods, with a non-positive support. Denote by $\mathbf{P}(p_{t+1}|p_t)$ the corresponding Markov kernel. This process reflects technological advances in the manufacturing of digital projectors, as well as learning-by-doing and scale economies, which exogeneously decreases the hardware adoption cost over time. It is publicly observable to all firms.[36]

Firm-specific adoption cost: The per-screen adoption cost for theater i in period t is the sum of two components:

$$p_t + \epsilon_{it} \tag{1.3}$$

where p_t is the aggregate adoption cost and ϵ_{it} is a theater-specific shock, drawn from a normal distribution $N(0, \sigma^2)$. This theater-specific shock is privately drawn at the beginning of each period and is independent across periods and theaters.

Before defining theaters' single-period profit function, the relationship between the share of digital movies, and s_t, the number of digital screens, must be specified. This is done by considering distributors' technology-choice problem.

Availability of digital movies: Every period, a continuum of mass M of movies is released. Movies are short-lived and are screened by theaters for one period. Let h_t^d denote the share of movies available exclusively in digital format in period t, and h_t^f denote the share of movies available exclusively in film in period t. Denote by $h_t^m = 1 - h_t - h_t^f$ the share of multi-homed movies (i.e., distributed on both film and digital).

The share of multi-homed movies, h_t^m, is not observed in the data (only $h_t^d + h_t^m$ is observed). In anticipation of the estimation section, and to keep the exposition concise, the rest of the model is derived under the "no multi-homing" assumption: $h_t^m = 0$ for all t. As a robustness check, the model is also derived and estimated

[36]The exogeneity of the price process can be relaxed. For example, the transition matrix of the process can depend on the stock of digital screens in the industry: $\mathbf{P}(p_{t+1}|p_t, s_t)$. The simpler specification is imposed due to the limited amount of data available to estimate this transition matrix (only time series information used).

under the polar "wide multi-homing" assumption, $h_t^d = 0$ for all t. The qualitative findings, included in Appendix A.3.2, are robust to this assumption.

For the rest of the analysis, define $h_t \equiv h_t^d$ so that $1 - h_t = h_t^f$. The share of digital movies, h_t, is assumed to depend on the share of digital screens in the industry at the beginning of the period.[37] Let

$$h_t = \Gamma(s_t/S) \tag{1.4}$$

denote distributors' reaction function giving the share of movies released in digital as a function of the industry-wide share of digital screens.[38]

Theaters' Single-Period Profit Function: The single-period profit of theater i (net of the adoption cost) in period t if it adopts a_{it} units of digital hardware is given by

$$\Pi(\mathbf{x}_{it}, p_t, h_t, a_{it}, \epsilon_{it}) = \pi(\mathbf{x}_{it}, h_t) - a_{it}(p_t + \epsilon_{it}) \tag{1.5}$$

where $\pi(\mathbf{x}_{it}, h_t)$ are theater i's operating profits (screenings, concessions, advertisements) in period t, which depends on the firm state $\mathbf{x}_{it}$ and the shares of digital movies h_t, and $a_{it}(p_t + \epsilon_{i,t})$ is the total cost of converting a_{it} screens in period t.

Operating profits, which are not observed in the data, are specified via reduced-form and estimated. Namely, $\pi(\mathbf{x}_{it}, h_t)$ is obtained by aggregating profits per movie screening across all movies the theater *is able* to screen given its stock of digital and film screens. The technologies being incompatible, a theater can screen a digital movie only if it has a digital screen. A theater lagging in its adoption of digital will have fewer movies to screen over time, because as the network of digital screens grows, an increasing share of movies is only available on digital.

[37]This approximation is made for tractability. Because distributors share box-office revenue with theaters, and this revenue differs across theaters, distributors might base their decision not only on the aggregate number of digital screens, but also on the identity of the converted theaters.

[38]This simple specification for the relationship determining h_t can be relaxed to account for other exogenous determinants of the share of digital movies. It is imposed due to the limited amount of time-series information necessary to fit the relationship.

It is assumed that, every period, theater i randomly draws a mass $M_i \subset M$ of movies from the mass M released in that period. By the law of large numbers, a share h_t of movies drawn are digital movies. In particular, every theater draws the same share h_t of digital movies per period.[39] Denote by R_i the number of screenings a theater can host if it is able to screen all M_i movies drawn. The number of screenings R_i is exogenous and depends only on the theater's type $\boldsymbol{\tau}_i$.

Let $\pi_d(\mathbf{x}_{it})$ and $\pi_f(\mathbf{x}_{it})$ be the single-period profits *per movie screening* in state $\mathbf{x}_{it}$ from screening a digital or film copy, respectively. Due to cost reductions from digital projection, these profits satisfy: $\pi_f(\mathbf{x}_{it}) \leq \pi_d(\mathbf{x}_{it})$. Operating profits are obtained by aggregating profits *per movie screening* across all screenings, and are given by

$$\pi(\mathbf{x}_{it}, h_t) = R(\boldsymbol{\tau}_i) \times \begin{cases} \frac{s_{it}}{S_i}\pi_d(\mathbf{x}_{it}) + (1-h_t)\pi_f(\mathbf{x}_{it}) & \text{if} \quad \frac{s_{it}}{S_i} \leq h_t \\ h_t\pi_d(\mathbf{x}_{it}) + (1-\frac{s_{it}}{S_i})\pi_f(\mathbf{x}_{it}) & \text{if} \quad \frac{s_{it}}{S_i} \geq h_t \end{cases} \tag{1.6}$$

where s_{it}/S_i is the share of digital screens in the theater in period t, and h_t is the share of movies released in digital.

Three points should be noted. First, equation (1.6) implies $\pi(\mathbf{x}_{it}, h_t)$ is strictly concave and piece-wise linear so the single-period marginal benefit from further adoption is decreasing. Ideally, theaters wishes to match the share of digital screens s_{it}/S_i to the share of digital movies released h_t.

Second, network effects are indirect: For a theater, the benefit from adopting a digital projector depends on the share of movies available in digital h_t (technology-specific software), which in turn depends on the share of digital screens in the industry through equation (1.4).

[39] This assumption reduces significantly the complexity of the model. Relaxing these assumptions would yield firm-specific shares of digital movies $\{h_{it}\}_{i\in I}$. Smaller theaters might be able to delay their conversion longer because they screen fewer movies overall. Ignoring this mechanism would affect the estimation results by predicting lower profits from adoption for smaller theaters.

Third, a theater's type (firm and market characteristics) impacts the theater's profits per screen, $\pi(\mathbf{x}_{it}, h_t)/S_i$, via two channels: the number of screenings per screen $R(\boldsymbol{\tau}_i)/S_i$ and the profit per screening $(\pi_d(\mathbf{x}_{it}), \pi_f(\mathbf{x}_{it}))$. The impact of theater size on profits per screen can be non-linear if, for instance, larger theaters have more screenings per screen ($R(\boldsymbol{\tau}_i)/S_i$ increasing in S_i), but profits per screening decrease with theater size ($\pi_d(\mathbf{x}_{it}), \pi_f(\mathbf{x}_{it})$ decreasing in S_i).

State space: In a Markov Perfect Equilibrium, firms use Markov adoption strategies and condition their adoption decision only on the current vector of state variables $\boldsymbol{\omega}_{it} \equiv (\mathbf{x}_{it}, p_t, \mathbf{y}_t, \epsilon_{it})$. Although theater i's single-period profits, $\Pi(.)$, do not directly depend on the industry state $\mathbf{y}_t$, the firm tracks this variable in order to form expectations about the future evolution of s_t which in turn determines, via equation (1.4), the (payoff-relevant) share of digital movies h_t.

In this setting, the large number of firms (due to network effects at the industry level) and the dimension of firms' state $\mathscr{X}$ generate a high-dimensional industry state space. For instance, ignoring firm heterogeneity and assuming all 399 firms are four-screen theaters (so $s_{it} \in \{0, 1, 2, 3, 4\}$), the total number of possible industry states $\mathbf{y}_t$ is 1,071,993,300. To alleviate the computational burden, two assumptions are imposed.

First, firms are assumed to condition their adoption decision on moments summarizing the industry state vector $\mathbf{y}_t$, rather than all possible realizations of the vector $\mathbf{y}_t$. Ifrach and Weintraub (2017) proposed an alternative approximation of MPE based on moments of the industry space, which is the approach followed here.[40] More precisely, firms are assumed to condition their adoption decisions on the un-normalized first moment of the distribution of $\{s_{it}\}_{i \in I}$, that is, the total number of digital screens in the industry s_t.

[40] Approximation methods to MPEs were initially proposed in the IO literature on dynamic games with complete information by Weintraub, Benkard, and Van Roy (2008), and subsequently refined in Benkard, Jeziorski, and Weintraub (2015).

Second, the paper assumes that firms do not keep track of the whole vector of competitors' states $\mathbf{z}_{it}$, but only of competitors' total number of digital screens denoted z_{it}.

In summary, firms are assumed to condition their adoption decisions on the *moment-based state* $\widetilde{\boldsymbol{\omega}}_{it} = (\widetilde{\mathbf{x}}_{it}, p_t, s_t, \epsilon_{it})$, where the moment-based firm state $\widetilde{\mathbf{x}}_{it}$ is defined by $\widetilde{\mathbf{x}}_{it} = (\boldsymbol{\tau}_i, s_{it}, z_{it})$.[41] The analysis focuses on equilibria in pure symmetric *moment-based* strategies, defined as mappings from the current *moment-based* state $\widetilde{\boldsymbol{\omega}}_{it}$ into actions (i.e., number of screens to be converted to digital).

Perceived transition kernel: Define $\widetilde{\boldsymbol{\gamma}}_{it} \equiv (\widetilde{\mathbf{x}}_{it}, p_t, s_t)$, as the moment-based state excluding the private firm-specific shock, so that

$$\widetilde{\boldsymbol{\omega}}_{it} = (\widetilde{\mathbf{x}}_{it}, p_t, s_t, \epsilon_{it}) = (\widetilde{\boldsymbol{\gamma}}_{it}, \epsilon_{it})$$

Similarly define $\boldsymbol{\gamma}_{it} \equiv (\mathbf{x}_{it}, p_t, \mathbf{y}_t)$ as the true underlying state excluding the private firm-specific shock, so that

$$\boldsymbol{\omega}_{it} = (\mathbf{x}_{it}, p_t, \mathbf{y}_t, \epsilon_{it}) = (\boldsymbol{\gamma}_{it}, \epsilon_{it})$$

As noted in Ifrach and Weintraub (2017), the moment-based state process $\{\widetilde{\boldsymbol{\gamma}}_{it}, t \geq 0\}$ is in general *not* Markov, even if the true state process $\{\boldsymbol{\gamma}_{it}, t \geq 0\}$ is. By aggregating information via moments, the moments obtained are not necessarily sufficient statistics for next period's moments: for instance, given s_t, next-period state s_{t+1} depends on whether the s_t digital screen are owned by small vs. large theaters, art vs. non art house theaters, or located in rural vs. large urban market. In this sense, many underlying industry distributions can yield the same current-period moment s_t, but different next-period moment s_{t+1}. Therefore, one has to construct a Markov approximation of the kernel transition matrix guiding the dynamics of the process

[41] This definition differs from Ifrach and Weintraub's (2017) definition of the moment-based industry state in the fact that it does not keep track of *dominant firms' states*, but only of an aggregate moment describing the states of firms.

$\{\widetilde{\gamma}_{it}, t \geq 0\}$. Assume theater i follows moment-based adoption strategy a' while all other theaters use moment-based adoption strategy a. The objective is to define a kernel $\widehat{\mathbf{P}}_{a',a}$, which is a good approximation of the non-Markov process $\{\widetilde{\gamma}_{it}, t \geq 0\}$. The process described by $\widehat{\mathbf{P}}_{a',a}$ correponds to firm i's perception of the evolution of its own state $\widetilde{\mathbf{x}}_{it} = (\boldsymbol{\tau}_i, s_{it}, z_{it})$, the adoption price p_t, and the industry moment s_t. To keep the exposition concise, the construction of the perceived transition kernel is detailed in Appendix A.3.1.

Value function and optimal adoption rule: Let $a(\widetilde{\gamma}_{it}, \epsilon_{it})$ be a pure moment-based adoption strategy. Firms aim to maximize expected discounted profits, by choosing the number of screens to equip with a digital projector at the current period, taking into account the effect on future operating profits and given their belief about future values of the state vector. The *moment-based* value function of firm i is defined as the solution of the following Bellman equation (where the subscript t is omitted and next-period variables are marked with a prime):

$$\begin{aligned} V(\widetilde{\gamma}_i, \epsilon_i) = \max_{a_i} \big\{ & \pi(\widetilde{\mathbf{x}}_i, h) - a_i(p + \epsilon_i) \\ & + \beta \sum_{z_i', p', s'} V(\boldsymbol{\tau}_i, s_i + a_i, z_i', p', s') \widehat{\mathbf{P}}_a(z_i', p', s' | \widetilde{\gamma}_i) \big\} \end{aligned} \tag{1.7}$$

where β is the discount factor, $\widehat{\mathbf{P}}_a$ is the perceived transition kernel previously defined, giving the perceived probability of one-period reachable states for competitors' digital screens, the adoption price, and the adoption moment vector s, given firm i's belief a about its competitors' actions. The share of digital movies can be derived using equation (1.4) as: $h' = \Gamma(s'/S)$. Moreover, $V(\boldsymbol{\tau}_i, , s_i + a_i, z_i', p', s')$ is firm i's *ex-ante* value function, that is, before observing next-period firm-specific shock ϵ_i'. It is given by: $V(\boldsymbol{\tau}_i, s_i + a_i, z_i', p', s') = \int V(\boldsymbol{\tau}_i, s_i + a_i, z_i', p', s', \epsilon_i') dF(\epsilon_i')$.

The optimal adoption rule can be expressed as a function of the *choice-specific* value functions. Let $W(a_i | \widetilde{\gamma}_i)$ denote the discounted expected value function when firm i converts a_i screens in the current period:

$$W(a_i|\widetilde{\gamma}_i) = \beta \sum_{z_i', p', s'} V(\boldsymbol{\tau}_i, s_i + a_i, z_i', p', s')\widehat{\mathbf{P}}_a(z_i', p', s'|\widetilde{\gamma}_i) \tag{1.8}$$

Define $\Delta W(k|\widetilde{\gamma}_i) \equiv W(k|\widetilde{\gamma}_i) - W(k-1|\widetilde{\gamma}_i)$ for $k \in \{1, 2, .., S_i\}$ as the difference in the choice-specific value functions of converting k and $k-1$ screens to digital. Firm i's optimal adoption rule is derived by noting that, in deciding the number of screens to convert to digital technology, the firm compares the choice-specific value functions *net* of the adoption cost. The adoption cost, in turn, depends on the current list price p_t, and firm i's idiosyncratic shock ϵ_{it}. The optimal adoption rule takes the form of a set of cut-offs in ϵ_{it}.[42] It can be expressed as

$$a_{it} = \begin{cases} 0 & \text{if} \quad \Delta W(1|\widetilde{\gamma}_{it}) - p_t \leq \epsilon_{it} \\ k & \text{if} \quad \Delta W(k+1|\widetilde{\gamma}_{it}) - p_t < \epsilon_{it} \leq \Delta W(k|\widetilde{\gamma}_{it}) - p_t \\ & \quad \text{and } 1 \leq k < S_i - s_{it} \\ (S_i - s_{it}) & \text{if} \quad \epsilon_{it} < \Delta W(S_i - s_{it}|\widetilde{\gamma}_{it}) - p_t \end{cases} \tag{1.9}$$

The optimal adoption rule can alternatively be recast in the form of conditional choice probabilities (CCP):

$$P(a_{it}|\widetilde{\gamma}_{it}) = \begin{cases} \int_{\Delta W(1|\widetilde{\gamma}_{it})-p_t}^{\infty} dF(\epsilon_{it}) & \text{if} \quad a_{it} = 0 \\ \int_{\Delta W(k+1|\widetilde{\gamma}_{it})-p_t}^{\Delta W(k|\widetilde{\gamma}_{it})-p_t} dF(\epsilon_{it}) & \text{if} \quad a_{it} = k \in \{1, 2, ..., S_i - s_{it} - 1\} \\ \int_{-\infty}^{\Delta W(S_i - s_{it}|\widetilde{\gamma}_{it})-p_t} dF(\epsilon_{it}) & \text{if} \quad a_{it} = S_i - s_{it} \end{cases} \tag{1.10}$$

[42]The cut-off rule can be derived by noting the following two points: (1) a_{it} is optimal in state $\widetilde{\gamma}_{it}$ *iff* $W(a_{it}|\widetilde{\gamma}_{it}) - a_{it}(p_t + \epsilon_{it}) \geq W(a'|\widetilde{\gamma}_{it}) - a'(p_t + \epsilon_{it})$ for all $a' \neq a_{it}$, and (2) $\Delta W(k|\widetilde{\gamma}_{it})$ is decreasing in k (W is concave - which stems from the strict concavity of the single-period profit function $\pi(\widetilde{\mathbf{x}}_i, h)$ in $a_{i,t-1}$). Combining (1) and (2) yield the cut-off rule.

Finally, the *ex-ante* value function (i.e., before the firm observes the idiosyncratic shock ϵ_{it}) can be derived by taking expectations with respect to ϵ_i in equation (1.7):

$$V(\widetilde{\boldsymbol{\gamma}}_i) = \pi(\widetilde{\mathbf{x}}_i, h) + \sum_{a_i} P(a_i|\widetilde{\boldsymbol{\gamma}}_i)\big(- a_{it}(p + E[\epsilon_i|\widetilde{\boldsymbol{\gamma}}_i, a_i]) + W(a_i|\widetilde{\boldsymbol{\gamma}}_i)\big) \tag{1.11}$$

The last equation expresses the ex-ante value function V, as a function of the choice-specific value function W and probabilities $P(a_i|\widetilde{\boldsymbol{\gamma}}_i)$. The latter are both functions of the ex-ante value function. An equilibrium ex-ante value function is a fixed-point of this mapping.

1.6.2 Market Equilibrium

In every period, the sequence of events is as follows: First, distributors observe the outstanding number of digital screens s_t in the industry and publicly make their distribution decision (film or digital) for movies released in that period. Second, theaters receive a private draw ϵ_{it} from the distribution of hardware costs, and decide whether to convert any screens to digital, given the share of movies released in digital h_t, their competitors' digital screens, and their private adoption cost. Third, theaters receive operating profits and pay the adoption cost. The state variables evolve as the adoption decisions are completed and new values of the exogenous variables are realized.

The analysis focuses on equilibrium in pure symmetric moment-based strategies. In a moment-based equilibrium, each theater's adoption decision is optimal in every (moment-based) state, given its beliefs about future states, and those beliefs are consistent with the adoption decisions of other theaters. The adoption strategy $\mathbf{a}^*$ is a moment-based equilibrium if:

$$V(\widetilde{\boldsymbol{\gamma}}_{it}; \mathbf{a}^*) \geq V(\widetilde{\boldsymbol{\gamma}}_{it}; a_i', \mathbf{a}_{-i}^*) \text{ for all firm states } \widetilde{\boldsymbol{\gamma}}_{it} \text{ and strategies } a_i' \tag{1.12}$$

where $V(\widetilde{\boldsymbol{\gamma}}_{it}; \mathbf{a}^*)$ is theater i's ex-ante value function at state $\widetilde{\boldsymbol{\gamma}}_{it}$, given that all theaters play strategy $\mathbf{a}^*$, and $V(\widetilde{\boldsymbol{\gamma}}_{it}; a_i', \mathbf{a}^*)$ is theater i's ex-ante value function when the theater unilaterally deviates to strategy a'_i. Given the large number of theaters, the mapping $\Gamma(.)$ is assumed to be the same under strategy profiles $\mathbf{a}^*$ and $(a'_i, \mathbf{a}^*_{-i})$ (it is not affected by unilateral deviations).

Due to network effects, the game has multiple equilibria, some of which can be found numerically.[43]

1.6.3 Remarks

The assumption that theaters make their adoption decisions independently, even within chains, is violated if theater chains coordinate adoption decisions across theaters. Two incentives to do so are: (1) to benefit from lower per-unit adoption cost when placing large orders of projectors and (2) to tip the industry by significantly increasing the share of digital screens in the industry. To alleviate concern (1), the model controls for chain effects in the profits from operating (firm type $\boldsymbol{\tau}_i$ include an indicator for the three largest theater chains). Regarding the second motive, the largest chain (Gaumont-PathÃ©) controlled 12.1% of screens and had a market share (box-office revenue) of approximately 20%: given its relatively small capital stock of screens, its ability to tip the market toward digital appears limited.

1.7 Estimation

This section discusses the identification and estimation of the structural model presented in section 1.6. The objective is to recover (1) the single-period profit functions per digital and film movie screening $(\pi_d(\widetilde{\mathbf{x}}_{it}), \pi_f(\widetilde{\mathbf{x}}_{it}))$ and (2) the variance of the firm-specific shock. To circumvent computational and equilibrium multiplicity issues, the

[43]The existence proof in Ifrach and Weintraub (2017) requires the industry state process to be irreductible and aperiodic (and is derived for the long-run perceived kernel, not the short-run kernel used here), so it is not directly applicable to this setting. One can, however, show that at least one degenerate equilibrium (e.g., with no adoption) exists.

model is estimated using a two-step CCP-based approach (Bajari, Benkard, and Levin (2007)). Estimation results point to cost reductions from digital projection, heterogeneity in profits per screen across theaters, and the presence of scale economies in operation (profits per screen increasing in theater size).

1.7.1 Identification

Standard results for the identification of discrete-choice models, as in Magnac and Thesmar (2002) and Bajari, Chernozhukov, et al. (2015), apply to this setting. Variable profits are identified, but fixed-cost components entering operating profits must be normalized to zero. The discount factor β is assumed to be known.

Differences in adoption times (and units of technology acquired) across firms allow the identification of the functions π_d and π_f's dependence on firm state $\widetilde{\mathbf{x}}_{it}$. Because the aggregate adoption cost, which is observed by the econometrician (up to a firm-specific private shock), is decreasing over time, differences in adoption times across firms reveal differences in the paid adoption costs, which translate into differences in firms' single-period profits: for example, firms that adopt earlier must be receiving higher single-period profits than firms that adopt later.

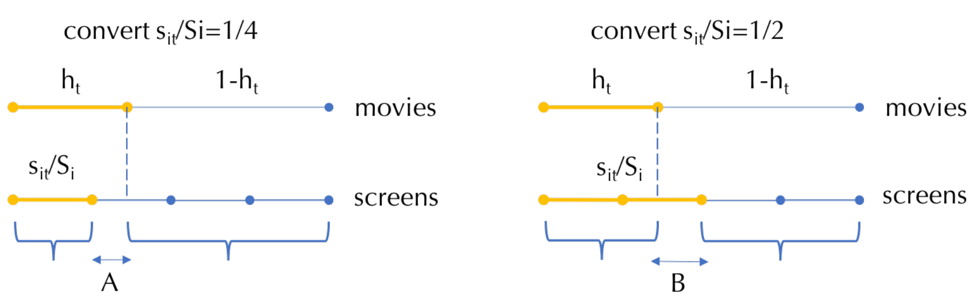
More precisely, to evaluate which elements of the single-period profit functions are identified, note that there are two motives for adopting a digital projector. A firm adopts a digital projector: (1) to gain access to digital movies it would otherwise not be able to screen and (2) for cost-reduction purposes (the profits from screening a digital movie may be higher than the profits from screening a film movie: $(\pi_f(\widetilde{\mathbf{x}}_{it}) \leq \pi_d(\widetilde{\mathbf{x}}_{it})))$.

Adoption for the first motive is informative about $\pi_d(\widetilde{\mathbf{x}}_{it})$. When $h_t > \frac{s_{it}}{S_i}$ (i.e., the share of digital movies released is greater than the share of digital screens owned by theater i), theater i adopts in order to gain access to new digital movies (the fraction

$h_t - \frac{s_{it}}{S_i}$ of un-screened digital movies). The marginal benefit from adoption per-movie screening is $\pi_d(\widetilde{\mathbf{x}}_{it})$. The variation in adoption times across different theaters, allows the identification of the function $\pi_d(\widetilde{\mathbf{x}}_{it})$.

Adoption for the second motive is informative about the difference $\pi_d(\widetilde{\mathbf{x}}_{it}) - \pi_f(\widetilde{\mathbf{x}}_{it})$, or the cost-reduction from screening a digital movie relative to a film movie. [44] This difference is identified by exploiting capital indivisibilities. Consider the example in Figure 1.7. Given h_t and the fact that the theater has not adopted yet, the four-screen theater contemplates the options of converting one screen (left panel) or two screens (right panel). In the first case, the theater forgoes profits of $A \times \pi_d(\widetilde{\mathbf{x}}_{it})$, whereas in the second case, the theater forgoes profits of $B \times \pi_f(\widetilde{\mathbf{x}}_{it})$ (where A and B are masses of movies). Ignoring continuation values, the theater chooses to convert two rather than one screen ($\frac{s_{it}}{S_i} = 1/2$) *iff* $A \times \pi_d(\widetilde{\mathbf{x}}_{it}) - B \times \pi_f(\widetilde{\mathbf{x}}_{it}) \geq p_t$. In cases where this inequality holds (which correspond to cases where theaters over-invest), different adoption times allow identification of the difference $(\pi_d(\widetilde{\mathbf{x}}_{it}) - \pi_f(\widetilde{\mathbf{x}}_{it}))$ (up to the known constants A and B).

