2017

Essays In Industrial Organization And Applied Microeconomics

Peichun Wang
University of Pennsylvania, peichun21@gmail.com

Essays In Industrial Organization And Applied Microeconomics

Abstract
This dissertation consists of three essays in the areas of Industrial Organization and Applied Microeconomics.

The first essay studies high-tech firms' product portfolio choices under competition. I develop a model of dynamic portfolio adjustments in the context of the Chinese smartphone market, using the product life cycle as an empirically tractable heuristic to capture firms' dynamic incentives in new product introductions. I first show that product life cycles endogenously arise in markets with rapid technological innovations, are heterogeneous across products, and are affected by the level of market competition. I then estimate smartphone demand and manufacturers' variable, maintenance and sunk introduction costs on a detailed monthly market-level dataset of Chinese smartphones between 2009 and 2014. Finally, I use a 2012 large-scale pro-competitive policy introduced by the Chinese government as an experiment to decompose the handset manufacturers' incentives to introduce new products and show that the increased competition reduces the average product's short-run profits by 5% but its lifetime profits by 41% by shrinking its product life cycle. These dynamic incentives have large implications for both consumer welfare gains from product variety and the speed of technology adoption in this market.

In the second essay, my co-authors and I explore the sensitivity of the U.S. government's ongoing incentive auction to multi-license ownership by broadcasters. We document significant broadcast TV license purchases by private equity firms prior to the auction and perform a prospective analysis of the effect of ownership concentration on auction outcomes. We find that multi-license holders are able to raise spectrum acquisition costs by 22% by strategically withholding some of their licenses to increase the price for their remaining licenses. We analyze a potential rule change that reduces the distortion in payouts to license holders by up to 80%, but find that lower participation could greatly increase payouts and exacerbate strategic effects.

The third essay studies whether liberalizations of gun permits in the U.S. deterred violent crimes. Setting off an ongoing controversy, Lott and Mustard (1997) famously contended that so-called shall-issue laws (SILs) deterred violent crime. In this controversy the weapon of choice has been the differences-in-differences (DD) estimator applied to state and county panel data spanning various intervals of time. By treating violent crime as a career choice, this essay brings to bear a more general method, a cohort panel data model (CPDM) that incorporates the fundamental dynamic insights regarding entering and exiting a career. Our model distinguishes among three key parameters that jointly determine the effect of SILs on crime, (i) a direct effect on entry decisions, (ii) a surprise effect on exit decisions by individuals who entered criminal careers prior to the passage of SILs, and (iii) a selection effect on exit decisions by those who entered in the presence of SILs. We find significant and time-invariant results that reject the deterrence hypothesis as well as the DD model specification. Our results suggest that passages of SILs contribute to about one third of total violent crimes in 2011, particularly through higher turnover of violent criminals.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

ESSAYS IN INDUSTRIAL ORGANIZATION AND APPLIED MICROECONOMICS

Peichun Wang

A DISSERTATION

in

Applied Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Ulrich Doraszelski, Professor of Business Economics and Public Policy

Graduate Group Chairperson

Catherine Schrand, Celia Z. Moh Professor, Professor of Accounting

Dissertation Committee

Katja Seim, Associate Professor of Business Economics and Public Policy

Michael Sinkinson, Assistant Professor of Business Economics and Public Policy

Dedicated to

my mother and father, Aiping and Tong,

for the values, knowledge, and love they have instilled in me.

ACKNOWLEDGEMENTS

I am deeply indebted to my thesis advisors, Ulrich Doraszelski, Katja Seim, and Michael Sinkinson, without whom this dissertation would not have been possible. I owe my development throughout graudate school, both professionally and personally, to them, and I aspire to continue my career as an economist after them.

I would further like to thank my undergraduate advisors from Duke University, Marjorie B. McElroy, Andrew Sweeting, and James W. Roberts, for their encouragement to pursue economics and their continued support since.

I am grateful for helpful feedback from many others in the Business Economics and Public Policy Department at the Wharton School, the Department of Economics at the University of Pennsylvania, and elsewhere: Thomas Wollmann, Aviv Nevo, Kent Smetters, Robert Miller, Peter Newberry, Sarah Moshary, Ying Fan, Greg Lewis, Ben Lockwood, Corinne Low, Yao Luo, Todd Sinai, Jessie Handbury, Kevin Williams, Mike Abito, Joe Mazur, Dan Sacks, Joseph Harrington, Ashley Swanson, Ana Gazmuri, Juan Pablo Atal, Matthew Grennan, David Schindler and Matthew Weinberg.

Many helpful conversations on China have also emerged from the research group with Ann Harrison, Marshall Meyer, Linda Zhao, and Minyuan Zhao. I especially thank Ann Harrison for her guidance since the early stage of my research.

Yifan He and Siyang Xu provided excellent research assistance. I thank the Mack Institute for Innovation Management and Penn Center for the Study of Contemporary China for their financial support.

I was fortunate to work with the Applied Economics Ph.D. Coordinators Fernando V. Ferreira and Joseph Harrington, staff members Diana Broach and Christine Palumbo, and

Gidget Murray from the Wharton Doctoral Programs, who collectively run an excellent graduate program.

Finally, I thank my friends and family for their unconditional love and encouragement throughout the writing of this dissertation. In particular, I thank my classmates Jin Soo Han, Xingtan Zhang, Ben Hyman, Lindsay Relihan, Boris Vabson, Andy Wu, Jeremy Yoo, Hongyu Xiao and Amanda Chuan, and my parents, Aiping and Tong.

ABSTRACT

ESSAYS IN INDUSTRIAL ORGANIZATION AND APPLIED MICROECONOMICS

Peichun Wang

Ulrich Doraszelski

This dissertation consists of three essays in the areas of Industrial Organization and Applied Microeconomics.

The first essay studies high-tech firms' product portfolio choices under competition. I develop a model of dynamic portfolio adjustments in the context of the Chinese smartphone market, using the product life cycle as an empirically tractable heuristic to capture firms' dynamic incentives in new product introductions. I first show that product life cycles endogenously arise in markets with rapid technological innovations, are heterogeneous across products, and are affected by the level of market competition. I then estimate smartphone demand and manufacturers' variable, maintenance and sunk introduction costs on a detailed monthly market-level dataset of Chinese smartphones between 2009 and 2014. Finally, I use a 2012 large-scale pro-competitive policy introduced by the Chinese government as an experiment to decompose the handset manufacturers' incentives to introduce new products and show that the increased competition reduces the average product's short-run profits by 5% but its lifetime profits by 41% by shrinking its product life cycle. These dynamic incentives have large implications for both consumer welfare gains from product variety and the speed of technology adoption in this market.

In the second essay, my co-authors and I explore the sensitivity of the U.S. government's ongoing incentive auction to multi-license ownership by broadcasters. We document significant broadcast TV license purchases by private equity firms prior to the auc-

tion and perform a prospective analysis of the effect of ownership concentration on auction outcomes. We find that multi-license holders are able to raise spectrum acquisition costs by 22% by strategically withholding some of their licenses to increase the price for their remaining licenses. We analyze a potential rule change that reduces the distortion in payouts to license holders by up to 80%, but find that lower participation could greatly increase payouts and exacerbate strategic effects.

The third essay studies whether liberalizations of gun permits in the U.S. deterred violent crimes. Setting off an ongoing controversy, Lott and Mustard (1997) famously contended that so-called shall-issue laws (SILs) deterred violent crime. In this controversy the weapon of choice has been the differences-in-differences (DD) estimator applied to state and county panel data spanning various intervals of time. By treating violent crime as a career choice, this essay brings to bear a more general method, a cohort panel data model (CPDM) that incorporates the fundamental dynamic insights regarding entering and exiting a career. Our model distinguishes among three key parameters that jointly determine the effect of SILs on crime, (i) a direct effect on entry decisions, (ii) a surprise effect on exit decisions by individuals who entered criminal careers prior to the passage of SILs, and (iii) a selection effect on exit decisions by those who entered in the presence of SILs. We find significant and time-invariant results that reject the deterrence hypothesis as well as the DD model specification. Our results suggest that passages of SILs contribute to about one third of total violent crimes in 2011, particularly through higher turnover of violent criminals.

Contents

[1]Joint with Ulrich Doraszelski, Katja Seim, and Michael Sinkinson, the Wharton School of the University

of Pennsylvania.

[2]Joint with Marjorie B. McElroy, Duke University.

List of Tables

List of Figures

Chapter 1

Product Portfolio Choices with Product Life Cycles

1.1 Introduction

High-tech products often exhibit product life cycles (Hughes, 1990)[3], or time paths of sales (or profits) from a product's initial release into the marketplace until its discontinuation[4], that are persistently bell-shaped (Cox, 1967; Day, 1981)[5], albeit with much variation between products (Polli and Cook, 1969; Wood, 1990)[6]. Such time paths have mostly been explained by the slow diffusion of information: consumers are unsure about the quality of new products, and learn through word-of-mouth, repeated purchases, or imitation (Bass, 1969; Harrell and Taylor, 1981; Kwoka, 1996); or are only partially aware of new product introductions, and rely on advertising to expand their choice sets (Goeree, 2008). These

[3]The origin of the term "product life cycle" can be traced to Schumpeter (1934). The concept was popularized in the 1960s in marketing (Levitt, 1965) by drawing an analogy with the life cycle of biological creatures (Tellis and Crawford, 1981).

[4]In this paper, I refer to product life cycle as the sales path of a particular product. Product life cycles can also refer to the evolution of the overall market size of an industry (Jovanovic, 1994; Klepper, 1996); the technological development of a new prototype before reaching the market (Terzi et al., 2010); or export patterns from developed to developing countries (Vernon, 1966; Segerstrom, Anant and Dinopoulos, 1990).

[5]In marketing, this is also described as the four stages of product life cycles: introduction, growth, maturity, and decline, also sometimes referred to as an S-curve.

[6]Criticisms of the rigid description of the birth-growth-maturity-death path of PLC (Wood, 1990; Michelle Grantham, 1997) are based on the fact that PLCs are often heterogeneous across products. These critics, however, concede that the PLC theory provides a useful framework for firms' new-product strategies (Dhalla and Yuspeh, 1976). This paper explicitly explains the heterogeneity of PLCs across products and how firms take them into account.

demand-side explanations are likely incomplete, however. Such time paths are also affected by technological innovations that push the frontier of the quality spectrum and drive down production costs. More importantly, firms compete in product portfolios rather than on a product-by-product basis. These features point to a different explanation for product life cycles: firms' strategic incentives for innovation and product introductions.

What are firms' incentives in product introductions? How does competition affect firms' product portfolio choices in technologically progressive markets? Firms' product portfolio competition is key to understanding consumer welfare gains from product variety. In high-tech industries, these incentives also determine the speed at which products with new technology are introduced to the market.

In a differentiated product market, firms' introductions of similar products steal business from each other while differentiated products may help expand the market (Spence, 1976; Berry and Waldfogel, 1999). Firms' portfolio adjustments involve more considerations: A multi-product monopolist can offer a menu of products to screen consumers with heterogeneous preferences (White, 1977; Crawford and Shum, 2007); The effect of competition on multi-product firms' portfolio choice is theoretically ambiguous even in duopoly markets (Johnson and Myatt, 2003; Chu, 2010)—the incumbent firm can respond to entry by either expanding or contracting its product portfolio depending on demand and cost characteristics. Furthermore, different from mature industries, high-tech markets typically exhibit fast-changing variable costs and quality availability with high product development costs, inducing firms to be forward-looking in product introductions. Finally, the fast pace of technological innovation also quickly expands the breadth of quality covered by the typical product portfolio, generating large gains to variety from heterogeneous demand. To address how competition affects firms' product portfolio choices and to decompose firms' incentives (both static and dynamic) in product introductions, this paper develops a tractable model of forward-looking firms' equilibrium product portfolio choices.

I study this question in the context of the $70 billion (annual) Chinese smartphone market. The Chinese smartphone industry experienced rapid technological growth between 2009 and 2014. During this period, the quality of smartphones improved significantly: For example, the frontier CPU clock speed improved from 1GHz to 2.7GHz. At the same time, production costs fell: Total component costs for a low-end 3G smartphone dropped from $90 in mid-2011 to $43 in mid-2013[7]. Handset manufacturers also carried many products and updated frequently. On average, a major manufacturer sold 15.4 products in a market-month, and an average product was available for sale for 21.9 months. Compared to the component costs, upfront development costs in this industry were much higher—for instance, estimated to be more than $200,000 for a domestic low-end handset in 2012[8].

I first provide a stylized model to show that the simple assumption of Moore's Law (Moore, 1965)—decreasing variable costs and expanding quality frontier[9]—is sufficient to endogenously generate bell-shaped product life cycles. This model also shows that the product life cycle peaks higher and tapers off more slowly for a higher-quality product and for the same product in a market with less competition at the time of its release. Descriptive evidence also shows that these characteristics (product quality, market competition) at the time of a product's initial release have strong predictive powers for the eventual realization of lifetime sales accumulated over the product's life cycle. These relationships form a basis for how firms can account for the variation in its product life cycle when introducing a new product. Interviews with smartphone product managers in China also suggest the use of product life cycle in portfolio adjustments.

With this completed, I then turn to addressing my two main research questions: how does competition affect firms' product portfolio choices and what are firms' static and dynamic incentives in product introductions? To address the former, I first develop a model of smartphone manufacturers' product portfolio choices. Every month, firms adjust their portfolios by introducing new products and/or discontinuing existing products, taking into account the static tradeoffs that include per-period maintenance and marginal costs, as well as the product portfolio competition that anticipates second-stage Bertrand price com-

[7]Nomura Global Markets Research, "China Smartphone chips: LTE changes the balance," https://www.nomura.com/events/china-investor-forum/resources/upload/China_Smartphone_chips.pdf, accessed March 27, 2016.

[8]http://www.cctime.com/html/2011-7-14/2011713151032991.htm, in Chinese, accessed October 15, 2016. The average handset testing fee of $32,000 was estimated to be about 15% of total development costs per product.

[9]The original observation in Moore (1965), referred to as Moore's Law, was that the number of transistors in an integrated circuit doubles approximately every two years. This observation has since been revised and rephrased in different ways by technology executives.

petition, based on consumers' preferences over prices and product characteristics. When introducing a new product, firms also evaluate the dynamic tradeoffs between the sunk introduction cost and the expected future stream of profits. In particular, I assume that firms have rational expectations for the future evolution of technologies and market structures, conditional on the observable characteristics of the new product and the market at its release time.

Addressing the latter question—what are firms' static and dynamic incentives in product introductions?—is the main empirical challenge in this paper. The dynamic game of firms' portfolio choices is intractable, given the proliferation of products. Forming expectations for the future evolution of technologies and market structures requires managers to track more than trillions of possible product configurations. This paper proposes a heuristic way to approximate firms' decisions about a new product's introduction by relating the pattern of its product life cycle to what firms observe at the time of its release. A closely related approach is proposed in Wollmann (2016) to simplify the dynamic product introduction problem using the hurdle rate, which is effectively a way of net present value calculation based on a fixed time path of profits across products and markets. I show that in more mature industries with slow innovation, such as the commercial truck industry studied by Wollmann (2016), product life cycles are much flatter and homogeneous, and therefore my model with product life cycles converges to one with hurdle rates. However, when bell-shaped product life cycles are salient, as is the case in smartphone or other high-tech markets, the heterogeneity of product life cycles recovers unbiased sunk costs for different products. More importantly, hurdle rates are assumed to be fixed in the counterfactual, while product life cycles in my model are endogenously affected by the level of competition, which changes in the counterfactual. This captures firms' dynamic incentive changes under different market structures.

To recover the parameters of how firms form beliefs about future profits, as well as their costs and the demand they face, this paper employs a proprietary dataset of monthly province-level mobile phone sales, prices, and characteristics in China between 2009 and 2014. During this time, mobile phone handsets were mainly sold through brick-and-mortar retailers, largely constraining consumers to shop within their province. Therefore, I define a market on the province-month level, yielding 2,201 markets from 71 months and 31 provinces of data. The rational expectation assumption allows me to estimate firms' beliefs about future profits using the observed technologies and market structures. The large variation in market competitiveness across Chinese provinces then identifies how competition affects product life cycles.

I estimate maintenance and sunk costs by forming bounds based on firms' revealed preferences[10]. Given the modular nature of smartphone production, scrap values are negligible in this industry[11]. The thin tails of product life cycles toward the product's end then allow me to interpret product discontinuation decisions statically, and also give me tight bounds on per-period maintenance costs. Sunk costs are estimated by comparing changes in both instantaneous and lifetime product profitability. Identification of maintenance and sunk costs comes from variations across provinces in market structure, demand characteristics, and timing differences in product introductions and discontinuations.

I illustrate the contributing factors to the speed of innovation via product introductions using a Chinese government experiment that allows me to study the effect of a large wave of small firm entry on incumbent firms' product portfolio choices by both directly changing the market competitiveness and indirectly changing the patterns of product life cycles. Specifically, in 2012, facing foreign high-tech firms that dominate the market, the Chinese government implemented a large-scale, pro-competitive policy to subsidize entries from small domestic smartphone manufacturers. Aimed at promoting market competition and domestic high-tech manufacturing, the policy cost billions of dollars[12] and induced a large

[10]Following the recent literature on using bounds (Pakes et al., 2015; Wollmann, 2016), instead of parametric assumptions, to treat fixed costs in the presence of multiple equilibria, I assume that the observed product configurations in markets are Nash equilibrium outcomes. In other words, no unilateral change in product portfolio choice can be profitable for any firm, yielding both upper and lower bounds for product maintenance and sunk introduction costs. Similar to Berry, Eizenberg and Waldfogel (2016) and Fan and Yang (2016), I make use of the computed bounds directly and also feed them into a inequality penalty function to obtain point estimates.

[11]Manufacturers also do not face significant capacity and liquidity constraints in this industry, given their large size and the presence of contract factories. This eliminates much of the opportunity costs of introducing and maintaining a product in this market. More details are provided in Section 1.2.

[12]The policy first lowered handset testing fees by the Ministry of Industry and Information Technology (MIIT). http://www.cctime.com/html/2011-7-14/2011713151032991.htm, in Chinese, accessed October 15, 2016. The report suggests that this policy cut about 4% of total development costs per product, decreasing the government's revenue by roughly $20 million, based on the number of new products in my data. The policy also urged state-owned telecom carriers to spend up to $10 billion in 2012 on marketing mainly for small domestic firms' products. http://tech.qq.com/a/20111229/000116.htm, in Chinese, accessed October 15, 2016.

inflow of fringe firms[13]: from 92 to 318 between 2012 and 2013, making the smartphone market more competitive[14]. Given the recency of the policy, I ask the question of how successful it was in promoting market competition[15]. On the surface, the average incumbent firm's portfolio size went up from 16.3 to 20.2 products, and the median handset price was consistently dropping from RMB 1,693 to 1,306. However, as technology was quickly improving in this market, the quality frontier was expanding while component costs were falling, promoting such market evolution even in the absence of an industrial policy.

With my estimates, undoing the Chinese competitive policy reveals two results. First, decomposition of firms' incentives to introduce new products shows that the dynamic consideration is important in firms' portfolio choices. Increased competition reduces the average product's short-run profits by 5% but its lifetime profits by 41% by shrinking its product life cycle. Second, the policy might not have been as successful as it seemed in increasing variety: Firms would have been able to self-regulate through product introductions. In the absence of the costly fringe entry, incumbent major manufacturers would have introduced seven more handsets per province-month (almost four times the observed number of new products), and the average handset price would have increased by only about $1 (0.5%) compared to what was observed after the policy. Annualized total welfare would have been only $0.36 billion (0.4%) less, compared to the billions of dollars spent. Product life cycles generated by technological innovations play an important role in firms' self-regulation: Not accounting for firms' future outlook would overstate the welfare benefit of the policy by $430 million.

This paper makes three main contributions. The first is its demonstration that product life cycles naturally arise in markets with rapid technological innovations. This expands the reduced-form view of product life cycles in the literature. In explaining the life cycle of products in various markets, previous studies have either included product age effects in their specifications of consumer utility (Moral and Jaumandreu, 2007; Ngwe, 2016) or estimated hazard rates of product turnover based on product age (Stavins, 1995; Greenstein

[13]These firms are considered fringe in this paper, given their relatively homogeneous and low-quality products, as well as their low market shares. More details are provided in Section 1.2.

[14]Fringe firms gained about 20% of total market share. The market was already relatively competitive: The average Herfindahl Index went from 1189 to 882 between January of 2012 and 2013.

[15]I do not attempt to address the other policy goal, of promoting domestic high-tech manufacturing.

and Wade, 1998).

The second contribution is to develop an empirically tractable model of equilibrium product portfolio choices of forward-looking firms. The empirical entry literature has cleanly characterized firms' static tradeoffs in product choices (Mazzeo, 2002; Seim, 2006). Advances in the estimation of dynamic games (Hotz et al., 1994) have enabled empirical work to capture firms' dynamic incentives in markets with smaller state space (Bajari, Benkard and Levin, 2007; Blonigen, Knittel and Soderbery, 2013; Collard-Wexler, 2013; Ryan, 2012; Sweeting, 2013). I contribute to these two literatures by accounting for firms' dynamic incentives for product introductions in the large product space—typical in high-tech markets—through the use of product life cycles as an empirically tractable heuristic. The results of this paper also contribute to studies on the effect of changes in market structure on welfare (Fan, 2013; Li et al., 2016; Wollmann, 2016), the value of technological innovations (Eizenberg, 2014; Nosko, 2014), and welfare gains from product variety (Berry, Eizenberg and Waldfogel, 2016; Fan and Yang, 2016).

The third contribution is its analysis of how changes in the level of competition affect product variety and welfare in the market. I decompose firms' static and dynamic incentives in product introductions and show that ignoring the dynamics would significantly overestimate the benefits of industrial policies aimed at promoting competition. This contributes to the discussion of the costs and benefits of industrial policies (Greenwald and Stiglitz, 2006; Aghion et al., 2015) adopted by many developing countries in recent decades.

The rest of the paper proceeds as follows. Section 1.2 introduces the empirical setting and provides descriptive evidence of product life cycle properties. Section 1.3 presents the empirical model. Section 1.4 discusses the estimation strategy and results. Section 1.5 describes the policy setting and compares counterfactual welfare estimates from different entry models. Section 1.6 concludes.

1.2 The Chinese smartphone industry: Data and evidence

In this section, I introduce the empirical setting and data. I then briefly discuss the intuition for product life cycle formation in high-tech markets, which I analyze in detail in Appendix A.1. Finally, I present descriptive evidence of the product life cycle and its properties.

Overview

The Chinese smartphone industry is a significant and ideal market for studying high-tech firms' product portfolio choices. Smartphone manufacturers sold more than 300 million handsets in 2014, with a total of $70 billion in revenue. The Chinese smartphone industry during this time was a typical high-tech market, with fast-improving technologies but decreasing retail prices (Figure 1.2.1). Several features of this market make it particularly attractive. Unlike the US smartphone market, smartphone manufacturers in China decide which models to introduce, as opposed to the carriers[16,17]; demand for smartphone handsets is also much more separable from carrier services[18]. As mentioned, dominant offline sales during this time[19] segments this market by geography. Finally, mobile phone manufacturing has become specialized during this time—manufacturers purchase components (e.g., chipsets, displays, cameras, etc.) from upstream firms[20], and only integrate them

[16]The three state-owned telecom carriers in China are (with subscribers as of May 2015, in millions): China Mobile (816); China Unicom (290); and China Telecom (191). Each operates a different network across different generations of telecommunications in China, allocated by MIIT. Network ownership during 2G: GSM (CM, CU) and CDMA (CT); 3G: TD-SCDMA (CM), WCDMA (CU), and EVDO (CT); and 4G: TD-LTE (CM, CU, CT) and FDD-LTE (CU, CT).

[17]One industry source suggests that the timing of different models/versions of a product—but typically not whether to introduce the product at all—can be influenced by carriers. In this paper, I focus on portfolio choices at the product level and collapse different models/versions of the same product into one. See Appendix A.2.2 for details. Moreover, I observe some cases in which products were compatible with some carriers but not others. This is often due to cost and demand, rather than contracting and bargaining. For example, CT has the smallest subscriber base and also operated CDMA in the 2G era, which is monopolized by Qualcomm with its intellectual property, and thus requires higher royalty payments from smartphone manufacturers to make and sell CDMA phones.

[18]The majority of smartphone handsets were sold contract-free. For the portion of sales through carriers, retail prices in the data also reflect handset prices, without accounting for promotions on carrier services. Carrier service quality is controlled for in the demand estimation with the networks they operate. I assume that residual variations in the carrier contracts and service quality are absorbed by time trend and market fixed effects of the outside good in the demand specification.

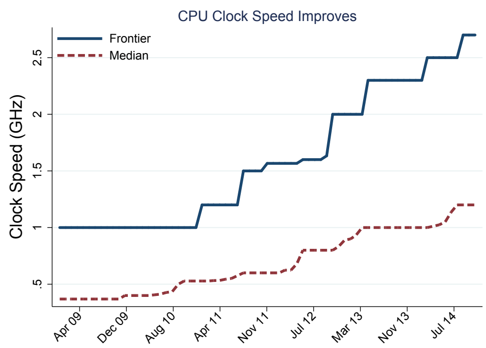
[19]About 10% of total handset sales were made online during 2009-2014. I drop all online sales in this paper, as I cannot observe their destinations.

[20]Several firms considered this paper are vertically integrated, e.g. Samsung. However, subsidiaries of these international conglomerates often remain separate entities in decision-making. For example, Samsung's handset manufacturer only ordered 60% of its batteries for its Galaxy Note 7 from Samsung's own battery

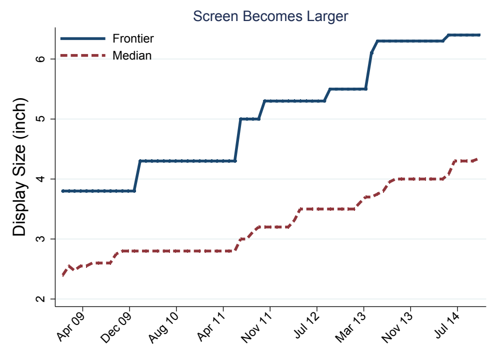
with their operating systems and software[21]. Such modular production allows handset manufacturers to easily adjust their product portfolios on a monthly basis, and also suggests that very little R&D is required for handsets below the technology frontier[22].

Figure 1.2.1: Typical high-tech industry dynamics: Improving quality and declining price

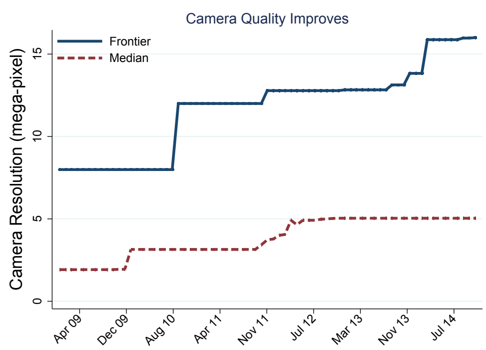
(a) Major characteristic (1): CPU clock speed



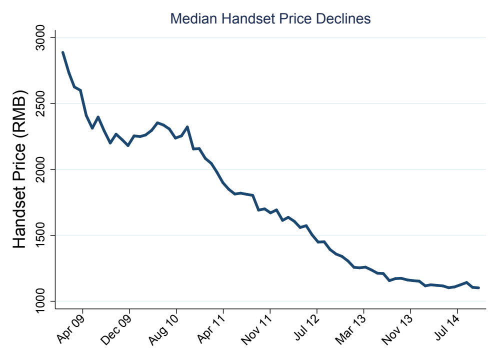
(b) Major characteristic (2): Display size



(c) Major characteristic (3): Camera resolution



(d) Median price



Notes: This figure plots the frontier and median of three main characteristics (CPU clock speed, display size, and camera resolution) of smartphone handsets, and the median selling price among all smartphone models from major manufacturers in China, between Jan. 2009 and Nov. 2014. This figure shows that the Chinese smartphone market during 2009-2014 exhibits the typical features of a high-tech industry: Product quality (technology frontier) is quickly improving; at the same time, prices are falling.

Data sources

Data for this paper come from four main sources. First, smartphone sales data come from GfK Market Research, and include the universe of mobile phone (feature phone and

manufacturer, SDI, and the other 40% from a Chinese manufacturer, ATL. I also account for these firms' potential cost advantages in production by including firm fixed effects in my cost estimations.

[21] The abundance of contract-based factories, design houses, and software developers in this industry (details in Appendix A.2.1) further specializes the handset production process.

[22] See, for example, https://www.bloomberg.com/news/articles/2015-07-13/how-1-000-buys-a-smartphone-brand-to-challenge-samsung, accessed November 6, 2016.

smartphone, separately) sales[23] in China between Jan. 2009 and Nov. 2014. There are 31 provinces and 71 months, giving me 2,201 markets. I observe unit sales, average prices, and handset characteristics at the handset model/province/year-month level[24]. Handset characteristics include all aspects of a smartphone[25]. For a parsimonious empirical specification, I include in my demand and cost estimation only three major characteristics—CPU clock speed, display size, and camera resolution—and a fourth characteristic, "Other", constructed from the first component of a principal component analysis of the other characteristics (summary statistics in Table 1.1).

I supplement the GfK data with hand-collected data from two electronics catalogue websites: GSMArena and its Chinese equivalent, ZOL[26], for each product's advertised product line name[27], to capture the demand effects of product line marketing as well as missing product characteristics[28]. These two websites, together with the GfK data, provide a complete set of characteristics for all 1,782 models of major smartphone handsets sold during this period, which I then collapse down to 691 unique products[29].

I match the smartphone manufacturers identified in GfK to the Annual Industrial Survey (AIS), which is a comprehensive firm/year level data of all medium and large enterprises in China. I construct firm-specific cost shifters using the differential government subsidies received by firms and the interest rates they pay on their loans each year[30].

[23]Available feature phone data allow me to specify demand that allows for substitutions between feature phones and smartphones.

[24]I also observe the types of retail channels (rather than the specific stores) a handset is sold through. I make use of the types of retailers to shift a product's maintenance cost per month in my estimation. Unless otherwise noted, I collapse data to the model/province/year-month level.

[25]These include CPU clock speed, display size, camera resolution, battery capacity, RAM, storage space, thickness of the phone, number of SIM card slots, near-field communication capability, compatibility with various carrier networks, etc.

[26]Websites: http://www.gsmarena.com/ and http://www.zol.com.cn/

[27]GfK collects information on product factory codes. The websites give me the matched advertised names of the products, allow me to put products into manufacturers' heavily advertised product lines, and capture the joint marketing effects in my demand estimation. For example, several models of the original Samsung Galaxy S are coded as Samsung I9008 in the GfK data.

[28]Several variables are missing observations from the GfK data. For example, while the GfK data does provide the model name of the chipsets and other characteristics, such as the number of cores, the variable CPU clock speed is mostly missing.

[29]Different versions of the same product are collapsed based on characteristics. See Appendix A.2.2 for details.

[30]All major smartphone manufacturers (defined in the section "Key players") that have legally registered for a separate entity in China are included in this data. With the exception of Apple, I match all other 11 major manufacturers with the AIS data. In the case of Apple, I can match it to its main manufacturer in China, Foxconn Technology Group, and argue that cost shifters of the contract manufacturer likely also affect Apple's

Data on the demographics of the markets come from the Dios Database. I observe the population of each province/year[31], which defines the size of the markets[32]. Dios data also provide information on the quintiles of annual consumption (henceforth referred to as income) separately for urban and rural residents, which I use to construct empirical distributions of income in each province/year[33].

Producers

This paper focuses on the product portfolio choices of the major smartphone manufacturers. I define a smartphone manufacturer as major if it obtains at least 5% national market share (in units) at any point during 2009-2014. This results in 12 major manufacturers, including, as loosely categorized in the industry in 2014, Apple and Samsung, who are the highest premium brands; Nokia, Motorola and HTC, who are on the decline in brand values; ZTE, Huawei, Coolpad, and Lenovo, who are mid-level domestic brands; and Xiaomi, Oppo, and Vivo, who are the new domestic high-end brands.

These 12 manufacturers constitute a relatively stable set of firms in the industry, accounting for more than 80% of total market share throughout this period (summary statistics in Table 1.2). While market shares move between these manufacturers (e.g., notably the fall of Nokia), only Xiaomi, Oppo, and Vivo are slight latecomers to the market[34], and the other nine firms are always present[35].

The rest of this market consists of small manufacturers I consider fringe[36]. They produce relatively more homogeneous and lower-quality products compared to the major

manufacturing costs.

[31]I only use the population between age 15 and 64, given their more likely use of smartphones.

[32]Similar to Nevo (2001), who defines market size as 1 cereal serving per person per day, I define market size as 1 handset per person per year.

[33]I assume that income distribution in each market is log-normally distributed, and draw from distributions estimated with the quintiles. See Appendix A.3.1 for more details.

[34]Oppo has been producing feature phones, but only started making smartphones in 2011. Vivo spun off from Oppo around the same time. Xiaomi is a completely new entrant that has grown to be one of the larger smartphone manufacturers. All three focus on higher-end handset production.

[35]There are also three major acquisition activities during this period: Motorola to Google (Aug. 2011), Nokia to Microsoft (Apr. 2014), and Motorola to Lenovo (Oct. 2014). Since the first two do not involve another incumbent manufacturer, and the third happened at the very end of my sample, I assume the same firm behavior before and after acquisition, and treat Lenovo and Motorola as separate firms throughout.

[36]These firms are often referred to as "white-box" manufacturers in the press, given the history of many of them as no-label phone manufacturers and the homogeneity of their products. As the technological and market barriers to entry were lowered around 2012, many more fringe firms entered the market and captured more market shares (up to 20% overall). See Section 1.5 for details of changes in 2012.

Table 1.1: Major smartphone handset characteristics

	Mean	Std	Min	Max	
CPU clock speed (GHz)	1.078	0.473	0.003	2.7	
Display size (inch)	4.131	0.954	1.5	6.4	
Main camera resolution (mega-pixels)	5.768	3.399	0	16.15	
					Correlations
Other characteristics (PCA 1^{st} comp.)	0	1.752	-3.837	6.769	
Battery capacity (mAh)	1,806	620.4	700	4,250	0.8741
Depth (mm)	11.27	3.286	4.9	28.2	-0.7448
RAM (mega-bytes)	792.9	657.9	1	3,072	0.9049
ROM (giga-bytes)	7.007	11.05	0.01	128	0.6881
# of SIM cards	1.391	0.488	1	2	0.2343
NFC	0.1225	0.3280	0	1	0.6364

Notes: Handset characteristics for smartphone handsets by major manufacturers. Data on model level after winsorizing on specialty phones but before collapsing on similar characteristics ($N = 1,755$). "Other" characteristic is the first component of the principal component analysis of the six other characteristics, with respective correlations shown in the last column. Display size is measured from the diagonal of phone screens (excluding edges that are not part of the screen); screen technologies (AMOLED, IPS, TFT, Retina, Super AMOLED, Super LCD, etc.) and resolutions (pixels per inch) are also observed, but are highly correlated with screen sizes (hence dropped). Main camera typically refers to the camera on the back of the phone; front camera and/or secondary camera on the back are also observed, but are dropped due to high correlation with the quality of the main camera. Depth measures the physical thickness of the handset in millimeters. RAM measures memory, and ROM measures storage of the handsets. Handsets in China also typically have 1-2 SIM card slots. NFC indicates near-field communication functionality.

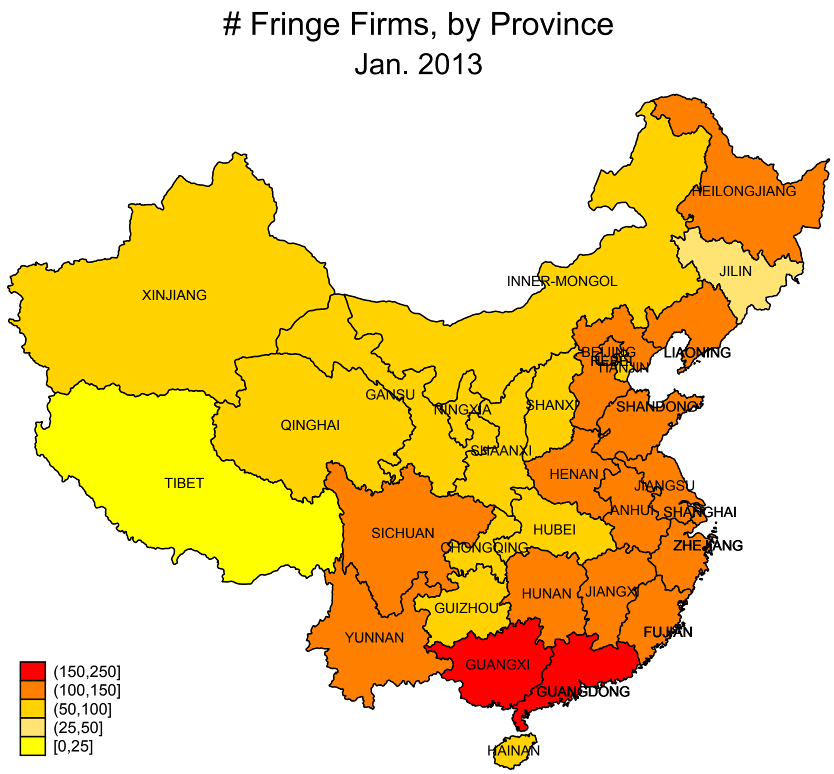
manufacturers, and mainly differ in their industrial design. Fringe firms also enter more into provinces with larger/wealthier populations, generating variations in market competitiveness across provinces. Figure 1.2.2 shows significant variation in the presence of fringe firms across markets in Jan. 2013. Figure 1.2.3 then documents the price variation of handsets by the major manufacturers across these markets. Other upstream, downstream, and related industries are summarized in Appendix A.2.1.

Table 1.2: Market shares of major manufacturers

	Market shares (in units)					
	2009	2010	2011	2012	2013	2014
Apple	0.5%	4.3%	7.2%	7.7%	6.0%	8.7%
Samsung	4.6%	4.2%	18.5%	21.5%	21.3%	17.1%
Nokia	77.6%	75.1%	36.3%	8.4%	2.4%	1.3%
Motorola	9.6%	7.2%	6.8%	4.0%	0.9%	0.2%
HTC	3.2%	2.6%	4.8%	5.1%	2.8%	1.8%
ZTE	0.1%	0.6%	6.5%	7.3%	5.5%	3.1%
Huawei	0.0%	0.8%	8.1%	8.9%	9.7%	10.3%
Coolpad	2.0%	1.7%	2.8%	8.3%	9.3%	10.0%
Lenovo	0.4%	0.7%	2.4%	9.4%	12.4%	10.5%
Xiaomi			0.0%	1.3%	2.6%	7.1%
Oppo			0.1%	1.8%	3.6%	5.1%
Vivo			0.0%	1.5%	3.5%	5.8%
Fringe total share	2.1%	2.9%	6.3%	15.0%	19.9%	19.0%
# Firms	41	44	107	347	471	495
Units sold (millions)	19.3	33.4	85.8	194	352	342
Value sold (billions of RMB)	43.7	73.5	176	343	546	520

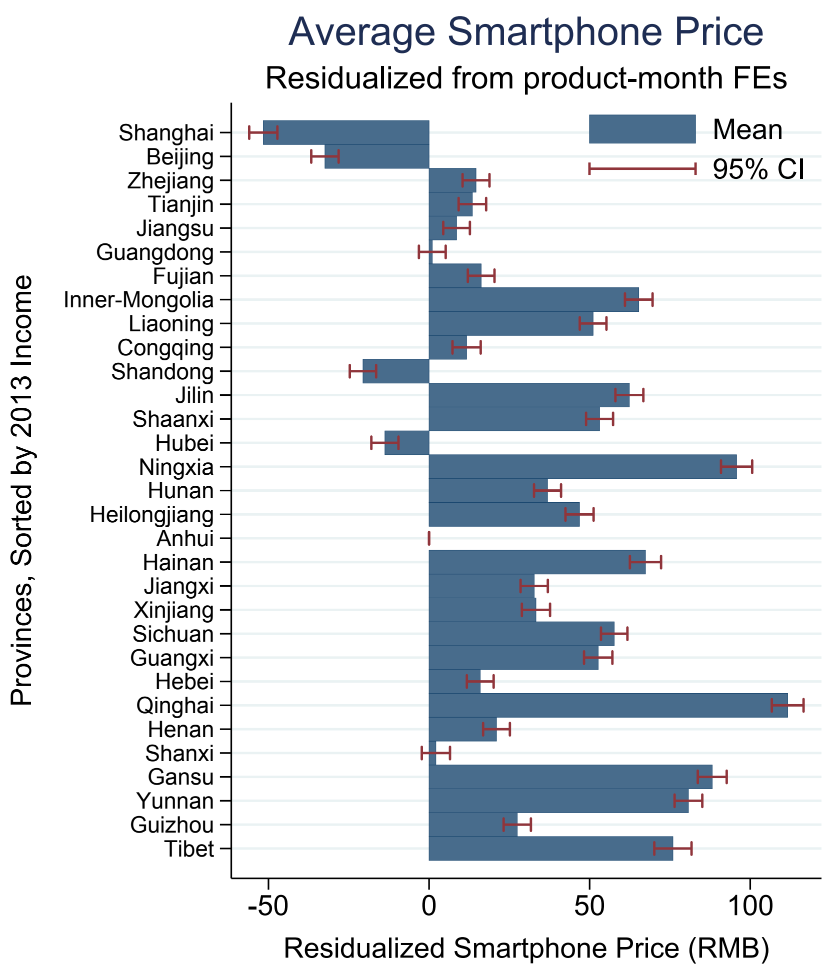
Notes: Data from GfK Market Research. The set of major manufacturers is relatively stable during my sample—only Xiaomi, Oppo, and Vivo are later entrants into this market. Firms are grouped by loose categorizations according to industry sources: Apple and Samsung command highest brand premiums; Nokia, Motorola, and HTC used to be dominant players, but have since been on the decline as iOS and Android gained popularity. ZTE, Huawei, Coolpad, and Lenovo are typically grouped together in industry reports, and are often referred to as "Zhong-Hua-Ku-Lian" in Chinese; these are domestic mid-level brands. Xiaomi, Oppo, and Vivo have focused more on higher-end products since their entry into the market. Total # of firms include the 12 major manufacturers. Total sales include both major and fringe smartphones (thus higher than the $70 billion revenue of major manufacturers mentioned before).

Figure 1.2.2: Geographical variation in market competition: Fringe presence



Notes: This figure shows geographic variations in market competition due to fringe firms' local concentration along the east coast of China. Plotted is the number of unique fringe firms in each province in Jan. 2013. Map coordinates follow Merryman (2008). Darker color indicates higher number of fringe firm presence in the province.

Figure 1.2.3: Major smartphone price variation across provinces



Notes: Plotted are avg. major smartphone handset price in each province, after being residualized from product-month fixed effects. Anhui Province is the omitted category. For example, this figure shows that, on average, the same handset sells in Tibet for almost 80 RMB higher than in Anhui in the same month. 95% confidence intervals are shown with error bars around the means. Strong price variation across provinces is still present after controlling for distances to east coast manufacturing hubs, and concentration of different types of retail channels within provinces. Provinces are listed on the y-axis, and sorted by average province income in 2013.

Products

This paper focuses on major smartphone manufacturers' portfolio choices of non-flagship products. A flagship product is defined, in this paper, as the highest-priced product of a manufacturer in each market-month[37]. The detailed monthly patterns of product introductions and discontinuations across markets allow me to identify entry and maintenance costs. Out of the 569 non-flagship products in the sample, I observe the initial national release dates of 513 products (due to left truncation). On average, they are eventually released into 25.6 markets (out of 31 provinces) with a standard deviation of 7.7 markets. Conditional on entry, Figure 1.2.4 shows that most products are released in all markets within 5 months, suggesting that entries across markets are likely not due to inventory management practices across markets[38]. I observe, for 8,319 product-markets, the discontinuation month of the product in that market, which serves as my sample for the estimation of maintenance costs. Of these, I observe both the entry and exit month of 7,405 product-markets that I will use to estimate firms' product life cycle beliefs and sunk introduction costs. Among the 7,405 observations, on average, a product lasts on the shelf for 21.9 months, with a standard deviation of 12.6 months. Moreover, the average firm portfolio at any time and market has 15.4 products, with a standard deviation of 11.8 products; covers a large range of products in terms of price (on average, more than half the market); and exhibits large variations across firms and time (summary statistics in Table 1.3).

Demand attributes

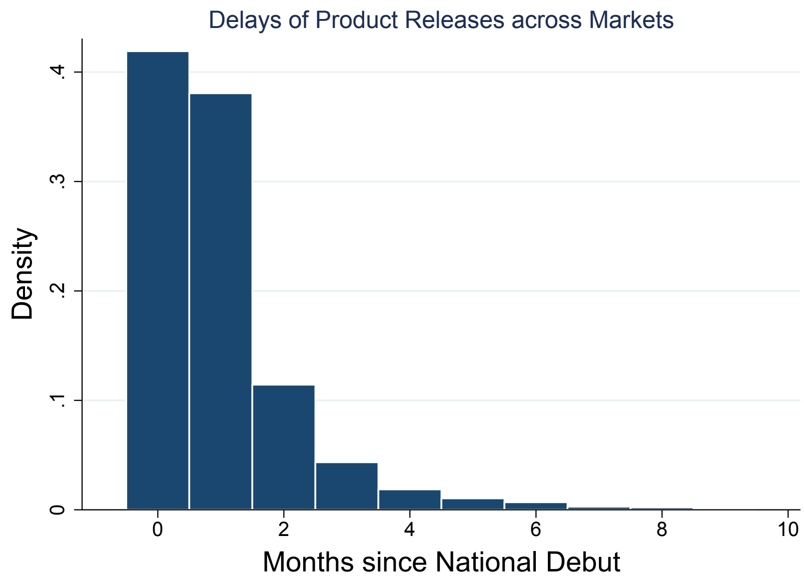
The mobile penetration rate has gone up substantially during this period in China. In 2009, there were only 56 mobile phone users per 100 population. This penetration rate has gone up to 95% by early 2015, with 1.29 billion mobile users[39]. During this time, smartphones are quickly replacing feature phones as the dominant type of mobile phone in China, as

[37] There are 122 unique flagship products in this sample. For example, this includes Samsung Galaxy S and Note. This also includes all of Apple's and Xiaomi's products, given their simple product lines.

[38] This practice is more likely to occur within markets in which manufacturers could move inventories among different retailers. In the global market, it is also common practice to ship obsolete products from more developed markets to emerging markets.

[39] http://www.miit.gov.cn/n11293472/n11293832/n11294132/n12858447/16505685.html, Ministry of Industry and Information Technology, in Chinese, accessed March 27, 2016.

Figure 1.2.4: Market-level product introduction delays



Notes: This figure shows that product introductions at the province level are unlikely due to inventory management practices, given that most products have short delays—within five months. These delays reflect firms' strategic product portfolio choices at the province level, given different local market conditions, and are used to identify products' sunk introduction costs. Plotted is the distribution of entry delays in months on the product-market level. Entry delays are defined as the lag of time between when a product is first introduced anywhere in China and when it is introduced to a particular province.

Table 1.3: Major manufacturers' product portfolio sizes

Avg. province-month	Portfolio size					
	2009	2010	2011	2012	2013	2014
Apple	1.857	2.197	2.411	2.720	3.454	4.317
Samsung	5.194	9.539	19.10	31.72	36.05	33.15
Nokia	21.59	29.48	29.90	31.09	25.97	17.95
Motorola	7.293	13.23	22.97	31.48	24.57	11.98
HTC	5.533	8.393	11.35	17.71	18.56	19.95
ZTE	1	1.585	3.838	10.70	18.70	21.42
Huawei	1	1.778	5.043	15.02	21.52	25.41
Coolpad	3.418	5.047	9.067	22.99	29.60	33.32
Lenovo	1.248	1.846	3.558	17.90	31.38	33.34
Xiaomi			1	1.038	2.290	4.331
Oppo			1.356	6.822	14.33	18.24
Vivo			1	6.483	15.55	18.47
Avg. # products (across firms)	5.348	8.122	9.216	16.31	20.16	20.16
Total # products (national)	94	153	238	361	443	489
Avg. range of price % rank	0.562	0.500	0.529	0.569	0.682	0.727

Notes: This table summarizes firms' product portfolio sizes. The number of products includes both flagship and non-flagship products, after collapsing on similar characteristics, winsorizing on specialty phones, and averaging across market-months. Firms are grouped based on loose categorizations according to industry standards. Total # products (national) only includes 12 major manufacturers' products; for fringe products, see Figure 1.5.1b. Range of price % rank is equal to a firm's highest-priced product's % rank in prices in the market minus its lowest, measuring the spread of the portfolio (ranges between zero and one). These statistics show that the average major manufacturer carries a product portfolio that covers more than half the quality spectrum in the market at any time.

shown in Figure 1.2.5. By early 2015, smartphone ownership has also reached 58% and varies substantially across income groups[40].