Figure 1.7: Identification of $\pi_d(\widetilde{\mathbf{x}}_{it}) - \pi_f(\widetilde{\mathbf{x}}_{it})$



Finally, the variance of the firm-specific shock, $V(\epsilon_{it}) = \sigma^2$, is identified from the

[44] A direct way of identifying this difference would exploit the presence of multi-homed movies, which are available on both formats, and therefore impact theaters' incentive to adopt only via the cost reduction motive. However, due to data limitation, this identification approach is not possible under the "no-multihoming" assumption ($h_t^m = 0$).

variation in adoption times and units of technology adopted between theaters in the same firm state $\widetilde{\mathbf{x}}_i$.

1.7.2 Estimation

1.7.2.1 Parameterization:

This section details the model parameterization. A theater's single-period operating profits are constructed as the product of the total number of screenings $R(\tau_i)$ and the expected profit per screening. The total number of screening $R(\boldsymbol{\tau}_i)$ and the profits per movie screening $(\pi_f(\widetilde{\mathbf{x}}_{it}), \pi_d(\widetilde{\mathbf{x}}_{it}))$ are parameterized, and are estimated separately.

The number of screenings $R(\boldsymbol{\tau}_i)$ is estimated outside the model, using data on screenings in 2015 for each active theater in the model. This variable is explained by a reduced-form model that includes theater and market characteristics part of theater type $\boldsymbol{\tau}_i$: theater size S_i, market size $market_i$, number of rival screens S_{-i}, art house status art_i, and interaction between these variables. This specification allows the number of screenings per screen to vary non-linearly with theater size, capturing potential scale economies. The assumption that the 2015-level for the dependent variable is representative of the diffusion period relies on the fact that the annual number of screenings per screen did not vary significantly over the diffusion period.[45]

For the single-period profit per movie screenings $\pi_d(\widetilde{\mathbf{x}}_{it})$ and $\pi_f(\widetilde{\mathbf{x}}_{it})$, a simple Breshnahan and Reiss (1991)-style reduced form is used:

$$\pi_f(\widetilde{\mathbf{x}}_{it}) = \alpha_0^f + \alpha_1^f S_i + \alpha_2^f \mathbf{1}\{art_i = 1\} + \alpha_3^f S_{-i} + \alpha_4^f z_{it} + \alpha_{market_i}^f + \alpha_{chain_i}^f \quad (1.13)$$

[45] The annual number of screenings per screen for multi/megaplexes (8 screens or more) was on average $1,787$ with a standard deviation of 40 over $2008-2015$, whereas for miniplexes, the average number of screenings per screen is $1,346$ with a standard deviation of 30 (based on CNC's "bilan 2015" and the report by the Digital Diffusion Observatory published in September 2016. Both reports are available (in French) at: `http://www.cnc.fr/web/fr/publications`)

$$\pi_d(\widetilde{\mathbf{x}}_{it}) = \alpha_0^d + \alpha_1^d S_i + \alpha_2^d \mathbf{1}\{art_i = 1\} + \alpha_3^d S_{-i} + \alpha_4^d z_{it} + \alpha_{market_i}^d + \alpha_{chain_i}^d + \alpha_5^d \frac{s_{it}}{S_i} \quad (1.14)$$

where S_i is the number of screens in theater i, s_{it} is the number of digital screens in theater i in period t, art_i is an indicator for art house theaters, S_{-i} is the total number of screens owned by theater i's competitors, z_{it} is the total number of digital screens owned by theater i's competitors in period t, and α_{market_i} and α_{chain_i} are dummies for market size and chain identifier.[46]

The cost reduction from screening digital movies compared to film movies is allowed to depend on the share of digital screens in theater i via the term $\alpha_5^d \frac{s_{it}}{S_i}$. The cost reduction is expected to be increasing in theater i's share of digital screens. This specification is imposed because, according to industry professionals, operating both technologies simultaneously within a given theater is relatively costly (e.g., limited ability to re-allocate movies across screens within the theater).

The specifications for $R(\boldsymbol{\tau}_i)$ and $(\pi_d(\widetilde{\mathbf{x}}_{it}), \pi_f(\widetilde{\mathbf{x}}_{it}))$ allow profits per screen to vary non-linearly with theater size via two channels: the number of screenings per screen and profits per screening. This specification will capture potential scale economies, for which empirical evidence is available (Verma (2001)).[47] The variables S_{-i} and z_{it} entering $(\pi_d(\widetilde{\mathbf{x}}_{it}), \pi_f(\widetilde{\mathbf{x}}_{it}))$ capture the effect of strategic interactions between competitors on profits.

The parameters of interest are the vector

$$\boldsymbol{\alpha} = \left(\{\alpha_i^j\}_{i=0\ldots5}, \alpha_{market=1\ldots6}^j, \alpha_{chain=0\ldots3}^j, j \in \{f, d\}\right)$$

entering the profit per digital and film screenings and the variance σ^2 of the firm-specific shock.

[46]See Table 1.1 for the different categories of market size.

[47]Only variable profits are identified, so potential scale economies are not due to fixed costs (concession stands, box-office etc.) but to decreasing average variable cost.

1.7.2.2 Estimation approach:

The paper follows the two-step method of BBL (2007) for two reasons. First, the high dimensionality of the model, due to network effects at the industry-level, renders a full-solution method impractical. Second, as is common in games of technology adoption under network effects, equilibrium multiplicity is severe. The two-step method, which avoids equilibrium computation, helps circumvent the multiplicity issue by directly estimating the equilibrium played in the data.

In a first step, the equilibrium policy rule and transition probabilites are estimated from the data, under the assumption that firms play a moment-based equilibrium, and then equilibrium value functions are approximated via simulation by using the estimated equilibrium policy functions and transition probabilities. In a second step, the parameters are estimated by imposing the optimality condition stating that the equilibrium value function yields a higher payoff than the value function from non-equilibrium deviations.

First-step estimation:

Movie theaters' adoption-policy function. The first element to estimate is the equilibrium policy function governing theaters' adoption of digital hardware. The policy function is a cut-off rule in the idiosyncratic shock ϵ_{it} given by equation (1.9). The estimation proceeds by first recovering the conditional choice probabilities (CCP) from the data. Next, the cut-offs forming the equilibrium policy function are obtained from the CCP by inverting equation (1.10). The conditional choice probabilities $P(a_{it}|\widetilde{\boldsymbol{\gamma}}_{it})$ are estimated using an ordered probit model, and in what follows are assumed to be known.[48]

[48] In the estimation part, the network size effect (share of digitally equipped screens across the industry) is separately identified from the adoption cost *in the policy rule* $P(a_{it}|\widetilde{\boldsymbol{\gamma}}_{it})$ via functional form restriction. Additionally, to simulate the value function at states $\widetilde{\boldsymbol{\gamma}}_{it}$ which are not visited in the data, knowledge of the CCP $P(a_{it}|\widetilde{\boldsymbol{\gamma}}_{it})$ is required, and the two-step estimation method relies on the functional form of the policy function.

The cut-offs are recovered by noting that for $a \in \{0, .., S_i - s_{it}\}$,

$$P(a_{it} \leq a|\widetilde{\gamma}_{it}) = \int_{\Delta W(a+1|\widetilde{\gamma}_{it})-p_t}^{\infty} dF(\epsilon_{it}) = 1 - \Phi\left(\frac{\Delta W(a+1|\widetilde{\gamma}_{it}) - p_t}{\sigma}\right) \quad (1.15)$$

where $\epsilon_{it} \sim N(0, \sigma^2)$ and Φ is the normal cumulative distribution. The cut-offs can be obtained by inverting equation (1.15):

$$\frac{\Delta W(a+1|\widetilde{\gamma}_{it}) - p_t}{\sigma} = \Phi^{-1}(1 - P(a_{it} \leq a|\widetilde{\gamma}_{it})) \quad (1.16)$$

If the firm idiosyncratic shock ϵ_{it} equals this (normalized) cut-off, firm i is indifferent between adopting a and $a+1$ digital screens, in state $\widetilde{\gamma}_{it}$.

Transition probabilities of the exogeneous hardware price process. To estimate the transition probabilities of the price process $\{p_t, t \geq 0\}$, the variable is first discretized. The number of discrete grid points is 15. Over the diffusion period, the price process was on a downward trend, possibly due to technological advances in hardware manufacturing, learning by doing, and scale economies. Every period, the price is assumed to either move to a lower grid point, or stay at the current state. The initial (and maximum) price at $t = 0$ is set at the actual price level observed in the data: €84,000. The minimum price level is set at €40,000. The probability that the price transitions to a lower grid point is estimated from the transitions observed in the data.

Distributors' policy function. As noted in the model, distributors' policy rule regarding the distribution format (film or digital) is aggregated, because the data provides only information on the aggregate share of movies released on digital format h_t.[49] The relationship between the share of digital movies h_t and the aggregate share of digital screens s_t, given by equation (1.4), is fitted directly from the data.[50]

[49]Under the "no multi-homing" assumption, $h_t^m = 0$ for all t.

[50]Note the industry share of digital screens accounts for (subsidized) small theaters and theaters that entered already equipped with digital projection. Their adoption is, however, assumed to follow a deterministic and exogeneous process. Theaters in the model take this process as given.

Value functions. The expected value function at a given state is forward-simulated: a large number of paths starting at the given state are simulated using the estimated policy rules, and the discounted sum of profits obtained from these simulated paths are averaged. The simulated paths, starting from a given state $\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}$, are generated using the following procedure:

1. Initialize the industry at the given state $\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}$.

2. Draw firm specific adoption shocks $\{\epsilon_{i0}\}_{i\in I}$ and corresponding adoption decisions dictated by the estimated policy rule.

3. Calculate the single period profit $\Pi(\widetilde{\boldsymbol{\gamma}}_{i0}, \epsilon_{i0})$ given by equation (1.5).

4. Update the current state $\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}$ according to the adoption decisions and transition of the exogeneous price process to next period state: $\{\widetilde{\boldsymbol{\gamma}}_{i1}\}_{i\in I}$.

5. Repeat steps 1-4 for T periods.

The equilibrium value function at state $\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}$ is obtained by averaging $L = 500$ simulated paths. The paths have length $T = 40$ periods (or 20 years). The discount factor used is $\beta = 0.975$.[51] An estimate of the equilibrium value function is obtained as

$$\frac{1}{L}\sum_{l=1}^{L}\left\{\sum_{t=0}^{T}\beta^t\Pi^h(\widetilde{\boldsymbol{\gamma}}_{it}, \epsilon_{it}|\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}, \boldsymbol{\alpha}, \sigma)\right\} \tag{1.17}$$

where $\widetilde{\boldsymbol{\gamma}}_{it} = (\widetilde{\mathbf{x}}_{it}, p_t, s_t)$ and $\Pi^h(\widetilde{\boldsymbol{\gamma}}_{it}, \epsilon_{it}|\{\widetilde{\boldsymbol{\gamma}}_{i0}\}_{i\in I}, \boldsymbol{\alpha}, \sigma)$ is the single period profit in simulation h at period t, when the firm follows the equilibrium adoptions strategy a^*, under the candidate parameters $(\boldsymbol{\alpha}, \sigma)$.[52]

[51] An alternative value for the discount factor, $\beta = 0.95$, is also considered.

[52] The linearity of the single-profit function in $(\boldsymbol{\alpha}, \sigma)$ is exploited to reduce the computational intensity of the procedure: in the second step, the value function at a given state can be simulated only once, instead of for every candidate parameter vector $(\boldsymbol{\alpha}, \sigma)$.

Second-step estimation: In the second step, the underlying parameters $(\boldsymbol{\alpha}, \sigma)$ are set such that equilibrium condition (1.12) is satisfied for every firm i, state $\widetilde{\boldsymbol{\gamma}}_i$ and non-equilibrium deviation a'_i. Denote by $\chi = \{i, \widetilde{\boldsymbol{\gamma}}_i, a'_i\}$ a particular equilibrium condition. Under parameter vector $(\boldsymbol{\alpha}, \sigma)$, define

$$g(\chi; \boldsymbol{\alpha}, \sigma) = \widehat{V}(\widetilde{\boldsymbol{\gamma}}_i | a^*_i, a^*_{-i}) - \widehat{V}(\widetilde{\boldsymbol{\gamma}}_i | a'_i, a^*_{-i}) \tag{1.18}$$

as the difference between the simulated value function at state $\widetilde{\boldsymbol{\gamma}}_i$, when firm i plays the estimated policy rule a^*_i and the deviation a'_i.[53] The equilibrium condition (1.12) is satisfied if $g(\chi; \boldsymbol{\alpha}, \sigma) \geq 0$. The objective of the second step is to find the parameter vector $(\boldsymbol{\alpha}, \sigma)$ such that this inequality holds for all possible equilibrium conditions indexed by χ. BBL (2007) demonstrate that one can restrict estimation to a sufficiently large subset that covers the space of inequalities. The estimation proceeds by selecting $N_\chi = 3,600$ equilibrium conditions. Deviations from the equilibrium adoption strategy are obtained by adding perturbations to the estimated cut-off points. The selected equilibrium conditions are combined to form the objective function:

$$Q(\boldsymbol{\alpha}, \sigma) = \frac{1}{N_\chi} \sum_{j=1}^{N_\chi} \left(\min\{g(\chi_j; \boldsymbol{\alpha}, \sigma)\}, 0\right)^2 \tag{1.19}$$

The estimator of the underlying parameters is the solution of

$$\min_{\boldsymbol{\alpha}, \sigma} Q(\boldsymbol{\alpha}, \sigma)$$

This function is not trivially minimized at the zero vector, because the adoption cost $a_{it} \times p_t$ enters in periods when the theater converts some of its screens to digital.

Although BBL (2007) derive the asymptotic formula for the variance-covariance matrix, implementing it remains computationally burdensome (as one needs to compute the cross-partial derivate of Q with respect to $(\boldsymbol{\alpha}, \sigma)$ and the first-stage parameters). Bootstrap sampling is therefore preferred to obtain standard errors. One

[53] In the above notation, dependence on the first-step parameters is omitted for ease of notation.

difficulty with non-parametric bootstrap is the presence of correlation in decisions across local markets, therefore, sampling market-histories with replacement, as is commonly done in dynamic oligopoly games, is not a valid approach. Instead, a parametric boostrap procedure is used.[54]

1.7.3 Estimation Results

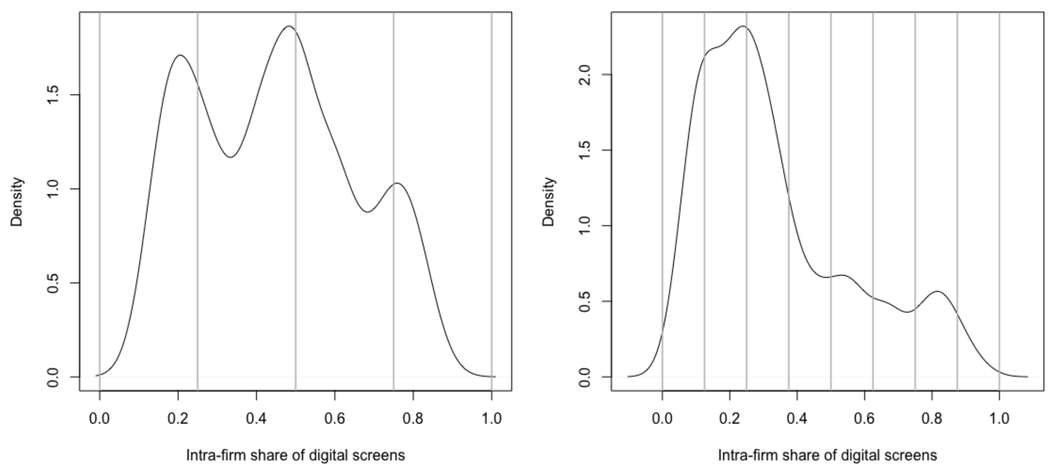
1.7.3.1 First-step estimates

Theaters' adoption-policy function. The conditional choice probabilities are estimated using a flexible reduced form, via an ordered probit model. To further control the size of the state space, theaters' strategy space (the number of screens that can be converted) is restricted to lie on a grid. More precisely, miniplexes (theaters with 4 to 7 screens) are assumed to adopt on the space $s_{it}/S_i \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$, whereas multi- and megaplexes (theaters with 8 screens or more) are assumed to adopt on the space $s_{it}/S_i \in \{0, \frac{1}{8}, \frac{2}{8}, ..., \frac{7}{8}, 1\}$. Figure 1.10 shows kernel density estimates of the intra-firm adoption rates for miniplexes (panel (a)) and multi/megaplexes (panel (b)), conditional on partial adoption ($s_{it}/S_i > 0$ and $s_{it}/S_i < 1$). For miniplexes, the density has three identifiable modes. The previous assumption restricting the strategy space to the set $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ appears non-restrictive. For multi/megaplexes, the density shows a mode around 0.2. The grid chosen, $\{0, \frac{1}{8}, \frac{2}{8}, ..., \frac{7}{8}, 1\}$, is sufficiently fine given the estimated density of s_{it}/S_i.

Because theaters cannot divest and roll back the old technology, a firm cannot transition to lower states. For instance, a four-screen theater with $s_{it}/S_i = 3/4$ can only transition to $s_{it+1}/S_i \in \{3/4, 1\}$. In this sense, next period's possible states de-

[54](1) draw a bootstrap sample of local markets (initial industry state), (2) *simulate* the diffusion process across all markets in the bootstrap sample using the (parametric) first-stage estimates, and (3) estimate the model following the two-step procedure using the bootstrap sample. Repeat the three steps N_b times.

Figure 1.8: Density estimate of the intra-firm rate of adoption by firm size



(a) Miniplexes (4-7 screens) (b) Multi/Megaplexes (8 screens or more)

Notes: Both density estimates correspond to the distribution of s_{it}/S_i *conditional on* $s_{it}/S_i > 0$ *and* $s_{it}/S_i < 1$

pend on the firm's adoption rate in the current period. This dependence is accounted for in constructing the likelihood (see Appendix A.4).

A theater's share of screens converted to digital between t and $t+1$, denoted a_{it}/S_i, is explained by the share of digital movies (or equivalently the aggregate share of digital screens in the industry, through equation (1.4)), the aggregate adoption cost, the number of screens in the theater (and its square), the share of digital screens in the theater in period t, whether the theater is an art house, competitors' digital screens in period t, and competitors' total number of screens. A second specification augments the model by including market dummies to control for market size. A third specification includes both market dummies and theater-chain dummies for the three major French theater chains (Gaumont-PathÃ©, CGR, and UGC). Finally, a fourth specification also controls for interactions between theater size S_i and all other variables.

Table 1.5 presents the estimates of the ordered probit model under the four specifications. As expected, across the four specifications, the share of digital screens in the industry (equivalently the share of digital movies) is positively related to the probability of adoption, whereas the adoption cost is negatively related to the probability of adoption. Larger theaters are more likely to adopt, but the marginal effect is decreasing. Art house theaters are less likely to adopt. The share of a theater's screens already converted to digital is negatively related to further adoption.[55]

Competitors' total number of screens and digital screens do not significantly impact a theater's likelihood of adoption. This estimate indicates that strategic interactions between firms are not a major determinant of adoption.[56] Theaters located in Paris are more likely to adopt than theaters located in the small urban areas with $20,000$ to $100,000$ thousands inhabitants. Among the chain dummies, CGR theaters are more likely to adopt than single theaters or theaters belonging to smaller chains. The rest of the analysis uses specification (4).

To check the goodness of fit, model predictions (from specification (4)) for the share of digital screens are compared to actual shares in the data. Tables 1.6, 1.7, and 1.8 present the comparison for all firms, miniplexes only, and multi/megaplexes, respectively. In each table, the aggregate share of digital screens, the share of adopters (theaters with at least one digital screen), and the average within-theater share of digital screens (among adopters) are shown from 2006 to 2013. Overall, given the limitations imposed by the parametric specification of the policy function, the model captures the main trends in the aggregate, inter-firm, and intra-firm diffusion rates, for all firms and by firm size (miniplexes vs. multi/megaplexes). Note that the aggregate

[55] This finding is expected because, given a share of digital movies, theaters that are lagging behind in terms of adoption (low s_{it}/S_i) have a greater incentive to adopt.

[56] The likelihood-ratio test of specification (3) and (4) against a specification without competitors' screens and digital screens fails to reject the null that both coefficients are zero at the 5% confidence level.

Table 1.5: Adoption policy function

	Dependent variable: Share of screens converted a_{it}/S_i							
	(1)		(2)		(3)		(4)	
	Estimate	s.e	Estimate	s.e	Estimate	s.e	Estimate	s.e
Industry share of d-screens	0.775	0.156	0.796	0.158	1.013	0.164	1.939	0.384
Adoption cost	-1.313	0.023	-1.317	0.025	-1.382	0.025	-1.514	0.038
Own screens	0.088	0.030	0.084	0.031	0.068	0.032	0.102	0.044
Own screens sqrd	-0.003	0.001	-0.003	0.001	-0.001	0.001	-0.005	0.002
Art house	-0.114	0.066	-0.130	0.069	-0.156	0.069	-0.321	0.218
Competitors' d-screens	0.001	0.009	0.000	0.009	-0.003	0.009	-0.003	0.009
Competitors screens	-0.002	0.003	-0.005	0.003	-0.001	0.003	0.016	0.008
Own share of d-screens	-1.422	0.113	-1.443	0.113	-1.699	0.118	-1.361	0.300
Market dummies								
Urban unit - <20k inhab and rural			0.102	0.139	0.104	0.140	-0.335	0.343
Urban unit - >100k inhab			0.024	0.073	0.009	0.071	-0.965	0.243
Paris - inner suburbs			-0.237	0.139	-0.064	0.140	-0.718	0.354
Paris - outer suburbs			-0.133	0.114	-0.000	0.060	-0.692	0.296
Paris			0.092	0.114	0.293	0.116	-0.568	0.308
Chain dummies								
Gaumont-Pathe					-0.146	0.079	0.211	0.254
CGR					0.309	0.093	-1.028	0.386
UGC					-0.891	0.129	-1.238	0.378
Interactions: own screens × variables							X	
Observations	4,788		4,788		4,788		4,788	
- log Likelihood	2,222.981		2,220.041		2,182.751		2,150.577	
AIC	4,461.962		4,466.082		4,397.502		4,357.154	

Note: For market dummies, the omitted category is "urban unit with 20–100 *thousands inhabitants." For the chain dummies, the omitted category is "single firm and small chains."*

share of digital screens was constantly lower for miniplexes than for multi/megaplexes, as reflected in the predictions as well. Additionally, the intra-firm rates' evolution over time is smoother in the prediction than in the data.[57]

Table 1.6: Predictions using the adoption policy function - All firms

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.003	0.000	0.021	0.001	0.127	0.161
2007	0.006	0.001	0.028	0.005	0.210	0.215
2008	0.043	0.010	0.056	0.028	0.182	0.256
2009	0.159	0.093	0.122	0.197	0.420	0.359
2010	0.236	0.270	0.431	0.498	0.459	0.470
2011	0.460	0.483	0.684	0.724	0.583	0.596
2012	0.791	0.709	0.841	0.878	0.880	0.761
2013	0.939	0.922	0.934	0.931	0.985	0.934

Note: The column labelled "Aggregate" corresponds to the share of digital screens across all firms in the industry. The column labelled "Inter-firm" corresponds to the share of theaters with at least one digital screen. The column labelled "Intra-firm" corresponds to the within-theater average share of digital screens among theaters with at least one digital screen. The predicted rates are obtained by averaging 500 *simulation paths.*

[57]In particular, for miniplexes, the intra-firm rate jumps to 41% as early as 2007, whereas the model predicts a slow increase between 2006 and 2009 to reach 40%. The prediction is also smoother in the case of multi/megaplexes, with an increase in the actual intra-firm rates from 13.5% in 2008 to 41.5% in 2009, whereas the model predicts a smoother transition.