Figure 1.2.6 shows the large heterogeneity in income across markets, as well as the relative level of income compared to the average price of feature phones and smartphones. For example, the average income in Beijing is consistently more than 3 times higher than that of Tibet during this period. Moreover, the national average income in 2009 is RMB 7,928, while the average smartphone price is 2,258 and feature phone price is 771. By 2014, the national average income has grown to 14,603, while the smartphone price has fallen to 1,458 and feature phone price to only 319 (also with quality improvements, as suggested in Figure 1.2.1). Though not depicted, within-market income heterogeneity is also large. On average, the highest quintile of urban residents consumes more than 2.8 times more than the lowest urban quintile annually, and the median quintile of urban residents consumes about 2.7 times more than the median quintile of rural residents annually.

Intuition and evidence: Product life cycles

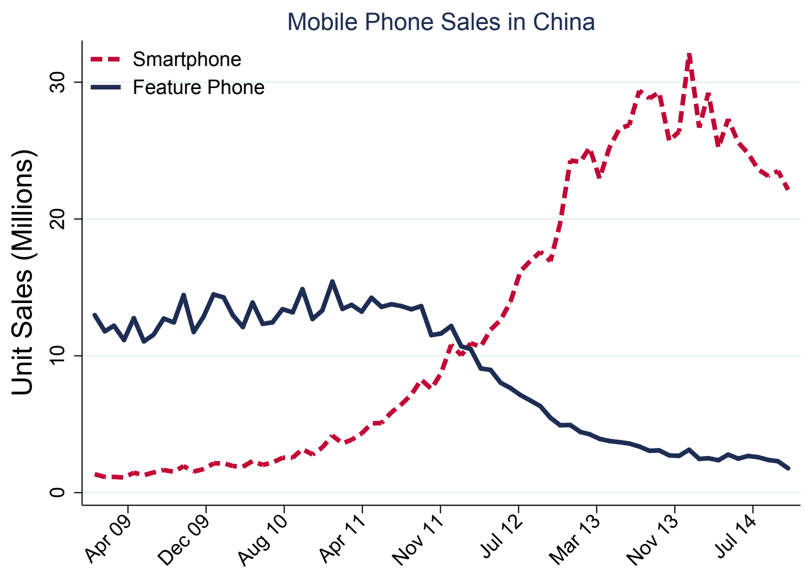
Figure 1.2.7 shows that product life cycles are pervasive in the Chinese smartphone market. I align the time paths of (unit) sales of all product-markets to their release times (age zero) and compute the 25^{th}, 50^{th}, and 75^{th} percentiles of sales within each age cohort, after normalizing by each product-market's first-month sales[41]. The resulting time paths of sales across all product-markets exhibit the typical bell shapes discussed in the early marketing product life cycle theory.

In Appendix A.1, I provide a stylized model of product life cycle formation in high-tech markets. I show that bell-shaped product life cycles endogenously arise in a model with a standard Logit demand system, decreasing production costs, and an expanding technology frontier. The intuition is as follows. In high-tech markets, Moore's Law suggests that production costs for new products fall faster than the average product in the market. As a result, the price of the new product also falls faster and generates more sales at the be-

[40] http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/, accessed March 27, 2016.

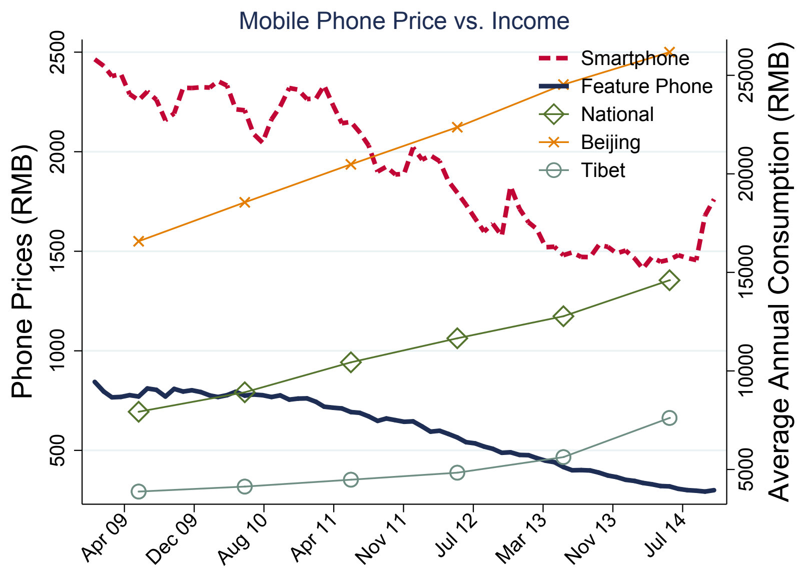
[41] The reason for this normalization is, one, for comparisons across products with different initial sales; and two, as evidenced in Figure 1.2.8 and specified in the model in Section 1.3, forward-looking managers forecast the ratio of lifetime and immediate payoffs.

Figure 1.2.5: Substitution between feature phone and smartphone



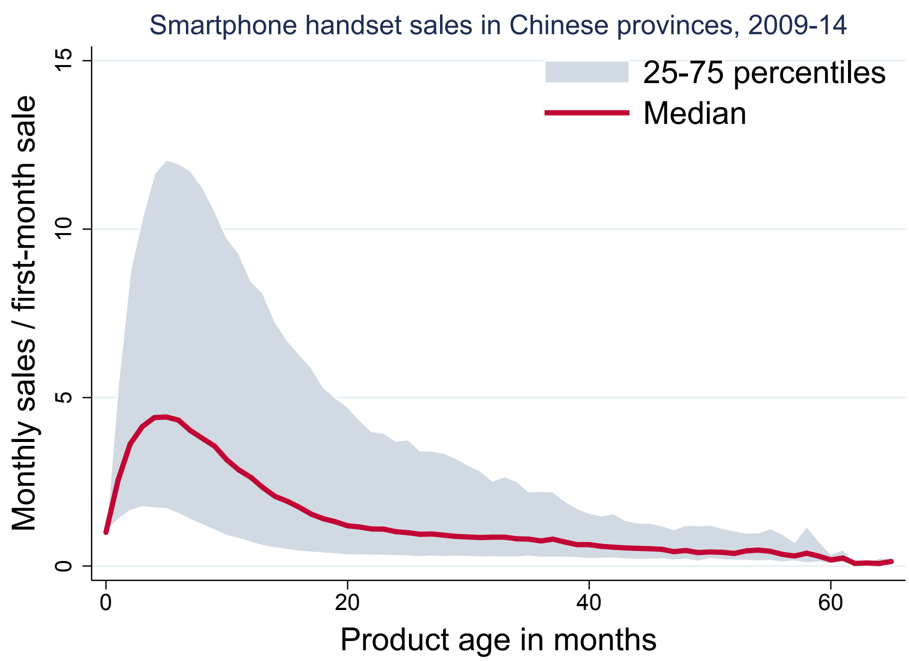
Notes: This figure shows that smartphones were quickly surpassing feature phones as the main type of mobile phone in China during this period. This figure also highlights the importance of accounting for substitutions between smartphones and feature phones in the demand estimation for the first half of the sample. Plotted are monthly national smartphone and feature phone unit sales (in millions of handsets) between Jan. 2009 and Nov. 2014.

Figure 1.2.6: Demand heterogeneity across markets and time



Notes: This figure shows demand heterogeneity across markets and time during this period in China. For example, the average income (consumption) in Beijing is consistently more than 3 times higher than that of Tibet. This gives rise to very different demand characteristics in these two markets, given the high prices of smartphones compared to income. These heterogeneities are also evolving over time, as mobile phone prices quickly fall, and income in China is fast growing during this period. Plotted are average smartphone handset prices (monthly, national); average feature phone handset prices (monthly, national); average national annual consumption (plotted at mid-year); and average annual consumption in Beijing and Tibet.

Figure 1.2.7: Product life cycles (PLCs)



Notes: This figure shows that bell-shaped product life cycles are pervasive in the Chinese smartphone market across provinces and products, and also exhibit much variation across product-markets. Plotted from monthly smartphone handset sales in China between Jan. 2009 and Nov. 2014; pooled across provinces and products; aligned by the release month of each product in each province. Monthly unit sales are normalized by release-month sales; the red line shows the median sale within age (in months) cohorts; the gray area shows the 25^{th} and 75^{th} percentiles.

ginning of the life cycle. As the speeds at which prices fall converge in a market, sales will also stabilize. As the quality frontier expands with new technology, sales eventually drop to zero (or below the threshold to justify per-period fixed costs), given the competition from more and better products. This nonmonotonic time path results in the bell shapes of product life cycles. Details of the stylized model are in Appendix A.1. Comparative statics in Figure A.1.2 further show that in the same market, a higher-quality product might have lower immediate sales due to high initial price, but higher lifetime sales compared to a lower-quality product; the same product has lower immediate sales, but much lower lifetime sales in a more competitive market.

I now present descriptive evidence for these properties of the product life cycle. I first quantify the dynamic realizations of product life cycles, or the area under the curve in Figure 1.2.7, by constructing a PLC multiplier, which is defined as the realized lifetime unit sales of a product normalized by its first-month sales[42]. This multiplier then captures the relationship between the immediate sales and the lifetime sales of a new product, which a forward-looking product manager would need to account for in her future outlook before adjusting her product portfolio. In theory, the PLC multiplier is a complicated object, such as the realization of a Markov perfect equilibrium of the dynamic product portfolio game. In the rest of this section, free from any model, I show that, as suggested by the stylized model, the PLC multiplier can be predicted based on static observable characteristics of the products and markets.

Figure 1.2.8a shows a binned scatter-plot of the log of the PLC multiplier against the relative quality of the product. To remain model-free, the quality here is measured by the percentile rank of a product's release price among all the products in the market at the time of its release[43]. This is thus a static measure of initial relative quality, and does not change over time. Figure 1.2.8a shows that, relative to a product's immediate sales, the product's lifetime sales increase in its initial relative quality in the market: Roughly, controlling for

[42]Only products for which I observe both the entry and exit months are included in this analysis for the entire realized paths of sales.

[43]The relationship between the PLC multiplier and the quality of the product remains the same with alternative measures of relative quality, such as using the first component of a principal component analysis of all handset characteristics. In my empirical specification, I construct a quality index using products' mean utilities from demand estimates.

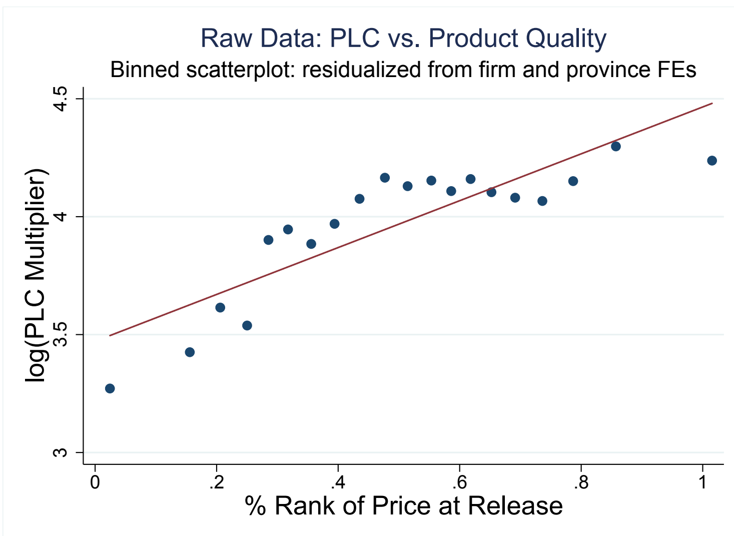
the same level of immediate sales, a product released at the top of the quality spectrum has about twice as large a product life cycle (lifetime sales) than a product released at the bottom of the quality spectrum. From a firm's perspective, in introducing a new product, for the same amount of immediate sales, a higher-quality product can justify much higher sunk set-up costs given its more durable expected product life cycle.

Figure 1.2.8b shows the relationship between the realized product life cycle and static market competitiveness at the release time of the product. I measure market competition using the number of products in the market when the product is first introduced. Variations in market competitiveness come from both provincial differences and fringe entries over time, as discussed in Section 1.2. Figure 1.2.8b shows that not only can we infer the immediate sales of a new product, considering the usual static tradeoffs faced by firms, given the competitiveness of the market, but we can also, to some extent, predict the lifetime sales of the product based on the expected immediate sales and the contemporaneous level of market competition. Moreover, this relationship has implications for counterfactuals: When markets become more (less) competitive, forward-looking firms not only expect less (more) immediate sales[44], but also revise down (up) their future outlook on their products' lifetime sales as their expected product life cycles become shorter (longer).

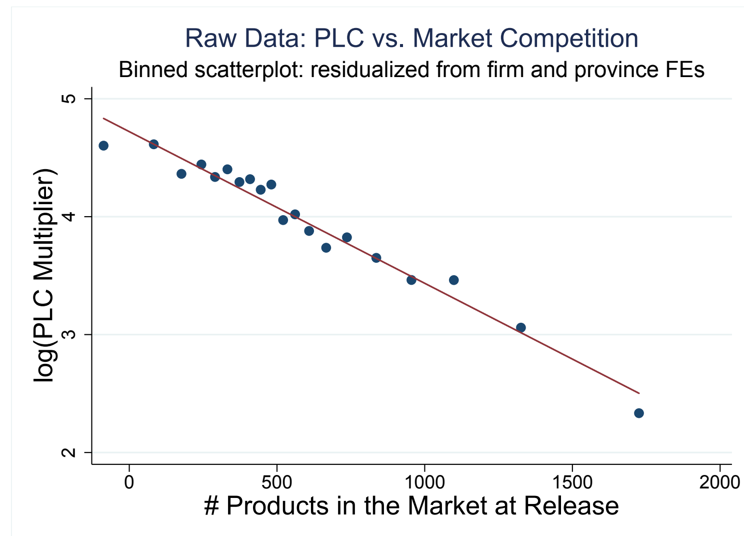
[44]The direction of change for static profitability remains theoretically ambiguous, depending on where new products are introduced into the product space and demand substitution patterns.

Figure 1.2.8: Lifetime sales vs. product and market characteristics at release

(a) Product Quality



(b) Market Competition



Notes: This figure shows that a product's eventual realization of lifetime sales systematically correlates with the characteristics of the product and the market at its release time, even after normalizing by its immediate sales. Plotted from model-free data. PLC multiplier is the ratio of a product's realized lifetime unit sales and its first-month sales. The relative quality of a product is measured as the percentile rank of its release price in the market at the time of release (thus this rank does not change over time). Market competitiveness is measured by the number of smartphone models (major and fringe) in the market at the time of the product's release (also static and constant over time). Data are residualized from firm and province fixed effects. Relationships shown are both statistically significant at the 1% level.

1.3 Model

This section presents a two-stage model over many periods that captures firms' static and dynamic incentives in product introductions, as well as how they are affected by competition. Given features of this industry and computational limits, I make several modeling choices. I take the set of smartphone manufacturers as given, and do not model firms' entry and exit decisions. I focus on the 12 major manufacturers' strategic incentives in product introductions, and take the hundreds of fringe firms' products as given[45]. More importantly, I model firms' portfolio adjustment decisions on the province-month (m, t) level—the modular nature of smartphone production allows firms to quickly adjust their product portfolios based on local market conditions[46]. Specifically, before each period (month), outside my model, I assume that firms are endowed with a pool of potential product designs $\mathbb{J}_{f,t}$, from which they could choose to introduce as new products into their portfolios J_{fmt}. I thus only recover market-level product introduction costs, as opposed to total product development costs, in my estimation. Finally, I only model firms' strategic incentives to introduce non-flagship products. The reason for this abstraction is that firms' choice of flagship products often relates to reasons beyond the costs and benefits analyzed in this paper, such as technological constraints and brand image[47]. Therefore, I treat flagship products to be exogenously (to my model) placed in firms' portfolios and do not attempt to estimate costs for these products[48].

[45]In other words, equilibria of the model allow for major manufacturers' strategic reactions in product portfolios to fringe entry, but assume that fringe firms exogenously introduce products. This is reasonable, given major manufacturers' dominant market share, and also because fringe firms mostly produce mid- to low-end handsets often due to technology constraints, rather than strategic incentives in product positioning. While I do not model fringe firms' strategic product offerings, my model allows consumers to choose fringe products, and major manufacturers to respond to fringe entry in their portfolio choices and pricing.

[46]When asked about product development, industry sources suggest that they have a pool of new designs every year, developed at the national level or higher, but product managers choose what to release based on local market conditions at their regional offices on a monthly, or even more frequent, basis. Moreover, modeling and estimating the actual development costs of these designs are computationally infeasible: The portfolio game is at least of the order of $2^{31\times 12}$ if firms optimize over which set of markets to enter and when, ex ante.

[47]Several of the major manufacturers in this market design and export smartphones globally. As discussed, I specify the entry game at the province level. As a result, I do not need to additionally assume the exclusivity of designs for the Chinese market. I do, however, assume that the timing of release is related to local market conditions only. This is true for non-flagship products, even when they are also sold outside China, but less so for flagship products, where many of the release dates are globally coordinated.

[48]I do not, for example, explain any of the portfolio choices of Apple or Xiaomi, given that they only carry flagship products.

As a result, I specify a full-information, discrete game of product portfolio competition, played between major smartphone manufacturers as follows:

1. Stage I: Product portfolio choice

   (a) Introduction $\left(SC_{jmt}\right)$ and maintenance $\left(F_{jmt}\right)$ cost shocks $\left(\mu_{jmt}, \eta_{jmt}\right)$ of each potential product $j \in \mathbb{J}_{f,t}$ are realized.

   (b) Firms form beliefs about the lifetime profitability of each potential product, $\mathbb{E}\Pi_{jm}$.

   (c) Firms simultaneously adjust product portfolios, by introducing new products (out of potential products $\mathbb{J}_{f,t}$), and/or discontinuing old ones (from existing products $J_{fm,t-1}$), based on beliefs about new products' lifetime profitability $\mathbb{E}\Pi_{jm}$, $j \in J_{fmt}, j \notin J_{fm,t-1}$, and expectations of the existing products' current profitability $\mathbb{E}\pi_{jmt}(J_{mt})$, $j \in J_{fm,t-1}, j \in J_{fmt}$, over Stage II shocks $\left(\xi_{jmt}, \omega_{jmt}\right)$, to maximize expected portfolio profits Π_{fmt}.

2. Stage II: Pricing

   (a) Marginal cost $\left(mc_{jmt}\right)$ and demand shocks $\left(\xi_{jmt}, \omega_{jmt}\right)$ of every active product $j \in J_{mt}$ are realized.

   (b) Given all of the product portfolios J_{mt}, firms simultaneously set prices p_{jmt}, given demand, to maximize profits.

   (c) Consumers choose, given characteristics and prices $\{X(J_{mt}), p(J_{mt})\}$ of the available products in the market.

Firms solve the game backwards by first computing Stage II payoffs for all possible configurations of product portfolios, then choosing their portfolios in Stage I to maximize expected profits. I also present the model in the same order in the rest of this section.

1.3.1 Demand

Demand for smartphone handsets in China is simpler than the US, given its separable nature from carrier services, as discussed in Section 1.2—I therefore describe the demand system for smartphone handsets with a Logit discrete choice model. Demand attributes described in Section 1.2 suggest several features of demand in this market to be accounted for in the model. The first is income effects. Given the price of handsets as a relatively substantial portion of consumers' annual consumption—especially for the earlier sample, lower-income provinces, and rural population—demand likely exhibits income effects that are often assumed away for smaller-item purchases. The second is the heterogeneity in price sensitivity across consumers with different income levels. I account for both by specifying a log function of utility for money in the demand model. The third is the likely heterogeneity in the quality of consumers' current handsets over time. While the dynamics are important for estimating demand for durable goods[49], I abstract away from this aspect, given data limitations. Instead, I include a flexible specification of the outside good in my demand estimation to capture this heterogeneity.

I now specify the demand system. A market is defined as a province-month, and consumers choose among the J alternatives of mobile phone handsets in a market—or the outside good of not purchasing any handset this month—to maximize utility. In the case of not purchasing, the consumer enjoys utility from consuming her current handset holding (or lack thereof). I then specify the utility of consumer i choosing product j in province m and month t as follows:

$$\begin{aligned} u_{ijmt} = {} & Major \cdot (\beta x_j + D^{NC}_{jmt} + \lambda_{f(j)} + \lambda_{l(j)}) + Fringe \cdot (\kappa_1 + \kappa_1^t t) + Feature \cdot (\kappa_2 + \kappa_2^t t) \\ & + \alpha log(y_i - p_{jmt}) - (\beta_0^t t + \lambda_m + \lambda_{q(t)} + \alpha log(y_i)) + \xi_{jmt} + \epsilon_{ijmt}. \end{aligned} \tag{1.3.1}$$

Available handsets on the market are categorized into major smartphones, fringe smartphones, and feature phones. I collapse all fringe and feature phones so that there is one of each in any market, with share-weighted average price within each type. Major smartphone characteristics x_j include CPU clock speed, camera resolution, display size, and the "Other" characteristic described in Section 1.2. The network compatibility term D^{NC}_{jmt} includes fixed effects for the generations of networks, as well as compatibility with each

[49] See Gowrisankaran and Rysman (2012). I thus also abstract away from firms' dynamic pricing behaviors induced by forward-looking demand.

carrier[50], which captures consumers' valuations of carrier service quality and the potential cost of switching between carriers[51]. Firm and product line fixed effects are included to capture brand preferences and marketing effects. For fringe and feature phones, I allow for separate intercepts, as well as time trends to capture the growth in their quality over time. The term $\alpha log(y_i - p_{jmt})$ is consumers' disutility for price, where α measures price sensitivity and y_i the income of the individual. I use the log functional form, which is supported by a Cobb-Douglas utility function in consumption similar to Berry, Levinsohn and Pakes (1995) and Petrin (2002), which reflects consumers' heterogeneous sensitivity to price due to income effects[52]. The outside option is chosen to be not buying a mobile phone this month after including both fringe and feature phones in the demand model. A time trend is included in the outside option to reflect the fact that consumers' current handset holdings are, on average, of higher quality over time, similar to Eizenberg (2014). A province fixed effect is included to allow for a different intercept for the different quality of current holdings in each province. $\lambda_{q(t)}$ is a month dummy (Jan. to Dec.) that absorbs the seasonality in sales. The base utility from income does not drop out of the equation, given the nonlinearity of the functional form. ξ is the unobserved (to the econometrician) characteristics that are observable to both firms and consumers. Finally, ϵ is a preference shock assumed to be *i.i.d.* type I extreme value distributed.

I can then express the market share of product j in market mt as,

$$s_{jmt} = \int \frac{exp[\delta_{jmt} + \alpha(log(y_i - p_{jmt}) - log(y_i))]}{1 + \sum_{l \in J_{mt}} exp[\delta_{jmt} + \alpha(log(y_i - p_{lmt}) - log(y_i))]} dP_y(y_i), \tag{1.3.2}$$

where δ is the linear part of the utility function and $P_y(y_i)$ is the empirical distribution of income in market mt[53].

[50]This term varies over time as I collapse different versions of the product that are released at different times.

[51]Switching costs mostly capture the fact that phone numbers are not portable if consumers decide to switch carriers in China. This policy was only temporarily relaxed in two provinces during this time.

[52]The log functional form also more naturally limits the size of the markets—consumers with annual consumption levels lower than the price of a handset will not purchase that product in the model.

[53]See Appendix A.3.1 for income distributions and evaluation of the integral in equation (1.3.2).

1.3.2 Pricing

In Stage II of the game, firms observe everyone's product portfolio decided in Stage I. They have also observed the shocks of demand and marginal costs, and therefore know their demand and cost. Firms then simultaneously set prices for all of the products in their portfolios, J_{fmt}, to maximize profits, for firm f in market mt,

$$\pi_{fmt} = \sum_{j \in J_{fmt}} (p_{jmt} - mc_{jmt}) \cdot s_{jmt}(p) \cdot M_{mt}, \tag{1.3.3}$$

where M_{mt} is the market size and s_{jmt} is given by the demand system and all of the prices in the market. I then first invert out the marginal costs without assuming any structures with the first-order conditions,

$$mc_{jmt} = p_{jmt} + s_{jmt} \int (T \cdot \Delta_i)^{-1} dP_y(y_i), \tag{1.3.4}$$

where $\Delta_i = \frac{\partial s_i}{\partial p}$ is the matrix of partial derivatives for consumer i and T is the product ownership matrix (i.e., $T_{lj} = 1$ if products l and j belong to the same firm and zero otherwise).