Table 1.7: Predictions using the adoption policy function - Miniplexes (4-7 screens)

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.002	0.000	0.008	0.000	0.267	0.229
2007	0.007	0.000	0.017	0.002	0.415	0.293
2008	0.007	0.003	0.017	0.007	0.415	0.338
2009	0.024	0.042	0.054	0.093	0.435	0.400
2010	0.112	0.165	0.243	0.317	0.443	0.463
2011	0.294	0.348	0.498	0.584	0.569	0.549
2012	0.653	0.597	0.745	0.798	0.857	0.702
2013	0.879	0.872	0.891	0.885	0.975	0.906

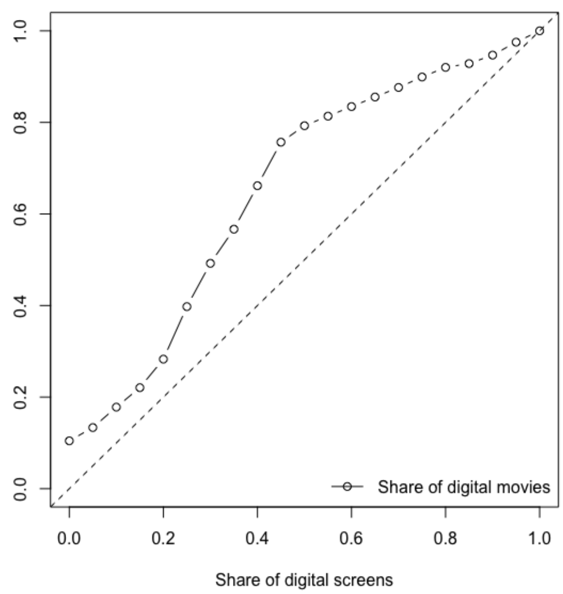
Note: The columns are defined in the same way as in Table 1.6, but the reference group is miniplexes instead of all firms. The predicted rates are obtained by averaging 500 *simulation paths.*

Table 1.8: Predictions using the adoption policy function - Multi/Megaplexes (8-23 screens)

	Aggregate		Inter-firm		Intra-firm	
Year	Data	Prediction	Data	Prediction	Data	Prediction
2006	0.003	0.000	0.037	0.001	0.087	0.156
2007	0.005	0.002	0.043	0.010	0.107	0.206
2008	0.015	0.013	0.106	0.056	0.135	0.240
2009	0.092	0.113	0.207	0.332	0.415	0.344
2010	0.307	0.320	0.670	0.707	0.466	0.475
2011	0.554	0.554	0.920	0.914	0.593	0.637
2012	0.870	0.775	0.963	0.978	0.903	0.825
2013	0.973	0.950	0.989	0.987	0.997	0.966

Note: The columns are defined in the same way as in Table 1.6, but the reference group is multi/megaplexes instead of all firms. The predicted rates are obtained by averaging 500 *simulation paths.*

Figure 1.9: Share of movies available in digital h_t as a function of share of digital screens s_t/S



Distributors' reaction function. Figure 1.9 shows distributors' reaction function (equation (1.4)) as fitted from the data. This relation gives the share of digital movies per period, as a function of the share of digital screens in the industry.[58] One issue with games of technology adoption under network effects is the way the network is initially seeded with the new technology to break the "chicken or the egg" problem. In the case of France, the problem was resolved by the US studios' initial commitment to distribute movies in digital (in Figure 1.9, the intercept is non-zero).

1.7.3.2 Second-step estimates

This section presents estimation results for the total number of screenings $R(\boldsymbol{\tau}_i)$ and profits per movie screening $(\pi_f(\widetilde{\mathbf{x}}_{it}), \pi_d(\widetilde{\mathbf{x}}_{it}))$. These components are combined, as in

[58] Under the no-multihoming assumption, h_t^m equals 0 for all t.

equation (1.6), to obtain theaters' single-period operating profits.

First, estimates for the total number of screenings are presented in Table 1.9. Specification (1) includes firm size S_i and market size $market_i$ (and their interaction) as explanatory variables. Specification (2) augments the model with the art house variable. Specification (3) also includes the number of competing screens S_{-i}.

Table 1.9: Annual number of screenings as function of theater type

	Dependent variable: number of screenings		
	(1)	(2)	(3)
Own screens	2,285*** (102)	2,547*** (129)	2,570*** (130)
Paris - outer suburbs	1,206 (1,346)	2,645* (1,415)	2,747* (1,411)
Urban unit - ≤ 20k inhab and rural	469 (1,733)	960 (1,726)	1,114 (1,723)
Urban unit - ≥ 100k inhab	2,989*** (914)	4,253*** (983)	3,926*** (1,089)
Paris - inner suburbs	1,855 (1,575)	2,778* (1,581)	2,889* (1,577)
Paris	−1,512 (1,219)	−11 (1,315)	−456 (1,456)
Art house		2,986*** (1,008)	3,002*** (1,005)
Competitors' screens			48 (36)
Own screens × Paris - outer suburbs	33 (158)	−216 (173)	−237 (173)
Own screens × urban unit - ≤ 20k inhab and rural	83 (249)	−69 (250)	−97 (250)
Own screens × urban unit - ≥ 100k inhab	−268** (114)	−501*** (132)	−450*** (138)
Own screens × Paris - inner suburbs	−5 (167)	−217 (177)	−236 (176)
Own screens × Paris	797*** (156)	546*** (172)	609*** (179)
Own screens × art house		−520*** (154)	−531*** (154)
Own screens × competitors' screens			−8* (4)
Constant	−5,093*** (753)	−6,691*** (962)	−6,796*** (962)
Observations	399	399	399
R^2	0.909	0.911	0.912
Adjusted R^2	0.906	0.908	0.909
Residual Std. Error	2,726.2 (df = 387)	2,692.6 (df = 385)	2,683.9 (df = 383)
F Statistic	349.7*** (df = 11; 387)	304.3*** (df = 13; 385)	265.7 *** (df = 15; 383)

Note: ***$p < 0.01$;**$p < 0.05$; *$p < 0.1$. *For market dummies, the omitted category is "urban unit with 20 to 100 thousands inhabitants."*

As expected, the annual number of screenings per theater is increasing in theater size, and in market size. Theaters located in Paris and other urban units with more than 100,000 inhabitants host more screenings. Art house theaters also host more screenings, all else equal. The effect of the number of competing screens is not significant. Finally, the negative intercept in all specification indicates potential non-linearities in the effect of size: the number of screenings per screen is increasing in the theater size.[59]

Next, estimates of the parameters entering the single-period profits *per movie screening* and the variance of adoption costs are presented in Table 1.10. The coefficients entering $\pi_f(\widetilde{\mathbf{x}}_{it})$ and $\pi_d(\widetilde{\mathbf{x}}_{it})$ are restricted to be equal $\alpha_i^f = \alpha_i^d$. Profits per screening are decreasing in theater size and are lower for art house theaters, although estimates are not precise. In light of the previous results regarding the total number of screenings, this finding indicates that the utilization rate (i.e., the share of seats occupied per screening) does not vary significantly with theater size. Theater size positively affects profits per screen mainly through its effect on the number of screenings per screen. Fixing market size, profits are decreasing in the number of competing screens, whereas competitors' digital screens do not affect profits (effect of the opposite sign of competitors' screens).

The market dummies are not significant. These estimate indicate that market size mainly affects the number of screenings but not profits per screening. This is the case if the utilization rate (i.e., share of seats occupied per screening) does not vary with market size. Next, cost reductions from digital projection are positive and significant as indicated by the coefficient on "own-share of digital screens." Finally, the standard deviation of adoption costs is €$2,995$, and is relatively smaller than the adoption cost, which is between €$40,000$ and €$84,000$. The estimate for the standard deviation of the adoption costs is close to the average price decrease per period, €3,308, indicating that the dispersion in adoption times for otherwise identical firms is close to one

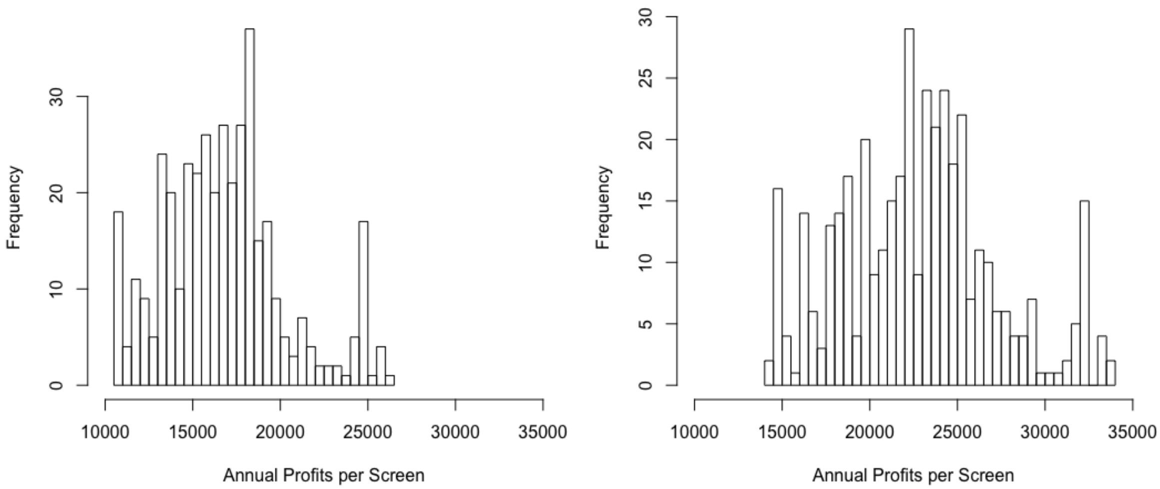
[59]For example, the annual number of screenings per screen for a monopolist non-art house theater in an urban unit with 20,000–100,000 inhabitants is 1,012 for a 4-screen theater, 1,648 for an 8-screen theater, and 1,860 for a 12-screen theater.

Table 1.10: Structural parameter estimates (in 2010 euros)

	Estimate	s.e
Profits per movie screening: $\pi_f(\widetilde{\mathbf{x}}_{it})$		
Constant	12.274	1.885
Own screens	-0.043	0.057
Art house	-0.657	0.827
Competitors screens	-0.015	0.053
Competitors' d-screens	0.015	0.074
Market dummies		
Urban unit - >100k inhab	-0.340	0.840
Urban unit - <20k inhab and rural	0.166	1.275
Paris - inner suburbs	-0.923	0.833
Paris - outer suburbs	-0.629	0.865
Paris	-0.779	1.098
Chain dummies		
Gaumont-Pathe	-0.778	0.823
CGR	3.310	2.241
UGC	-2.295	0.909
Profits per digital movie screening:		
own share of d-screens	2.420	1.054
Adoption cost		
Firm shock: standard deviation σ	$2,995$	$1,441$

Note: Standard errors are calculated using $N_b = 600$ bootstrap samples.

Figure 1.10: Predicted distribution of annual profits per screen (in euros) across theaters: (a) Before the diffusion of digital cinema, (b) after the diffusion of digital cinema



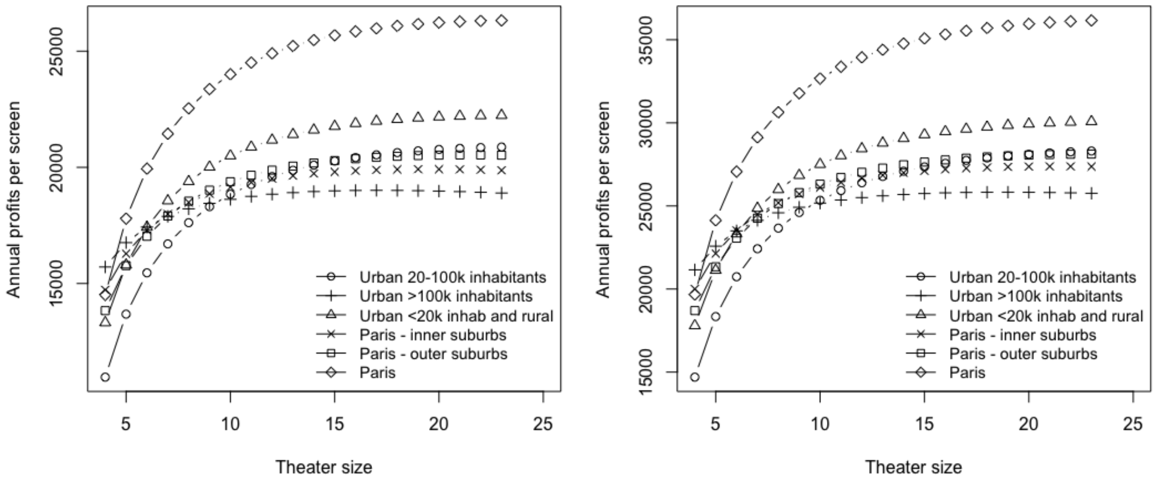
(a) (b)

period.

Profits levels implied by the structural model have the correct order of magnitude. The distribution of annual profits per screen across theaters, predicted by the model, is presented in Figures 1.10a (before the diffusion of digital cinema) and 1.10b (at the end of the diffusion).

Annual profits per screen are approximately between €10,680 and €26,130 before diffusion, and between €14,620 and €33,900 after diffusion. These results are contrasted with estimates using theater chains' income statements, obtained for the years 2014 to 2016 (i.e., after the diffusion period). The chains' annual total profits are divided by the total number of screens. Estimates range between €22,000 and €30,000. Therefore, profits levels implied by the structural model are economically

Figure 1.11: Predicted annual profits per screen as a function of firm size



(a) Before the diffusion (b) After the diffusion

Notes: Predicted profits are calculated fixing other characteristics to: monopolist, non art house theater, not horizontally integrated.

plausible and reasonable.

Finally, combining estimates for the total number of screenings with profits per screening, the model predicts economies of scale in operation. Figures 1.11a and 1.11b show predicted annual profits per screen as a function of theater size and market size. Other firm characteristics are set to: monopolist, non art house, and not horizontally integrated. The combined effect of theater size on the number of screenings per screen and profits per screening implies that profits per screen are increasing in theater size. The marginal effect is decreasing. An increase from 5 to 10 screens increases profits per screen from €15,000 to €19,000 (to €23,000 in Paris), before the conversion to digital.

1.8 Counterfactual Simulations

This section uses the calibrated model to conduct counterfactual simulations and examine the role of the intra-firm margin at the industry and local market levels. First, the intra-firm margin is found to significantly contribute to industry-level diffusion: both the introduction time and the overall dispersion in adoption times across screens depend on within-theater adoption rates. Second, the relationship between local market structure (in terms of theater size and market concentration) and technology adoption is impacted by theaters' ability to adopt at the margin (intra-firm margin effect), controlling for the role of other factors (economies of scale and strategic interactions).[60]

1.8.1 Intra-Firm Margin and Aggregate Diffusion

The first simulation exercise decomposes the diffusion of digital projection over the industry capital stock (screens) into an inter-firm margin (diffusion across firms) and an intra-firm margin (diffusion within firms). Additionally, the counterfactual simulation is used to evaluate the impact of the intra-firm margin on the introduction time, that is, the expected time to first adoption. The analysis indicates both aspects of aggregate diffusion—duration and introduction time—are significantly impacted by the intra-firm margin, pointing to the importance of this margin of adoption in understanding industry diffusion.

A firm's equilibrium adoption behavior is decomposed into an inter-firm (or extensive) margin, in which the firm decides whether to begin using the technology, and an intra-firm (or intensive) margin, in which the firm decides what fraction of its capital stock to convert. Whereas firms' equilibrium behavior reflects both margins, the ex-

[60]This paper does not discuss social welfare. A discussion of potential welfare benefits of digital cinema, in particular via new product offerings, is presented in section 1.3.

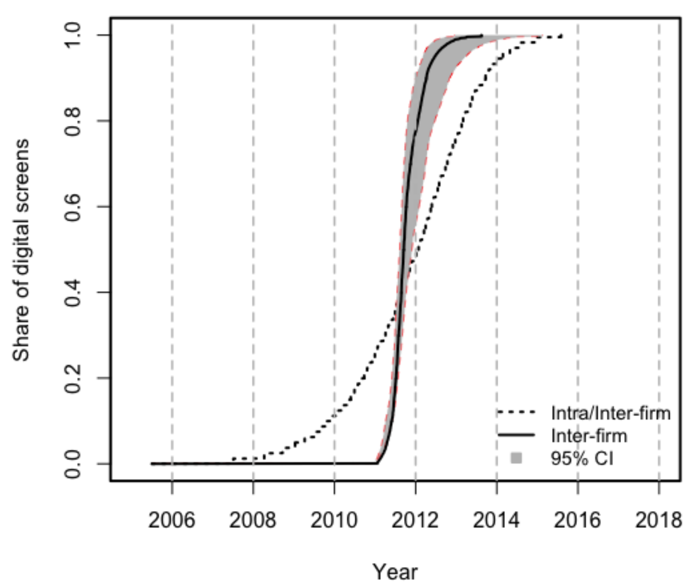
tensive margin can be isolated by simulating firms' adoption behavior restricting the strategy space from $s_{it} \in \{0, 1, ..., S_i\}$ to $s_{it} \in \{0, S_i\}$. In this counterfactual, firms are restricted to convert their whole capital stock at once, conditional on adoption. In this sense, the intra-firm (or intensive) margin is shut down. Because the objective is to decompose the firm's *equilibrium* adoption behavior, only a counterfactual *best response* (to the equilibrium played in the data) is necessary—that is, fixing the price process $\{p_t, t \geq 0\}$ and the share of digital movies $\{h_t, t \geq 0\}$ over time to their values in the equilibrium played in the data. The best response is computed using the value function iteration algorithm of Pakes and McGuire (1994).[61]

Figure 1.12 presents the diffusion curves (industry-wide share of screens equipped with a digital projector over time) under the equilibrium adoption strategy (inter/intra-firm), and the counterfactual adoption best response (inter-firm only). The simulation shows that the introduction time, or expected time to first adoption, is delayed from June 2007 to January 2011. In the counterfactual case, the diffusion is complete by June 2013, whereas in the equilibrium case, diffusion is completed by June 2015. These findings imply that the inter-firm margin, or dispersion in adoption times across firms, explains 31% of the aggregate diffusion (or dispersion in adoption times across units of capital). In other words, 69% of the dispersion in adoption across screens is due to dispersion in adoption within firms.

The introduction lag—defined as the difference in introduction times—between the equilibrium and counterfactual diffusion paths is $1,297$ days and corresponds to 44% of the equilibrium diffusion duration. The introduction time is significantly impacted by firms' ability to gradually convert their capital of screens to the new digital technology. This finding is expected as firms delay their adoption until sufficiently many movies are released in digital, because they are constrained to make a binary

[61] Note that, while a counterfactual *equilibrium*—in which the share of digital movies is endogenously determined—is of separate interest, obtaining such equilibrium is complicated by equilibrium multiplicity due to network effects.

Figure 1.12: Aggregate adoption rate with and without the intra-firm adoption margin



Note: The diffusion curves are obtained by generating 500 *sample paths with a length of* 20 *years. The sample average of these paths are reported. The* 95% *confidence interval is obtained by using the structural parameters corresponding to* 5*th and* 95*th percentiles of the distribution of time to full adoption.*

adoption decision.[62] While the previous result concerns introduction times at the industry level, the next section focuses on introduction times at the local market level (urban and rural unit).

1.8.2 Introduction Lag and Market Structure

This section analyzes the role of the intra-firm adoption margin in explaining differences in adoption times across firms. Such differences have been historically attributed to two important factors: firm size (economies of scale) and market concentration (strategic incentives).[63] This section distinguishes the effect of the intra-firm margin from the latter two factors, and shows that this margin accounts for an important share of the differences in adoption times across firms.

1.8.2.1 Firm size

The adoption data indicates large theaters converted faster than small theaters to digital projection. For example, by 2010, 24% of all miniplex theaters (4–7 screens) had at least one digital screen, against 64% of all multi/megaplex theaters (8–23 screens).[64] Fixing theater and local market characteristics, the presence of scale economies can in part explain the delay in adoption of a small theater relative to a large theater. If profits per digital screen increase with theater size, large theaters will adopt earlier than small theaters. This subsection argues that, in addition to the aforementioned factor, the intra-firm margin plays an important role in explaining this delay: large theaters introduce the technology faster because they are able to convert

[62] Fixing the counterfactual share of digital movies h_t and the adoption cost p_t as in the data is crucial for the introduction lag computed. If, for instance, firms expected h_t to reach 1 in 2008, the counterfactual introduction time would be earlier than 2008.

[63] See Hall and Khan (2002).

[64] See Tables 1.7 and 1.8.

a smaller fraction of their capital stock. It is optimal to do so due to the presence of indirect network effects: the benefit from adopting depends on the availability of digital movies, and initially, only a small fraction of movies is released in digital.

To separate the contribution of the intra-firm margin from that of scale economies, the introduction time—defined as the expected time to first adoption—is simulated in a given local market (set to an urban unit with more than $100,000$ inhabitants), with one monopolist theater owning S_i screens, under (1) the equilibrium adoption strategy (equilibrium played in the data) and (2) the counterfactual best response with no intra-firm margin. Theater characteristics are set to non-art house, not part of a theater chain.

Denote by $T^E_{S_i,m}$ the introduction time in this local market with a monopolist theater with size S_i, under the equilibrium adoption strategy. Similarly, define $T^C_{S_i,m}$ as the introduction time in the same local market with a monopolist theater with size S_i, under the counterfactual best-response strategy. As in the previous section, the theater is best responding to the equilibrium played in the data. In particular, the hardware price and the share of digital movies follow the same processes as in the equilibrium played in the data. By varying the monopolist's size S_i, differences in introduction times between small and large firms, $T^E_{S_i=small,m} - T^E_{S_i=large,m}$ and $T^C_{S_i=small,m} - T^C_{S_i=large,m}$, are obtained.

In the counterfactual with no intra-firm margin, differences in introduction times across theaters with varying size ($T^C_{S_i=small,m} - T^C_{S_i=large,m}$) will reflect differences in period profits stemming from economies of scale. In the equilibrium, these differences in introduction times ($T^E_{S_i=small,m} - T^E_{S_i=large,m}$) will reflect both scale economies as well as the intra-firm margin effect. The results are shown in Table 1.11.[65] The reference firm is a four-screen theater ($small = 4$), and introduction lags are computed relative

[65]The percentiles of $T^E_{4,m} - T^E_{S_i,m}$ and $T^C_{4,m} - T^C_{S_i,m}$ are obtained by using all bootstrap estimates (for the policy rule and the structural parameters) to simulate these introduction lags, and derive their empirical distribution.

to larger theaters ($S_i \in \{8, 10, 12\}$). The introduction lag in the equilibrium case is between 398 and 532 days, whereas the introduction lag in the counterfactual case is between 278 and 304 days. The results indicate the intra-firm margin accounts for 30% to 42% of the introduction lag. The remainder is due to scale economies.

Table 1.11: Introduction lag (in days) by firm size

	Equilibrium: $T^E_{4,m} - T^E_{S_i,m}$ intra and inter-firm margins			Counterfactual: $T^C_{4,m} - T^C_{S_i,m}$ inter-firm margin		
Firm size	Mean	5th	95th	Mean	5th	95th
$S_i = 8$	398.3	314.2	480.4	278.3	141.9	404.8
$S_i = 10$	468.0	383.9	557.8	294.3	146.3	416.8
$S_i = 12$	532.9	434.9	633.4	304.8	150.4	427.5

Note: Summary statistics of the introduction lag between the reference firm ($S_i = 4$) and larger firms, are presented. The introduction lags are computed by averaging 500 *sample paths for each firm. The means of* $T^E_{4,m}$ *and* $T^C_{4,m}$ *correspond to March* 2011 *and September* 2012 *respectively.*

1.8.2.2 Market concentration

The previous exercise focuses on the difference in introduction times between a small and large monopolist theater, controlling for theater and market characteristics. This subsection performs the same decomposition exercise, varying local market concentration instead of theater size. That is, the total capital stock of screens in the local market is kept fixed, while varying the number of (equally-sized) theaters competing in this local market. Differences in introduction times between local markets with different levels of market concentration will stem from scale economies (as the average theater size increases with market concentration), strategic interactions (as a theater's adoption decision depends on its rivals' film and digital screens), and the intra-firm margin effect outlined in the previous subsection. The objective is to evaluate the contribution of the latter to differences in introduction times between local market with different levels of concentration.

Introduction times are computed under (1) the equilibrium adoption strategy (equilibrium played in the data) and (2) the counterfactual best response with no intra-firm margin. Theater characteristics are set to non-art house, not part of a theater chain. The total stock of screens is set to 24 screens. The market size is set to "urban unit with more than 100,000 inhabitants."

Denote by $T^E_{n,m}$ the introduction time in this local market with n equally sized theaters, under the equilibrium adoption strategy (equilibrium played in the data). Similarly, define $T^C_{n,m}$ as the introduction time in this local market with n equally sized theaters, under the counterfactual best-response strategy, with no intra-firm margin. By varying the number of theaters in the market, n, differences in introduction times between markets with different level of concentration are obtained: $T^E_{n_1,m} - T^E_{n_2,m}$ and $T^C_{n_1,m} - T^C_{n_2,m}$, with $n_1 > n_2$.