I then project the implied marginal costs at equilibrium onto characteristics of the handsets with flexible functional forms[54]. In particular, I assume that marginal costs are convex in the major characteristics of the handsets, but I allow for the speeds at which the production costs of components of different quality fall over time to differ nonparametrically,

$$mc_{jmt} = \sum_{k=1}^{4} c^k(t) exp(x_j^k) + \lambda_{f(j)} + \lambda_m + \lambda_t + \Gamma_{jmt} + \omega_{jmt}, \tag{1.3.5}$$

where, again, major characteristics include CPU clock speed, camera resolution, display size, and the "Other" characteristic. Industry reports[55] show that chipsets, displays, and cameras constitute most of the production costs (bill of materials, or BOM) across handsets. I therefore allow the per-quality cost function to be very flexible for the three major characteristics, but remain constant for the "Other" characteristic. Additionally, marginal costs are allowed to have different intercepts for different firms, provinces (distribution

[54] I do not have sufficient bill of materials tear-down data to construct component costs.
[55] Sample cost breakdowns from Teardown.com.

costs from major manufacturing centers on the east coast), and time. Finally, I include the cost shifters (government subsidies and loan interest rates) and network compatibility (both in the Γ_{jmt} term), as well as an *i.i.d.* cost shock ω_{jmt}.

1.3.3 Product portfolio choice

In Stage I of the game, firms adjust their product portfolios. In particular, managers need to decide which new products to introduce, and which existing products to take down, to maximize expected portfolio profits. Forward-looking firms in fast-changing high-tech markets anticipate not only Stage II static profits $\pi_{fmt}(J_{mt})$, but also lifetime profits of their new products $\mathbb{E}\Pi_{jm}(J_{mt})$ in the market:

$$\Pi_{fmt} = \max_{J_{fmt} \subseteq \mathbb{J}_{f,t}} \sum_{j \in J_{fmt}, j \in J_{fm,t-1}} \left[\mathbb{E}\pi_{jmt}(J_{mt}) - F_{jmt} \right] + \sum_{j \in J_{fmt}, j \notin J_{fm,t-1}} \left[\mathbb{E}\Pi_{jm}(J_{mt}) - SC_{jmt} \right], \tag{1.3.6}$$

where SC_{jmt} is the sunk introduction cost of product j to market m in month t, reflecting one-time marketing costs, product launch events, renegotiation with local retailers, etc.; F_{jmt} is a monthly fixed cost to maintain a product in the firm's portfolio, reflecting any channel fixed costs and per-period marketing costs (billboards, TV airtime rates, etc.). Therefore, the first term is the expected total static profits of maintaining existing products, and the second term is the expected lifetime profits of introducing new products.

This specification of firms' objective function, rather than a Bellman equation, makes several important simplifying assumptions. First, I assume zero scrap values—if a product is taken down, firms make zero profits[56]. I then interpret firms' product discontinuation decisions as static[57], which allows me to identify maintenance costs[58]. Firms therefore evaluate only the product's static profits—as well as its static impact on other products in the portfolio if it is maintained on the shelf—against the maintenance cost. Second, I interpret the sunk introduction costs SC_{jmt}'s to also include opportunity costs of waiting

[56]This is reasonable, given the lack of capacity (see Section 1.2) and liquidity (major manufacturers are mostly large multi-sector firms) constraints in this market.

[57]The static assumption here is also based on managers' practice in this industry: Once a product is introduced, they do not actively think about its life cycle, but only monitor sales so that its presence in the market is always justified. This is reasonable, in the sense that the impact of any single product's introduction on other products' life cycle paths is likely second order compared to its static impacts.

[58]This is also reasonable, given the observed thin tails of product life cycles toward their end.

to launch the product in the future. The opportunity costs are likely small in this market, given the fast-changing technology—an available design today will quickly become obsolete in a few months. As a result, the dynamic game of product introductions collapses to equation (1.3.6), where firms weigh expected lifetime profits of a new product against its sunk introduction cost.

However, as alluded to earlier, the expectation of future sales of a product, $\mathbb{E}\Pi_{jm}(J_{mt})$, taken over the future evolution of technology mc_{jmt}'s, and product portfolio J_{mt}'s, requires managers to keep track of trillions of states, and is thus intractable. While I have shown in Section 1.2 that the static observables have predictive power for future sales of a product, and thus can be used by managers in new product introduction, the question of what managers actually do in the industry remains.

Is the concept of the product life cycle and the prediction of its magnitude used by managers in this industry? This was Theodore Levitt's concern (Levitt, 1965)[59]. Since then the concept of the product life cycle has been written into an abundance of business review articles[60] and introductory marketing textbooks[61], and has become one of the most familiar concepts among executives around the world. Interviews with product managers of smartphone manufacturers and industry analysts in China suggest the prevalent use of the product life cycle to forecast product sales after introduction. For example, the Head Product Manager of Samsung Mobile in China said:

> Every month we determine PLCs and EOPs [end-of-products] to adjust our product lines. Everyone tries their best to make predictions of PLCs given the competition. We used to be able to sell our mid-level handsets for 18 months, but can barely maintain 12 months now with the amount of competition.

Combined with the descriptive evidence shown in Section 1.2, I model firms' product introduction decisions by weighing a product's sunk cost of introduction against the firm's rational expectation of the lifetime profitability of the product based on static observ-

[59] Levitt (1965) famously said, "The concept of the product life cycle is today at about the stage that the Copernican view of the universe was 300 years ago: a lot of people knew about it, but hardly anybody seemed to use it in any effective or productive way."

[60] Chambers, Mullick and Smith (1971) discuss various forecasting methods for the product life cycle; Sampere (2014) mentions Xiaomi's product strategy around the length of its product life cycle.

[61] E.g., Buzzell (1972) and Kotler and Armstrong (2010).

able product and market characteristics. Specifically, I let firms approximate the expected lifetime profits of a new product, by relating its lifetime profits to its short-run profits at release,

$$\mathbb{E}\Pi_{jm}(J_{mt_0^{jm}}) = \mathbb{E}_{(\xi_{jmt_0^{jm}}, \omega_{jmt_0^{jm}})} \pi_{jmt_0^{jm}}(J_{mt_0^{jm}}) \cdot \widehat{PLC}_{jm}, \tag{1.3.7}$$

where firms first form beliefs about the magnitude of the product life cycle, based on characteristics of the product and the market at launch-time, for product j, released at t_0^{jm} in province m,

$$\widehat{PLC}_{jm} = \theta^{PLC} X_{jmt_0^{jm}}, \tag{1.3.8}$$

where the parameters θ^{PLC} is what firms use to make their best linear predictions, and remain to be estimated.

The dynamic product portfolio game specified in equation (1.3.6) can be then reformulated as follows: Firms simultaneously choose a set of non-flagship products (given their flagship products exogenously) to maximize their own expected profits, given the other firms' portfolios, or the market product configuration J,

$$\begin{aligned}\Pi_{fmt} = \max_{J_{fmt} \subseteq \mathbb{J}_{f,t}} & \sum_{j \in J_{fmt}, j \in J_{fm,t-1}} [\mathbb{E}_{(\xi_{jmt}, \omega_{jmt})} \pi_{jmt}(J_{mt}) - F_{jmt}] \\ & + \sum_{l \in J_{fmt}, l \notin J_{fm,t-1}} [\mathbb{E}_{(\xi_{lmt}, \omega_{lmt})} \pi_{lmt}(J_{mt}) \cdot \widehat{PLC}_{lm} - SC_{lmt}],\end{aligned} \tag{1.3.9}$$

where the expected static profits are integrated over Stage II shocks, and lifetime profits are approximated with firms' rational beliefs about product life cycles—both of which are determined by the market product configuration J, as a result of firms' portfolio competition[62].

Necessary equilibrium conditions of this game then require every firm f to consider all possible subsets of the potential-product pool $\mathbb{J}_{f,t}$ (or the power set) in each market-

[62]This game specified in equation (1.3.9) also assumes no economies of scope for either introducing a new product, or maintaining an existing one. This is fairly standard in the literature. Empirically, product entries plausibly do not exhibit economies of scope, given the setup of large manufacturers' regional offices and low transportation costs of smartphones. I also do not model the actual development of products, but only their introductions to the market, after they are developed. Product maintenance could potentially exhibit economies of scope. I argue that, with small product configuration changes in the counterfactual, this effect is likely small.

month (mt), and have no incentive to deviate from the chosen product portfolio J_{fmt}. These conditions are fairly weak and typically yield many equilibria in positioning games such as equation (1.3.9). In the estimation to follow, I only build off these necessary conditions to make inference on sunk and maintenance cost parameters. In the counterfactual analysis, I rely on firms' best-response dynamics to select equilibrium.

Finally, sunk costs $SC_{jmt} = SC(\theta^{SC}, X_{fm}|\mu_{jmt})$ also vary with the smartphone manufacturer, and which market the product is introduced into, with an $i.i.d.$ shock μ_{jmt} observed by firms at the beginning of Stage I. Maintenance costs $F_{jmt} = F(\theta^{F}, X_{jm}|\eta_{jmt})$ are shifted by observable characteristics of the product, the type of retail channels, and the market, with an $i.i.d.$ shock η_{jmt} also observed by firms at the beginning of Stage I.

1.4 Estimation and results

The estimation proceeds in five steps. Similar to solving the game, I work backward in estimating the model.

1.4.1 Demand

I estimate demand similarly to Berry, Levinsohn and Pakes (1995), using the Generalized Method of Moments. The main source of endogeneity concerned here is that major smartphone manufacturers choose prices after observing demand shocks ξ's. I construct three sets of moments. Following the literature[63], I first make a timing assumption that firms make product choices prior to observing demand shocks, and therefore $E[\xi_{jmt}|x_j] = 0$. Given the large number of firms and products, I also assume that fringe and feature phones are priced competitively, and thus $E[\xi_{jmt}|p_{jmt}] = 0$ for product j's that are not major smartphones. The timing assumption also validates the second set of moments, which uses characteristics of other firms' products to shift markups (i.e. $E[\xi_{jmt}|x_{-f}] = 0$)—I use the differentiation IVs in Gandhi and Houde (2015), which measures competition from prod-

[63]Notably, Fan and Yang (2016), Wollmann (2016), Berry, Levinsohn and Pakes (1995), and many other studies based on BLP.

ucts with similar characteristics[64]. The last set of moments is constructed with instruments that shift marginal costs of production. Specifically, as mentioned in Section 1.2, I use the level of government subsidies and the interest rates firms pay on their loans each year as cost shifters. As shown in Aghion et al. (2015), government subsidies and low-interest loans are two important types of industrial policies in China that are shown to be correlated with state ownership of firms. Therefore, I argue that while these are likely correlated with firms' production costs, they are plausibly uncorrelated with demand shocks.

Table 1.4 presents the demand estimates. I first compare estimates from linear specifications of OLS and IV with no income heterogeneity for a descriptive look at the data. I simply use the average income in a market for all consumers in that market, reducing equation (1.3.1) to a linear Logit model, which is then estimated following Berry (1994). This is shown in the first two columns. The sign of the price sensitivity coefficient α is negated due to the functional form. Notably, the price sensitivity coefficient is much larger when price endogeneity is accounted for; at the same time, consumers' tastes for handset characteristics are also much stronger. When I move to the third column, where I incorporate income heterogeneity and estimate the model by minimizing the GMM objective, the price sensitivity coefficient is much larger, which suggests the large degree of heterogeneity in consumers' sensitivity to prices. The other estimates are relatively stable from the linear IV estimates, with stronger preferences for the major characteristics. Other estimates are sensible: Consumers prefer 3G and 4G phones; China mobile compatibility is valued the most due to its largest subscriber base; Apple commands a much larger brand value, followed by Oppo, Xiaomi, and Samsung, which are the other higher-end brands; both fringe and feature phone qualities are increasing over time; and finally, the quality of the outside good is also quickly increasing over time, reflecting the fact that consumers are holding better handsets over time.

[64] Demand estimates using the standard "BLP" instruments—sum of characteristics of all competing products in the market—are similar, with slightly weaker estimates of price sensitivity.

Table 1.4: Demand estimates

	OLS	IV	IV
Income heterogeneity	No	No	Yes
$log(y_i - p_{jmt})$	1.3457***	9.2285***	25.921***
Major smartphones			
CPU clock speed (GHz)	0.0004***	0.3496***	0.4581***
Display size (inch)	0.7255***	1.3004***	1.3978***
Camera resolution (MP)	-0.0981***	0.0225***	0.1031***
"Other" characteristic	0.1441***	0.1824***	0.2464***
2G phone	(omitted)	(omitted)	(omitted)
3G phone	1.7602***	2.1035***	2.2079***
4G phone	1.9071***	2.2145***	2.3663***
Compatible with CMCC	1.1973***	1.1869***	1.1262***
Compatible with CU	0.8886***	0.6621***	0.5981***
Compatible with CT	0.8225***	0.7961***	0.8254***
Apple	(omitted)	(omitted)	(omitted)
Samsung	-2.2775***	-5.2575***	-5.6323***
Coolpad	-0.9288	-7.5883***	-8.3547***
HTC	-5.2996***	-6.4191***	-6.6120***
Huawei	-1.5449***	-6.7430***	-7.4443***
Lenovo	-0.3630	-7.7284***	-8.3827***
Motorola	-3.3387***	-7.0755***	-7.5790***
Nokia	-2.4202***	-5.6584***	-6.1392***
Oppo	-1.8160***	-4.7697***	-4.9703***
ZTE	-1.0991	-8.0059***	-8.9878***
Xiaomi	-1.1026***	-4.1727***	-5.4851***
Vivo	-2.4004***	-5.4351***	-5.9169***
Feature/fringe phones			
Feature phone	7.2663***	2.8261***	1.7808***
Feature phone trend	0.0334***	0.0958***	0.1011***
Fringe smartphone	-0.3097***	-2.4251***	-2.4218***
Fringe smartphone trend	0.1680***	0.1969***	0.1974***
Outside good			
Time trend	0.0611***	0.1375***	0.1642***
Constant	-8.6051***	-3.1578***	0.0033***
N	330,866	330,866	330,866
R^2/GMM objective	0.542	0.215	3,212

Notes: This table summarizes raw demand coefficient estimates. Columns 1 & 2—without income heterogeneity—are estimated with average province income which reduces equation (1.3.1) to a linear specification, which can be estimated using Berry (1994). Column 3 shows results from the full demand model estimated with GMM. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% level, respectively. Product line, month, and province fixed effects are not reported.

1.4.2 Marginal costs

With the demand estimates, marginal costs are first inverted out using equation (1.3.4) and then projected onto product characteristics to estimate the evolution of marginal costs over time using equation (1.3.5). I make a similar timing assumption that product choices happen prior to the realization of marginal cost shocks so that $E[\omega_{jmt}|x_j] = 0$. I estimate the per-quality cost function $c^k(t)$ for each of the three major characteristics over time non-parametrically with year fixed effects. Table 1.5 presents the results. Columns 2-4 report the coefficients for each major characteristic $exp(x_j^k)$ over time. The coefficients are falling, at a progressively slower pace over time. The pattern very well reflects Moore's Law, where production costs exponentially decay. Figure 1.4.1 shows this pattern graphically. Although I do not have data on component-specific costs, I plot the estimated component costs of CPUs, displays, and cameras using observed product characteristics by year. Cost schedules along the quality dimension flatten over time. Equivalently, component costs at higher-quality levels fall much faster. Back to Table 1.5, the other estimates are also sensible: Cost of the "Other" characteristic is smaller but comparable; cost shifters indeed move production costs in the expected directions.

While Table 1.5 and Figure 1.4.1 show sensible slope estimates, Table 1.6 presents evidence that the estimated levels of marginal costs are in line with industry estimates of the bill of materials. As I do not have detailed data on marginal costs of the handsets, I follow the industry standards of "entry-level" 3G smartphones to select two low-end 3G smartphones and compare their estimated marginal costs (predicted marginal costs $E_{\omega_{jmt}} mc_{jmt}$) with the industry estimates of comparable handsets. The industry BOM estimates also fall quickly over time, and my predicted marginal costs for the Nokia X5 and the Samsung I5508 fall reasonably close to the range of estimates.

1.4.3 Maintenance costs

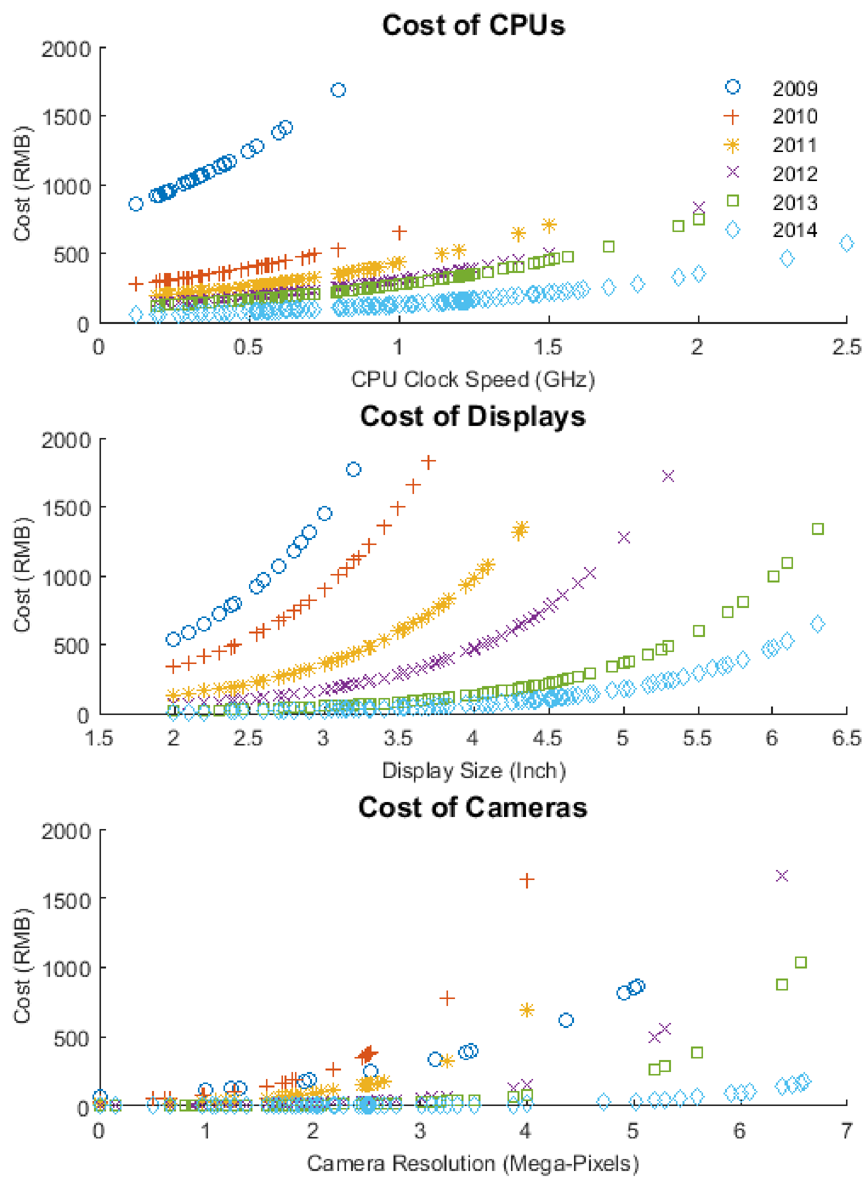
I now turn to the estimation of maintenance costs. Solving the game in equation (1.3.9) is difficult: The full entry game is of the order 2^N, where N is the total number of actual and potential products in the market. In the case of Jan. 2013, Beijing, $N \geq 199$, which is

Table 1.5: Marginal cost coefficient estimates

	Estimate	SE
exp(CPU clock speed)		
2009	755.6***	14.43
2010	242.2***	8.980
2011	159.1***	5.119
2012	112.6***	2.617
2013	100.4***	1.855
2014	46.87***	1.495
exp(Display size)		
2009	72.03***	0.6775
2010	45.34***	0.3338
2011	17.90***	0.1698
2012	8.574***	0.0722
2013	2.460***	0.0268
2014	1.189***	0.0164
exp(Camera resolution)		
2009	69.77***	0.7661
2010	30.15***	0.4274
2011	12.66***	0.1512
2012	2.791***	0.0314
2013	1.463***	0.0261
2014	0.2381***	0.0108
exp("Other" characteristic)	12.59***	0.2280
Subsidies	-27.29***	0.7000
Interest rates	5.050***	0.9064
Constant	1272***	24.46
R^2	0.751	
N	261,051	

Notes: This table shows that component cost (almost) exponentially decays over time without any functional form assumptions. The convex cost schedule for each major characteristic flattens out over time. Cost shifters (government subsidies and loan interest rates) used for demand estimation are included and indeed shift production costs in expected directions. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% level, respectively. Firm, time, and province fixed effects, and network compatibility are not reported. The number of observations drops after I eliminate flagship products for cost estimation. CPU clock speed is measured in GHz, display size in inches (diagonal), and camera resolution in half mega-pixels.

Figure 1.4.1: Estimated component cost reductions



Notes: This figure plots predicted component costs at estimated parameters, against the observed component quality of smartphone handsets in my sample. While the convex-shaped cost schedule across quality is assumed from the exponential functional form, rates at which these cost schedules fall over time are estimated nonparametrically with year dummies. This figure shows that marginal cost schedules flatten over time, i.e., a new high-quality component comes out at a higher price but the price also falls much faster than that of a lower-quality component (Moore's Law). In Appendix A.1 I show that this feature of high-tech markets naturally generates bell-shaped product life cycles and their properties shown in Figure 1.2.8.

the observed number of non-flagship products in the market. Therefore, I instead use the necessary conditions (for any observed product configuration to be a Nash equilibrium) to first construct bounds on maintenance costs[65]. The argument is as follows. For the current configuration J in market mt to be a Nash equilibrium, it has to be the case that any unilateral change by any firm f to remove a product j cannot be profitable for that firm, i.e.,

$$\mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J) - F_{jmt} \geq \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J \setminus j), \tag{1.4.1}$$

which provides an upper bound on the maintenance cost F_{jmt}. Similarly, any unilateral change by any firm f to add a product j also cannot be profitable for that firm, i.e.,

$$\mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J) \geq \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J \cup j) - F_{jmt}, \tag{1.4.2}$$

which provides a lower bound on the maintenance cost F_{jmt}. Combining equation (1.4.1) and equation (1.4.2) gives us an interval that F_{jmt} must fall between.

To empirically compute the bounds, I rely on monthly EOPs (product discontinuations[66]), as well as variations in market structures and demand characteristics across markets for identification. I first restrict the sample to the 8,319 product-markets mentioned in Section 1.2 in which I observe the discontinuation months (products discontinued at least one month before the sample ends in Nov. 2014). To construct the upper bounds, I remove one product at a time from the last month before it is discontinued in a market, and compute equation (1.4.1), i.e., $\mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J) - \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J \setminus j)$, where I compute Stage I expected profits by integrating out the estimated empirical distribution of $\left(\xi_{jmt}, \omega_{jmt}\right)$ from demand and marginal cost estimation[67]. To construct the lower bounds, I add one product at a time to the first month it is discontinued in a market, and similarly compute

[65]The conditions used to construct the bounds are similar to Fan and Yang (2016). I differ in the empirical implementation and interpretation, as I separately estimate product maintenance and introduction costs.

[66]Given the fine level of my market definition (province-month) and the retail transaction nature of the data, in some cases, it would look like the products were temporarily taken down and then started selling again. I let the demand model rationalize such patterns instead of modeling it as a discontinuation decision by managers.

[67]The procedure is similar to Fan and Yang (2016): I draw $\left(\hat{\xi}_{jmt}, \hat{\omega}_{jmt}\right)$ from their respective empirical distributions within market-month, compute equilibrium prices and variable profits, and then take the average profits for each firm.

equation (1.4.2), i.e., $\mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J \cup j) - \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J)$. The resulting bounds are shown in Figure 1.4.2. The bounds are fairly tight, given the monthly data.

For the estimation of firm beliefs about product life cycles and sunk introduction costs to follow, the rest of this section shows how I obtain point estimates of maintenance costs[68]. I specify per-period maintenance costs to be shifted by product, retail channel, and market characteristics,

$$log(F_{jmt}) = \theta^F X_{jm} + \eta_{jmt}, \tag{1.4.3}$$

where the characteristics X_{jm} include a quality index of the product constructed from the demand estimates[69], $q_j = \beta x_j$; average income of the market; overall shares of the sales of product j in province m through different channels (carriers, general electronics stores, telecommunication stores, and others[70]), to reflect the different inventory and shelving costs at different retail channels; and firm fixed effects. In particular, I assume that the maintenance cost for a product-market is, on average, constant over time[71].

With the computed bounds and the specification in equation (1.4.3), I obtain point estimates of maintenance costs by minimizing the following simulated inequality objective function,

$$Q(\theta) = \frac{1}{s}\sum_{s,j,m}[max\{F(\theta, X_{jm}|\eta^s_{jmt}) - UB_{jm}, 0\}^2 + max\{LB_{jm} - F(\theta, X_{jm}|\eta^s_{jmt}), 0\}^2], \tag{1.4.4}$$

the rationale of which is to penalize the parameters for going beyond the bounds. Com-

[68]Given the tight bounds of estimated maintenance costs shown in Figure figure 1.4.2, I could follow Berry, Eizenberg and Waldfogel (2016) and Fan and Yang (2016), and directly use these bounds for welfare analyses. Different from those studies, the dynamic nature of my model requires feeding estimated maintenance costs into the estimation of sunk costs—which are also estimated with bounds—resulting in "bounds on bounds." This is currently in progress.

[69]See Fan and Yang (2016) for a similar quality index measure.

[70]I observe six types of retail channels in the GfK data: general electronics stores (e.g., GOME and Suning—large chain electronics and appliances stores similar to BestBuy in the US); carrier retail channels; telecommunication stores, which are small storefronts that only sell cell phone-related products (both chain and independent); online retail; and others (general-purpose markets similar to Walmart in the US). After dropping online sales, I combine chain and independent telecommunication stores in this analysis.

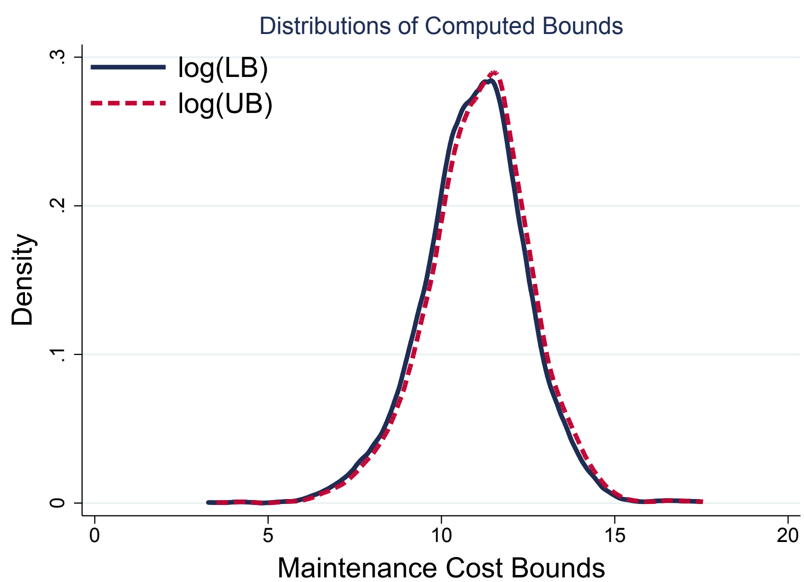
[71]This is consistent with smartphone product managers' and industry analysts' expectations as well: Maintenance cost is typically small and fixed over time. Moreover, after controlling for product quality, firm identity, market wealthiness, and types of channels, I argue that it is reasonable to assume that any idiosyncratic shock each month is unlikely to be correlated with observables, especially within a two-year window—the typical life span of a smartphone handset in this market. It is worth noting that this assumption allows me to separately identify sunk introduction costs from the per-period maintenance costs.

Table 1.6: Estimated productions costs vs. industry BOM estimates

	2010	2011	2012	2013	2014
Nokia X5	1,209	688	332	210	
Samsung I5508		746	426	287	209
Industry BOM est. for "entry-level" 3G smartphones					
		553-618	378-410	254-285	

Notes: This table shows that, for selected "entry-level" 3G smartphones (as defined by industry standards), my estimates are in line with industry BOM (bill of materials, or total production costs) estimates. Marginal cost predictions are mean estimates without cost shocks. Industry estimates from Nomura Global Markets Research; see footnote 148 for link. Currency in RMB; industry estimates converted from USD using each year's respective exchange rates. Industry numbers are based on midyear estimates. Reported marginal cost predictions are as of July of the corresponding year.

Figure 1.4.2: Maintenance cost bounds



Notes: This figure plots distributions of upper (red dashed) and lower (navy solid) bounds (in logs), computed from equation (1.4.1) and equation (1.4.2). This figure shows that computed bounds used to estimate maintenance costs are very tight given thin tails of product life cycles and monthly data.

putational details are in Appendix A.3.2.

Table 1.7 shows the point estimates: Maintenance costs are higher for higher-quality products, in wealthier markets, at carrier channels, and for generally higher-end brands. In terms of levels of maintenance costs, Figure 1.4.3 shows the distribution of estimated maintenance costs in the sample market of Jan. 2013, Beijing. In this market, maintenance costs are estimated to average about RMB 106K and range between 20K and 262K. Alternatively, these are equivalent to USD of 17K, 3.3K, and 42K, respectively. Among the products in this market, the higher-end Samsung Galaxy Premier would cost RMB 232K to be maintained on the shelf each month, while a mid-level Samsung handset—the Galaxy Trend 2—would cost only about half that, or 128K, each month. For a much lower-end handset from the old HTC sub-brand Dopod[72], Model 566 would only cost 37K to remain on retailers' shelves each month.

1.4.4 Firm beliefs about product life cycles

The rational expectation assumption allows me to construct firms' beliefs about the product life cycle $\widehat{PLC}$ from the realized technologies (marginal costs and product qualities in $E\pi$) and market structures J_{mt}, and expected variable profits and maintenance costs each period internally consistent with the model,

$$PLC_{jm} = \frac{\sum_{t=t_0^{jm}}^{T^{jm}} [\mathbb{E}_{(\xi_{jmt},\omega_{jmt})} \pi_{jmt}(J_{mt}) - \hat{F}_{jm}]}{\mathbb{E}_{(\xi_{jmt_0^{jm}},\omega_{jmt_0^{jm}})} \pi_{jmt_0^{jm}}(J_{mt_0^{jm}})}, \tag{1.4.5}$$

where T^{jm} denotes the month product j is actually discontinued from province m, and $\hat{F}_{jm}$ is the expected per-period maintenance cost prior to observing the cost shock η's. The term PLC_{jm} then represents the relationship between product j's static payoff to its lifetime payoff in province m—the denominator is its expected flow profits prior to observing demand and marginal cost shocks, whereas the numerator is its expected lifetime profits.

To relate firms' expectations of dynamic payoffs to their static profits in the model, I estimate their expected magnitude of product life cycles, $\widehat{PLC}$, by projecting the realized

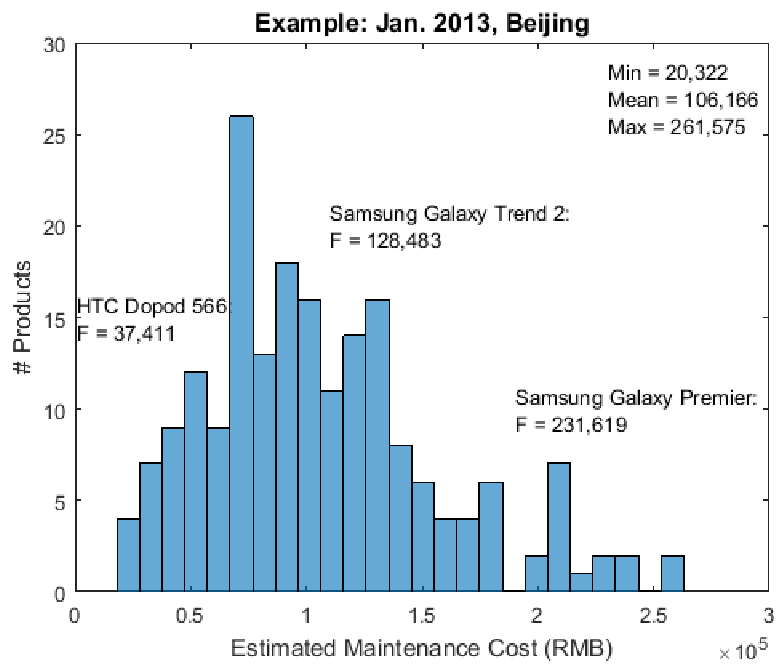
[72] Similar to Samsung's sub-brand Anycall in China, HTC's sub-brand Dopod was discontinued since 2010 when all of its models began to be simply branded under HTC.

Table 1.7: Maintenance cost coefficient point estimates

	Estimate	SE
Product, market, retailer chars.		
Product quality (index)	0.264***	0.006
Avg. province income	0.324***	0.014
Share of carrier sales	0.479***	0.110
Share of GES sales	-0.037	0.122
Share of TCS sales	-0.186*	0.113
Share of Other sales	(omitted)	
Manufacturer-specific costs		
Samsung	(omitted)	
Coolpad	-0.321***	0.030
HTC	-0.862***	0.034
Huawei	-0.170***	0.034
Lenovo	0.208***	0.033
Motorola	-0.451***	0.027
Nokia	0.370***	0.031
Oppo	0.305***	0.046
ZTE	-0.374***	0.034
Vivo	0.099**	0.044
Std. dev.(σ_η) of cost shock (η)	1.03(-4)	0.003
Constant	9.12***	0.112
N	8,319	

Notes: This table shows point estimates of maintenance cost coefficients. Product quality is constructed from demand estimates of four major characteristics of smartphone handsets (CPU clock speed, display size, camera resolution, and the "Other" characteristic). Avg. province income is included as a proxy for wealthiness of the market, as it likely shifts advertising rates and shelving cost at retailers. Sales through four types of retail channels are defined on product-market level, and do not vary over time—this reflects different costs of inventory and shelf space at different retailers. Manufacturer-specific costs reflect firms' differential costs in marketing, their relationships with local retailers, whether they have regional offices, and etc. Apple and Xiaomi are excluded from all cost estimations, given that they only carry flagship products. Standard deviation of the cost shock also implicitly incorporates unexpected shocks in firms' future profitability, to rationalize contradictions of maintenance cost bounds. (x) indicates $\times 10^x$. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Figure 1.4.3: Estimated maintenance costs in sample market: Jan. 2013, Beijing



Notes: This figure plots estimated maintenance costs in Jan. 2013, Beijing, with summary statistics in this market, and three sample products. Dopod is the old HTC subbrand in China before 2010. The Galaxy Premier is a higher-level product than the Galaxy Trend 2 by Samsung. Plotted are estimated mean maintenance costs without cost shocks. Currency in RMB. In USD: min = 3.3K; mean = 17K; max = 42K.

PLC's onto contemporaneous characteristics of the products and markets at the time of release, to estimate firms' belief parameters θ^{PLC}:

$$log(PLC_{jm}) = \beta^{PLC} x_{jmt_0^{jm}} + \lambda_{f(j)} + \lambda_m + \phi_{jm}, \tag{1.4.6}$$

where $x_{jmt_0^{jm}}$ includes major characteristics of the product and market at the time of release informed by the stylized model: the quality (index) of the product, the frontier quality (index) of the market, the number of products in the market, and the average income of the market (as proxy for price sensitivity). I also allow for different intercepts for each firm and province. The PLC shock ϕ is assumed to be *i.i.d.*[73].

Estimated demand and marginal and maintenance costs are used to compute equation (1.4.5), which is then used to estimate firms' PLC beliefs in equation (1.4.6). The sample is further restricted to the 7,405 product-markets whose introduction and discontinuation are both observed. Table 1.8 presents the results, which confirm the practice of the managers, evidence in Section 1.2, and intuition from the stylized model: PLCs are larger for higher-quality products and in less competitive markets; moreover, PLCs are also depressed by the quality frontiers of markets and strengthened by consumer price sensitivity[74]. These relationships are even clearer in Table 1.9 with a 2-by-2 example. Managers can expect about 4 times more lifetime profits of a product compared to its first-month profits in the less competitive Tibet market than in Beijing (although the level of lifetime profits of any product might very well be higher in Beijing). At the same time, Samsung's managers of the Galaxy Ace can expect about one-third more lifetime profits, compared to its profits in the release month, than Nokia's managers of the C5.

[73]This shock includes many future uncertainties faced by managers at the time of decision. In particular, this shock could be correlated with today's actions, in that managers' portfolio decisions would likely affect future market structures. I argue, however, that this is of second-order importance to the managers whose heuristic predictions likely do not cover this effect.

[74]Average consumer price sensitivity is negatively proxied by the wealthiness of the market. The intuition behind this relationship is that PLCs in the model in Appendix A.1 are driven by price changes (due to cost drops), which are magnified into profits in markets with higher price sensitivities.

Table 1.8: Firm belief coefficient estimates

	Estimate	SE
Chars. at release time		
Product quality (index)	0.609***	0.015
Frontier quality (index)	-0.344***	0.019
# Products in market	-0.001***	5.65(-5)
Avg. province income	-1.470***	0.192
Manufacturer-specific effects		
Coolpad	(omitted)	
HTC	0.651***	0.054
Lenovo	0.225***	0.049
Motorola	0.601***	0.051
Nokia	0.781***	0.052
ZTE	-0.070	0.047
Oppo	0.489***	0.078
Samsung	0.664***	0.047
Huawei	0.269***	0.046
Vivo	0.556***	0.079
Constant	3.423***	0.142
R^2	0.513	
N	7,405	

Notes: This table shows coefficient estimates of how firms form rational expectations over lifetime profits of their new products at release time. Dependent variable is the sum of the expected future stream of profits, computed with estimated demand, marginal cost, and from firms playing Stage II pricing game under realized market structures in the future. Inclusion of the four continuous variables are informed by the descriptive evidence in Figure 1.2.8 and the stylized model provided in Appendix A.1. The sample of 7,405 observations consists of all product-markets in which I observe both introduction and discontinuation of the product in the market. (x) indicates $\times 10^x$. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% level, respectively. Province fixed effects not reported. Summary statistics for the continuous variables: product quality (mean: 5.83, sd: 1.33); frontier quality (mean: 8.64, sd: 1.80); # products in market (mean: 591, sd: 472); and avg. province income (mean: 1.15, sd: 0.47).

1.4.5 Sunk introduction costs

The estimation of sunk introduction costs follows a revealed preference argument similar to that of the maintenance costs. I also integrate out estimated ϕ_{jm}'s, and use firms' expected product life cycles to relate their static profits to dynamic profits. Specifically, removing a product from the first month it is introduced in a market results in an upper bound of the introduction cost, i.e.,

$$\mathbb{E}_{(\xi_{jmt},\omega_{jmt})}\pi_{jmt}(J)\cdot\widehat{PLC}_{jm} - SC_{jmt} + \sum_{l\in J_{fmt},l\neq j}\mathbb{E}_{(\xi_{lmt},\omega_{lmt})}\pi_{lmt}(J) \geq \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J\setminus j), \quad (1.4.7)$$

and similarly, adding a product to the month before it is first introduced in a market results in a lower bound of the introduction cost, i.e.,

$$\mathbb{E}_{(\xi_{jmt},\omega_{jmt})}\pi_{jmt}(J\cup j)\cdot\widehat{PLC}_{jm} - SC_{jmt} + \sum_{l\in J_{fmt}}\mathbb{E}_{(\xi_{lmt},\omega_{lmt})}\pi_{lmt}(J\cup j) \leq \mathbb{E}_{(\xi,\omega)}\pi_{fmt}(J). \quad (1.4.8)$$

Given the nature of the introduction cost, I allow it to vary flexibly with the identity of the firm and which market the product is being launched into, with an *i.i.d.* shock μ each month,

$$log(SC_{jmt}) = \lambda_{f(j)} + \lambda_m + \mu_{jmt}. \quad (1.4.9)$$

The rest of the introduction-cost estimation follows closely the logic of the maintenance costs using objective function 1.4.4 to obtain point estimates. Identification of the sunk costs comes from several sources. First, while I estimate maintenance costs with bounds, I obtain point estimates with the inequality objective function to avoid bounds on bounds[75]. Second, although I estimate maintenance costs using end-of-products, I assume monthly maintenance costs are, on average, fixed throughout each product's lifespan. With these assumptions, the identification of sunk costs comes from variations in local market conditions and the timing of introductions across markets (Figure 1.2.4).

Results are presented in Table 1.10. Instead of the actual coefficient estimates, I report the implied introduction cost of an average product from each brand into a select sample

[75]Given the estimated tight bounds of maintenance costs, bound estimates of sunk costs based on maintenance-cost bounds might still be reasonable—this is currently in progress.

of representative provinces for interpretation. Overall, these introduction costs range between 2 and 15 million RMB—roughly USD 320K and 2.4 million—per product-market[76]. Introduction costs are higher for higher-end brands and wealthier and larger (in population) markets.

1.5 Counterfactual policy analysis

With the model and estimates in hand, I now illustrate firms' product portfolio responses to the industrial policy introduced at the beginning of this paper with counterfactual analyses. Spending up to $10 billion USD in subsidies[77], this policy induced entry of a large fringe[78] both in the number of fringe models and their total market share, shown in Figure 1.5.1. This policy started roughly in Jan. 2012, and I evaluate the impact of fringe entry on incumbent firms' product portfolio choices one year later, in Jan. 2013. I do so in a counterfactual experiment by considering what would have happened to product variety, prices, and welfare in the absence of the fringe competition.

Addressing this question requires explicitly solving the incumbent firms' product portfolio choice game in equation (1.3.9), which presents several challenges. First, the presence of multiple equilibria in positioning games is the rule rather than exception. Enumerating all equilibria of the portfolio game is impossible for any decent size of potential prod-

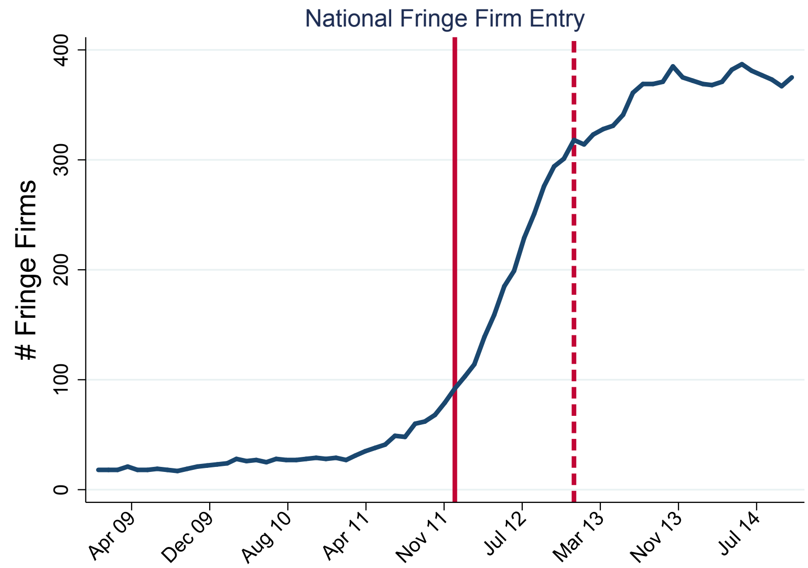
[76]These estimates are substantially higher than the estimates of development costs cited from press in Section 1.1 for several potential reasons. First, MIIT's estimates are on the model level, while I collapse similar models onto the product level. Second, MIIT's estimates are based on a low-end domestic handset—target of the industrial policy—while my estimates are for the major smartphone manufacturers' products. Third, MIIT's estimates likely focus on design and engineering costs of handset development, while my estimates include market-level marketing and retail negotiation efforts, as well as the opportunity costs of waiting to launch the product in the future.

[77]These subsidies were delivered in the form of cutting product testing fees and fringe-product marketing, by state-owned carriers. See footnote 12 for more details.

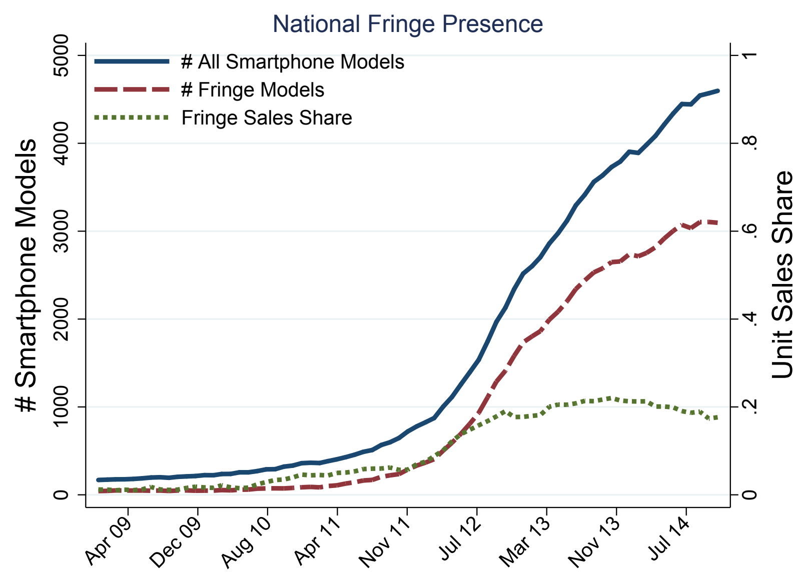
[78]Two technological barriers were also lifted around the same time. Both MediaTek and Qualcomm started releasing reference designs that made integrating their chipsets into smartphone handsets much easier for handset manufacturers. Android also started gaining popularity in China around 2011-2012, starting with its version 2.2 (Froyo), which gave rise to a large support industry of software developers. These two combined largely reduced the need for smartphone manufacturers to have teams of engineers and R&D to produce mid- to low-level handsets, allowing them to focus on low-cost manufacturing. Therefore, results in this section can also be interpreted as the effects of technology sharing and platform entry on product varieties. However, I do not estimate firms' product development costs on the country/year level, and therefore cannot separate out the effect of the competitive policy and the technology-sharing inventions. This paper thus does not aim to provide a detailed cost/benefit analysis of the competitive policy, but only studies the effect of competition on product variety.

Figure 1.5.1: Chinese industrial policy and fringe presence

(a) 2012 Chinese pro-competitive policy in the smartphone market



(b) Fringe presence



Notes: These two figures show the effect of the 2012 pro-competitive policy implemented by the Chinese government—the increasing presence of fringe firms in the Chinese smartphone market after 2012. Plotted in Figure 1.5.1a is the total number of unique fringe smartphone manufacturers in China. Solid vertical line, at Jan. 2012, indicates the approximate beginning of the policy. Dashed vertical line, one year later (Jan. 2013), indicates the month in which I evaluate the effect of this policy in counterfactual analysis. In Figure 1.5.1b, solid blue line and dashed red line are stacked (gap represents products by major manufacturers).

ucts for the twelve major manufacturers to consider[79]. I therefore use an iterated best-response algorithm and randomize the order of play between firms across simulations[80]. Second, the pool of potential product designs available to firms every month is unobserved, changes over time, and determines the dimensionality of the game. To tractably answer the question of firms' product portfolio responses to fringe entry, I define the set of potential products in Jan. 2013 to include all non-flagship products that were introduced in China between January and February of 2013. This allows me to examine which major manufacturer's product introduction was promoted early or delayed due to fringe entry. Finally, I further constrain the dimensionality of the game by only considering new product introductions in the counterfactual[81].

Table 1.11 first summarizes the set of potential products. There are a total of 17 products to be considered by seven firms, with Samsung having the most products at four. Column 3 shows that this set is a significant portion ($\sim$10%) of firms' existing product portfolios. Examples show that these products are important—the Samsung Galaxy Grand eventually became a blockbuster handset and the ZTE Grand S was near the market's frontier quality in all major characteristics. The final column shows how many product introductions were observed, on average, in each market. Table A.1 in Appendix A.2.3 presents the full list of potential products.

A final piece needed for the counterfactual simulation is the construction of sunk and variable costs, firms' PLC beliefs, as well as consumers' mean utility for products, which

[79]The size of the product introduction game alone is the power set of the set of all potential products. For example, in my actual counterfactual exercise, I consider 17 potential products among seven incumbent firms. Solving for all equilibria of this game is of the size 2^{17}—in order to consider all possible product configurations. Then to solve for equilibrium prices in Stage II under each product configuration, 10 simulations of this exercise would roughly take 3 months of runtime on 100 high-power computing cores in parallel on the Wharton High-Performance Computing Cluster.

[80]The iterated best-response algorithm is equivalent to an equilibrium selection procedure based on a best-response dynamic suggested in Lee and Pakes (2009) and also implemented in Wollmann (2016). Randomizing the order of play across simulations produces an "average" equilibrium in the counterfactual. Details of the iterated best-response algorithm are in Appendix A.3.3.

[81]In the counterfactual, I keep firms' product discontinuation decisions at the observed level. Allowing firms to re-adjust product discontinuation timings also makes the portfolio game computationally infeasible—for instance, the game of product discontinuation alone is of order 2^{199} in Jan. 2013, Beijing. This further simplification is justified for two reasons. First, for a product still near the peak of its product life cycle, it is unlikely that this change in competition will alter the firm's portfolio profitability to the extent that the product will be immediately discontinued. Second, for a product approaching the tail of its product life cycle, it is more likely that its EOP might be affected by the change in competition. However, given the thin tail of the product life cycle, the impact of its discontinuation (or lack thereof) on the firm's new-product-introduction decisions is likely orders of magnitudes smaller than the profitability change of the new product.

are not observed in the data in a given market-month (Jan. 2013), but could have been offered, and might now be offered in the more concentrated world[82]. I follow the timing of the model described in Section 1.3, and let firms play the Stage I portfolio choice game: First draw sunk cost shocks from the estimated distribution $\mu_{jmt} \sim N(0, \hat{\sigma}_\mu)$; then construct sunk costs from the point-estimated coefficients, using equation (1.4.9); firms form beliefs $\widehat{PLC}_{jm}$, using equation (1.4.6); then take expectations for Stage II marginal costs and demand, over shocks $(\omega_{jmt}, \xi_{jmt})$ (from equation (1.3.5) and (1.3.1)); and finally, firms take turns in maximizing expected portfolio profit specified in equation (1.3.9), conditional on other firms' up-to-date portfolios, and iterate until no firm has an incentive to deviate.