The number of competitors in the market, n, takes values in $\{1, 2, 3, 6\}$. The reference market is the less concentrated market with six four-screen theaters ($n_1 = 6$), and is compared to more concentrated markets: monopoly, duopoly, and three-firm oligopoly ($n_2 \in \{1, 2, 3\}$). The introduction time is simulated for all markets, and differences (or introduction lags) are shown in Table 1.12.[66] The introduction lag under the equilibrium adoption strategy is between 205 and 256 days, whereas the introduction lag under the counterfactual is between 78 and 116 days. These results

[66]The percentiles are obtained as in Table 1.11.

indicate that the intra-firm margin accounts for 43% to 69% of the introduction lag between the reference market ($n_1 = 6$) and more concentrated markets ($n_2 \in \{1, 2, 3\}$). The rest of the lag is explained by economies of scale (as average theater size decreases with n) and strategic interactions between theaters.

Table 1.12: Introduction lag (in days) by number of firms

	Equilibrium: $T^E_{6,m} - T^E_{n_2,m}$ intra and inter-firm margins			Counterfactual: $T^C_{6,m} - T^C_{n_2,m}$ inter-firm margin		
Market	Mean	5th	95th	Mean	5th	95th
$n_2 = 3$	205.5	136.6	269.5	116.7	64.5	184.2
$n_2 = 2$	210.7	106.7	298.1	113.1	50.3	189.7
$n_2 = 1$	256.2	-14.7	528.2	78.5	5.0	160.8

Note: Summary statistics of the introduction lag between the reference market ($n_1 = 6$) and more concentrated markets, are presented. The introduction lags are computed by averaging 500 sample paths for each market. The means of $T^E_{6,m}$ and $T^C_{6,m}$ correspond to August 2009 and February 2012 respectively.

1.9 Conclusion

This paper investigates the role of network effects in intra-firm technology adoption. The within-firm share of capital equipped with the new technology increases with the availability of complementary software. In turn, software availability increases as the hardware technology diffuses across and within firms. The study focused on the digitalization of the movie distribution and exhibition industries because they offer an ideal setting for studying this mechanism.

The analysis shows that when software availability matters for adoption, the intra-firm margin explains a significant share of the aggregate diffusion phenomenon: in terms of dispersion in equilibrium adoption times, as well as time to first adoption. Second, because firms with varying size differ in their ability to gradually roll out the technology, intra-firm adoption dynamics and the presence of capital indivisibilities amplify the positive relationship between firm size and early adoption; i.e., they explain a significant share of the delay in adoption of smaller firms.

Two implications can be derived. First, the results underline the fact that designing policies (e.g., technology subsidies) that encourage faster technological diffusion within firm may be as important as designing policies that encourage faster diffusion across firms. Second, in a network industry or hardware-software system, new technologies are adopted earlier if individual firms are larger, because they are able to adopt at the margin. This sheds new light on potential effects of antitrust and merger control on firms' ability to adopt innovations.

The current study could be extended in several directions, one important is noted. An avenue of research would be to analyze the role of vertical relations between the software and hardware industries (here, distribution and exhibition) and the effect of such vertical relations on firms' incentives to adopt. These vertical relations are particularly relevant if the payoff from adopting is asymmetric between the software and hardware markets and involves transaction costs.

CHAPTER 2

A Study of Umbrella Damages from Bid-Rigging

2.1 Introduction

When a cartel does not include every firm competing in an industry, non-cartel firms can set their own prices higher than they would otherwise have been able to under competitive conditions. This is in particular the case in markets where contracts are awarded via first-price procurement auctions. When the results of such procurement procedures (bids and bidders identities) are not concealed from non-cartel firms, they may serve as an indication of the prevailing price level when future contracts are procured for. Consequently, non-cartel bidders benefit from the protection of the cartel's inflated bidding, and operate "under the cartel's umbrella." Purchasers from non-cartel bidders will still pay a price that exceeds what the market price would be in the absence of collusion. In this sense, damages inflicted by non-cartel bidders broaden the scope of cartel damages. Nonetheless, empirical research investigating the importance of such damages remains scarce. This paper conducts a detailed study of umbrella damages by examining the bidding behavior of non-cartel bidders facing the Texas school milk cartel between 1980 and 1992.

Bid-rigging was a pervasive phenomenon in auctions for the supply of milk to schools, at least until the early 1990s. According to Porter and Zona (1999), investigations were conducted in more than twenty states across the US, more than $90 millions of fines were levied, while about 90 people were sent to jail for sentences lasting 6 months on average. The Texas milk cartel is well-suited for analyzing damages inflicted by non-cartel bidders. First, two essential conditions are met: the cartel was not all-inclusive and the auction format was first-price sealed bid.[1] In particular, bids and bidder identities are publicly announced. This public information enables non-cartel firms to learn and adjust their bids to the supra-competitive levels sustained by the cartel. Second, all firms involved in the cartel were convicted, which allows to isolate non-cartel bidders and focus on their bidding behavior. Third, the dataset collected by the Antitrust Divison of the Department of Justice spans markets with and without cartel operations, which enables to identify the effect of the cartel's price umbrella on non-cartel bidders. Finally, the dataset is rich enough to allow an assessment of damages inflicted by the cartel per se, as well as damages inflicted through non-cartel bidders.

Reduced form analysis of the bid data reveals that, controlling for auction and bidder observed heterogeneity, the largest non-cartel firm bid significantly higher when facing the cartel. Further investigation of cartel and umbrella damages and inefficiencies requires estimation of a structural model. Damages to the auctioneer are decomposed into (1) *cartel damages*, defined as damages in auctions won by the cartel and ($2a$) *outsider damages*, defined as damages in auctions won by the non-cartel firm, when it is the lowest cost bidder and ($2b$) *misallocation damages*, defined as damages in auctions won by the non-cartel firm, when the cartel bidder is the lowest cost bidder. Case ($2b$) arises because partial collusion introduces asymmetry among bidders in the first-price procurement setting: the cartel bidder has a stronger incentive to inflate his bid above his cost than the non-cartel bidder. As a result, the winner is not necessarily the lowest cost bidder, and the auction is no longer efficient. ($2a$) and ($2b$) form what is defined as umbrella damages. Because the cartel's internal structure is unknown,

[1] In ascending or second price auctions, bidding one's private valuation is a dominant strategy irrespective of the existence of the cartel, therefore umbrella damages do not arise. Note for umbrella damages to arise in first-price procurement auction, one need not assume that non-cartel firms know the existence of the cartel per se, but only that they know the equilibrium bid distribution of their lowest bidding opponent, which can be inferred from past auctions.

bounds on outsider and misallocation damages are derived assuming two extreme cases for the cartel mechanism: either the mechanism is efficient (i.e., the lowest cost cartel member is the cartel bidder in the target auction), or the mechanism is inefficient (i.e., the cartel bidder is selected randomly among cartel members).[2]

The structural analysis shows that per contract, outsider damages (conditional on the non-cartel firm winning) are at least 47% of cartel damages (conditional on the cartel winning). This lower bound is obtained with an efficient cartel. If the cartel is inefficient, outsider damages can be as large as cartel damages. Misallocation damages are estimated to be as large as 64% when the cartel is efficient, and the auction is asymmetric.[3] With respect to inefficiencies due to the asymmetry across bidders, losses are estimated to be 3.7% or $2,909 per contract. The structural estimates show that conditional on the non-cartel firm winning, prices are inflated by 2.9% to 8.5% relative to the competitive winning bid. These bounds are consistent with the reduced form estimates of a 6% overcharge. These results points to a cartel mechanism that was not fully efficient, but far from inefficient.[4]

From a competition law perspective, the complexity of proving umbrella claims has been one of the argument stifling the recognition by US and European civil courts of the ability of the purchasers from non-cartel firms to pursue treble damages from colluders. This paper sheds new light on this debate by providing a case study of umbrella damages, emphasizing the type of data and methodology that render the estimation of damages possible and far from speculative. The US and European

---

[2]While the definition of an efficient cartel mechanism is straightforward, it is less obvious for an inefficient cartel mechanism. In the inefficient case, the cartel bidder could be selected randomly as is assumed in this paper. But one could think of other inefficient mechanisms such as the one in which the cartel bidder is the least cost efficient member.

[3]If the cartel is inefficient, there is no asymmetry between bidders under the maintained assumptions and therefore no efficiency losses.

[4]This is consistent with Pesendorfer (2000) which shows that although the cartel didn't use sidepayments, it managed to retain quasi-efficient collusive rents through market division.

competition laws have recently taken divergent paths in their treatment of whether the civil liability in damages of the cartel members extend to umbrella damages. On June 5^{th} 2014, the ECJ handed down the awaited judgement in the elevator cartel case (*Kone AG and Others v. ÖBB-Infrastruktur AG*), stating the right of plaintiffs to compensation for umbrella damages.[5] In its press release No 79/14, the ECJ states:

> [...] where it has been established that the cartel is, in the circumstances of the case and, in particular, the specific aspects of the relevant market, liable to result in prices being raised by competitors not a party to the cartel, the victims of this price increase must be able to claim compensation for loss sustained from the members of the cartel.

At the same time, the ECJ emphasizes the "high hurdles in terms of the burden of proof that await" any umbrella claimants.[6]

In the US, competition law is inconsistent in its standing vis-a-vis umbrella claims. The ability of purchasers from non-cartel firms to recover treble damages from the conspirators under section 4 of the Clayton Act is uncertain because of the Supreme Court decision in *Illinois Brick Co. v. Illinois.*[7] In the former case, the Court ruled that indirect purchasers (downstream buyers) may not sue on a theory that a price-fixing overcharge has been passed on to them by intermediate sellers purchasing from upstream colluders. Although Illinois Brick did not deal with the standing of purchasers from nonconspiring competitors of the antitrust violator, that case was

[5] The elevator cartel, which involved the conclusion of anticompetitive agreements between major European manufacturers of elevators and escalators, more specifically, Kone, Otis, Schindler and ThyssenKrupp, operated in several Member States of the European Union over a period of many years. The European Commission uncovered that cartel in 2003 and, in 2007, imposed fines for the elevator cartel's practices in the Belgian, German, Netherlands and Luxembourg markets.

[6] Opinion of the Advocate General Kokott.

[7] Section 4 of the Clayton Act provides that "any person ... injured in his business or property" by an antitrust violation may bring an action for treble damages.

relied on heavily by the Third Circuit to deny standing to purchasers from non-cartel firms.[8]. Again, one of the policy reasons underlying the *Illinois Brick* doctrine revolves around the complexity of tracing out the causal link between the antitrust violation and non-cartel firms' response.[9] This paper shows that in the particular case of bid-rigging of procurement auctions, bid data can be leveraged to estimate the size of damages when buyers do not contract directly with colluders.

This paper also relates to the debate around auction format choice. A commonly advanced argument in favor of sealed bid auctions is that open auctions are more prone to collusion because conspirators can immediately punish any deviations.[10] This reasoning does not account for the fact that given (partial) collusion, the auctioneer will suffer damages of a greater scope in sealed bid auctions. As non-cartel firms adjust their bidding strategy to the cartel, damages to the auctioneer extend to contract won by non-conspiring firms. While in ascending auctions, the auctioneer suffers damages only when the cartel is able to suppress the second highest bid (the auctioneer may benefit in some cases from the cartel's overbidding as found by Asker (2010)). This work shows that damages in procurement auctions won by non-conspiring firm can potentially form a non-negligible fraction of overall cartel damages. The potential merit of sealed bid relative to open auctions is therefore nuanced with regard to these findings.

[8]Judgments finding in favour of liability: United States Court of Appeals (Seventh Circuit), United States Gypsum Co. v. Indiana Gas Co., 350 F.3d 623, 627 (2003); United States Court of Appeals (Fifth Circuit), In re Beef Industry Antitrust Litigation, 600 F.2d 1148, 1166 (1979), State of Washington v. American Pipe Construction Co., 280 F. Supp. 802 (D. Haw. 1968), Pollock v. Citrus Associates, Inc. 512 F. Supp. 711 (S.D.N.Y. 1981). Judgements finding against such liability, on the other hand, include: United States Court of Appeals (Third Circuit), Mid-West Paper Products Co. v. Continental Group Inc., 596 F.2d 573, 597 (1979); United States District Court (District of Columbia), Federal Trade Commission v.Mylan Laboratories, 62 F.Supp.2d 25, 39 (1999).

[9]Many states, including California, have enacted Illinois Brick-repealer legislation providing indirect purchasers standing to sue for antitrust violations.

[10]See the discussion in Athey, Levin, and Seira (2011) for timber auctions.

The US milk cartels were the subject of various papers in the empirical literature on bid-rigging. Pesendorfer (2000) examines the Florida and Texas school milk cartels, and shows that the data (in particular market shares and incumbency rates) is consistent with a strong cartel in Florida (in the sense that sidepayments were used between cartel members) and a weak cartel in Texas (no sidepayments). This paper differs from the former in two dimensions: first, Pesendorder focuses on the Dallas-Fort Worth (DFW) area while the dataset used in this paper includes in addition the Waco and San Antonio areas in which the cartel was not operating, providing a set of competitive auctions which are useful to predict counterfactuals; second, this paper focuses primarily on outsiders' response to the cartel's bidding behavior. Hewitt, McClave, and Sibley (1996) demonstrate that the high incumbency rates in Texas (the supplier of a given school district doesn't change from year to year in many cases) can only be explained by collusion. Lee (1999) finds evidence of complementary bidding and high incumbency premia in the DFW school milk market. Porter and Zona (1999) test for the presence of collusion in the Ohio school milk market by comparing defendants firms in Cincinnati to a control group of non-defendants and compute estimate of cartel damages. Lanzillotti (1996) provides a review of US milk cartel cases, and shows that several features of the bids in Kentucky are indicative of collusive behavior.

This paper is more broadly related to the empirical literature on bidding rings. A first strand in this literature aims at providing econometric tests of collusion: Baldwin, Marshall, and Richard (1997) for timber auctions, Porter and Zona (1993) for highway construction, Bajari and Ye (2003) for contracts in the seal coat industry, Athey, Levin, and Seira (2011) for timber auctions. A second strand in the literature focuses on the internal organization of bidding rings. Asker (2010) studies equilibrium bidding and sidepayments in knock-out auctions held privately by the New York stamp cartel before the actual auction. Kwoka Jr (1997) analyzes bids and sidepayments in knock-out auctions held by a real estate ring. Finally, a more recent strand of the literature, closer to this paper, studies the interaction of partial cartels and non-cartel bidders. Harrington, Hüschelrath, and Laitenberger (2016) analyzes how the German cement cartel controlled the expansion of non-cartel supply (from Eastern European countries) by sharing the collusive rents with German importers.

The empirical section of the paper relies on results from the literature on the structural estimation of auctions. In particular, the non-parametric estimation of the bidders' underlying cost distribution follows the methodology introduced in Guerre, Perrigne, and Vuong (2000). Observed auction heterogeneity is also controlled for by using the first-stage regression technique developed in Haile, Hong, and Shum (2006). Finally, the empirical section makes use of some of the numerical methods developed in the computational literature on asymmetric auctions. Recent contributions in this literature are: Li and Riley (2007), Gayle and Richard (2008), and Fibich and Gavish (2011).

The next section describes the theoretical model of first-price procurement auctions with asymmetric bidders and analyzes the effect of collusion, cartel size, and cartel mechanism on the non-cartel firm profits and bidding behavior. Section 2 presents the school milk market and the relevant factors affecting the cost structure. Section 3 describes the dataset. Section 4 examines the largest non-cartel firm's bidding behavior through a reduced-form approach, and shows that all else equal, the largest non-cartel bidder overbids in school district where the cartel is operating. An assessment of cartel and umbrella damages is presented in section 5 using a structural approach.

2.2 A Theoretical Model

This sections investigates properties of a non-cartel firm's bidding behavior when facing a cartel bidder.[11] The section builds on the theoretical literature on asymmetric first price auctions. In particular, Maskin and Riley (2000) and Lebrun (1999, 2006) derive existence and uniqueness results, as well as comparative statics results for the equilibrium bid functions.[12] This section is also related to the theoretical literature on bidding rings in first-price auctions: McAfee and McMillan (1992) characterize the optimal mechanism for strong and weak cartels, while Marshall and Marx (2007) compares first and second price auction formats when a cartel is not all-inclusive and may or may not be able to control the bids of its members.

The questions of interest are: when does non-cartel firms benefit from their competitors colluding ? How are non-cartel firm's bidding and profits affected by the cartel size and the cartel mechanism (relative to a situation with no cartel)?

The case of a single non-cartel firm is considered for simplicity, and in anticipation of the empirical part. Risk-neutral firms are bidding for a single contract in a first-price procurement setting. There is no reserve price. Denote by 1 the cartel bidder, and by 2 the non-cartel bidder.[13] For $i \in \{1, 2\}$, firm i's cost c_i is drawn from a distribution F_i with support $[\underline{c}, \overline{c}]$. F_i has a continuous density f_i strictly positive on $(\underline{c}, \overline{c}]$. Cost are drawn independently across bidders, and are private. Existence and uniqueness of an equilibrium in strictly increasing strategies was proved by Maskin and Riley (2000) and Lebrun (1996,1999,2006) among others.

[11] The cartel bidder is defined as the cartel member selected to bid in the auction on behalf of the cartel.

[12] Other references include Vickrey (1961), Griesmer and Levitan (1967) who study asymmetric uniform distributions, Plum (1992) who study a class of power distributions and more recently: Bajari (2001) and Cheng (2006). Athey (2001) derives more general existence results.

[13] Note that this setting does not rule out the possibility that other cartel members submit "non-serious" or "complementary" bids. If that is the case, the non-cartel firm knows that it is facing only one "serious" bid from the cartel.

Denote by β_i bidder i's equilibrium bidding strategy, and by $\phi_i = \beta_i^{-1}$ the corresponding inverse bid function. Note that if bidder j bids according to ϕ_j and bidder i submits a bid b, then the latter wins if and only if $c_j > \phi_j(b)$. Thus, bidder i's expected profit from bidding b is given by:

$$\pi(b; c_i) = (b - c_i)\Pr(c_j > \phi_j(b)) = (b - c_i)(1 - F_j(\phi_j(b))) \tag{2.1}$$

The first-order condition with respect to b (at $c_i = \phi_i(b)$) is :

$$\frac{f_j((\phi_j(b))}{1 - F_j((\phi_j(b))}\phi_j'(b) = \frac{1}{b - c_i} \quad \text{with boundary condition } \phi_i(\bar{c}) = \bar{c}. \tag{2.2}$$

Combining optimality conditions for the two bidders, the equilibrium bid functions solve the following system of differential equations:

$$\begin{cases} \frac{1}{b-\phi_1(b)} = \frac{f_2(\phi_2(b))}{1-F_2(\phi_2(b))}\phi_2'(b) \\ \frac{1}{b-\phi_2(b)} = \frac{f_1(\phi_1(b))}{1-F_1(\phi_1(b))}\phi_1'(b) \end{cases} \tag{2.3}$$

with right-boundary conditions : $\phi_1(\bar{c}) = \phi_2(\bar{c}) = \bar{c}$. Maskin and Riley (2000) show that inverse bid functions must satisfy the additional condition that the minimum bid of all bidders is the same: $\phi_1(\underline{b}) = \phi_2(\underline{b}) = \underline{c}$ for some unknown $\underline{b}$. Although analytical solutions of problem (2.3) are in general not available, one can derive properties on the equilibrium bid functions.

The first property, proved by Maskin and Riley (2000) and Pesendorfer (2000), can be used to compare the cartel and non-cartel firms bid functions. Assume that the bidders' cost distributions can be ordered according to hazard rate dominance, i.e

$$\frac{f_1(c)}{1 - F_1(c)} > \frac{f_2(c)}{1 - F_2(c)} \quad \text{for all } c \tag{2.4}$$

This implies that, conditional on having a cost above c, the cartel bidder is more likely to have a low cost than the non-cartel bidder.

Proposition 1. *Under the hazard rate dominance assumption:*

1. *The cartel firm bids higher than the non-cartel firm: $\beta_1(c) > \beta_2(c)$ for all $c \in (\underline{c}, \overline{c})$*

2. *Denoting by G_i bidder i's equilibrium bid distribution (and g_i the corresponding density function), the cartel firm's bid distribution dominates the non-cartel firms bid distribution (in the hazard rate sense): $\frac{g_1(b)}{1-G_1(b)} > \frac{g_2(b)}{1-G_2(b)}$*

Proof. See Propositions 3.3 and 3.5 in Maskin and Riley (2000) or Krishna (2006). □

Part 1 of the proposition implies that the non-cartel firm may win despite not having the lowest cost among the bidders. This result in an inefficient allocation. Part 2 of the proposition implies in particular that the non-cartel firm's equilibrium bid distribution first-order stochastically dominates the cartel firm's equilibrium bid distribution. As noted by Pesendorfer (2000), the hazard rate dominance condition will be satisfied for instance when all firms (cartel and non-cartel firms) are ex-ante symmetric and the cartel mechanism is efficient, i.e when the cartel bidder has the lowest cost among cartel members. In this case, $F_1(c) = 1 - (1 - F(c))^n$ where F is the ex-ante symmetric distribution of bidders, and n is the number of cartel members. One can see that $\frac{f_1(c)}{1-F_1(c)} = \frac{nf(c)}{1-F(c)} > \frac{f(c)}{1-F(c)} = \frac{f_2(c)}{1-F_2(c)}$.

Studying the magnitude and determinants of umbrella damages requires the comparison of competitive equilibrium bid functions, i.e bid functions when the firms do not collude, with the collusive bid functions derived above. Denote by n the number of cartel members, or cartel size. For simplicity, assume that all bidders, whether or not in the cartel, are $ex-ante$ symmetric, and that there is a single non-cartel firm. All firms draw their cost from a distribution F with support $[\underline{c}, \overline{c}]$. F has a continuous density f strictly positive on $(\underline{c}, \overline{c}]$. Under these assumptions, the competitive auction in which firms do not collude is a symmetric auction with $n + 1$ bidders. There is a unique equilibrium in strictly increasing strategies (see Krishna (2006)).

Assumption 1 (*Ex-ante* Symmetric IPV)**.** *Across bidders, costs are symmetric (identically distributed according to the distribution F), independent, and private.*

Denote by β the symmetric equilibrium bidding strategy, and by $\phi = \beta^{-1}$ the corresponding equilibrium inverse bid function. In this case, if a bidder's cost is c_i, bidding b yields expected profits given by:

$$\pi(b; c_i) = (b - c_i)\,\Pr(c_j > \phi(b), \forall j \neq i) = (b - c_i)\,(1 - F(\phi(b)))^n \tag{2.5}$$

The first-order condition with respect to b (at $c_i = \phi(b)$) is :

$$\frac{nf(\phi(b))}{1 - F(\phi(b))}\phi'(b) = \frac{1}{b - c_i} \quad \text{with boundary condition } \phi(\bar{c}) = \bar{c}. \tag{2.6}$$

Let $G(b) = F(\phi(b))$ denote the distribution of a firm's equilibrium bid. Let $g(b)$ denote the corresponding density function. Since bidders are symmetric, this distribution doesn't depend on i. The first-order condition can be rewritten:

$$c_i = b - \frac{1 - G(b)}{ng(b)} \tag{2.7}$$

Equation (2.7) expresses the individual private cost c_i as a function of the individual equilibrium bid b, and the distribution of equilibrium bid G. This mapping is at the core of the structural estimation of the cost distribution F from the observed distribution of bids G (see Guerre, Perrigne, and Vuong (2000) and step 2 in section 2.6.2).

2.2.1 Effect of collusion on the non-cartel firm's profits

This subsection examines the effect of collusion between a strict subset of bidders on the bidding and profits of the non-colluding bidder—or, outsider. First, the outsider's interim payoff (i.e., the expected profits conditional on his cost) always increases when his competitors form a cartel.

Define the outsider's equilibrium interim payoff when facing n competitors (no collusion):

$$\pi(c) = \max_b (b-c)\left[1 - F(\phi(b))\right]^n$$

Similarly, define the outsider equilibrium interim payoff when facing the cartel bidder (the n competitors collude):

$$\widetilde{\pi}(c) = \max_b (b-c)\left[1 - F_1(\phi_1(b))\right]$$

Assume first that the cartel mechanism is efficient, i.e $F_1(c) = 1-(1-F(c))^n$. Then we have the following lemma:

Lemma 1. *If the cartel mechanism is efficient, the outsider's interim payoff is strictly larger when his competitors collude:*

$$\widetilde{\pi}(c) > \pi(c) \quad \textit{for all } c \in [\underline{c}, \overline{c})$$

Proof. Note that if ϕ solve the symmetric $n+1$ bidders procurement auction, it also solves the two bidder procurement auction in which bidders' cost are drawn from $1-(1-F(c))^n$. When the n firms collude efficiently, we obtain a 2 bidders asymmetric auction in which bidders' costs are drawn from F for the non-cartel firm and $1-(1-F(c))^n$ for the cartel bidder. Since F and $1-(1-F(c))^n$ can be ordered stochastically in the hazard rate sense, Corollary 1 of Lebrun (1998) applies to obtain the strict inequality for the non-cartel firm profits. □

In general, the cartel mechanism might not be efficient. Assume that the cartel bidder's cost distribution F_1 (given a mechanism) satisfies the *conditional stochastic dominance relation*[14]

$$\frac{d}{dc}\left(\frac{1-F_1(c)}{(1-F(c))^n}\right) > 0 \quad \text{for all } c \in (\underline{c}, \overline{c}] \tag{2.8}$$

[14]Note that this relation can also be rewritten in terms of hazard rate dominance.