I first decompose the static and dynamic incentives behind firms' product introduction incentives. Increased fringe competition affects the incumbent firms' static portfolio profitability as well as their expected patterns of product life cycles in the future. I illustrate these two mechanisms by considering a scenario in which all 17 potential products were introduced in all 31 markets. In Table 1.12, I compare how the average static and dynamic profitability of these 527 product-markets were affected by fringe entry. Ignoring impacts of these introductions on other products, the average product's short-run profits decrease by RMB 38K, or 5.3% with fringe competition. This is the typical static entry argument, which predicts additional firm/product entry, induced by higher profitability in less competitive environments, and vice versa. The dynamic channel considered in this paper is through changes in expected product life cycles: On average, additional competition also depresses the average magnitude of expected product life cycles by 38%, and when put together with the static profit changes, this implies an average lifetime profitability reduction of 41%, or 6.9 million RMB, per product-market. Firms' dynamic product introduction incentives thus will amplify the effect of competition on firms' portfolio choices in equilibrium.

I now turn to equilibrium results. In equilibrium, however, the effect of competition on firms' static portfolio choices is ambiguous (Johnson and Myatt, 2003; Chu, 2010): Incumbent firms could "fight" by expanding their portfolios toward the low-end fringe entry,

[82] I observe 78 instances of product introductions in the 31 markets in Jan. 2013, out of the $17 \times 31 = 527$ possible product introductions.

or “accommodate” by segmenting the market, depending on the heterogeneous demand, costs, existing products, and strategic interactions among the incumbent firms. Furthermore, these different static incentives will be amplified by the dynamic incentives of changing product life cycles from the competitive pressure.

To illustrate the effect of competition on firms’ pricing, and static and dynamic product introduction incentives in equilibrium, I compare predictions of three different models in the counterfactual to the observed product variety, prices, and welfare: In the counterfactual without fringe competition, 1) product portfolios are fixed at the observed level[83], and firms only adjust prices; 2) firms re-optimize product portfolios, but (naively) hold their PLC beliefs constant; 3) firms re-optimize product portfolios, and rationally adjust PLC beliefs.

Findings: portfolio sizes and welfare predictions

As a baseline, I observe 2.52 new product introductions per market after fringe entry (Table 1.13). Average market-month welfare totals 1.497 billion RMB, or an annualized 557 billion RMB, for China, including consumer surplus from access to the mobile phone handset market[84] and major manufacturers’ variable profits[85].

[83] Another comparison is a model that does not allow for any new product introduction at all, which is more consistent from an ex ante policy analysis perspective. This is currently in progress. Compared to that model, the current model, with portfolios fixed at the observed level, provides a lower bound on the power of product introductions in constraining markups and prices in the counterfactual.

[84] Consumer surplus is computed as the simulated compensating variation with income effects (McFadden, 2012).

[85] In the following welfare calculations (only in the calculations of firm profits in the fourth row in Table 1.13), for comparison purposes, I do not include firms’ sunk and maintenance costs, and accordingly also do not account for expected future profits of products.

Table 1.9: Predicted PLC multipliers, example

$\widehat{PLC}$		Beijing	Tibet
	*Quality**Comp.*	competitive	concentrated
Nokia C5	low	29.04	115.4
Samsung Galaxy Ace	high	37.46	158.3

Notes: This table shows a 2-by-2 example of predicted PLC multipliers $\widehat{PLC}$ for the (lower-end) Nokia C5 and the (higher-end) Samsung Galaxy Ace in Beijing (more competitive) and Tibet (less competitive), at estimated parameters shown in Table 1.8. $\widehat{PLC}$ is defined as the ratio between expected lifetime profits of a product with its instantaneous profits in the first month after its release. This table shows that expected PLCs are much larger/longer for a higher-quality product, or in a market with less competition at the time of release.

Table 1.10: Estimated average introduction costs

	Avg. introduction cost (RMB, millions)
Brand	
avg. across products and markets	
Vivo	10.22
Oppo	9.303
Samsung	7.577
Motorola	7.053
Nokia	6.686
Lenovo	6.137
Huawei	5.033
Coolpad	4.324
HTC	3.969
ZTE	2.917
Province (entry cost into)	
avg. across products and brands	
Guangdong	15.19
Jiangsu	10.26
Shanghai	8.552
Henan	7.753
Beijing	7.720
Inner-Mongolia	5.713
Tianjin	5.010
Tibet	4.143
Jiangxi	2.859
Hainan	1.937

Notes: This table shows that the average introduction cost for a new product into one province is about $1 million, with large heterogeneity across brands and markets. It is estimated to be more costly to introduce a new product from a higher-end brand into a larger/wealthier market. Reported introduction costs are not actual coefficient estimates, but rather implied average introduction costs for the respective brands into the respective provinces. Reported province introduction costs are a select subset of the estimated introduction costs for 31 provinces.

Table 1.11: Potential products in Jan. 2013

Firm	# Potential product	Existing avg. portfolio size	Example	# Introduction (Jan.)
Samsung	4	35.52	Grand	0.32
Huawei	3	20.58	Y310	1.29
Lenovo	3	29.29	LePhone	0.23
Coolpad	3	27.77	8070	0.29
ZTE	2	14.42	Grand S	0
Oppo	1	12.00	R2	0.39
Vivo	1	12.58	S11	0
Total	17	237.2		2.52

Notes: This table summarizes the set of potential products and players in the counterfactual game. Potential products are defined as the set of all products newly introduced to any of the 31 provinces between Jan. and Feb. of 2013. This table suggests that these products are important (e.g., the Samsung Galaxy Grand was a blockbuster handset; ZTE's Grand S was close to the technology frontier at the time), and significant in magnitude (about 10% of existing portfolios). Columns 2 and 4 are averaged across markets. Column 4 shows, in Jan., how many product introductions each market sees, on average, in the data. Table A.1 in Appendix A.2.3 presents the full list of potential products with their characteristics.

Table 1.12: Decomposition of static and dynamic product-introduction incentives

	Before policy w/o fringe	After policy w/ fringe	Change
Short-run profits (RMB, 000's)	757.2	719.2	-5%
Expected PLC magnitude (Lifetime / short-run profits)	22.21	13.73	-38%
Lifetime profits (RMB, millions)	16.82	9.875	-41%

Notes: This table illustrates the mechanisms under how market competition affects firms' product profitability and in turn their product introduction incentives. I decompose firms' product introduction incentives into direct profitability concerns and indirect concerns through their expectations of the product life cycle in the following exercise: let all 17 potential products be introduced into all 31 markets and compare the average profitability/product life cycle of each product when the fringe is present from the policy to when it is eliminated in the counterfactual. This table shows that the increased competition reduces an average product's immediate profitability by 5% but its expected lifetime profitability by 41% through the 38% reduction in firms' PLC multiplier belief calculations. All reported numbers are averaged across the 17×31 product-markets.

Table 1.13: Effects of fringe entry: Portfolio sizes and welfare predictions

	Observed	Counterfactual: no fringe		
	w/ fringe (*baseline*)	fixed products (*price effects only*)	w/o ΔPLC (*static*)	w/ ΔPLC (*dynamic*)
# New products	2.52	2.52	7.32	9.58
Avg. handset price (RMB)	1,501	1,523	1,512	1,508
Consumer surplus (RMB, millions)	897	830	839	842
Firm profits (RMB, millions)	601	635	645	650
Annualized welfare (RMB, billions)	557	545	552	555

Notes: This table summarizes results of the counterfactual simulations, and decomposes the price effect, effect on firms' static product-introduction incentives, and effect on firms' dynamic product-introduction incentives, due to the fringe competition. Specifically, this table presents major handset manufacturers' product portfolio choices in equilibrium, and welfare estimates, both observed after fringe entry, and predicted in the absence of fringe competition, using three different models, in Jan. 2013. Column 1 presents observed product introductions, and welfare estimates based on those, after fringe entry. In column 2, firms' portfolios are fixed at the observed level, and can only adjust prices. In column 3, firms can adjust both their product portfolios and prices, but (naively) hold their future beliefs constant. In column 4, firms optimally adjust their portfolios and prices, based on both instantaneous profitability changes, as well as their future-belief changes, at the estimated parameters, in the counterfactual with no fringe competition. All numbers, except welfare, are averaged across 31 markets and 100 simulations in one month. Avg. handset price is share weighted. Consumer surplus is computed by simulated compensating variations with income effects for access to the entire set of mobile phone handsets following McFadden (2012). Firm profits exclude introduction and maintenance costs and expected future profits. Welfare is annualized and summed over all provinces in China.

Column 2, Table 1.13, shows the price effects of fringe entry if we do not let firms adjust their product portfolios in the counterfactual: Increased competition is predicted to have constrained prices by 1.4%; increased consumer surplus by 8.0% (both lower prices and more product variety from fringe firms); and decreased producer markups (profits) by 5.4%. The net welfare benefit of this policy is predicted to be an annualized 11.5 billion RMB in China.

In column 3, I let firms re-optimize product portfolios based on considerations of static tradeoff changes and a fixed future outlook (equivalent to a fixed hurdle rate model[86]): Incumbent manufacturers would have introduced almost three times more new products without fringe competition. Firms' additional product introductions in the less competitive environment self-regulate anti-competitive price effects, markups, and loss of consumer surplus: Welfare is predicted to have been only 4.8 billion RMB lower, without the fringe entry.

Finally, column 4 presents predictions using my model with product life cycles. Without fringe competition, firms fully re-optimize product portfolios, taking into account both the change in short-run profits and the change in future outlook for product life cycles. As alluded to in Table 1.12, the additional channel of firms' dynamic product-introduction incentives is important: Firms would have introduced 2.3 more handsets per province-month, had we accounted for their now larger magnitudes of product life cycles in the less competitive environment. Firms' dynamic considerations of product life cycles in portfolio choices help further self-regulate markets: Without fringe entry, the average price would have been only 7 RMB higher, and total welfare would have fallen by only about 2 billion RMB, or 19% of what was predicted in column 2, without allowing for portfolio adjustments. Compared to results in column 3, not accounting for firms' dynamic incentives in product introductions can understate firms' portfolio responses to the industrial policy by 24%, and overstate the welfare benefit of the policy by about 3 billion RMB, or more than twice as much.

[86]See Wollmann (2016). This still differs slightly, in that I use sunk cost estimates from my model with PLCs in the counterfactual.

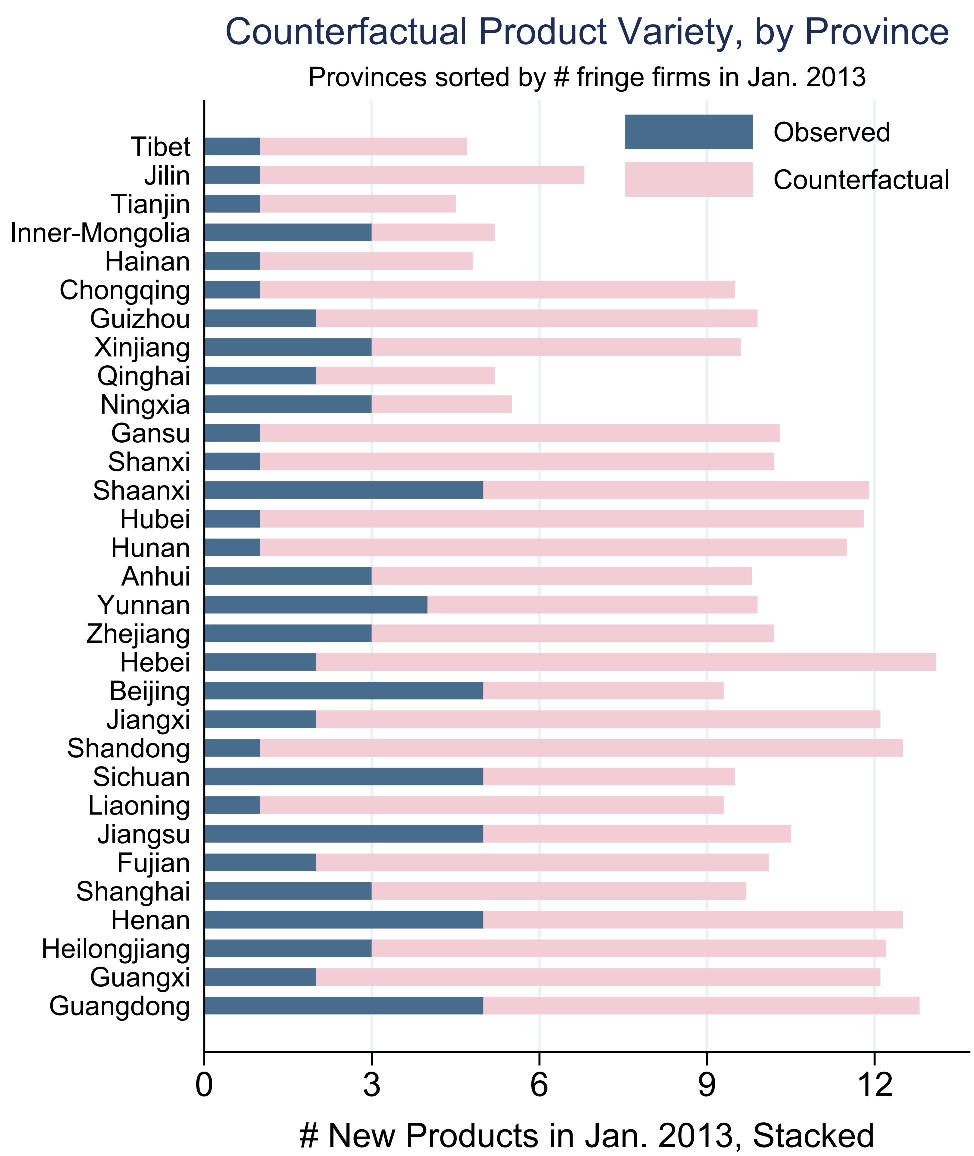
Findings: heterogeneous effects and product positioning

Table 1.13, however, also masks significant heterogeneity in the effect of competition on firms' portfolio choices. As previously shown in Figure 1.2.2, after the implementation of the industrial policy, fringe entry also varied significantly across markets given heterogeneous demand characteristics and market sizes. As a result, removing fringe competition in the counterfactual, Figure 1.5.2 unveils heterogeneous effects of the industrial policy on product variety across markets. Expectedly, larger and wealthier markets were able to justify more fringe entry after the policy, which, in turn, depressed incumbent firms' new product introductions the most: For instance, both Shanghai and Inner-Mongolia saw three new product launches in Jan. 2013—while Shanghai would have seen, on average (across simulations), 9.7 products, but Inner-Mongolia would have only seen 2.2 more products in the absence of fringe competition.

More interestingly, the kind of products that would have been introduced is also different. Figure 1.5.3 shows the effect of competition on firms' product positioning choices by comparing the distribution of incumbent firms' new product characteristics with and without fringe competition. I summarize product characteristics in a vertical measure of product quality—constructed from demand estimates, i.e. $\beta x_j + D^{NC}_{jmt} + \lambda_{f(j)} + \lambda_{l(j)}$ from equation (1.3.1)—for exposition: For example, Lenovo's LePhone 2802 was introduced at the bottom of the quality spectrum, while Samsung's Galaxy Grand was near the quality frontier at the time.

As discussed, the effect of low-end fringe entry on incumbent firms' portfolio choices is theoretically ambiguous—Figure 1.5.3 shows that incumbent smartphone manufacturers responded to fringe entry by expanding their product portfolios (in range, not size), and introducing more "fighting brands," such as the Lenovo LePhone 2802, to compete with fringe products. Figure 1.5.3 thus reveals another unintended consequence of the industrial policy: Not only increased competition depresses new product introductions overall, but also, to a larger extent, reduces mid- to high-end market competition between major manufacturers, effectively slowing down the speed of product innovation in the market, driven by firms' strategic incentives for product introductions. Bottom two panels

Figure 1.5.2: Market-level analysis of the effect of competition on product variety



Notes: This figure breaks down results presented in Table 1.13 by markets. Dark blue bars indicate observed market-level new-product introductions, in Jan. 2013, from major manufacturers. Light pink bars denote additional (stacked) new-product introductions—ones that would have happened, had fringe competition been eliminated. Provinces are sorted by the level of fringe presence in Jan. 2013 (number of fringe firms in province). This figure shows that the depression of new-product-introduction incentives is stronger in larger/wealthier markets, where fringe firms were justified to enter, with significant heterogeneity across markets.

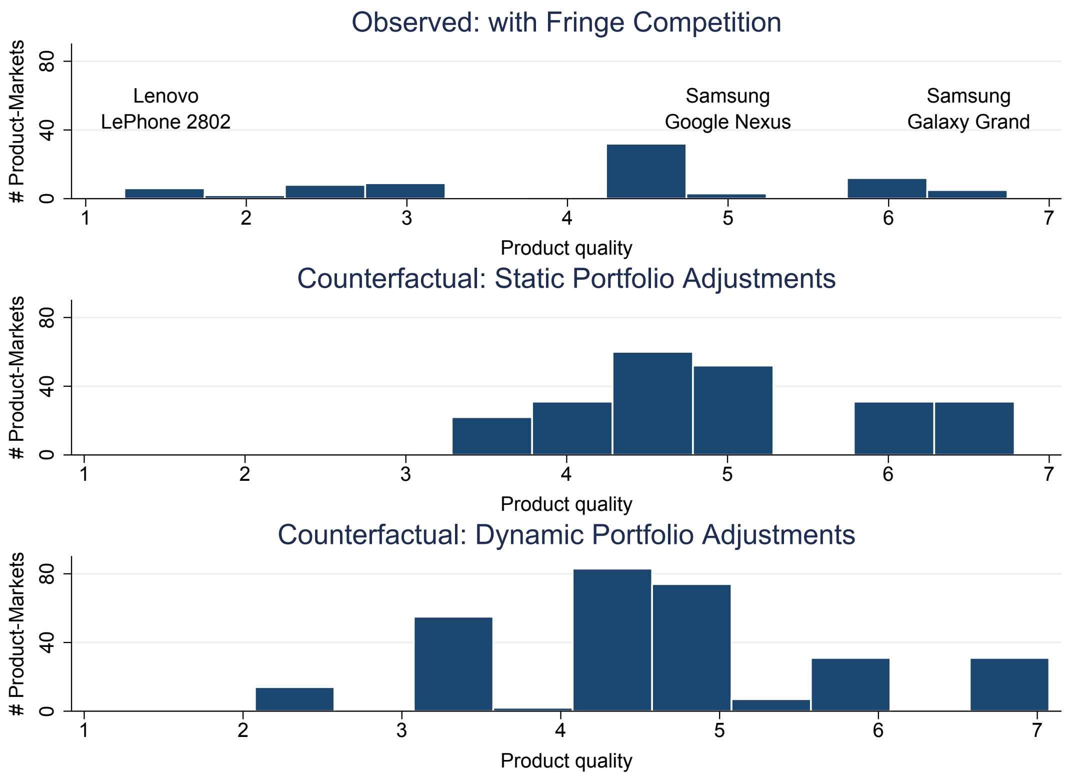
of Figure 1.5.3 also show different predictions of firms' product portfolio responses based on static and dynamic incentives. Interestingly, my model of dynamic product portfolio choice predicts a more spread out product configuration than a static model—suggesting that, in the absence of fringe competition, even the low-end market would not have been hurt as much as what would be predicted by a static model.

1.6 Conclusions

The bell-shaped time path of sales called the product life cycle is pervasive in many markets and is one of the oldest concepts in marketing. Although descriptive theories about product life cycle shapes have been developed over the past decades, little is known about how firms take product life cycles into account in introducing new products, and even less is known about the welfare implications of such PLC-based new-product strategies.

This paper first shows that product life cycles endogenously arise in markets with rapid technological innovations and exhibit bell-shaped sales paths for many high-tech products. The stylized model of product life cycle formation also provides useful comparative statics of the magnitudes of lifetime product sales compared to static observables of the product and market. These comparative statics are supported by the data and by the practices of product managers of smartphone manufacturers. I then use the product life cycle as an empirically tractable heuristic to model firms' dynamic product portfolio choices. When the level of market competition changes, firms not only expect immediate changes in their product profitability, but also adjust their future outlook for the product's profitability. As a result, not accounting for firms' dynamic incentives in product introductions can understate firms' product portfolio responses to industrial policies aimed at promoting competition and overstate policy impacts on product varieties and welfare in the market.

Figure 1.5.3: Effect of fringe entry on incumbent firms' product quality choices



Notes: This figure shows that, in the face of fringe entry, incumbent firms expanded their product portfolios and introduced more "fighting brands"—this also softened portfolio competition in the mid- to high-end market, and effectively slowed down the speed of product innovation in the market. Bottom two panels also show that the dynamic model of product portfolio choices predicts a more spread out product configuration in the absence of fringe entry, suggesting that even the low-end market would not have been hurt as much as what would be predicted under a static model. Product characteristics are summarized in this figure by the "estimated quality" measure constructed from demand estimates, i.e., $\beta x_j + D_{jmt}^{NC} + \lambda_{f(j)} + \lambda_{l(j)}$.

Chapter 2

Ownership Concentration and Strategic Supply Reduction[87]

2.1 Introduction

In 2010, the U.S. Federal Communications Commission (FCC) proposed to acquire spectrum from broadcast TV license holders and sell it to wireless carriers to be repurposed into mobile broadband spectrum. The so-called incentive auction combines a reverse auction for broadcast TV licenses with a forward auction for selling the thus-acquired spectrum to wireless carriers. Between the two auctions lies a repacking process where remaining broadcasters are reassigned channels to clear a contiguous nationwide block of spectrum for wireless use. Prior to the currently ongoing auction, estimates put the expected revenue from the forward auction at up to $45 billion, in excess of the payouts to broadcast TV license holders in the reverse auction, with the balance going towards the costs of repacking spectrum into a contiguous block and to the government.[88,89] In this paper, we study the potential for strategic behavior in the reverse auction.

[87] Joint with Ulrich Doraszelski, Katja Seim, and Michael Sinkinson, the Wharton School of the University of Pennsylvania.

[88] See Expanding Opportunities for Broadcasters Coalition (EOBC) Notice of Oral Ex Parte Filing with the FCC, June 13, 2014, available at http://www.tvtechnology.com/portals/0/EOBC0614.pdf, accessed on November 15, 2015.

[89] The Congressional Budget Office (CBO) estimates the net proceeds from the incentive auction to fall between $10 billion and $40 billion, with an expected value of $25 billion, the middle of that range. "Proceeds From Auctions Held by the Federal Communications Commission", CBO Report 50128, April 21, 2015, available at https://www.cbo.gov/publication/50128, accessed on November 15, 2015.

We document that following the announcement of the incentive auction, a number of private equity firms acquired broadcast TV licenses in several local media markets, often purchasing multiple licenses in the same market. Newspaper articles and industry reports claimed that these purchases were undertaken with the goal of trying to "flip" broadcast TV licenses for profit in the reverse auction.[90] Politicians also raised concerns about speculation.[91] Yet, reselling broadcast TV licenses does not necessarily entail efficiency losses.

We argue that in addition to speculative motivations behind reselling broadcast TV licenses, there is the potential for strategic bidding. In a prospective analysis of the incentive auction, we show that owners of multiple licenses have an incentive to withhold some of their licenses from the auction, thereby driving up the closing price for the remaining licenses they own and affecting a large transfer of wealth from the government - and ultimately taxpayers - to themselves. Apart from affecting closing prices, this behavior causes efficiency losses as the set of broadcast TV licenses surrendered in the auction is not the socially optimal set and may reduce the total amount of spectrum that will be repurposed.

Repurposing spectrum from broadcast TV to mobile broadband usage is no doubt an extremely valuable - and complex - undertaking. The incentive auction was very carefully designed and has many desirable properties such as strategy proofness (Milgrom et al., 2012; Milgrom and Segal, 2014). If broadcast TV licenses were separately owned, then it is optimal for an owner to bid a station's value as a going concern in exchange for relinquishing the broadcast license; we refer to this as naive bidding. Our paper points out the sensitivity of the incentive auction to multi-license ownership. In particular, the rules of the auction leave room for strategic supply reduction for firms that own multiple broadcast TV licenses. Such firms can withhold a subset of their licenses from the auction to raise the closing price for the remaining licenses. This behavior is purely rent-seeking, as these firms are attempting to increase their share of existing wealth without creating any

[90]See "NRJ Wins Bidding For WSAH New York," TVNewsCheck, November 29, 2011, "Small TV Stations Get Hot," The Wall Street Journal, September 3, 2012, "Speculators Betting Big on FCC TV Spectrum Auctions," Current.org, February 26, 2013, "TV Spectrum Speculation Nears $345 Million," TVNewsCheck, March 1, 2013, "Broadcast Incentive Spectrum Auctions: Gauging Supply and Demand," SNL Kagan Broadcast Investor, November 20, 2013, and "TV Station Spectrum Deals Expand Into Major Network Affiliates as Players Stake Out Positions Pre-Auction," SNL Kagan Broadcast Investor, December 4, 2013.

[91]See "Rep. LoBiondo Seeks FCC Info On Possible Spectrum Speculation," Broadcasting & Cable, February 12, 2014.

new wealth.

We use a simple model to illustrate how strategic supply reduction works in the context of the reverse auction and under what circumstances it is a profitable strategy for multi-license owners. Our model implies that certain types of broadcast licenses are more suitable for a supply reduction strategy and that certain types of local media markets are more vulnerable to this type of behavior. We begin by showing that the ownership patterns in the data are broadly consistent with the implications of the model.

In a second step, we analyze the reverse auction in more detail and quantify increased payouts and efficiency losses due to strategic supply reduction. To do so, we undertake a large scale valuation exercise to estimate reservation values for all currently held UHF broadcast TV licenses. We combine various data sources to estimate a TV station's cash flows and from them infer its value as a going concern. This allows us to simulate the auction outcome for all participating license holders, accounting for the repacking process at a regional level, and then to assess the impact of potential strategic bidding.

We compare the outcome under naive bidding with the outcome when we account for the ownership patterns in the data and allow multi-license owners to engage in strategic supply reduction. Our approach solves for all equilibria of a simplified version of the reverse auction that accounts for the repacking process at the regional - though not at the national - level. We then show that across markets, strategic supply reduction has a large impact on closing prices and payouts to broadcast TV license holders and causes sizable efficiency losses. For a nationwide clearing target of repurposing 126 MHz of spectrum, the initial starting point of the incentive auction when it commenced on March 29, 2016, strategic behavior by multi-license owners increases payouts by 22%. The payout increases, as well as payouts from the auction in general, are concentrated in a small number of markets, however, with nearly 99% of payout increases occurring in markets with two or more owners of multiple licenses. This reflects two factors. First, there is significant variation in station cashflows and thus willingness-to-sell due to stations' differential success in attracting advertising revenue. This results in steep increases in closing prices as more licenses are acquired. Second, the FCC's need to clear spectrum is particularly pronounced in large markets where the expected demand by wireless carriers means that even high-value TV licenses can successfully sell into the auction and that strategically withholding a low-value station can drive up closing prices significantly. We illustrate these issues in a case study of the Philadelphia, PA, market.

Netting out the firms' reservation values from their auction payouts, the strategic supply reduction strategy translates into surplus increases of several billion dollars across broadcast TV license owners. This reflects that a multi-license owner who withholds a license from the auction creates a positive externality for other market participants by raising the closing price in the market. The multi-license owner, by selling his remaining licenses in the market, captures some of this externality, but not all of it: while in aggregate, payouts increase by 22%, they increase by 25% for multi-license owners and 19% for single-license owners.

We propose a partial remedy that imposes a constraint on the ordering of bids of multi-license owners. The rule change reduces the effect of strategic behavior by roughly eighty percent. This result is directly policy relevant as it suggests ways of mitigating the impact of strategic supply reduction. Our hope is that this proves useful in designing future auctions in the U.S. and other countries as they strive to alleviate the "spectrum crunch" resulting from the rapid growth in data usage by smartphones in recent years.

We also illustrate how a firm may extend a supply reduction strategy by leveraging technological constraints on the repacking of spectrum across local media markets. While a complete analysis of multi-market strategies at the national level is computationally infeasible, we highlight a particular case of a firm owning licenses in adjacent markets and the potential effect of reducing supply in one market on the closing price in another targeted market. In this case, we find a six-fold increase in the impact of strategic bidding.

Finally, we quantify the effect of potential low participation by non-commercial (public) and religious broadcasters. A reduction in participation among such licenses leads to a significant number of auction failures, and large increases in payouts when auctions complete. We suggest that modifying must-carry regulations could be a useful tool in increasing auction participation.

There is a rich literature on strategic bidding in multi-unit auctions. Substantial theoretical work (Wilson, 1979; Back and Zender, 1993; Menezes, 1996; Engelbrecht-Wiggans and Kahn, 1998; Jun and Wolfstetter, 2004; Riedel and Wolfstetter, 2006; Ausubel et al., 2014) and experimental evidence (List and Lucking-Reiley, 2000; Kagel and Levin, 2001; Engelmann and Grimm, 2009; Goeree, Offerman and Sloof, 2013) point to the potential for strategic demand reduction. In addition, case studies of past spectrum auctions have documented strategic demand reduction (Weber, 1997; Grimm, Riedel and Wolfstetter, 2003). Our paper is most closely related to the empirical literature examining market power in wholesale electricity markets (Wolfram, 1998; Borenstein, Bushnell and Wolak, 2002; Hortacsu and Puller, 2008), where firms bid supply schedules and have strategic incentives to alter their bids and raise closing prices for inframarginal units. In contrast to electricity markets, however, in our setting the acquisition of significant market power is easier: a small number of licenses outstanding in a given local market, combined with the discreteness of broadcast spectrum units (6 MHz), implies that a single additional license can confer a significant increase in market power in the spectrum market to its owner.

Significantly, our paper departs from much of the auction literature in that it does not invert the first-order conditions to recover valuations from observed bids. Instead, we use auxiliary data to directly estimate valuations. One reason is the lack of bidding data because, by congressional order, the FCC is unable to release details of the auction proceedings until two years after completion of the incentive auction.

However, even if data were available, extending the standard first-order conditions approach to our case of multi-unit auctions with personalized prices is less than straightforward and may entail challenges to identification as discussed by Cantillon and Pesendorfer (2007) in the context of heterogeneous multi-unit first-price auctions. More importantly, the descending clock nature of the ongoing incentive auction exacerbates identification concerns. While the value of a broadcast license can be inferred from the price at which it drops out of the auction, payout price provides at most an upper bound on the value of a broadcast license that instead sells into the auction. Further, observing that a broadcast license is withheld from the auction is uninformative about its value if a multi-license owner chooses to withhold it from the auction for strategic reasons. This is in contrast to work on wholesale electricity markets where complete supply schedules are observed. We also do not adopt the moment inequalities approach in Fox and Bajari (2013) that - rather than assuming full optimality of bids - only assumes that the configuration of licenses that a multi-license owner sells into the auction dominates any alternative configuration due to the typically small number of alternatives in our setting given that licenses are not perfect substitutes for one another. This approach also only identifies relative valuations but not the levels of valuations that are required for the analysis of welfare effects of ownership concentration.

Our work is also related to the extensive literature on collusion in auctions (Asker 2010, Conley and Decarolis 2016, Kawai and Nakabayashi 2015, and Porter and Zona 1993, among others), including spectrum auctions (Cramton and Schwartz, 2002). A multi-license owner in our setting internalizes the profit implications of all licenses he controls as is the case with colluding, but otherwise independent, single-license owners. Finally, we contribute to the literature on distortions induced by incentive schemes and regulation in various settings such as employee compensation (Oyer, 1998), environmental regulation (Fowlie, 2009; Bushnell and Wolfram, 2012), health care (Duggan and Scott Morton, 2006), and tax avoidance (Goolsbee, 2000).

The remainder of this paper is organized as follows: Section 2 describes the setting and sets out a simple model of strategic supply reduction, Section 3 presents data and descriptive evidence, Sections 4 and 5 describe the main analysis and results, and Section 6 concludes.

2.2 The FCC incentive auction

The rapid growth in data usage by smartphones has significantly increased the demand for mobile broadband spectrum in recent years.[92] At the same time, some previously allotted spectrum is no longer used intensively. In particular, each of approximately 8,500 currently

[92] According to FCC Chairman Tom Wheeler, "America has gone mobile. Most Americans would have a hard time imagining life without their smartphones, and tens of millions are similarly in love with their tablets. The problem is that spectrum, the lifeblood of all wireless technologies, is finite. That wasn't a problem before the mobile web, when most consumers were mostly watching videos or surfing the web at home. If we don't free up more airwaves for mobile broadband, demand for spectrum will eventually exceed the supply. If you've ever been frustrated by websites that loaded slowly or videos that wouldn't download to your phone, you have a sense what that world could look like." See "Channel Sharing: A New Opportunity for Broadcasters," Official FCC Blog, available at https://www.fcc.gov/news-events/blog/2014/02/11/channel-sharing-new-opportunity-broadcasters, accessed on November 15, 2015.

operating TV stations owns a license for a 6 MHz block of spectrum covering a particular geographical area for over-the-air transmission of programming. Yet, as of 2010 only about 10% of U.S. TV households used broadcast TV, with a rapidly declining trend.[93]

In its 2010 National Broadband Plan, the FCC under then-chairman Julius Genachowski proposed, and was authorized by Congress in 2012, to conduct a incentive auction to reallocate spectrum from TV stations located in the higher frequency UHF band to wireless providers. The incentive auction consists of a reverse auction in which TV stations submit bids to relinquish spectrum rights in exchange for payment and a forward auction in which wireless operators bid for the newly available spectrum.

While the FCC has conducted spectrum auctions in the past, the incentive auction is the first time an auction to sell spectrum is combined with an auction to purchase spectrum from existing licensees.[94] Designing this auction is complicated not only by incumbent claims on spectrum, but also by technological constraints for mobile data and broadcast TV uses. Originally projected for early 2014, the incentive auction was repeatedly postponed due to legal and technological challenges, first to the middle of 2015 and then again to its ultimate starting date of March 29, 2016.[95]

The format of the ongoing auction was made public in May 2014.[96] The forward auction to sell spectrum to wireless carriers uses an ascending-clock format similar to previous spectrum auctions. The reverse auction uses a descending-clock format in which the price offered to a TV station for its spectrum usage rights declines with each successive round

[93] "Connecting America: The National Broadband Plan", FCC, 2010, Chapter 5, p. 89.

[94] "Let's start with the concept of an incentive auction. While it has never been tried before, its power lies in how it addresses the root of all issues: economics. If it is possible to marry the economics of demand with the economics of current spectrum holders, it should be possible to allow market forces to determine the highest and best use of spectrum. In mid-2015 we will run the first ever incentive auction. Television broadcasters will have the opportunity to bid in a reverse auction to relinquish some or all of their spectrum rights, and wireless providers will bid in a forward auction on nationwide, 'repacked' spectrum suitable for two-way wireless broadband services." See FCC Chairman Tom Wheeler's prepared remarks at the "Wireless Spectrum And The Future Of Technology Innovation" Forum, available at https://apps.fcc.gov/edocs_public/attachmatch/DOC-326215A1.pdf, accessed on November 15, 2015.

[95] See "The Path to a Successful Incentive Auction," Official FCC Blog, December 6, 2013, available at https://www.fcc.gov/news-events/blog/2013/12/06/path-successful-incentive-auction-0, accessed on November 15, 2015, and "F.C.C. Delays Auction of TV Airways for Mobile," The New York Times, October 24, 2014.

[96] See https://apps.fcc.gov/edocs_public/attachmatch/FCC-14-50A1.pdf, accessed on November 15, 2015. An excellent and detailed explanation of the mechanism is available from the FCC and greatly informs our analysis. See Appendix D of FCC Public Notice in matter FCC-14-191 "Comment Sought On Competitive Bidding Procedures For Broadcast Incentive Auction 1000, Including Auctions 1001 And 1002," released December 17, 2014.

of bidding. A TV station faces a price for its broadcast license that is personalized to it (see Section 2.2.2 for details). If a TV station chooses to participate in the reverse auction, it has several options for relinquishing its spectrum usage rights: going off the air, moving channels from a higher frequency band (UHF channels 14-36 and 38-51 or high VHF channels 7-13) to a lower frequency band (respectively, VHF channels 2-13 or low VHF channels 2-6) to free up more desirable parts of the spectrum, or sharing a channel with another TV station.

Between the reverse and forward auctions, a repacking process takes place in which the remaining TV stations are consolidated in the lower end of the UHF band to create a contiguous block of spectrum in the higher end of the UHF band for wireless use.[97] The process is visually similar to defragmenting a hard drive on a personal computer. However, it is far more complex because many pairs of TV stations cannot be located on adjacent channels, even across markets, without causing unacceptable levels of interference. As a result, the repacking process ties together all local media markets. In practice, the reverse auction is therefore at the national level. A further consequence of interference is that even though each TV station owns a license for a 6 MHz block of spectrum covering a particular geographical area, far more than $6n$ MHz of spectrum are likely required to accommodate n remaining TV stations in a market.

The auction rules integrate the reverse and forward auctions in a series of stages. For the first stage, initial commitments from stations and repacking constraints determined an initial maximum nationwide clearing target of 126 MHz. Each stage of the incentive auction begins with multiple rounds of the reverse auction, followed by multiple rounds of the forward auction. The reverse auction determines the cost of acquiring a set of licenses that allow the repacking process to meet the clearing target. There are many different feasible sets of licenses that could be surrendered to meet a particular clearing target given the complex interference patterns between stations; the reverse auction is intended to identify the low-cost set.

After the reverse auction determines the cost of acquiring an amount of spectrum, the

[97]Congress' authorization of the incentive auction required the FCC to make all reasonable efforts to preserve the coverage area and population served by TV stations involved in the repacking.

forward auction determines the willingness-to-pay of wireless operators for this amount of spectrum. If willingness-to-sell in the reverse auction outpaces willingness-to-pay in the forward auction, the clearing target is decreased, so that a smaller set of lower value TV stations have to be acquired. The process repeats until a "final stage rule" is satisfied that ensures that proceeds in the forward auction (more than) cover payouts in the reverse auction and the cost of repacking spectrum.[98]

At the time of this writing, the FCC has completed three stages of the auction. The reverse auction phase of the auction's first stage concluded on June 29, 2016. The total payout required to meet the initial 126 MHz clearing target amounted to $86.4 billion, which was not met in the forward auction in attempting to sell the freed spectrum to wireless providers; it yielded a payout of only $23.1 billion. Since this first stage, the clearing target has been lowered twice and the FCC opened the fourth stage for a clearing target of 84 MHz on December 13, 2016.

We next provide additional details on the reverse auction and illustrate the potential for strategic supply reduction.

2.2.1 The reverse auction

The reverse auction uses a descending-clock format. A TV station that participates in the reverse auction is offered a personalized price at which it can either remain in the auction, indicating that it is prepared to accept this price to cease operating and surrender its broadcast license, or leave the auction, indicating that the price is too low and that it prefers to continue operating and potentially be repacked to a new UHF channel. In the subsequent analysis, we abstract from the options to relocate from a higher to a lower frequency band or to share a channel with another station. We discuss this simplification further in Section 2.4.2.

Broadcast licenses are assigned by the FCC to a local media market, which is the designated market area (DMA) as defined by Nielsen Media Research based on the reach and

[98]Specifically, the final stage rule requires that proceeds in the forward auction are at least $1.25 per MHz per population for the largest 40 wireless service market areas and not only cover payouts in the reverse auction but also the FCC's administrative costs, the reimbursements of channel relocation costs incurred by TV stations, and the funding of the First Responder Network Authority's public safety operations.

viewing patterns of TV stations. A DMA is defined as a group of counties such that the home market TV stations hold a dominance of total hours viewed. There are 210 DMAs in the U.S. that vary in size from New York, NY, with over 7 million TV households, to Glendive, MT, with 4,230 TV households. Appendix Table B.1 lists the top ten DMAs based on their 2012 rank.

Across these 210 markets, a total of 2,166 broadcast licenses are eligible for the auction.[99] They can be classified by type of service into UHF and VHF stations, by type of use into commercial and non-commercial stations, and by power output into full-power (primary and satellite[100]) and low-power (class-A) stations. Appendix Table B.2 summarizes the auction-eligible broadcast licenses.

Formally, in round τ of the reverse auction, a currently active TV station j is offered the price

$$p_{j\tau} = \varphi_j P_\tau,$$

where P_τ is the base clock price and φ_j is the station's broadcast volume. The base clock price P_τ begins at $900 and decreases with each successive round of bidding. The broadcast volume

$$\varphi_j = M\sqrt{CoveragePop_j \cdot InterferenceCnt_j} \tag{2.2.1}$$

is a known function of the station's population reach $CoveragePop_j$ and the interference count $InterferenceCnt_j$, defined as the number of TV stations that station j can potentially interfere with in the repacking process. Finally, $M = 17.253$ is a scaling factor that is chosen to set the maximum φ_j across the U.S. to one million.

The broadcast volume is an important concept: the FCC uses it to personalize the base clock price to a TV station based on its value as a broadcast business (as proxied by pop-

[99] See http://www.fcc.gov/learn, accessed on November 15, 2015. The FCC excludes approximately 10,000 low-power, translator, multi-cast signal, and cable stations from the reverse auction. In 2016, the FCC updated the list of auction-eligible stations, see http://transition.fcc.gov/Daily_Releases/Daily_Business/2015/db0609/DA-15-679A2.pdf, accessed on February 10, 2016. In this paper, we work with the earlier list of 2,166 auction-eligible stations as it underlies the FCC's repacking simulations (see Section 2.3.1).

[100] A satellite station is a relay station that repeats the broadcast signal of its parent primary station.

ulation reach) and the difficulty of repacking the station in case it does not surrender its license (as proxied by the interference count). The broadcast volume thus reflects that the FCC is willing to incentivize a TV station to surrender its license if the alternative of having to repack the station is particularly challenging. Importantly, the broadcast volume for all TV stations is known in advance to all auction participants.

The design of the reverse auction is partly dictated by the FCC's obligation to guarantee a 6 MHz block of spectrum to any TV station that chooses to remain on air rather than surrender its license. At the initial base clock price of $900, most, if not all, TV stations would be prepared to surrender their licenses. Hence, any remaining TV stations can be trivially repacked. The auction mechanism then preserves the feasibility of repacking as it unfolds as follows: as the base clock price descends, licenses withdraw from the auction, deciding that the price is too low and that they would prefer to continue broadcasting. When this happens, the feasibility of repacking every single remaining license in the auction must be asserted one-by-one given the interference patterns of the withdrawing and the remaining stations. If a remaining license can no longer be repacked, the price it sees is "frozen" and it is declared to be "provisionally winning," in that the FCC will accept its bid to surrender its license. In this case, the withdrawing station effectively sets the price of the frozen station. The base clock price then falls and the process of feasibility checking repeats with each new withdrawal. The reverse auctions ends if all TV stations have either withdrawn from the auction or are provisionally winning. Note that there is no single market price at which a station sells; different stations obtain different closing prices for their spectrum depending on the exact base clock price when, given the implied set of withdrawn stations, they could no longer be repacked.

2.2.2 Strategic supply reduction

Clock auctions are strategy proof (Milgrom et al., 2012; Milgrom and Segal, 2014). Hence, if a TV station is independently owned, its owner optimally remains in the reverse auction until

$$p_{j\tau} = \varphi_j P_\tau < v_j,$$

where v_j is the reservation value of TV station j that reflects its value as a going concern. We henceforth refer to this strategy as naive bidding.

Clock auctions are not only strategy proof but also "group-strategy proof" (Milgrom and Segal, 2014). This means that no coalition of bidders has a joint deviation from naive bidding that is strictly profitable for all members of the coalition. However, as Milgrom and Segal (2014) explicitly acknowledge, their results do not apply if bidders are "multi minded," a concept that includes bidders with multiple objects for sale.

We show that a firm owning multiple broadcast licenses may indeed have an incentive to deviate from naive bidding. Note that this does not contradict group-strategy proofness as it suffices that the deviating group, i.e., the multi-license owner, is better off as a whole. Withdrawing a license from the auction could increase the price for the remaining broadcast TV licenses that a firm owns. However, the firm is then left with a TV station that it may have been able to sell into the auction. Therefore, this supply reduction strategy is only profitable if the gain from raising the closing price for other stations exceeds the loss from continuing to own a TV station instead of selling it into the auction.

For concreteness and simplicity, consider a situation where all stations are perfectly interchangeable in the repacking process. A firm owns TV stations a and b. The FCC intends to acquire k broadcast licenses and stations are ordered in ascending order of the ratio $\frac{v_j}{\varphi_j}$. If $\frac{v_a}{\varphi_a} < \frac{v_k}{\varphi_k}$ and $\frac{v_b}{\varphi_b} < \frac{v_k}{\varphi_k}$, then under naive bidding the reverse auction closes at base clock price $\frac{v_{k+1}}{\varphi_{k+1}}$ and both licenses sell into the auction, yielding the firm a profit of $(\varphi_a + \varphi_b)\left(\frac{v_{k+1}}{\varphi_{k+1}}\right) - (v_a + v_b)$. On the other hand, if the firm withholds station a from the auction, then the closing base clock price rises to $\frac{v_{k+2}}{\varphi_{k+2}}$ and its profit is $v_a + \varphi_b \frac{v_{k+2}}{\varphi_{k+2}} - v_b$. It is therefore profitable to engage in strategic supply reduction and withhold TV station a from the auction if the gain in profit from selling the license of TV station b outweighs the loss in profit from not selling the license of TV station a, or

$$\varphi_b \left(\frac{v_{k+2}}{\varphi_{k+2}} - \frac{v_{k+1}}{\varphi_{k+1}}\right) > \varphi_a \left(\frac{v_{k+1}}{\varphi_{k+1}}\right) - v_a. \tag{2.2.2}$$

The left-hand side implies that strategic supply reduction is more likely to be profitable if φ_b is large and if the increase in the closing base clock price $\frac{v_{k+2}}{\varphi_{k+2}} - \frac{v_{k+1}}{\varphi_{k+1}}$ is large. The right-hand side implies that it is more likely to be profitable if φ_a is small and v_a is large. In short, strategic supply reduction is more likely to be profitable if the "leverage" of increasing the closing base clock price is large and the opportunity cost of continuing to own a TV station is small.

Strategic supply reduction has been explored in earlier work on multi-unit auctions in wholesale electricity markets (e.g. Wolfram, 1998): if a firm's bid for one of its generators has a chance to set the price, then the firm has an incentive to raise that bid if it will earn the price increases on its inframarginal generators.[101] Other electricity market papers consider this the exercise of market power, and note that the effects can be large when demand or supply is inelastic (Borenstein, Bushnell and Wolak, 2002). Unlike in wholesale electricity, a broadcast TV license is indivisible, leading to sharper behavior in our setting: while there is a maximum of 28 UHF licenses outstanding in a given DMA market, the median DMA market has 7 UHF licenses, and the mean is 8.2; for the median market, the market share increase from owning two licenses, rather than a single one, represents a market share jump from 14.3% to 28.6%. Furthermore, because of interference constraints, licenses are not homogeneous in the repacking process. Both facts may amplify the impact of strategic supply reduction. At the same time, unlike the short-run demand for electricity, the FCC's demand for licenses is not inelastic, and we account for this in our subsequent analysis of the reverse auction.

Equation 2.2.2 implies that certain types of DMAs are more vulnerable to a supply reduction strategy and that certain types of broadcast TV licenses are more suitable for this type of behavior. First, ideal markets from a supply reduction perspective are DMAs in which the FCC will likely need to acquire a positive number of broadcast licenses and where, at the expected demand levels, closing prices for selling stations are likely to change significantly should a lower value station be removed from the auction. This maximizes the impact of withholding a license from the auction on the closing price (the left-hand

[101] This mechanism is also similar to the upper bound of the "bidder exclusion effect" considered by Coey, Larsen and Sweeney (2015) in the case of a non-random merger of auction participants.

side of equation 2.2.2). Second, suitable groups of licenses consist of some stations with higher broadcast volume to sell into the auction and some with lower broadcast volume to withhold. We return to these implications of the model below when discussing the data and our results.

2.3 Data and descriptive evidence

We begin by describing the various sources of data used in the analysis and then turn to providing descriptive evidence in support of the model from Section 2.2.2.

2.3.1 Data sources

We use several sources of data to construct firm valuations, determine how non-selling TV stations would be repacked, and summarize the likely spectrum demand in a given market in the forward auction. First, we rely on the MEDIA Access Pro Database (2003 - 2013) from BIA Kelsey (henceforth BIA) and the Television Financial Report (1995 - 2012) from the National Association of Broadcasters (NAB) to estimate a TV station's cash flows and from them infer its reservation value going into the auction.