Proposition 2. *If (2.8) holds, the outsider's interim payoff is strictly larger when his competitors collude:*

$$\widetilde{\pi}(c) > \pi(c) \quad \textit{for all } c \in [\underline{c}, \overline{c})$$

Proof. F_1 and $1 - (1 - F(c))^n$ satisfy the assumption of conditional stochastic dominance, therefore Corollary 1 in Lebrun (1998) implies that the outsider's interim payoff is strictly larger when facing the cartel with the mechanism yielding F_1 than when facing an efficient cartel. Combining this observation with the previous lemma gives the result. □

2.2.2 Effect of the cartel size and cartel mechanism on the non-cartel firm's bidding

In this subsection, comparative statics for the non-cartel firm equilibrium bid function are presented. In particular, two features are investigated: the cartel size and the cartel mechanism. Comparative statics in first-price asymmetric auctions have been studied by Lebrun (1998). The results presented here are an application of the main theorem proved in the latter paper to the specific case of a cartel bidding in a procurement auction.

First, let the cartel mechanism be fixed. Using the same notations introduced above, denote by 1 the cartel bidder, and by 2 the non-cartel bidder (F_i, for $i \in \{1, 2\}$ the corresponding cost distributions, which satisfy the assumptions of the model). The dependency of the cartel bidder's cost distribution on the cartel size n is explicitly represented as: $F_1(.|n)$. In particular under Assumption 1, if the cartel is efficient: $F_1(c|n) = 1 - (1 - F(c))^n$ (minimum cost among n symmetric bidders). If the cartel is inefficient: $F_1(c|n) = F(c)$ (cartel bidder selected randomly).

Assume the cartel mechanism is such that $F_1(.|n)$ satisfies the following *conditional stochastic dominance condition*note1

$$\frac{d}{dc}\left(\frac{1-F_1(c|n)}{1-F_1(c|n+1)}\right) > 0 \quad \text{for all } c \in (\underline{c}, \overline{c}] \tag{2.9}$$

Proposition 3. *Assume that (2.9) holds. Let the bid functions and their inverses at the unique equilibrium when the cartel size is n be denoted $\beta_1(.|n)$,$\beta_2(.|n)$ and $\phi_1(.|n)$,$\phi_2(.|n)$ respectively. Then*

1. *As the cartel size increases, the non-cartel firm bids more aggressively:*

$$\beta_2(c|n) > \beta_2(c|n+1) \quad \textit{for all } c \in [\underline{c}, \overline{c})$$

2. *The larger cartel's bid distribution is stochastically dominated by the smaller cartel's bid distribution:*

$$F_1(\phi_1(b|n)|n) < F_1(\phi_1(b|n+1)|n+1) \quad \textit{for all } c \in [\underline{b}, \overline{c})$$

where $\underline{b} = \beta_1(\underline{c}|n) = \beta_2(\underline{c}|n)$

3. *As the cartel size increases, both the cartel and non-cartel bidders' interim profits (conditional on their private cost) decrease*

Proof. Follows directly from Theorem 1 and Corollary 1 in Lebrun (1998). □

Condition 2.9 holds when the cartel mechanism is efficient. However, the condition doesn't hold if the cartel mechanism is inefficient (F_1 is independent of n).

Next, let the cartel size be fixed. The effect of the cartel mechanism on the non-cartel firm bidding can be analyzed. Consider two mechanisms implying a cartel bidder's cost distribution of either F_1 or $\widetilde{F}_1$. Assume the two distribution satisfy the *conditional stochastic dominance relation*note1

$$\frac{d}{dc}\left(\frac{1-F_1(c)}{1-\widetilde{F}_1(c)}\right) > 0 \quad \text{for all } c \in (\underline{c}, \overline{c}] \tag{2.10}$$

This condition implies in particular that $\widetilde{F}_1$ is first order stochastically dominated by F_1. For instance, if $F_1(c) = F(c)$ (inefficient cartel) and $\widetilde{F}_1(c) = 1 - (1 - F(c))^n$ (efficient cartel), the condition holds.

Proposition 4. *Assume (2.10) holds. Let the bid functions and their inverses at the unique equilibrium when the cartel bidder's distribution is F_1 (resp. $\widetilde{F}_1$) be denoted (β_1,β_2) (resp. $(\widetilde{\beta}_1,\widetilde{\beta}_2)$) and (ϕ_1,ϕ_2) (resp. $(\widetilde{\phi}_1,\widetilde{\phi}_2)$). Then*

1. *The non-cartel firm bids more aggressively when the cartel bidder's cost distribution is $\widetilde{F}_1$ than when it is F_1:*

$$\beta_2(c) > \widetilde{\beta}_2(c) \quad \textit{for all } c \in [\underline{c}, \overline{c})$$

2. *The cartel's equilibrium bid distribution can be ordered according to first-order stochastic dominance:*

$$F_1(\phi(b)) < \widetilde{F}_1(\widetilde{\phi}_1(b)) \quad \textit{for all } c \in [\underline{b}, \overline{c})$$

   where $\underline{b} = \beta_1(\underline{c}) = \beta_2(\underline{c})$

3. *Cartel and non-cartel bidders' interim profits (conditional on their private cost) are lower under $\widetilde{F}_1$ than under F_1.*

Proof. Follows directly from Theorem 1 and Corollary 1 in Lebrun (1998). □

The intuition for Proposition 3 and 4 is as follow. In both cases, as the cartel bidder is made stronger (either by increasing the size of the cartel or by making the cartel mechanism more efficient), the non-cartel firm responds by bidding more aggressively (part (1) of the propositions). As a consequence, the cartel bidder's interim payoff decreases. In equilibrium, the cartel best response is such that it is more likely to bid lower: the new equilibrium bid distribution is first-order stochastically dominated by its initial equilibrium bid distribution (part (2) of the propositions). This results in the non-cartel firm interim payoff being lower as well.

2.3 The School Milk Market

Most public school districts in Texas use procurement auctions to allocate contracts for the supply of school milk. Every year, between May and August, each district sets a first-price sealed bid procurement auction, specifying contract characteristics such as the estimated quantities (in half pint) of milk to supply by milk categories, the number of delivery points, the contract period, the delivery times, and whether a cooler has to be provided.[15] Firms have one month to prepare their bid, which is a price per half pint for each category of milk. A bid can be escalated or fixed. Escalated bids are indexed to the price of raw milk, to insure the milk supplier against potentially large fluctuations of the price of raw milk over the contract period. Bidders have to sign a non-collusive affidavit, stating that they did not partake in any communication with other bidders regarding prices or participation, and that they will not give or receive any sidepayments. On the day of the letting, all submitted bids are opened and the bidders identities and bids are publicly announced. Dairy distributors have to deliver the packaged milk to various customers (e.g., retail stores, government agencies, and schools). Retail stores are the main revenue source for distributors. School milk contracts typically form 10% to 20% of a distributors' revenue.

By nature, the school milk market is remarkably exposed to collusion. Firms compete only on prices as the contracts terms (quantity and quality) are fixed and the product homogenous. There are many small contracts to be gained facilitating market division. Bids and bidder identities are publicly announced which helps detecting price cuts by cartel members and increase transparency of prices. Firms frequently interact as auctions are not held on the same day, which permits retaliation in case of cheating. The demand for milk is inelastic so prices increases will yield higher profits and are unlikely to face any buyer resistance. Finally, the market is relatively

[15] The main categories of milk being: whole white, whole chocolate, low-fat white, low-fat chocolate.

concentrated helping coordination.

Some of the aforementioned market features also enables potentially large umbrella damages. Indeed, these damages stem from non-cartel firms adjusting their bidding behavior to the supra-competitive levels sustained by the cartel. This is feasible since all bids and bidder identities are publicly announced, which results in non-cartel firms learning the price level in the rigged districts and adjusting to it. The latter channel is reinforced by the high frequency of interactions, due to the large number of contracts every year.

Next, the cost structure of milk processors is described. In anticipation of the structural analysis, it is useful to decompose a milk processor's cost when bidding for a specific contract. This cost can be decomposed into:

1. A component common to all firms bidding for a contract: this component may depend on *observed auction characteristics* such as the quantity of milk to supply, whether bids can be escalated, whether coolers and straws have to be provided, the number of school within the district (which affects the quantity to supply and the number of delivery points), as well as the number of deliveries per week.

   Additionally, this cost component depends on the price of raw milk, which is regulated by federal order, as well as processing, packaging and labor costs which, according to industry experts, are constant across firms in the market.[16]

2. An idiosyncratic component specific to the milk processor: this cost first depends on the distance between the milk processing plant and the school district. More precisely, it depends on how close the school district is to the firm's dis-

[16]A federal milk order sets a uniform minimum price for raw milk in the area. This price is typically increasing in the distance from the Midwest. In this paper, the marketing area was known as the Texas Milk Marketing Order. Within the marketing area, price differ by a fixed proportion from one zone to the other. The price of raw milk is around 7 cents per half-pint.

tribution route. This distribution route depends on the firm's current portfolio of clients (e.g., government agencies, military bases, and schools). Second, the idiosyncratic component depends on the firm current capacity utilization. If the firm is near capacity, winning an extra contract may signify employing an additional truck and driver which increases largely the cost of fulfilling the contract. Finally, idiosyncratic costs include a firm's efficiency in packaging, loading trucks, managing the machinery etc.

If the common component is controlled for, the cost structure falls in the independent and private value framework. Further discussion of this assumption is found in section 2.6.3.2.

A description of the conspiracy as well as the main actors in the industry is provided next. In 1992 and 1993, nine milk processors accused of collusion in the DFW market area reached settlement with the State.[17] The cartel included the main suppliers with plants in the DFW market area: Borden, Foremost, Schepps, Cabell, Oak Farms, Metzger, Vandervoort, Gandy, and Preston. Indictments suggest that collusion began at least as early as 1975. This paper focuses on the period from 1980 to 1992. The cartel went through the following structural changes: in 1983, Borden acquired Metzger; in 1985, Preston entered the school milk market and joined the conspiracy; in 1986, Schepps acquired Foremost; in 1990, Cabell acquired Oak Farms.

Pure Milk Co. is the largest non-cartel school milk supplier in the dataset. The firm's main plant was located in Waco, TX (i.e in a different federal order zone than the DFW cartel). The company was founded in the 1960s and was successful in establishing a strong local customer base in Central Texas by marketing higher quality dairy products. Its main raw milk supplier was Dairy Farmers of America, a national milk marketing cooperative. School contracts made up to 20% of the firm's business.

[17] As in Pesendorfer (2000), collusion is thought of as an explicit or implicit scheme designed to limit competition and increase profits.

While Pure Milk bid primarily for contract near its plant in the Waco market area, it also participated in a non-negligible number of auctions for school districts in the DFW market area. According to the General Manager of the firm at that time, these were typically larger contracts justifying "going the extra mile". In these occasions, Pure Milk bid against the cartel. The paper focuses on these contracts in which Pure Milk bid in district where the cartel was operating. Pure Milk's location further from the epicenter of the conspiracy may have played a non-negligible role in it not being part of the cartel. Most of its business was conducted in the Waco market area, while the cartel was only active in Dallas-Fort Worth.

2.4 Data and Descriptive Statistics

The paper studies school milk contracts awarded annually between 1980 and 1992 in three large market areas in Northeastern and Southern Texas: Dallas-Fort Worth, Waco and San Antonio. Contracts are awarded at the school district level. Figure 2.1 shows school districts in the dataset by market area. Each school district contains around four to five schools. Initially, the dataset contains information on 1620 auctions, 4444 bids.

The main dataset is the auction data. This dataset was collected by the Antitrust Division of the US Department of Justice during its investigation of the Texas milk cartel. For each contract awarded, the following characteristics are observed: the county and school district awarding the contract, the identity of the bidding firms and corresponding bids for each milk category (whole white, whole chocolate, low-fat white, low-fat chocolate, skim milk), the quantities required per milk category, whether a cooler has to be provided, the number of meals served in the school district, the school district enrollment, the number of deliveries per week, the number of schools in the district, whether a bid was fixed or escalated, and the identity of the winner[18].

Three auxiliary datasets complement the auction data. Two are obtained from the US Department of Agriculture's Marketing Service.[19] First, a dataset on prices of Class I fluid milk for the period of interest.[20] This is the price of raw milk sold by milk cooperatives (such as Dairy Farmers of America) to milk processors and distribution firms.[21] Second, a dataset giving the processing plants locations of the firms bidding for the school milk contracts. Third, the longitudes and latitutes of school districts is added to the main auction data.

The reduced form and structural analysis are conducted on the data after the following preparation. Auctions with more than one winner are dropped. Auctions with only one participant are dropped. Prices are deflated using the CPI deflator into 1982 dollars. Finally, a distance variable is constructed for each observed bidder-auction pair: this variable measures the great-circle distance between the school district and the bidder's closest plant. The variable is constructed using the latitude and longitude coordinates of firms' plants and school districts. After this procedure, the dataset contains information on $1,033$ auctions, $3,488$ bids.

Table 2.1 shows that the majority of firms win between 20% and 30% of auctions in which they participate. Borden and Oak Farms are the largest firms in terms of contract won over this period.[22] In terms of contracts won, Pure Milk (the largest non-cartel firm) is in the second-tier of the distribution. It won 28% of the contracts it bid on.

[18]In some large and densely populated district, several distributors shared the contract.

[19]Southwest Market Area (Federal Milk Order 126).

[20]Class I Milk is raw milk destined to be used as a beverage, in opposition to milk processed into yogurts, cheese etc.

[21]Note that such prices differ from one region to the other within the Texas Milk Marketing Order.

[22]Both firms are national, with a larger distribution network.

Table 2.1: Number of bids and wins by firm 1980 – 1992

Vendor	# of bids	# of wins	% of wins
BORDEN	836	232	0.278
CABELL	418	140	0.335
FOREMOST	139	45	0.324
GANDY	23	5	0.217
KNOWLTON	12	1	0.083
LILLY	7	0	0
METZGER	21	4	0.190
OAK FARMS	530	184	0.347
PRESTON	333	98	0.294
PURE	255	72	0.282
SCHEPPS	528	115	0.218
SUPERIOR	46	30	0.652
VANDERVOORT	340	107	0.315
Total	3,488	1,033	

Table 2.2 gives summary statistics by market area. Over the period of interest, 735 contracts were awarded in Dallas-Fort Worth, 143 in San Antonio, and 179 in Waco. Only 15 school districts awarded contracts in San Antonio, against 30 for Waco. Indeed, school milk contracts in San Antonio are for larger quantities. The average winning bid (for a half-pint of whole white milk) is greater in Dallas-Fort Worth, followed by San Antonio, and lastly Waco[23]. The average cost of a contract is the highest in San Antonio, reflecting again the larger size of contracts. The average contract cost in DFW is $85,115 against $18,779 in Waco. This reflects differences in contract sizes, raw milk prices, contract specification across the two market areas but is also potentially related to inflated cartel prices in the DFW area.

2.5 Reduced Form Analysis

In this section, the bidding behavior of the largest non-cartel bidder (Pure Milk Co.) is examined. This preliminary analysis provides evidence that Pure Milk bid significantly less aggressively when facing cartel bidders, and allow a preliminary assessment of the overbidding.

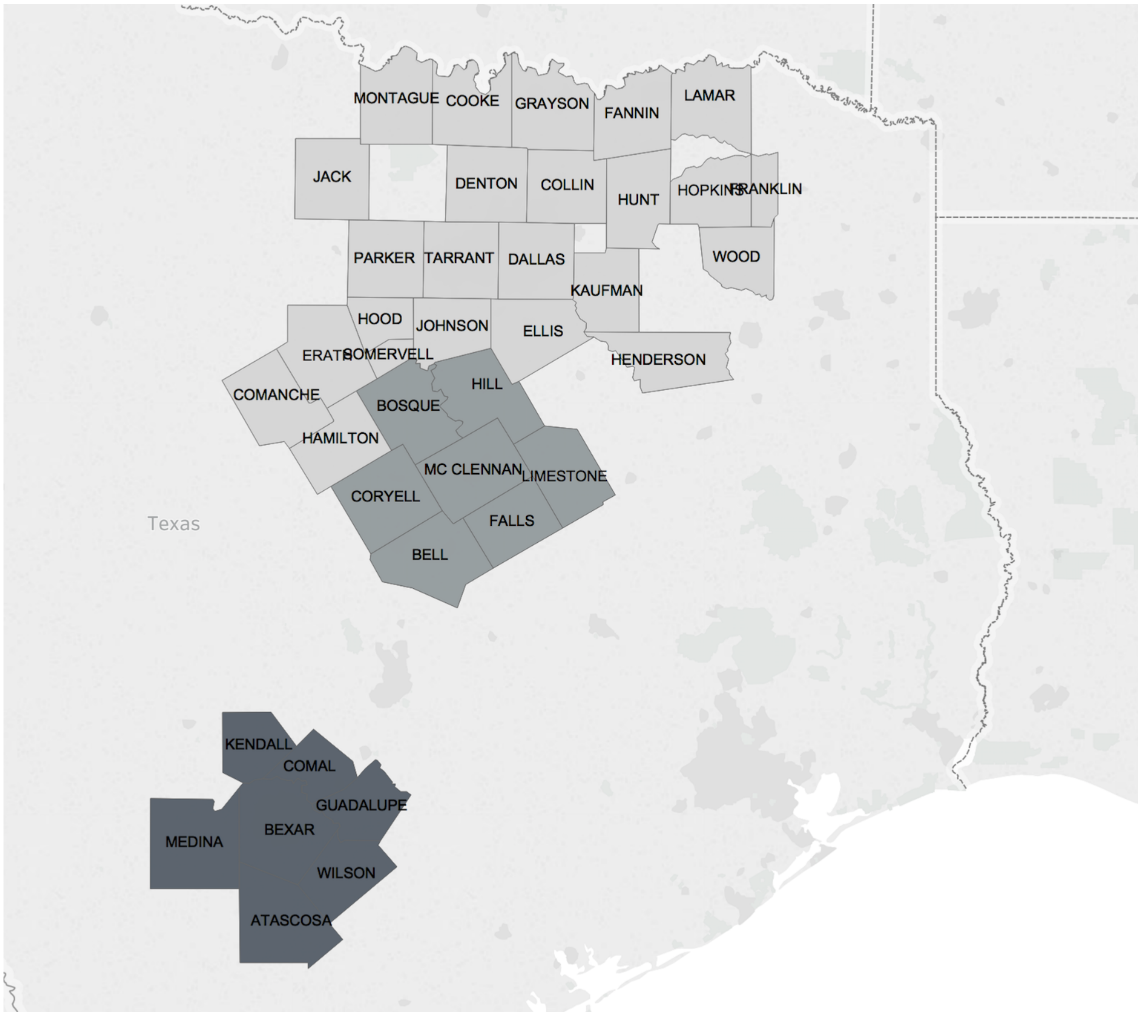
[23]Note that for these winning bid may reflect different auction characteristics from one market area to the other (quantities, input prices etc)

Table 2.2: Descriptive statistics by market area

Market Area		DFW	San Antonio	Waco
Number of bids		2411	555	591
Number of contracts		735	143	179
Number of contracts/year		56.5	11	13.7
Number of counties		25	7	7
Number of school districts		115	15	30
winning bid/half pint	Mean	0.1442	0.1417	0.1365
	SD	0.0214	0.0283	0.0226
# of meals per year-school district	Mean	522,578.2	1,273,001.4	339,188.0
	SD	655,877.1	1,436,814.3	620,308.4
Contract total cost per school district	Mean	85,115.37	168,938.34	18,779.90
	SD	112,916.49	202,501.19	24,883.69

Note: Bids are for a half-pint of whole white milk. Bids and contract cost are in 1982 *dollars.*

Figure 2.1: Map of the counties in the dataset, by market area



Pure Milk's processing plant is located near the city of Waco, in Mc Lennan county. The firm bid mainly in Mc Lennan and neighboring counties (see Figure 2.1). However, in 10% of the cases, the firm bid in further counties, located in the cartel area of activity. Auctions in which Pure Milk bid are divided by counties into two separate types:

- Collusive auctions: these are in counties contiguous or close to Dallas-Fort Worth in which the cartel presence was established by the DoJ. The counties are Johnson, Hood, Erath, Dallas, and Comanche. Such auctions form around

10% of Pure Milk's bids.

- Competitive auctions: these are in counties outside the Dallas-Fort Worth cartel territory (they are located in the Waco or San Antonio market areas). The counties are McLennan and directly contiguous counties: Coryell, Falls, Limestone, Bell, Hill, and Bosque.

This classification is based on the factual statement of the milk cartel prosecution. It is assumed that the cartel bid in the collusive counties, while all firms bid competitively in the competitive counties.

Table 2.3: List of variables

Type	Variable	Description
Auction specific	FMO price	raw milk price in 1982 dollars
	meals	number of school lunches per school district per year
	escalated	equals 1 if the bid is escalated
	deliveries	number of deliveries per week
	number of schools	number of schools in the school district
	cooler	equals 1 if a cooler needs to be provided
	number of bids	number of bids submitted in the auction
Bidder specific	distance	great-circle distance between closest plant and school district
	bid	bid submitted for whole white category
	incumbency	equals 1 if bidder won the contract in previous year

Note: Indicators for missing values of cooler, deliveries and escalated are also included

The logarithm of Pure Milk's bid for whole white milk is regressed on the variable listed in Table 2.3. All continuous variables are in logarithm.[24] Quadratic terms for the number of meals and distance are included. In the first specification, no fixed

[24]The log-log specification allows the interpretation of coefficients as elasticities. Additionally, because bids are positive, errors are positively skewed. By logging the observed variable, errors are made more symmetric.

effects are used. In the second specification, year fixed effects and dummy variables for "collusive" and "competitive" counties are used. In the third specification, year fixed effects and county dummy variables are used. Results are shown in Table 2.4.

Table 2.4 shows that bids move closely with the price of raw milk. As expected, escalated bids are lower than fixed bids (since they are indexed to the price of raw milk and therefore shield bidders against future fluctuations of their input price). Bids are increasing in the number of schools to supply within the district. Bids are decreasing in the number of bidders. In specifications (2) and (3), bids are convex in the number of meals served (equivalently in the quantity of milk to supply). This would be consistent with an optimal utilization rate for milk distributors. In specifications (2) and (3), bids are concave in the distance between the processing plant and the school district. Distance increases the cost of fulfilling a contract, but with diminishing effects.

The regression provides evidence that Pure Milk overbid when facing the cartel. In specification (2), the coefficient on the dummy for the group of collusive auctions defined above is significantly different from zero and positive. Pure Milk bid on average 6% higher in the collusive auctions (facing the cartel), relative to the competitive auctions. Specification (3) breaks down the effect at the county level. As shown in Figure 2.2, coefficients are significantly positive in collusive counties, while the coefficients are significantly negative in counties with competitive auctions. All coefficients are measured with respect to the average bid in Bell county (competitive auction).

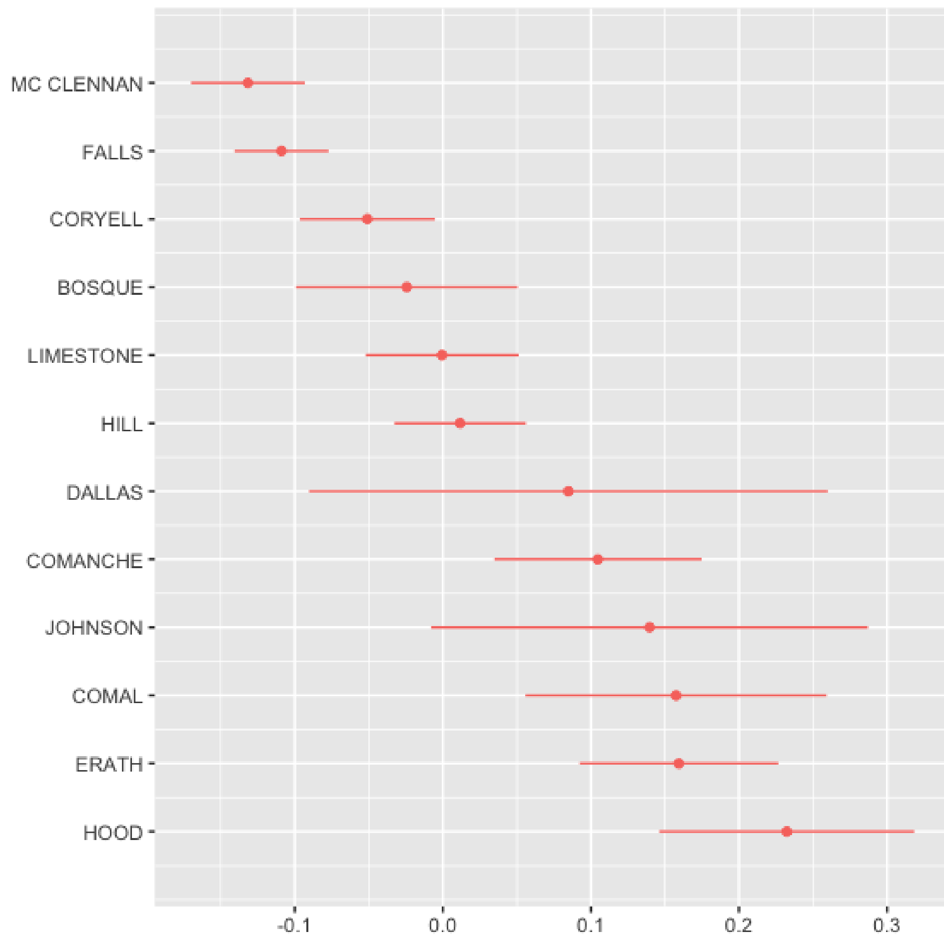
The reduced form analysis demonstrates that the largest non-cartel bidder overbid when facing the cartel. This approach is, however, limited if one is interested in the magnitude of umbrella damages relative to cartel damages, or in assessing the size of inefficiencies introduced by the cartel agreement. Therefore, more structure is imposed on the data. This structure is derived from the theoretical model of section 2.2. The structural analysis allows not only the estimation of damages and inefficiencies,

Table 2.4: Determinants of Pure Milk's bids

	(1) Coeff. (SE)	(2) Coeff. (SE)	(3) Coeff. (SE)
Incumbency	−0.013 (0.015)	−0.029*** (0.009)	−0.001 (0.009)
FMO price	1.175*** (0.068)	1.017*** (0.067)	1.062*** (0.060)
Meals	−0.064 (0.233)	−0.363** (0.141)	−0.339** (0.163)
Meals sqrd	0.002 (0.010)	0.015** (0.006)	0.013* (0.007)
Escalated	−0.041** (0.016)	−0.016* (0.010)	−0.019** (0.008)
Escalated missing	−0.015 (0.018)	−0.016 (0.011)	−0.012 (0.010)
Cooler	0.004 (0.073)	0.011 (0.045)	0.004 (0.056)
Cooler missing	−0.047 (0.067)	−0.014 (0.041)	0.013 (0.053)
Deliveries	0.004 (0.014)	−0.010 (0.008)	−0.006 (0.009)
Deliveries missing	0.063 (0.051)	0.007 (0.032)	0.040 (0.037)
Number of schools	0.006* (0.003)	0.003 (0.002)	0.008** (0.003)
Number of bids	−0.016 (0.013)	−0.016** (0.008)	−0.014* (0.008)
Distance	0.042 (0.027)	0.082*** (0.018)	0.115*** (0.022)
Distance sqrd	−0.005 (0.005)	−0.014*** (0.004)	−0.032*** (0.006)
Collusive auction		**0.066***** (0.024)	
Constant	1.587 (1.372)	3.002*** (0.848)	3.228*** (0.963)
Year FE	No	Yes	Yes
County FE	No	No	Yes
Observations	217	217	217
R^2	0.697	0.908	0.934
Adjusted R^2	0.676	0.896	0.921
Residual Std. Error	0.095 (df = 202)	0.054 (df = 190)	0.047 (df = 180)
F Statistic	33.214*** (df = 14; 202)	72.363*** (df = 26; 190)	71.104*** (df = 36; 180)

Note: A dummy for competitive auctions is omitted in specification (2)*. Omitted county is BELL in specification* (3)*. Counties corresponding to collusive auctions in bold in specification* (3)*. All continuous variables in log. Prices in* 1982 $*.* ****Significant at the 1 percent level.* ***Significant at the 5 percent level.* **Significant at the 10 percent level.*

Figure 2.2: Estimates and 95% confidence intervals for the county dummies



but also a finer decomposition of umbrella damages into damages stemming simply for the non-cartel firm overbidding, and damages originating in the inefficiency of the asymmetric auction.

2.6 Structural Analysis

In this section, the theoretical model of section 2.2 is estimated from the data. The objective is: first, to quantify the average damages caused to the school districts by the outsider firm (umbrella damages), and second, to quantify the loss in efficiency due to asymmetries between bidders introduced by the cartel agreement.

The estimation approach begins by recovering the underlying cost distribution of bidders from observed *competitive* bids. As auctions differ in their specifications (in the quantity to be supplied, whether bids can be escalated, number of deliveries per week, number of schools in the district), bids will reflect auction specific heterogeneity. A first step is to account for this observed auction heterogeneity to obtain a set *normalized* bids. Costs are then estimated non-parametrically using the empirical distribution of normalized bids following Guerre, Perrigne, and Vuong (2000)—GPV hereafter.

Using the estimated cost distribution, counterfactual bids are obtained by solving the auction in which the outsider firm faces the cartel. Two scenarios are considered: (1) assuming that the cartel mechanism is efficient (in the sense, that the cartel member with the lowest cost bids on behalf of the cartel), or (2) assuming that the cartel mechanism is inefficient (the cartel bidder is selected randomly from the cartel members). The cartel bidder is "stronger" when the mechanism is efficient (scenario (a)) and therefore the outsider will tend to bid more aggressively (see section 2.2). As a result, estimates under this scenario will give a lower bound on umbrella damages. Estimates under scenario (b) will accordingly provide an upper bound on umbrella damages.

Under scenario (a), cartel bidder and outsider draw their costs from different distributions. Equilibrium bid function in such asymmetric auction cannot be solved for analytically. The bid functions are obtained by numerical resolution of the system of differential equations (2.3). The method used is the fixed point iteration, introduced by Fibich and Gavish (2011).

Once the counterfactual bids are obtained, auction heterogeneity is added back to the bids, to reflect characteristics of auctions in which the cartel bid against the outsider firm. Damages to the seller (school district) and inefficiencies are estimated.

2.6.1 Data Limitations

Although the dataset used is rich in many dimensions, a few difficulties must be dealt with before laying out the estimation approach. A first issue is due to auctions differing in their observed characteristics, such as quantity of milk to supply. This auction observed heterogeneity is addressed by imposing additional structure on firms' costs (Assumption 2). A second issue is related to the relatively small sample size of auctions with a non-cartel firm bidding against the cartel.[25] This issue is compounded by the fact that the cartel mechanism is unknown. This section shows how assumptions restricting the structure of the model can be leveraged to estimate the variables of interest (cost distribution, equilibrium bid functions, and damages).

Denote by x_d, auction d' s observed characteristics (such as the quantity to be supplied, whether bids can be escalated, number of deliveries per week, number of schools in the district etc.).

Assumption 2 (Multiplicative Separability)**.** *Bidder i's cost in auction d, denoted*

[25]The data contains 250 bids by the non-cartel firm, 30 of which are against the cartel

c_{id}, can be written:

$$c_{id} = \tilde{c}_{id}\Gamma(x_d)$$

for some function $\Gamma(.)$. $\tilde{c}_{id}$ represents bidder i's idiosyncratic ("normalized") cost in auction d, which is independent from x_d.

Assumption 3 (Independence of costs across auctions)**.** *Costs are drawn independently across school districts and years.*

A detailed discussion of these assumptions' validity is delayed to Section 2.6.3.2.

The main issue is related to the estimation of the cost distribution from observed bids. The structural model developed in Section 2.2 requires (at least) two participants bidding *non-cooperatively* in each auction. Based on the list of firms prosecuted by the DoJ, all firms located in the Dallas-Fort Worth area were found colluding for contracts in that area. As a consequence, it is safe to assume that the majority of bids for auctions in the Dallas-Fort Worth area (in which all participants were cartel members) were either (1) a winning bid submitted by the cartel bidder to match the seller's (underlying) reserve price or (2) cooperative "phony" or "complementary" bids submitted by other cartel bidders. The structural model cannot be used to infer underlying costs from these complementary bids (and reserve prices), as most of them do not necessarily map to a firm's true cost, but were merely designed to simulate competition among bidders (or extract all of the seller's surplus). Auctions in which the cartel participates are referred to as "collusive" auctions.

Since the goal is to quantify the size of umbrella damages as a fraction of cartel damages, a natural way to solve the previous issue would be to use only the subset of collusive auctions in which the outsider firm bid against the cartel. In these auctions, at least two participants (the outsider and the cartel bidder) are bidding non-cooperatively. Unfortunately, two reasons make this alternative unappealing. First, the small sample size of this set of auctions (with the outsider bidding against the cartel) renders any non-parametric estimation of the distribution of costs impossible.[26] Second, the cartel mechanism (in particular the choice of the cartel bidder) is a priori unknown.[27] One could assume a mechanism for the selection of the cartel bidder, and estimate underlying costs from observed bids. However, such an assumption will be hard to justify. Additionally, if the mechanism is misspecified, this approach will lead to biased estimates of underlying costs. A more conservative approach is therefore preferred.

In the preferred approach, auctions from the two other market areas in the dataset (Waco and San Antonio) are used. As no firm was prosecuted in counties around Waco and San Antonio, it is assumed that firms were bidding competitively in these counties (denoted hereafter "competitive" auctions). By Assumption 2, observed auction heterogeneity can be separated from the idiosyncractic part of the bid. By Assumption 3, the set of competitive "normalized" bids can be used to recover the distribution of "normalized" costs (the $\tilde{c}_{id}$). Using the estimated distribution of costs, counterfactual bids are then simulated by solving the asymmetric auctions in which the outsider firm bid against the cartel, for different specification of the cartel mechanism (efficient cartel or inefficient cartel). Finally, by Assumption 2 (in particular the independence of "normalized" bids/costs and observed auction characteristics), observed auction characteristics drawn from the set of collusive auction with the outsider firm bidding against the cartel, are added back into the bids. By following this procedure, a set of competitive auction, and a set of corresponding collusive auctions with the outsider bidding against the cartel, are obtained.

Advantages of this approach are twofold. First, estimation of the cost distribution does not hinge on the correct specification of the cartel mechanism. Indeed, cost

[26]The outsider firm bid against the cartel in about 10% of the auctions in which it participates, which gives a set of 33 auctions.

[27]Pesendorfer (2000) finds evidence that the Texas cartel was quasi-efficient even though sidepayments were not used, by comparing it to the Florida cartel which was efficient.

are obtained from competitive auctions in which no bidders collude. Second, the cost distribution is estimated non parametrically, as the sample size of competitive auctions allows it.

One drawback of this approach is that it does not recover the specific cost corresponding to observed bids for collusive auctions in which the outsider participates. Instead, it computes upper and lower bounds on umbrella damages using (1) the cost distribution estimated from the set of competitive auctions and (2) auction observed characteristics drawn from collusive auctions in which the outsider participates.

2.6.2 Estimation Approach

The steps followed in the estimation are described below:.

- **Step 1: Observed auction heterogeneity**

  The estimation procedure assumes that the data available is from auctions of $ex - ante$ identical contracts. This is not the case for school milk contracts. Indeed, contracts differ in various dimensions, which are public information and observed by the bidders before submitting their bids. This public information will enter not only a bidder's private cost of realizing the contract, but also his belief about other bidder's costs.

  Haile, Hong, and Shum (2006) propose one method to account for auction-specific observed heterogeneity. The paper shows that multiplicative separability (Assumption 2) of idiosyncratic costs and auction characteristics carries out to bids and auction characteristics.

  Lemma 2. *Assume the multiplicative separable structure:*

  $$c_{idt} = \tilde{c}_{idt}\Gamma(x_{dt})$$

  where c_{idt} is bidder i's cost for contract d at time t, x_{dt} are contract-time specific characteristics. Then the equilibrium bid function has the multiplicative *separable form:*

$$\beta(c_{idt}) = \beta(\tilde{c}_{idt})\Gamma(x_{dt})$$

This result can be used to account for observed auction heterogeneity and homogenize the bids. Assume the following parametric specification: $\Gamma(x_{dt}) = \exp(x'_{dk}\beta)$. The first-stage regression is:

$$\ln b_{idt} = x'_{dt}\beta + \eta_t + \gamma_c + \kappa_i + n_{dt}\delta + \sigma_{idt} \tag{2.11}$$

where b_{idt} denotes the bid of bidder i for contract d at time t, η_t is a time specific dummy, γ_c is a county dummy, κ_i is a bidder specific dummy, and n_{dt} is the number of bidders participating in contract d at time t, σ_{idt} is the error term. x_{dt} include variables for: the price of raw milk, the number of meals served (and its square), whether bids can be escalated, whether a cooler has to be provided, the number of deliveries per week, and the number of schools in the school district. All continuous variables are in logarithm. Note that the use of time dummies enable us to capture seasonality: for instance, common packaging, processing and labor costs, that might change over time, but are common to all bidders.

The first-stage regression is ran on the sample of *competitive auctions*, that is, auctions in which the cartel did not participate.

Normalized bids are constructed from the results of regression (2.11), as $\ln \tilde{b}_{idt} = \ln b_{idt} - x'_{dt}\hat{\beta} - \hat{\eta}_t - \hat{\gamma}_c$.

Since equilibrium bid functions depend on the number of bidders participating in the auction, the rest of the estimation is conducted on auctions fixing the number of participants to three bidders.[28]

[28]Note that auctions with two participants do not give rise to potential umbrella damages, while auctions with more than four bidders are somewhat scarce in the dataset.

The estimation approach abstracts from auction unobserved heterogeneity. The latter would be relevant if bidders were to observe auction characteristics that are unobserved by the econometrician. In the context of school milk auctions, there do not seem to be other factors relevant to firms'costs aside from the ones controlled for in this first estimation step, i.e input prices, quantities, and auction specifications (escalated bids, coolers, number of deliveries etc).

- **Step 2: Estimation of the underlying cost distribution (GPV estimator)**

  Following the procedure presented in Guerre, Perrigne, and Vuong (2000), the underlying distribution of costs can be estimated using the distribution of normalized bids, obtained in the previous step. The cumulative distribution of bids is estimated using the empirical distribution function, while the density is estimated using a kernel with finite support.[29] Bid data are trimmed in order to control for the asymptotic bias at the boundaries of the support of the bid distribution as suggested in Guerre, Perrigne, and Vuong (2000).

- **Step 3: Derivation of the asymmetric equilibrium bid functions**

  As shown in Section 2.2, the cartel bidder and non-cartel firm equilibrium (inverse) bid functions, denoted ϕ_1 and ϕ_2 respectively, are the solutions of:

$$\begin{cases} \frac{1}{b-\phi_2(b)} = \frac{f_1(\phi_1(b))\phi_1'(b)}{1-F_1(\phi_1(b))} \\ \frac{1}{b-\phi_1(b)} = \frac{f_2(\phi_2(b))\phi_2'(b)}{1-F_2(\phi_2(b))} \end{cases} \tag{2.12}$$

  where F_i (resp. f_i) is the cumulative distribution (resp. density function) of costs of the cartel ($i = 1$) and the outsider firm ($i = 2$). Along with the boundary conditions $\phi_1(\underline{b}) = \phi_2(\underline{b}) = \underline{c}$ (lower bound of support of bids) and $\phi_1(\overline{c}) = \phi_2(\overline{c}) = \overline{c}$, equation (2.12) forms a nonlinear boundary value problem

[29] The epanechnikov kernel is used. The bandwidth selection method is likelihood cross-validation.

(BVP), in which the location of the left-boundary $\underline{b}$ is unknown. In addition to this latter "non-standard" feature, the numerical resolution of the BVP is furthermore complicated (see below) by the fact that the mapping M in $(\phi_1'(b), \phi_2'(b)) = M(b, \phi_1(b), \phi_2(b))$ is not Lipschitz-continuous in $(\phi_1(b), \phi_2(b))$ at the right boundary $\overline{c}$.

The BVP defined in equation (2.12) cannot be solved analytically. However, several numerical solutions have been proposed in the literature: the backward-shooting method was first used to solve this problem by Marshall, Meurer, et al. (1994) and used and refined in various subsequent papers.[30] While this method is currently the standard for computing equilibrium bids in asymmetric auctions, it suffers from large instability at the right boundary (see Fibich and Gavish (2011) for a detailed analysis). As a consequence, an alternative numerical method, proposed in the latter paper, is preferred. Their idea is to recast the BVP as a system of differential equations in ϕ_1 and b (instead of ϕ_2) as functions of ϕ_2. This allows to transform the BVP with unknown left boundary into a BVP with known boundaries, and apply standard numerical techniques such as fixed point iteration on a grid. The details of their numerical method, adapted to this setting are presented in Appendix B.1.

Since the internal organization of the cartel is unknown, in particular regarding how the cartel bidder is selected for a given contract, two extreme cases for the cartel internal mechanism are considered. These two scenarios will provide upper and lower bounds on umbrella damages.

1. Efficient cartel mechanism: the cartel is able to select its lowest cost member to bid on behalf of the cartel at the auction. Such mechanism can be

[30] Gayle and Richard (2008) use local Taylor series expansions of the solution and the distribution and Li and Riley (2007) use an adaptive step size for the numerical backward integration to allow better control of the error.

sustained for instance if the cartel uses sidepayments. Denote by F the true distribution of costs for each bidder (assuming symmetry), and by n the number of cartel members submitting bids in the auction. Then the cost distributions of the non-cartel firm and of the cartel bidder are:

$$F_2(c) = F(c) \qquad f_2(c) = f(c) \qquad \forall c \in [\underline{c}, \overline{c}]$$

$$F_1(c) = 1 - (1 - F(c))^n \qquad f_1(c) = nf(c)(1 - F(c))^{n-1} \qquad \forall c \in [\underline{c}, \overline{c}]$$

Note that F_1 is simply the distribution of the minimum of n random variables drawn independently from F.

2. Inefficient cartel mechanism: for each contract, the cartel bidder is selected randomly among cartel members. In this case:

$$F_2(c) = F_1(c) = F(c) \qquad f_2(c) = f_1(c) = f(c) \qquad \forall c \in [\underline{c}, \overline{c}]$$

Estimators of (F_1, F_2, f_1, f_2) are constructed from the non-parametric estimators of (F, f) obtained in Step 2 (GPV estimation).[31], and passed on to the fixed point iteration algorithm. The result of the numerical method are estimators of the true asymmetric equilibrium inverse bid functions: $\hat{\phi}_1$ and $\hat{\phi}_2$.

- **Step 4: Reincorporation of the observed auction heterogeneity**

  Observed auction heterogeneity is added back into the set of normalized bids by drawing from the empirical distribution of auction characteristics in cases where the outsider firm bid against the cartel. This is motivated by Assumptions 1 and 2.

[31] The estimator of the cumulative distribution F is kernel-smoothed in order to facilitate the numerical resolution.

- **Step 5: Damage Assessment**

  Estimates of damages are constructed by combining the set of competitive winning bids and corresponding counterfactual set of collusive winning bids.[32] Estimate of efficiency losses are constructed by comparing the winner's cost in the competitive and corresponding collusive auction.

2.6.3 Results

2.6.3.1 Assessment of Damages and Inefficiencies

Figure 2.3 shows the estimated bid functions in the case of three bidders auctions.[33] Each panel shows the bid function in the competitive auctions, along with the counterfactual outsider and cartel's bid functions in the case of an efficient cartel mechanism (left panel), and an inefficient cartel mechanism (right panel). In both cases, the collusive bidding functions lie above the competitive bidding function as both the cartel bidder and the outsider bid less aggressively. Note that the overbidding increases the smaller is the bidder's cost. Overbidding when the cartel is inefficient is larger than when the cartel is efficient (see section 2.2).

Table 2.5 presents damage estimates obtained from the structural analysis detailed in section 2.6.2.[34] Estimates of damages to the auctioneer (school district) correspond to the difference between the winning bid of the collusive auction in which the outsider firm bids against the cartel and the winning bid of the competitive auction. Damages

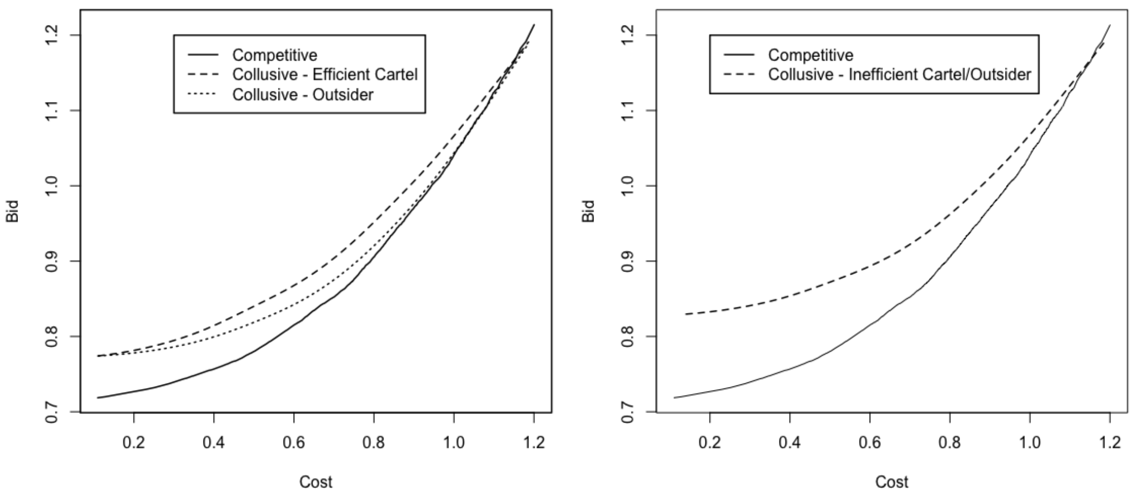
[32] Collusive auctions are the auctions in which the outsider firm bid against the cartel.

[33] Normalized bid functions are represented.

[34] $ are 1982 dollars. UB corresponds to an inefficient cartel. LB corresponds to an efficient cartel.[1] only cartel damages in the case of an efficient mechanism are reported. For an inefficient mechanism, these damages are the same as the UB on outsider damages. [2] per contract (whole white only): doesn't include damages for the other milk categories. [3] per contract: computed by applying the overcharge estimated to the total quantity purchased (whole white and other categories). Proportions of auctions in the efficient cartel case: 33% with outsider damages, 9% with misallocation damages, 58% with cartel damages. Confidence intervals based on 2500 bootstrap iterations.

Figure 2.3: Estimated bid functions for auctions with three bidders with an efficient and inefficient cartel mechanism

(a) Efficient cartel mechanism (b) Inefficient cartel mechanism



are classified into three types, depending on the identity of the winner in each of the competitive and collusive auction:

1. *Cartel damages*: these are damages in collusive auctions won by the cartel. From the estimated bid function, the cartel also wins the competitive auction.[35] These are the typical damages antitrust authorities try to assess.

2. *Misallocation damages*: these are damages in collusive auctions won by the outsider, when a cartel member would have won the competitive auction. When the cartel mechanism is efficient, it introduces asymmetry in the auction. As a result, the outcome of the auction is no longer efficient. The outsider wins the collusive auction even if they are not the lowest cost bidder. The reason is that the cartel bids less aggressively than the outsider.

[35]The cartel bids less aggressively in the collusive auction, so if the cartel wins, it must be lowest cost bidder.

3. *Outsider damages*: these are damages in collusive auctions won by the outsider, when the outsider would have won the competitive auction as well. The outsider is the lowest cost bidder. In this situation, damages to the auctioneer come merely from the fact that, by best replying to the cartel's bidding, the outsider is able to bid less aggressively than in the competitive auction.[36]

Importantly, these damages are computed conditional on the cartel (in the case of cartel damages) or the non-cartel firm (for misallocation and outsider damages) winning the collusive auction. When the cartel is efficient, auctions leading to cartel damages form 58% of the sample, auctions leading to misallocation damages form 9% of the sample, while auctions leading to outsider damages form 33% of the sample. When the cartel is inefficient, auctions leading to cartel damages form 50% of the sample, while auctions leading to outsider damages form 50% of the sample. Upper bounds (UB) and lower bounds (LB) are derived for outsider damages in the case of an inefficient and efficient cartel mechanisms respectively.

In auctions where the outsider firm bid against the cartel, the mean damages per contract in the whole white milk category are between $1,052 and $2,019. If the same overcharge was applied to all milk categories, damages per contract would be between $3,906 and $7,495. For auctions won by the outsider as the lowest cost bidder (outsider damages) the winning bid is 2.9% to 8.5% above the competitive winning bid, depending on the cartel mechanism. As a fraction of the winning competitive mark-up, these damages are between 12% and 42%. For auctions won by the outsider, while the cartel was the lowest cost bidder, misallocation damages are 3.5% of the competitive winning bid, or 26% of the competitive winning mark-up. For auctions won by the cartel, damages are between 5.9% and 8.4% of the competitive winning bid, or between 30% and 42% of the competitive winning mark-up.

[36]The situation in which the outsider wins the competitive auction while the cartel wins the collusive auction does not arise because the outsider bids more aggressively than the cartel.

Table 2.5: Estimates of damages

	Point Estimate	LB of 90% CI	UB of 90% CI
Mean damage per half-pint ($) - LB	0.00518	0.00447	0.00579
Mean damage per half-pint ($) - UB	0.00982	0.00890	0.01069
Mean damage per contract (whole white only) ($) - LB	1,052.26	852.62	1,264.88
Mean damage per contract (whole white only) ($) - UB	2,019.22	1,663.09	2,395.20
Mean damage per contract ($) - LB	3,906.65	3,068.04	4,765.71
Mean damage per contract ($) - UB	7,495.49	5,966.52	9,097.94
Mean damage per half-pint ($)			
Outsider damages - LB	0.00316	0.00246	0.00379
Outsider damages - UB	0.00919	0.00815	0.01019
Misallocation damages	0.00404	0.00294	0.00516
Cartel damages[1]	0.00647	0.00577	0.00713
Mean damage as fraction of competitive winning bid			
Outsider damages - LB	0.02996	0.02306	0.03648
Outsider damages - UB	0.08538	0.07443	0.09614
Misallocation damages	0.03538	0.02541	0.04637
Cartel damages	0.05968	0.05279	0.06631
Mean damage as fraction of competitive winning mark-up			
Outsider damages - LB	0.12353	0.09412	0.15096
Outsider damages - UB	0.42288	0.37230	0.47661
Misallocation damages	0.26609	0.19664	0.33934
Cartel damages	0.30961	0.27065	0.34453
Mean damage per contract (whole white only)[2] $			
Outsider damages - LB	655.47	449.78	877.98
Outsider damages - UB	1,902.96	1,425.27	2,378.19
Misallocation damages	812.32	350.07	1,340.04
Cartel damages	1,330.21	1,065.17	1,599.61
Mean damage per contract[3] $			
Outsider damages - LB	2,438.30	1,539.83	3,385.41
Outsider damages - UB	7,074.81	4,911.70	9,181.49
Misallocation damages	3,000.26	952.90	5,331.35
Cartel damages	4,937.94	3,808.12	6,100.61
Mean inefficiency due to Misallocation (efficient cartel)			
per half-pint ($)	0.00394	0.00249	0.00560
in percentage loss	0.03723	0.02494	0.06895
per contract ($)	2,909.34	747.15	5,677.51
per contract (whole white only) ($)	788.55	294.23	1,430.84

Inefficiencies introduced by the (efficient) cartel agreement are measured by the difference in the winner's cost in the competitive and collusive auctions. These two costs differ only in the case of misallocation damages, when the outsider firm wins the collusive auction, while the cartel would have won the competitive auction. Inefficiencies amount to an increase of 3.7% of the winner's cost, equaling 788$ per contract in the whole white category, or 2,909$ per contract.[37]

Damages caused by the outsider's bidding behavior form a non-negligible fraction of cartel damages. These umbrella damages can be decomposed into outsider damages and misallocation damages, as defined above. Per contract, outsider damages (conditional on the outsider winning) are estimated to be at least 47% of cartel damages (conditional on the cartel winning). In other words, the ratio of expected damages to the auctioneer conditional on the non-cartel firm winning over the expected damages to the auctioneer conditional on the cartel firm winning, is estimated to be at least 47%. This lower bound is obtained with an efficient cartel. If the cartel is inefficient, outsider damages are as large as cartel damages. Misallocation damages are estimated to be as large as 64% when the cartel is efficient, and the auction is asymmetric. The estimates found for outsider damages, between 2.9% and 8.5% of the competitive winning bid, are consistent with the overcharge of 6% estimated in the reduced form section.

2.6.3.2 Robustness of the Assumptions underlying the Estimation Approach

The validity of the assumptions made so far is discussed here. The model is cast within the symmetric IPV framework (Assumption 1). A bidder's cost can be decomposed

[37] If the cartel is inefficient, misallocation damages might arise. However they are due to the random selection of the cartel bidder rather than asymmetries between bidders. If the competitive auction would have been won by a cartel member that is not selected as the cartel bidder, there will be loss of efficiency. This type of misallocation is left aside as it is not stemming from asymmetries.

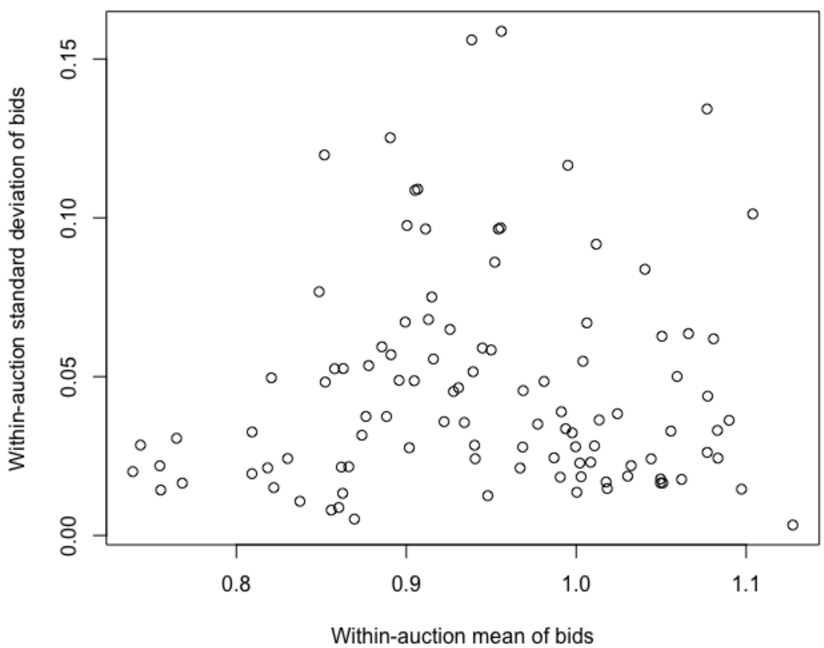
into : (1) a component common to all bidders, which includes not only observed auction characteristics, but also common processing, packaging and labor costs[38] (2) an idiosyncratic part which is bidder specific. The symmetric IPV assumption is imposed on the idiosyncratic part of costs (i.e once the common component of costs has been filtered out). A bidder's idiosyncratic cost depends on how close the school district is to its current delivery route. This delivery route depends on the bidder's current portfolio of clients (which include government agencies, hospitals, military bases etc). In the same line, a bidder's idiosyncratic cost depends on its current capacity utilization. Finally, idiosyncratic costs include a bidder's efficiency in packaging, loading trucks, managing the machinery etc[39]. Clearly, such factors are private to each bidder, as they do not affect its competitors' costs. Moreover, these factors are fairly independent across bidders. However, one might argue that firms could monitor each other's capacities and portfolio of clients, and therefore derive information about how far a competitor's route is from a school district of interest. If this was indeed the case, one way of controlling for such public information would be to add ex-ante asymmetries across bidders: the cost distribution of a bidder depends of its plant-school district distance. Nonetheless, the symmetry assumption is imposed. A first justification is that in the vast majority of cases, firms favour closer school districts. Second, information about competitors' distance from the school district of interest will be more important to infer competitors' participation decision rather than to get a precise estimate of their cost.

The multiplicative separability assumption of firms' idiosyncratic costs and auction characteristics (Assumption 2) can be tested. By Lemma 2, bids inherit the multiplicative structure. As noted in Asker (2010), this implies in particular that within auction, the standard deviation of bids depends on auction characteristics.

[38]See evidence found in Porter and Zona (1999).

[39]According to interviewed bidders.

Figure 2.4: Within-auction mean of bids against standard deviation for auctions with 3 bidders



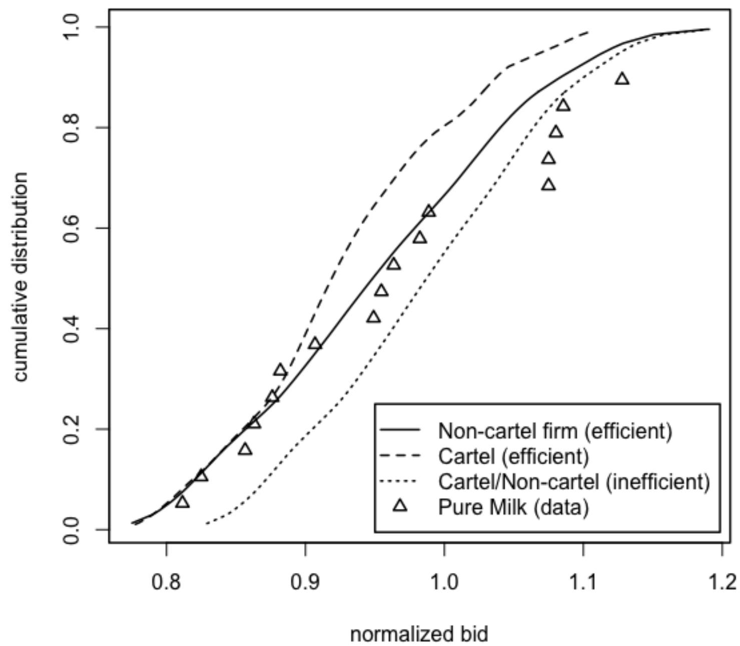
With an additive separable structure this will not be the case. As within-auction average bids also depends on auction characteristics, one expect a positive correlation between within auction average bid and standard deviation. Figure 2.4 shows these two variables plotted. The within-auction standard deviation varies with the average bid.

In order to conduct the structural estimation of the cost distribution, independence of idiosyncratic costs across school districts and years is needed (Assumption 3). The use of time fixed effects in the first step of the estimation approach help capture any within-year correlation between cost drawn. The use of county dummies controlls for within-county correlation of costs.

Another crucial assumption for the structural analysis is that the equilibrium derived is played in the data (in the competitive auctions from which we estimate the cost distribution). In particular, bidders know the number of firms participating in the auction. While this assumption is usually standard, additional evidence was found by interviewing managers who participated in the school milk procurements. According to interviewed bidders, the number and the identity of the firms who are likely to bid for a given contract is usually known in advance. This is due to the fact that firms bid for the same set of contracts every year and therefore develop a good understanding of which competitors will be interested in a given contract. Information about which competitors are the closest to the school district of interest also helps in refining their estimate of the number of bidders.

A last issue to deal with concerns endogenous selection in the case of the outsider. Indeed, it is assumed that the cost distribution of the outsider is the same in the competitive and the collusive auctions. This need not be the case if there is selection on observable auction characteristics (i.e the characteristics of the competitive and collusive auctions in which the outsider participates are systematically different) or on idiosyncratic costs (for instance if the outsider bids against the cartel only when the cost drawn is favourable enough). The first type of selection is dealt with by the separability assumption. Indeed, damages are estimated by drawing from the empirical distribution of *collusive* auction characteristics (cf step 4). The second type of selection, on idiosyncratic costs, seems less of a concern: according to the interviewed non-cartel bidder, the main factor driving participation in collusive auctions was the size of the contract. Figure 2.5 shows the empirical distribution of Pure Milk's bids when facing the cartel (triangle markers). This empirical distribution can be compared to the bid distributions predicted by the structural model. In particular, the bid distributions of : (1) the non-cartel firm facing an efficient cartel (solid line), (2) an efficient cartel (dashed line), and (3) an inefficient cartel/non-cartel firm (dotted line), are plotted. Although of a limited sample size, Pure Milk's bid distribution

Figure 2.5: Cumulative distributions of normalized bids for three-bidder auctions



estimated from the data is close to the distribution predicted for the non-cartel firm facing an efficient cartel.

2.7 Conclusion

This paper examines how non-cartel firms' bidding behavior can be affected by the existence of a cartel in a first-price procurement auction. In the case of the Texas school milk cartel, the analysis shows that the largest non-cartel firm bid significantly higher when facing the cartel (relative to when facing non-colluding firms).

The structural model shows that conditional on the non-cartel firm winning against the cartel, damages to the auctioneer (in the form of inflated winning bids) are a non-negligible fraction of the damages caused when the cartel wins. These results provide new evidence on the potential severity of umbrella damages, i.e damages caused to buyers by non-cartel firms adapting to the cartel supra-competitive price. As shown, umbrella damages broaden the scope of cartel damages in a non negligible way. The recent decision by the ECJ allowing ”umbrella claimants” to pursue treble damages against cartels seems to recognize the latter fact, albeit at the same time pointing to the difficulty of proving such claims. Assessing umbrella damages is nonetheless feasible, as shown here in the case of procurement auctions, as long as claimants have access to prices when the cartel competes against outsiders.

A number of open questions remain. First the paper focuses on the school milk industry in which contracts are awarded via first-price procurement auctions. But umbrella damages might not be restricted to auction environments. Alternative environments include industries where firms compete in quantities (an example would be the vitamin market and other similar commodities markets). Investigating the prevalence of such damages in alternative environment is left for future research. Second, in auction environments, the framework presented does not address the potential effect of a cartel existence on outsiders' participation decision. If entry is endogeneous and selective, in the sense that only firms with a cost realization (or signal) below a certain threshold enter, a cartel will induces more entry of outsiders (relative to a competitive environment with the same number of bidders).[40] This is because conditional on entry, the outsider's profits are strictly larger when his competitors collude (see section 1.1). This raises interesting predictions, which could be tested in environments where participation decisions are better observed.

[40] The two extreme models of entry being Levin and Smith (1994) in which firms pay an entry cost to learn their private cost, and Samuelson (1985), in which firms know their cost before making their entry decision. The former features no selection, while the latter features selection.

APPENDIX A

Appendix to Chapter 1

A.1 Panel of Digital-Projector Acquisitions

Table A.1 presents the observation dates for the panel of digital projector adoption by data sources. The table also shows the periodic subsample selected. The periods selected are such that there is a previous observation period 6 months earlier (in some exeption it is 5 or 7 months). For instance, “May 2012” is selected because the industry is observed on November 2011. The observation periods selected are represented in blue in Figure A.1.

Figure A.1: Share of digitally equipped screens and observation times

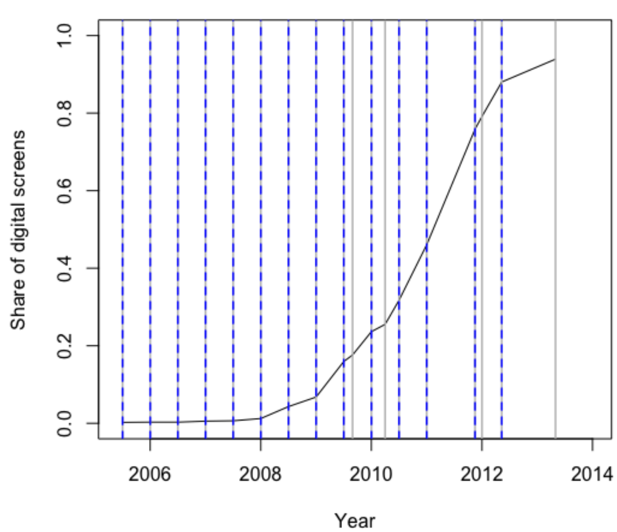


Table A.1: Observation times by data source

Date	Source	Periodic sample
July 2005	Media Salles	X
January 2006	Media Salles	X
July 2006	Media Salles	X
January 2007	Media Salles	X
July 2007	Media Salles	X
January 2008	Media Salles	X
July 2008	Media Salles	X
January 2009	Media Salles	X
July 2009	Media Salles	X
September 2009	Cinego	
Jarnuary 2010	Media Salles	X
April 2010	Cinego	
July 2010	Media Salles	X
January 2011	Media Salles	X
November 2011	Cinego	
January 2012	Media Salles	
May 2012	Cinego	X
June 2013	Cinego	

A.2 Reduced Form Analysis

This section details the construction of the variable "share of digital screens" for the art house and commercial networks. Denote by p_i the share of art house movies screened by theater i in 2015. Recall that art house theaters are defined as theaters with $p_i \geq 0.8$, while commercial theaters are defined as theaters with $p_i \leq 0.2$. The share of art house digital screens in period t is defined by:

$$\frac{s_{t,a}}{S_a} = \frac{\sum\limits_{i \in I} p_i s_{it}}{\sum\limits_{i \in I} p_i S_i} \tag{A.1}$$

Similarly the share of commercial digital screens is defined by:

$$\frac{s_{t,c}}{S_c} = \frac{\sum\limits_{i \in I} (1 - p_i) s_{it}}{\sum\limits_{i \in I} (1 - p_i) S_i} \tag{A.2}$$

note that in both cases the share is computed with respect to the total number of art house (resp. commercial) screens, not the total number of screens. This is because the analysis focuses on specialized art house theaters ($p_i \geq 0.8$) and commercial theaters ($p_i \leq 0.2$).

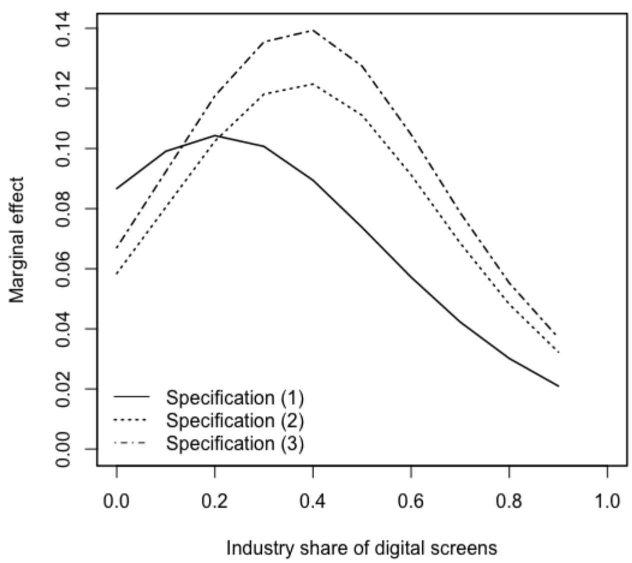
Table A.2 shows the reduced-form regression under the logit specification for the error term. The effect of a 10% increase in the industry share of digital screens in shown in Figure A.2. The predicted magnitude of network effect is $10\% - 14\%$ at a 50% share of digital screens. The results are similar to the predictions under the probit specification.

Table A.2: Share of screens converted s_{it}/S_i conditional on $s_{i(t-1)}/S_i = 0$ (ordered logit)

	Share of screens converted s_{it}/S_i conditional on $s_{i(t-1)} = 0$					
	(1)		(2)		(3)	
	Estimate	s.e	Estimate	s.e	Estimate	s.e
Industry share of d-screens	4.174	2.276	4.878	2.383	5.596	3.545
Adoption cost	-6.160	2.302	-6.437	2.361	-6.873	2.646
Own screens	0.142	0.107	0.170	0.110		
Seats	0.301	2.396	0.346	2.633		
Art house	0.133	0.257	0.130	0.273		
Competitor d-screens	0.024	0.019	0.024	0.020	0.029	0.024
Competitor f-screens	-0.007	0.008	-0.011	0.009	-0.008	0.010
Year FE	Yes		Yes		Yes	
Region FE	No		Yes		No	
market size FE	No		Yes		No	
Chain FE	No		Yes		No	
Box-office FE	No		Yes		No	
Theater RE	No		No		Yes	
Observations	$1,563$		$1,563$		$1,562$	
-log Likelihood	392.490		373.389		385.241	
AIC	818.980		804.777		800.482	

Note: *** *0.1% ;* * *10%. D-screen = screen equipped with a digital projector. f-screen = screens equipped with a film projector. For market dummies, the omitted category is " urban unit - 20 to 100k inhabitants". For the chain dummies, the omitted category is "single firm and small chains".*

Figure A.2: Effects of the industry share of digital screens and adoption cost on the probability of a



Note: The effect of a 10% *increment in the industry share of digital screen, on the probability of adoption, evaluated at the mean, as a function of the initial industry share of digital screens is represented.*

A.3 Industry Model

A.3.1 Perceived Transition Kernel

Let $\widehat{\mathbf{P}}_{a',a}$ be defined as follows (where the subscript t is omitted and next-period variables are marked with a prime):

$$\widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p', s'|\widetilde{\mathbf{x}}_i, p, s) = \widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p'|\widetilde{\mathbf{x}}_i, p, s)\widehat{\mathbf{P}}_{a,a}(s'|p, s) \tag{A.3}$$

The first sub-kernel $\widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p'|\widetilde{\mathbf{x}}_i, p, s)$ gives firm i's assessment of its next-period state (including competitors' total number of digital screens) and the exogenous price process. The second sub-kernel $\widehat{\mathbf{P}}_{a,a}(s'|p, s)$ gives firm i's assessment of next-period industry moment. Note this definition of the perceived kernel $\widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p', s'|\widetilde{\mathbf{x}}_i, p, s)$ implicitly assumes that firm i ignores its own impact on the evolution of the industry moment: $\widetilde{\mathbf{x}}_i'$ and s' are independent conditional on $(\widetilde{\mathbf{x}}_i, p, s)$. This assumption is realistic if the number of firms is very large, and therefore a single firm has a negligible impact on the aggregate industry state (and corresponding moment).

The perceived transition kernel for the moment s, $\widehat{\mathbf{P}}_{a,a}(s'|p, s)$, is first defined. The analysis focuses on short-run dynamics (i.e., the diffusion phenomenon) rather than the (adoption) steady state reached by the industry. The perceived kernel is meant to capture the short-run dynamics of the industry moment starting from the initial industry state $(\mathbf{y}_0, p_0)$. It is defined to coincide with the average observed transitions from the current moment-based state (s, p) to the next, the average being taken over many finite and short trajectories that start from the initial state of the industry. The perceived kernel corresponds to the observed frequencies of these transitions under adoption strategy a. Following Ifrach and Weintraub (2017) (Appendix A), $\widehat{\mathbf{P}}_{a,a}(s'|p, s)$ is defined as follows:

$$\widehat{\mathbf{P}}_{a,a}(s'|p,s) = \frac{1}{L}\sum_{l=1}^{L} \frac{\sum_{t=1}^{T} \mathbf{1}\{(p_t^l, s_t^l) = (p,s), s_{t+1}^l = s'\}}{\sum_{t=1}^{T} \mathbf{1}\{(p_t^l, s_t^l) = (p,s)\}} \tag{A.4}$$

where T is fixed to the time horizon of interest (in this case, 20 years covering the diffusion duration, or 40 periods), and $\{(p_t^l, s_t^l), T \geq t \geq 0)\}_{l=0}^{L}$ is a random sample of size L drawn from the distribution of the process $\{(p_t, s_t), t \geq 0)\}$ generated by the adoption strategy a, and initiated at the true initial industry state and exogeneous price $(\mathbf{y}_0, p_0)$.

The sub-kernel $\widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p'|\widetilde{\mathbf{x}}_i, p, s)$ can be further expressed as

$$\widehat{\mathbf{P}}_{a',a}(\widetilde{\mathbf{x}}_i', p'|\widetilde{\mathbf{x}}_i, p, s) = \mathbf{P}_{a',a}(\boldsymbol{\tau}_i, s_i'|\widetilde{\mathbf{x}}_i, p, s)\widehat{\mathbf{P}}_{a',a}(z_i'|\widetilde{\mathbf{x}}_i, p, s)\mathbf{P}(p'|p) \tag{A.5}$$

where $\mathbf{P}_{a',a}(\boldsymbol{\tau}_i, s_i'|\widetilde{\mathbf{x}}_i, p, s)$ is firm i's assessment of its next-period state, $\widehat{\mathbf{P}}_{a',a}(z_i'|\widetilde{\mathbf{x}}_i, p, s)$ is firm i's assessment of competitors' next-period digital screens, and $\mathbf{P}(p'|p)$ is the exogenous hardware price process. Equation (A.5) makes explicit three elements: (1) because firms use moment-based strategies, firm i's assessment of its next-period state is correct (so $(\widetilde{\mathbf{x}}_i, p, s)$ is sufficient to determine the transition probabilities of s_i, given the strategy profile (a', a)); (2) competitors' aggregate state (total number of digital screens) is approximated because only the first moment is tracked (3) the hardware price process is exogenous. The perceived kernel for competitors' next-period number of digital screens is defined similarly to the industry moment sub-kernel, to coincide with the average observed transitions over many finite and short trajectories starting from the initial industry state $(\mathbf{y}_0, p_0)$:

$$\widehat{\mathbf{P}}_{a',a}(z_i'|\widetilde{\mathbf{x}}_i, p, s) = \frac{1}{L}\sum_{l=1}^{L} \frac{\sum_{t=1}^{T} \mathbf{1}\{(z_i^l) = z_i', (\widetilde{\mathbf{x}}_i^l, p^l, s^l) = (\widetilde{\mathbf{x}}_i, p, s)\}}{\sum_{t=1}^{T} \mathbf{1}\{(\widetilde{\mathbf{x}}_i^l, p^l, s^l) = (\widetilde{\mathbf{x}}_i, p, s)\}} \tag{A.6}$$

where T is defined as in equation (A.4), and $\{(z_{it}^l, p_t^l, s_t^l, \widetilde{\mathbf{x}}_{it}^l), T \geq t \geq 0)\}_{l=0}^L$ is a random sample of size L drawn from the distribution of the process $\{(z_{it}, p_t, s_t, \widetilde{\mathbf{x}}_{it}), t \geq 0)\}$ generated by the adoption strategies (a', a) and initiated at the true initial industry state and exogenous price $(\mathbf{y}_0, p_0)$.

Equations (A.4) and (A.6) are used to define the perceived kernel for all states *visited* over the L simulation runs. The perceived kernels are defined arbitrarily outside this set. In particular, for non-visited states, the paper uses "status-quo" perceptions, as in Ifrach and Weintraub (2017), assuming the current state of the variable remains the same in the next period.

A.3.2 Multi-homing

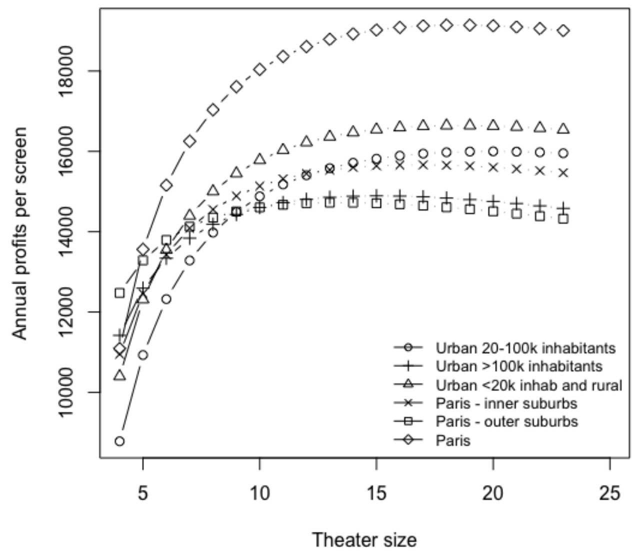
This appendix presents the estimation and counterfactual results under the "wide multi-homing" assumption. This assumption stands as the polar case to the "no multi-homing" assumption detailed in the main text. Setting $h_t^d = 0$ for all t, and defining $h_t \equiv h_t^m$ (so that $1 - h_t = h_t^f$), the single-period operating profits under "wide multi-homing" are given by:

$$\pi(\widetilde{\mathbf{x}}_{it}, h_t) = R(\boldsymbol{\tau}_i) \times \begin{cases} \frac{s_{it}}{S_i}\pi_d(\widetilde{\mathbf{x}}_{it}) + (1 - \frac{s_{it}}{S_i})\pi_f(\widetilde{\mathbf{x}}_{it}) & \text{if} \quad \frac{s_{it}}{S_i} \leq h_t \\ h_t\pi_d(\widetilde{\mathbf{x}}_{it}) + (1 - \frac{s_{it}}{S_i})\pi_f(\widetilde{\mathbf{x}}_{it}) & \text{if} \quad \frac{s_{it}}{S_i} \geq h_t \end{cases} \tag{A.7}$$

Under "wide multi-homing", theaters adopt the digital projection technology solely for cost-reduction purposes. Therefore, only the difference between profits from a digital and a film screening $(\pi_d(\widetilde{\mathbf{x}}_{it}) - \pi_f(\widetilde{\mathbf{x}}_{it}))$ can be identified from theaters' adoption times and units of technology adopted. Profits per film screening $\pi_f(\widetilde{\mathbf{x}}_{it})$ are normalized to zero.

The model predicts profits per screen $\pi_d(\widetilde{\mathbf{x}}_{it})/S_i$ (or equivalently cost-reductions per screen) between €$7,917$ and €$19,890$. Figure A.3 shows predicted profits per screen as a function of theater size and market size. Profits per screen are increasing in theater size, with a decreasing marginal effect. As in the "no multi-homing" case,

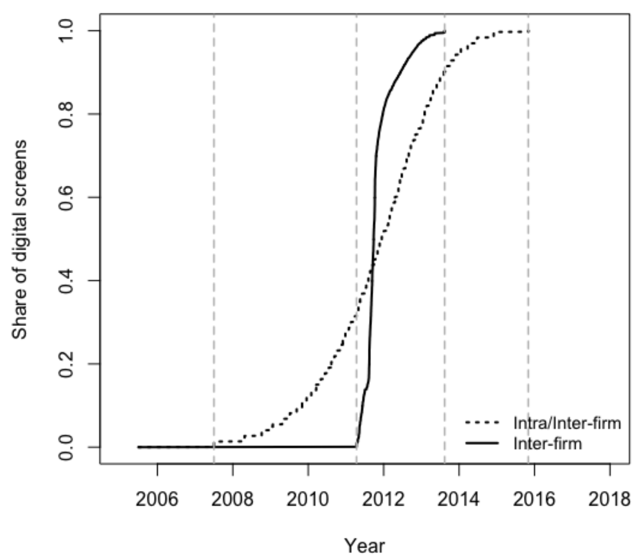
Figure A.3: Predicted annual profits per screen as a function of firm size



Note: Predicted profits are calculated fixing other characteristics to: monopolist, non art house theater, not horizontally integrated

estimates point to the presence of economies of scale in operation. The counterfactual exercise of section 1.8.1 is conducted using the estimated model. Figure A.4 presents the diffusion paths under the equilibrium played in the data and the counterfactual best-response (with no intra-firm margin). The qualitative results are similar to the ones obtained under "no multi-homing". The results are robust to the multi-homing assumption imposed because they are driven by heterogeneity in profits across theaters (which stem from differences in adoption times and units adopted across theaters), not by the absolute level of profits. The assumption on h_t only affects the absolute level of profits.

Figure A.4: Aggregate adoption rate with and without the intra-firm adoption margin



Note: The diffusion curves are obtained by generating 500 *sample paths with a length of* 20 *years. The sample average of these paths are reported.*

A.4 Estimation: Adoption Policy Rule (1st Step)

First step estimates for the adoption policy rule are obtained by estimating an ordered probit model. Denote by P_{ij} the probability that theater i transitions to state j. Possible states are $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ in the case of miniplexes ($4-7$ screens), and $\{0, \frac{1}{8}, \frac{2}{8}, ..., \frac{7}{8}, 1\}$ in the case of multi/megaplexes (8 screens or more). In constructing the likelihood, one has to account for the fact that theaters cannnot divest the new technology, and therefore cannot transition to lower states: the dependent variable s_{it}/S_i satisfies $s_{it} \geq s_{i(t-1)}$. The log likelihood is constructed as follows:

$$lnL = \sum_{i:4\leq S_i<8} \sum_{j=s_{i(t-1)}/S_i}^{1} d_{ij} lnP_{ij} + \sum_{i:8\leq S_i} \sum_{j=s_{i(t-1)}/S_i}^{1} d_{ij} lnP_{ij} \quad \text{(A.8)}$$

where d_{ij} is an indicator for firm i transitioning to state j.

APPENDIX B

Appendix to Chapter 2

B.1 Algorithm for solving the asymmetric auction (Step 3)

The details of the numerical method used to solve the asymmetric auctions are presented here. Recall that the cartel and non-cartel bidders' equilibrium inverse bid functions, denoted ϕ_1 and ϕ_2 respectively, are the solutions of:

$$\begin{cases} \frac{d\phi_1}{db} = \frac{1-F_1(\phi_1(b))}{f_1(\phi_1(b))}(b-\phi_2(b)) \\ \frac{d\phi_2}{db} = \frac{1-F_2(\phi_2(b))}{f_2(\phi_2(b))}(b-\phi_1(b)) \end{cases} \tag{B.1}$$

where F_i (resp. f_i) is the cumulative distribution (resp. density function) of costs of the cartel ($i = 1$) and the outsider firm ($i = 2$). Along with the boundary conditions $\phi_1(\underline{b}) = \phi_2(\underline{b}) = \underline{c}$ (for some $\underline{b}$, lower bound of the support of bids) and $\phi_1(\overline{c}) = \phi_2(\overline{c}) = \overline{c}$. The location of the left-boundary $\underline{b}$ is unknown. As shown in Fibich and Gavish (2011), the system of differential equations B.1 can be recast as a system in ϕ_1 and b (instead of ϕ_2) as functions of ϕ_2. After this change of variable, the new BVP is given by:

$$\begin{cases} \frac{d\phi_1}{d\phi_2} = \frac{1-F_1(\phi_1)}{f_1(\phi_1)}\frac{f_2(\phi_2)}{1-F_2(\phi_2)}\frac{b-\phi_1}{b-\phi_2} \\ \frac{db}{d\phi_2} = \frac{f_2(\phi_2)}{1-F_2(\phi_2)}(b-\phi_1) \end{cases} \tag{B.2}$$

with the boundary conditions: $\phi_1(\phi_2 = \overline{c}) = b(\phi_2 = \overline{c}) = \overline{c}$, and $\phi_1(\phi_2 = \underline{c}) = \underline{c}$. This BVP is defined on a known domain $\phi_2 \in [\underline{c}, \overline{c}]$. Fibich and Gavish (2011) propose fixed point iterations as one possible method for solving B.2. Iterations are given by:

$$\begin{cases} \left(\frac{d}{d\phi_2} + \frac{1-F_1(\phi_1^{(k)})}{f_1(\phi_1^{(k)})}\frac{f_2(\phi_2)}{1-F_2(\phi_2)}\frac{1}{b^{(k)}-\phi_2}\right)\phi_1^{(k+1)} = \frac{1-F_1(\phi_1^{(k)})}{f_1(\phi_1^{(k)})}\frac{f_2(\phi_2)}{1-F_2(\phi_2)}\frac{b^{(k)}}{b^{(k)}-\phi_2} \\ \left(\frac{d}{d\phi_2} - \frac{b^{(k)}-\phi_1^{(k+1)}}{b^{(k)}-\phi_2}\frac{f_2(\phi_2)}{1-F_2(\phi_2)}\right)b^{(k+1)} = -\frac{b^{(k)}-\phi_1^{(k+1)}}{b^{(k)}-\phi_2}\frac{f_2(\phi_2)}{1-F_2(\phi_2)}\phi_2 \end{cases} \tag{B.3}$$

with the boundary conditions: $\phi_1^{(k+1)}(\phi_2 = \overline{c}) = b^{(k+1)}(\phi_2 = \overline{c}) = \overline{c}$, and $\phi_1^{(k+1)}(\phi_2 = \underline{c}) = \underline{c}$. In the case of this particular empirical application, the initial guess used are: $\phi_1^{(0)}(\phi_2) = \phi_2$ and $b^{(0)}(\phi_2) = 0.9 + (\overline{c} - 0.9)/\overline{c} * \phi_2$. Although convergence of the fixed point iterations is not guaranteed, since a unique solution exists, if the algorithm converges to a function satisfying the BVP, this function is the solution to the BVP.

Bibliography

Ackerberg, D. A. and G. Gowrisankaran (2006). "Quantifying equilibrium network externalities in the ACH banking industry". *The RAND Journal of Economics* 37.3, pp. 738–761 (cit. on p. 8).

Aguiar, L. and J. Waldfogel (2018). "Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music". *Journal of Political Economy* 126.2, pp. 492–524 (cit. on p. 14).

Asker, J. (2010). "A study of the internal organisation of a bidding cartel". *American Economic Review* 100.3, pp. 724–762 (cit. on pp. 78, 79, 116).

Athey, S., J. Levin, and E. Seira (2011). "Comparing open and Sealed Bid Auctions: Evidence from Timber Auctions". *The Quarterly Journal of Economics* 126.1, pp. 207–257 (cit. on pp. 78, 79).

Athey, S. (2001). "Single crossing properties and the existence of pure strategy equilibria in games of incomplete information". *Econometrica* 69.4, pp. 861–889 (cit. on p. 81).

Bajari, P. (2001). "Comparing Competition and Collusion: A Numerical Approach". *Economic Theory* 18.1, pp. 187–205 (cit. on p. 81).

Bajari, P., L. Benkard, and J. Levin (2007). "Estimating dynamic models of imperfect competition". *Econometrica* 75.5, pp. 1331–1370 (cit. on pp. 5, 44).

Bajari, P., V. Chernozhukov, et al. (2015). "Identification and Efficient Semiparametric Estimation of a Dynamic Discrete Game" (cit. on p. 44).

Bajari, P. and L. Ye (2003). "Deciding Between Competition and Collusion". *Review of Economics and Statistics* 85.4, pp. 971–989 (cit. on p. 79).

Baldwin, L. H., R. C. Marshall, and J.-F. Richard (1997). "Bidder Collusion at Forest Service Timber Sales". *Journal of Political Economy* 105.4, pp. 657–699 (cit. on p. 79).

Battisti, G. and P. Stoneman (2005). "The intra-firm diffusion of new process technologies". *International Journal of Industrial Organization* 23.1-2, pp. 1–22 (cit. on p. 9).

Benkard, L., P. Jeziorski, and G. Weintraub (2015). "Oblivious equilibrium for concentrated industries". *RAND Journal of Economics* 46.4, pp. 671–708 (cit. on p. 38).

Bordwell, D. (2013). *Pandora's Digital Box: Films, Files, and the Future of Movies.* Irvington Way Institute Press (cit. on p. 14).

Cheng, H. (2006). "Ranking sealed high-bid and open asymmetric auctions". *Journal of Mathematical Economics* 42.4-5 SPEC. ISS. Pp. 471–498 (cit. on p. 81).

Clements, M. T. and H. Ohashi (2005). "Indirect Network Effects and the Product Cycle: Video Games in the U.S., 1994-2002". *The Journal of Industrial Economics* 53.4, pp. 515–542 (cit. on p. 8).

Corts, K. S. and M. Lederman (2009). "Software exclusivity and the scope of indirect network effects in the U.S. home video game market". *International Journal of Industrial Organization* 27.2, pp. 121–136 (cit. on p. 8).

Davis, P. (2006a). "Measuring the Business Stealing , Cannibalization and Market Expansion Effects of Entry in the U . S . Motion Picture Exhibition Market". *The Journal of Industrial Economics* 54.3, pp. 293–321 (cit. on pp. 10, 13).

— (2006b). "Spatial competition in retail markets: Movie Theaters". *Rand Journal of Economics* 37.4, pp. 964–982 (cit. on p. 10).

Dubé, G. J. Hitsch, and Chintagunta (2010). "Tipping and Concentration in Markets with Indirect Network Effects". *Marketing Science* 29.2, pp. 216–249 (cit. on p. 8).

Economides, N. and C. Himmelberg (1995). "Critical mass and network size with application to the US fax market". *NYU Stern School of Business working paper* EC-95-11.August, pp. 1–40 (cit. on p. 8).

Einav, L. (2007). "Seasonality in the U. S. motion picture industry". *The RAND Journal of Economics* 38.1, pp. 127–145 (cit. on p. 10).

Ericson, R. and a. Pakes (1995). "Markov-perfect industry dynamics: A framework for empirical work". *The Review of Economic Studies* 62.1, pp. 53–82 (cit. on p. 32).

Fibich, G. and N. Gavish (2011). "Numerical simulations of asymmetric first-price auctions". *Games and Economic Behavior* 73.2, pp. 479–495 (cit. on pp. 80, 103, 109, 132).

Fuentelsaz, I., J. Gomez, and Y. Polo (2003). "Intrafirm Diffusion of Technologies: An Empirical Application". *Research Policy* 32, pp. 533–551 (cit. on p. 9).

Gayle, W.-R. and J.-F. Richard (2008). "Numerical Solutions of Asymmetric, First-Price, Independent Private Values Auctions" (cit. on pp. 80, 109).

Gerarden, T. (2017). "Demanding Innovation : The Impact of Consumer Subsidies on Solar Panel Production Costs â", pp. 0–56 (cit. on p. 5).

Gil, R. and R. Lampe (2014). "The adoption of new technologies: Understanding Hollywood's (slow and uneven) conversion to color". *Journal of Economic History* 74.4, pp. 987–1014 (cit. on p. 10).

Gil, R. (2008). "Revenue Sharing Distortions and Vertical Integration in the Movie Industry". *The Journal of Law, Economics, & Organization* 25.2 (cit. on p. 10).

Gil, R., Houde, and Y. Takahashi (2015). "Preemptive Entry and Technology Diffusion in the Market for Drive-in Theaters" (cit. on p. 10).

Goettler, R. L. and B. R. Gordon (2011). "Does AMD Spur Intel to Innovate More?" *Journal of Political Economy* 119.6, pp. 1141–1200 (cit. on p. 10).

Goolsbee, A. and P. J. Klenow (1999). "Evidence on Learning and Network Externalities in the Diffusion of Home Computers" (cit. on p. 8).

Gowrisankaran, G. and J. Stavins (2004). "Network Externalities and Technology Adoption: Lessons from Electronic Payments". *The RAND Journal of Economics* 35.2, p. 260 (cit. on pp. 8, 26).

Griesmer, J. H. and R. E. Levitan (1967). “Toward a Study of Bidding Processes Part IV - Games with Unknown Costs”. *Naval Research Logistics Quarterly* 14.4, pp. 415–433 (cit. on p. 81).

Guerre, E., I. Perrigne, and Q. Vuong (2000). “Optimal Nonparametric Estimation of First-price Auctions”. *Econometrica* 68.3, pp. 525–574 (cit. on pp. 80, 84, 102, 108).

Haile, P., H. Hong, and M. Shum (2006). “Nonparametric Tests for Common values in First-Price Sealed-Bid Auctions”. *Econometric Society World Congress 2000 Contributed* (cit. on pp. 80, 106).

Hall, B. H. and B. Khan (2002). “Adoption of New Technology”. November, pp. 1–38 (cit. on p. 68).

Harrington, J. E. J., K. Hüschelrath, and U. Laitenberger (2016). “Rent Sharing to Control Non-Cartel Supply in the German Cement Market” (cit. on p. 80).

Hewitt, C., J. McClave, and D. Sibley (1996). “Incumbency and Bidding Behavior in the Dallas-Ft. Worth School Milk Market.” *Mimeo, University of Texas at Austin* (cit. on p. 79).

Hollenstein, H. (2004). “Determinants of the adoption of Information and Communication Technologies (ICT) An empirical analysis based on firm-level data for the Swiss business sector”. *Structural Change and Economic Dynamics* 15.3, pp. 315–342 (cit. on p. 9).

Hollenstein, H. and M. Woerter (2008). “Inter- and intra-firm diffusion of technology: The example of E-commerce. An analysis based on Swiss firm-level data”. *Research Policy* 37.3, pp. 545–564 (cit. on p. 9).

Ifrach, B. and G. Weintraub (2017). “A framework for dynamic oligopoly in concentrated industries”. *Review of Economic Studies* 84.3, pp. 1106–1150 (cit. on pp. 5, 38, 39, 43, 126, 128).

Igami, M. (2017). “Estimating the Innovatorâs Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981â1998”. *Journal of Political Economy* 125.3, pp. 798–847 (cit. on p. 10).

Igami, M. and K. Uetake (2017). “Mergers, Innovation, and Entry-Exit Dynamics: The Consolidation of the Hard Disk Drive Industry, 1996-2015” (cit. on p. 10).

Jeon, J. (2017). “Learning and Investment under Demand Uncertainty in Container Shipping”, pp. 1–59 (cit. on p. 5).

Karaca-Mandic, P. (2003). “Network Effects in Technology Adoption: The Case of {DVD} Players” (cit. on p. 8).

Kopp, P. (2016). *Le cinema a l'epreuve des phenomenes de concentration*. (Cit. on p. 15).

Krishna, V. (2006). *Auction Theory*. Elsevier Science (cit. on p. 83).

Kwoka Jr, J. E. (1997). “The Price Effects of Bidding Conspiracies: Evidence from Real Estate Auction Knockouts”. *Antitrust Bulletin* 42, p. 503 (cit. on p. 79).

Lanzillotti, R. F. (1996). “The Great School Milk Conspiracies of the 1980s”. *Review of Industrial Organization* 11.4, pp. 413–458 (cit. on p. 79).

Lebrun, B. (1998). “Comparative Statics in First Price Auctions”. *Games and Economic Behavior* 25.1, pp. 97–110 (cit. on pp. 85–88).

Lee, I. K. (1999). “Non-Cooperative Tacit Collusion , Complementary Bidding and Incumbency Premium”. *Review of Industrial Organization*, pp. 115–134 (cit. on p. 79).

Lee, R. S. (2013). “Vertical Integration and Exclusivity in Two-Sided Markets”. *American Economic Review* 103.7, pp. 2960–3000 (cit. on p. 8).

Levin, D. and J. L. Smith (1994). “Equilibrium in Auctions with Entry”. *The American Economic Review* 84.3, pp. 585–599 (cit. on p. 120).

Levin, S. G., S. L. Levin, and J. B. Meisel (1992). "Market Structure , Uncertainty , and Intrafirm Diffusion : The Case of Optical Scanners in Grocery Stores". *The Review of Economics and Statistics* 74.2, pp. 345–350 (cit. on p. 9).

Li, H. and J. G. Riley (2007). "Auction choice". *International Journal of Industrial Organization* 25.6, pp. 1269–1298 (cit. on pp. 80, 109).

Magnac, T. and D. Thesmar (2002). "Identifying Dynamic Discrete Decision Processes". *Econometrica* 70.2, pp. 801–816 (cit. on p. 44).

Majumdar, S. K. and S. Venkataraman (1998). "Network effects and the adoption of new technology: evidence from the U. S. telecommunications industry". *Strategic Management Journal* 19.11, pp. 1045–1062 (cit. on p. 8).

Mansfield, E. (1963). "Intrafirm Rates of Diffusion of an Innovation". *The Review of Economics and Statistics* 45.4, pp. 348–359 (cit. on p. 9).

Marshall, R. C. and L. M. Marx (2007). "Bidder collusion". *Journal of Economic Theory* 133.1, pp. 374–402 (cit. on p. 81).

Marshall, R. C., M. J. Meurer, et al. (1994). "Numerical Analysis of Asymmetric First Price Auctions". *Games and Economic Behavior* 7.2, pp. 193–220 (cit. on p. 109).

Maskin, E. and J. Riley (2000). "Asymmetric Auctions". *Review of Economic Studies* 67.3, pp. 413–438 (cit. on pp. 81–83).

McAfee, P. R. and J. McMillan (1992). "Bidding Rings". *The American Economic Review* 82.3, pp. 579–599 (cit. on p. 81).

Nabseth, L. and G. F. Ray (1974). "The Diffusion of new industrial processes: an international study." *Cambridge: University Press* (cit. on p. 9).

Pesendorfer, M. (2000). "A study of collusion in first-price auctions". *Review of Economic Studies* 67, pp. 381–411 (cit. on pp. 76, 79, 82, 83, 91, 105).

Plum, M. (1992). "Characterization and Computation of Nash Equilibria for Auctions with Incomplete Information". *International Journal of Game Theory* 20, pp. 393–418 (cit. on p. 81).

Porter, R. H. and J. D. Zona (1993). “Detection of Bid Rigging in Procurement Auctions”. *Journal of Political Economy* 101.3, pp. 518–538 (cit. on p. 79).

— (1999). “Ohio School Milk Markets: An Analysis of Bidding”. *The RAND Journal of Economics* 30.2, pp. 263–288 (cit. on pp. 74, 79, 116).

Rao, A. and W. R. Hartmann (2015). “Quality vs. Variety: Trading Larger Screens for More Shows in the Era of Digital Cinema”. *Quantitative Marketing and Economics* (cit. on p. 10).

Romeo, A. A. (1975). “Interindustry and Interfirm Differences in the Rate of Diffusion of an Innovation”. *The Review of Economics and Statistics* 57.3, pp. 311–319 (cit. on p. 9).

Ryan, S. P. and C. Tucker (2012). “Heterogeneity and the dynamics of technology adoption”. *Quantitative Marketing and Economics* 10.1, pp. 63–109 (cit. on p. 8).

Saloner, G. and A. Shepard (1995). “Adoption of Technologies with Network Effects : An Empirical Examination of the Adoption of Automated Teller Machines”. *The RAND Journal of Economics* 26.3, pp. 479–501 (cit. on p. 8).

Samuelson, W. (1985). “Competitive bidding with entry costs”. *Economics Letters* 16.1, pp. 8–37 (cit. on p. 120).

Schmidt-Dengler, P. (2006). *The Timing of New Technology Adoption: The Case of MRI.* 2006 Meeting Papers 3. Society for Economic Dynamics (cit. on p. 10).

Takahashi, Y. (2015). “Estimating a war of attrition: The case of the US movie theater industry”. *American Economic Review*, pp. 2204–2241 (cit. on p. 10).

Verma, N. (2001). “Larger Cinemas Outperform Smaller Ones”. *Quarterly Bulletin from the Culture Statistics Program* (cit. on p. 47).

Vickrey, W. (1961). “Counterspeculation, Auctions, and Competitive Sealed Tenders.” *The Journal of Finance* 16.1, pp. 8–37 (cit. on p. 81).

Waldfogel, J. (2016). “Cinematic Explosion: New Products, Unpredictabilty and Realized Quality in the Digital Era”. *Journal of Industrial Economics* 64.4, pp. 755–772 (cit. on p. 10).

Weintraub, G., L. Benkard, and B. Van Roy (2008). “Markov Perfect Industry Dynamics With Many Firms”. *Econometrica* 76.6, pp. 1375–1411 (cit. on p. 38).