BIA contains the universe of broadcast TV stations. It provides station, owner, and market characteristics, as well as TV stations' transaction histories covering the eight most recent changes in ownership. The BIA's revenue measure covers broadcast-related revenue in the form of local, regional, and national advertising revenue, commissions, and network compensation. We refer to BIA's revenue measure as advertising revenue in what follows. For commercial stations, advertising revenue is missing for 30.9% of station-year observations, which we impute as detailed in Appendix B.1.1.3. For non-commercial stations, advertising revenue is missing for 99.7% of station-year observations and we do not impute it.

The BIA data excludes non-broadcast revenue, most notably, retransmission fees. These are fees TV stations charge pay-TV providers to use their content, which the trade press mentions as a small but growing source of revenue for many TV stations.[102] Outside esti-

[102]See, e.g., "SNL Kagan raises retrans fee forecast to $9.8B by 2020; Mediacom's CEO complains to FCC",

mates suggest that advertising revenue accounts for a declining share of a typical station's revenue, estimated at 69% in 2016, with the remaining revenue coming from retransmission fees (24%) and online revenues (7%).[103] Consequently, the variation in advertising revenue across stations is a major, but not the sole, driver of the variation in cash flows and thus willingness-to-sell in the reverse auction.

To get at non-broadcast revenue and ultimately profitability, we rely on NAB as a second source of data. For commercial full-power stations, NAB collects financial information. Revenue is broken down into detailed source categories from which we are able to construct advertising revenue and non-broadcast revenue. NAB further covers expenses related to programming, advertising, and other sources, and profitability as measured by cash flows. However, for confidentiality reasons, NAB reports only the distributions of these measures (the 25^{th}, 50^{th}, and 75^{th} percentiles, as well as the mean) at various levels of aggregation, resulting in "tables" such as "ABC, CBS and NBC affiliates in markets ranked 51-60 in 2012" or "CBS affiliates in markets ranked 1-50 in 2012." Appendix Table B.4 lists the set of 66 tables for 2012; other years are very similar. In Section 2.4 we describe a method to combine the station-level data on advertising revenue from BIA with the aggregated data from NAB to estimate a TV station's cash flows.

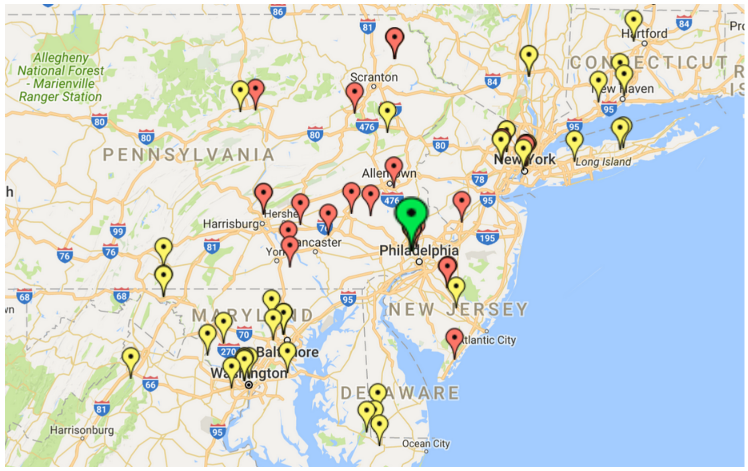
To simulate the repacking process required to construct a contiguous spectrum block out of acquired TV broadcast spectrum, we use three inputs available from the FCC: a TV station interference file, a TV station domain file, and a repacking feasibility checker that takes these two files as inputs.[104] The first file lists, for every TV station and every channel, sets of other TV stations that cannot be located on the same channel, or alternatively cannot be located on adjacent channels, due to interference constraints given the stations' facility locations. Looking only at the UHF channels that would exist if the 126 MHz clearing target had been met (channels 14-30), this file lists 2.5 million pairwise restrictions between broadcast TV stations. As an example, Figure 2.3.1 shows the set of the 102 TV stations that have interference constraints with WCAU, the Philadelphia affiliate of NBC.

FierceCable, July 7, 2015.

[103] "Retrans Revenue Share Expands In Latest U.S. TV Station Industry Forecast", Justin Nielson, S&P Global Market Intelligence, Jul 14, 2016.

[104] All three files are available at `http://data.fcc.gov/download/incentive-auctions/Constraint_Files/`.

Figure 2.3.1: Interference constraints for NBC Philadelphia (WCAU)

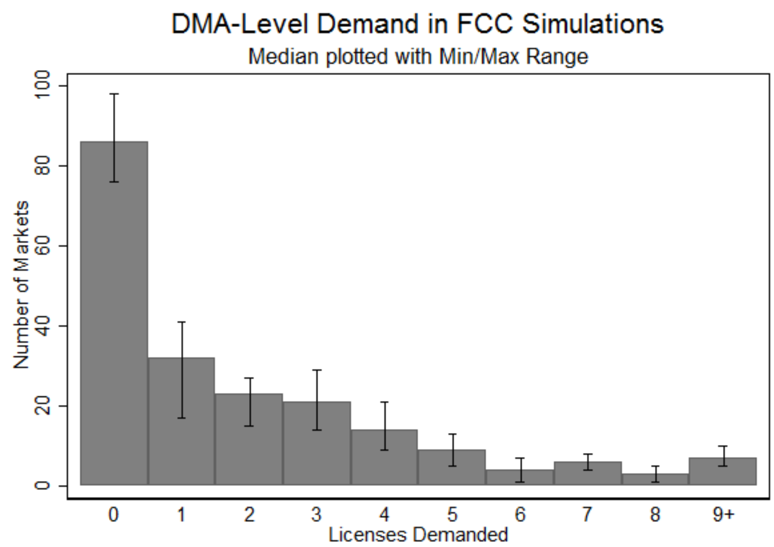


Notes: Each pin represents the facility location of a TV station. WCAU (NBC Philadelphia) is denoted by a green pin. TV stations denoted by red pins have adjacent-channel interference constraints, while those denoted by yellow pins have same-channel interference constraints. A total of 102 broadcast TV stations have some interference constraint with WCAU.

Of those, 37 have adjacent-channel constraints, meaning that they cannot even be located one channel above or below WCAU in the repacked region of spectrum, while the rest have same-channel constraints. Interference is influenced by several factors, including geography, broadcast tower height, and the transmitter's power output. The second file, the domain file, provides a list of channels that a given station may be assigned to in repacking. For most UHF stations, the set of valid channels is the set of all UHF channels, although some have fewer due to international broadcasters or military broadcasting. Relying on these two inputs, the so-called SATFC feasibility checker ascertains whether or not a set of stations can be feasibly repacked into a set of channels given interference constraints and constraints on station channel locations. SATFC is run as part of the reverse auction and uses optimized approaches to NP-complete problems to limit the space of problems to be considered.

Last, the simplified reverse auction model above pointed to DMA-level spectrum demand as a determinant of the likely success of strategic supply reduction. To construct a simple demand estimate, we rely on output of a large-scale simulation exercise conducted by the FCC to determine the likely number of UHF stations it has to acquire in each DMA

Figure 2.3.2: Demand across DMAs



Notes: This histogram indicates how many DMAs need a given number of licenses to be surrendered in order to meet the overall clearing target to be met in the FCC's simulations. Data are the median, minimum, and maximum of the FCC simulated repacking scenario outcomes that assume 100% participation for the 120 MHz clearing target.

to meet a nationwide clearing target of 120 MHz.[105] The FCC performed 100 simulations that differ in the identity of the TV stations that do not relinquish their licenses and require repacking after the auction. We restrict attention to the 27 repacking simulations that assume full participation by UHF auction-eligible licenses. For our initial descriptive analysis, we label a DMA as a positive demand DMA if at the median across these simulations the FCC expected to acquire at least one license. Figure 2.3.2 shows that in many DMAs no licenses need to be acquired in expectation to meet this clearing target; hence, payouts from the reverse auction are expected to be concentrated in a small number of DMAs.

The various data sets use a number of different station identifiers. Below, we identify a station either by its call sign (e.g., WCAU, continuing with the NBC Philadelphia example from above) or by the FCC identifier of the facility from which it broadcasts (e.g., WCAU's broadcast facility ID is 63153).

[105]The FCC also conducted a similar simulation exercise to derive the likely number of stations to be cleared in each market to satisfy an 84 MHz clearing target. We focus on 120 MHz as it is closer to the Stage 1 clearing target the FCC used in the actual auction, which forms the basis for our simulations. See FCC's Public Notice Appendix, "Analysis of Potential Aggregate Interference," available at https://apps.fcc.gov/edocs_public/attachmatch/DA-14-677A2.pdf, accessed on March 10, 2016.

2.3.2 Descriptive evidence

Our data reveal significant ownership concentration, both within and across DMAs, consistent with the idea of "chains" of broadcasters. We focus on the 1,672 UHF licenses that the FCC denoted as eligible for the reverse auction in November 2014 and used in its own simulation exercises.[106] In 2012, the 1,672 UHF licenses are held by 514 unique owners. Of these 514 owners, 330 hold a single license, 60 hold two licenses, and 37 hold three licenses. The remaining 87 owners hold at least four licenses. Of the 204 DMAs with UHF broadcasters, 79 DMAs have only single-license owners while the remaining 125 DMAs have at least one owner that owns multiple licenses in the DMA.

Ownership concentration has traditionally been a concern of regulators. The FCC Local TV Ownership Rules permit ownership of up to two full-power commercial stations in the same DMA if either the two stations' service areas do not overlap or at least one of the two stations is not ranked among the top four stations in the DMA, based on the most recent audience market share, and at least eight independently owned full-power stations broadcast in the DMA in addition to any jointly owned stations.[107] These rules are oriented towards the business of running TV stations that primarily earn revenues through advertising and have a limited effect in preventing a firm from accumulating market power in the reverse auction. Waivers for the rules can be - and have been - granted for failing or financially distressed stations. The rules also do not apply to satellite, public, and low-power stations. However, these types of stations still hold licenses to 6 MHz of spectrum and are eligible for the auction.

Table 2.1 summarizes ownership patterns, first for all 204 DMAs and then for the 121 DMAs with positive demand under a clearing target of 120 MHz in the FCC's simulations. On average, a positive demand DMA has 9 broadcast TV licenses that are held by 7.15 owners. The number of multi-license owners is 1.36 on average for positive demand DMAs compared to 1.20 for all DMAs. The counts of ownership configurations in the bottom panel of the table reinforce that ownership is more concentrated in positive demand

[106]The FCC excludes 6 DMAs without UHF stations from its repacking simulations. These DMAs are Bangor, ME, Glendive, MT, Juneau, AK, Lafayette, IN, Mankato, MN, and Presque Isle, ME.
[107]See Title 47 of Code of Federal Regulations, Chapter I.C, Part 73. H, Section 73.3555.

DMAs. In 81 out of 121 or 67% of positive demand DMAs at least one firm owns multiple licenses compared to 125 out of 204 or 61% of all DMAs. Taken together, this suggests that multi-license ownership is a broad concern for the reverse auction.

In addition, news reports have pointed out that at least three private equity firms - LocusPoint Networks, NRJ TV, and OTA Broadcasting - spent almost $345 million acquiring 39 broadcast TV licenses from 2010 to early 2013, mostly from failing or insolvent stations in distress, and mostly low-power licenses (25 low-power versus 14 full-power licenses).[108,109] Since those press mentions and through the end of our data set in late 2013, an additional 4 license purchases by the three private equity firms were recorded, for a total of 43 license purchases. Of the 43 transactions, 25 are for licenses that cover the same DMA as that of another purchased license and may thus be indicative of attempts to accumulate market power in the reverse auction. Most of the stations are on the peripheries of major DMAs, ranging from Boston, MA, to Washington, DC, on the Eastern Seaboard and from Seattle, WA, to Los Angeles, CA, along the West Coast.

Table 2.1 illustrates that ownership is especially concentrated in the 18 DMAs in which the three private equity firms have been active (henceforth, private equity active DMAs). The number of multi-license owners is 2.67 on average for private equity active DMAs, and in 15 out of 18, or 83%, of these DMAs at least one firm owns multiple licenses. Moreover, at a 120 MHz clearing target, the FCC anticipated to purchase 6.67 licenses on average in private equity active DMAs compared to 3.42 licenses in positive demand DMAs. In line with the model in Section 2.2.2, the three private equity firms appear to focus on DMAs with robust demand for spectrum.

Section 2.2.2 discusses what types of TV stations are best suited for a supply reduction strategy. Table 2.2 summarizes the characteristics of TV stations transacted from 2003 to 2009 in the first column and those of TV stations transacted from 2010, when the incentive auction was proposed, to 2013 in the remaining columns. The latter are further separated

[108] See, e.g., http://www.tvnewscheck.com/article/65850/tv-spectrum-speculation-nears-345-million or http://current.org/2013/02/speculators-betting-big-on-fcc-tv-spectrum-auctions/, accessed on November 15, 2015.

[109] According to FCC filings, the Blackstone Group LP owns 99% of LocusPoint Networks. NRJ TV LLC is a media holding company funded through loans from Fortress Investment Group LLC according to a recent U.S. Securities and Exchange Commission filing. Lastly, OTA Broadcasting is a division of MSD Capital, L.P., which was formed to manage the capital of Dell Computer founder Michael Dell.

Table 2.1: Ownership concentration

	All DMAs ($n = 204$)	Positive demand DMAs ($n = 121$)	Private equity active DMAs ($n = 18$)
DMA averages:			
Number of licenses	8.20	9.00	15.94
Number of owners	6.51	7.15	12.22
Number of multi-license owners	1.20	1.36	2.67
Expected number of licenses demanded	2.03	3.42	6.67
Counts of DMAs with j multi-license owners:			
$j = 0$	79	40	3
$j = 1$	53	31	3
$j = 2$	42	30	2
$j = 3$	17	11	3
$j = 4+$	13	9	7

Notes: An observation is a DMA. Table displays average number of licenses, owners, and multi-license owners present in each DMA, together with average of median DMA-level FCC simulated demand at the 120 MHz clearing target. Positive demand DMAs are DMAs where the FCC expects to purchase at least one license (at median) under the 120 MHz clearing target. Private equity active DMAs are DMAs where one of the three private equity firms holds at least one license. Multi-license owners refers to firms owning more than one auction-eligible license within one DMA.

into transactions in the 121 DMAs with positive expected demand under a clearing target of 120 MHz and transactions involving the three private equity firms.

Consistent with the model, the three private equity firms have acquired TV stations with high broadcast volume. They also typically have low valuations, as evidenced by the low prices paid and the fact that very few stations are affiliated with a major network. Even compared to transactions in positive demand DMAs, the TV stations acquired by these firms are particularly high in population reach, interference count, and broadcast volume. Private equity firms also concentrate predominantly on DMAs expected to have positive demand and above average levels of demand: at a 120 MHz clearing target, 98% of their transactions fall into positive expected demand DMAs with average demand of 9 licenses compared to 60% positive demand DMAs with average demand of 3 licenses for all transactions between 2010 and 2013. We caution that most differences between the subsamples are not statistically significant in light of the small sample sizes and large variances of many of the outcomes. In Section 2.5.1, we return to the model implications and investigate the attributes of licenses that multi-license owners choose to strategically withhold from and bid into the auction.

2.4 Analysis

We first estimate the reservation value of a TV station going into the auction. Then we simulate the auction and compare the outcome under naive bidding with the outcome when we account for the ownership pattern in the data and allow multi-license owners to engage in strategic supply reduction.

2.4.1 Reservation values

The reservation value of TV station j in a particular DMA going into the reverse auction held at time t_0 is the greater of its cash flow value $V_{jt_0}^{CF}$ and its stick value $V_{jt_0}^{Stick}$:

$$v_{jt_0} = \max\left\{V_{jt_0}^{CF}, V_{jt_0}^{Stick}\right\}. \tag{2.4.1}$$

The industry standard for valuing a broadcast business as a going concern is to assess its cash flow CF_{jt_0} and scale it by a cash flow multiple $Multiple_{jt_0}^{CF}$. Hence, the cash flow value of the TV station is

$$V_{jt_0}^{CF} = Multiple_{jt_0}^{CF} \cdot CF_{jt_0}. \tag{2.4.2}$$

This is the price a TV station expects if it sells itself on the private market as a going concern. The stick value $V_{jt_0}^{Stick}$, on the other hand, reflects solely the value of the station's broadcast TV license and tower, not the ongoing business. This is the valuation typically used for unprofitable or non-commercial broadcast licenses. It is computed from the station's population reach $CoveragePop_{jt_0}$ and the stick multiple $Multiple_{jt_0}^{Stick}$ as

$$V_{jt_0}^{Stick} = Multiple_{jt_0}^{Stick} \cdot 6\,MHz \cdot CoveragePop_{jt_0}. \tag{2.4.3}$$

For example, a TV station reaching 100,000 people with a license for a 6 MHz block of spectrum and a stick multiple of \$0.30 per MHz per population (henceforth MHz-pop) is

Table 2.2: Broadcast TV license transactions

	2003-2009	2010-2013		
		All	In positive demand DMAs	Involving private equity firms
Number of Licenses Transacted	518	329	199	43
Average Transaction Price ($ million)	47.74	23.00	23.58	8.54
Average Interference Count	68.33	71.74	87.95	92.58
Average Interference-Free Population (million)	1.78	1.92	2.68	3.89
Average Broadcast Volume (thousand)	163.43	177.92	232.97	285.01
Percentage major network affiliation (%)	48.84	52.58	40.70	4.65
Positive Demand DMA (%, at 120 MHz)	58.11	60.49	100	97.67
Average Demand (120 MHz)	2.62	2.94	4.85	8.93

Notes: An observation is a transaction of an auction-eligible station. “Interference Count” is the number of stations with which a given station would interfere if they were located on adjacent channels. “Demand” is the median number of licenses expected to be purchased in the DMA according to the FCC repacking scenarios for the 120 MHz clearing target. “Positive Demand DMAs” refers to transactions that involve a license located in a DMA with a positive median demand level under the 120 MHz clearing target in FCC repacking simulations. For the 507 acquisitions that involve multiple stations, some of which may not be eligible for the auction, we use the average price per license as a transaction price.

worth \$180,000 based on its fixed assets alone.

While we observe a TV station's covered population, its cash flow is only available at various levels of aggregation in the NAB data. Moreover, we observe neither the cash flow multiple nor the stick multiple. Below we explain how we estimate these objects and infer the station's reservation value v_{jt_0}.

Cash flows. We specify a simple accounting model for cash flows.[110] The cash flow CF_{jt} of TV station j in a particular DMA in year t is

$$CF_{jt} = \alpha\left(X_{jt};\beta\right) AD_{jt} + RT\left(X_{jt};\gamma\right) - F\left(X_{jt};\delta\right) + \epsilon_{jt}, \tag{2.4.4}$$

where $\alpha\left(X_{jt};\beta\right) AD_{jt}$ is the contribution of advertising revenue to cash flow, $RT\left(X_{jt};\gamma\right)$ is non-broadcast revenue (including retransmission fees), $F\left(X_{jt};\delta\right)$ is fixed cost, and $\epsilon_{jt} \sim N\left(0,\sigma^2\right)$ is an idiosyncratic, inherently unobservable component of cash flow. Only advertising revenue AD_{jt} and station and market characteristics X_{jt} are directly observable in the BIA data. To estimate the remaining components of cash flow, we specify flexible functional forms of subsets of X_{jt} for $\alpha\left(X_{jt};\beta\right)$, $RT\left(X_{jt};\gamma\right)$, and $F\left(X_{jt};\delta\right)$ and estimate the parameters $\theta = (\beta,\gamma,\delta,\sigma)$ drawing on the aggregated data from NAB.

We proceed using a simulated minimum distance estimator as detailed in Appendix B.1.2. The parameters $\theta = (\beta,\gamma,\delta,\sigma)$, together with our functional form and distributional assumptions in equation 2.4.4, imply a distribution of the cash flow CF_{jt} of TV station j in a particular DMA in year t. We first draw a cash flow error term ϵ_{jt} for each TV station covered by the aggregated data from NAB. Then we match the moments of the predicted cash flow and non-broadcast revenue distributions to the moments reported by NAB for different sets of TV stations and DMAs. In particular, we match the mean, median, 25^{th} and 75^{th} percentiles of cash flow and the mean of non-broadcast revenue for each NAB table in each year, yielding a total of 3,313 moments.

The correlation between the moments of the predicted distributions at our estimates and the moments reported by NAB is 0.98 for cash flow and 0.84 for non-broadcast rev-

[110] In doing so, we follow the Well Fargo analyst report, "Broadcasting M&A 101 Our View of the Broadcast TV M&A Surge," J. Davis Herbert and Eric Fishel, June 26, 2013.

enue. The estimates indicate that major network affiliates are most profitable; that non-broadcast revenue has grown significantly in recent years; and that there are economies of scale in fixed cost. Appendix B.1.2.4 gives details on parameter estimates and fit measures.

Multiples. To estimate the multiples $Multiple_{jt}^{CF}$ and $Multiple_{jt}^{Stick}$, we begin with the 350 transactions for an individual broadcast TV station in the eleven years from 2003 to 2013 as recorded by BIA.[111] We extract 136 transactions based on cash flows and 201 transactions based on stick values between 2003 and 2013.[112] We infer the cash flow multiple and stick multiple from the transaction price using equations 2.4.2 and 2.4.3, respectively. Because the transacted stations may be a selected sample, we incorporate industry estimates of the range of the multiples. Using these estimates as priors, we estimate a Bayesian regression model to project multiples on station and market characteristics X_{jt}. This allows us to predict multiples for any TV station, not just those that were recently transacted. Appendix B.1.3 provides further details. The resulting posteriors, shown in Appendix Figure B.1.3, are a normal distribution for the cash flow multiple and a log-normal distribution for the stick multiple.

Reservation values. We use our estimates to infer a TV station's reservation value for its broadcast license going into the auction. Not all the 1,672 UHF stations that the FCC includes in its simulation exercise are covered in the aggregated data from NAB that we use to estimate the cash flow model in equation 2.4.4. The omissions are 386 low-power UHF stations and 290 non-commercial UHF stations. We therefore extrapolate from our estimates as follows. First, we assume that low-power stations are valued in the same way as full-power stations conditional on station and market characteristics X_{jt}. Second, we

[111]BIA records 877 transactions with full transaction prices, as opposed to station swaps, stock transfers, donations, etc. We focus on the 350 transactions involving a single license in order to evaluate the trading multiples as a function of station and market characteristics. Of these 350 transactions, 26 involve the three private equity firms.

[112]Because 2012 is the last year of availability for the NAB data, we cannot estimate a TV station's cash flow for 2013. To classify transactions, we proceed as follows: We first define a TV station to be a major network affiliate if it is affiliated with ABC, CBS, Fox, or NBC. We then classify a transaction as based on stick value if it is for a non-major network affiliate with a cash flow of less than $1 million. Regardless of network affiliation, we also classify a transaction as based on stick value if the TV station has a negative cash flow. Finally, we classify a transaction that would have implied a stick value greater than $4 per MHz-pop to be based on cash flow and a transaction that would have implied a cash flow multiple greater than 30 to be based on stick value. Together, we drop 13 transactions that do not fit the criteria.

assume that non-commercial stations are valued by stick value, consistent with industry practice.

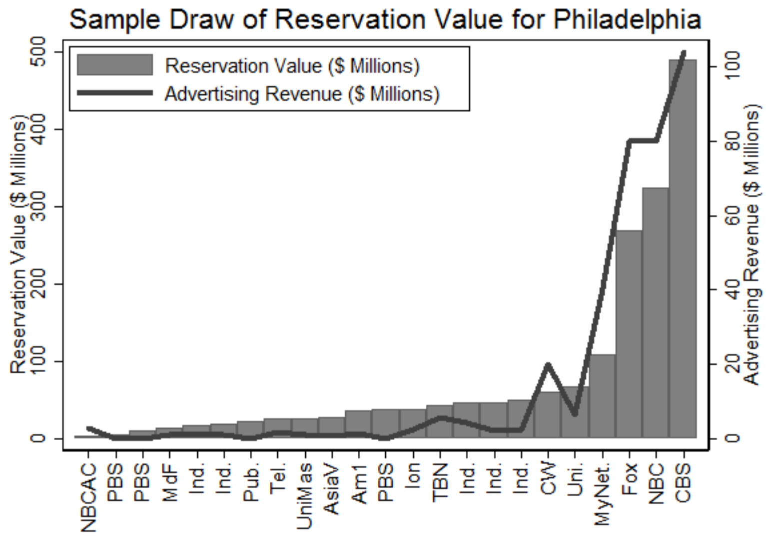
To infer the reservation value of TV station j in a particular DMA going into the reverse auction, we set $t_0 = 2012$ and draw from the estimated distribution of the cash flow error term ϵ_{jt_0} to get $\widehat{CF}_{jt_0}$. We draw from the respective posterior distributions of the multiples to get $\widehat{Multiple}_{jt_0}^{CF}$ and $\widehat{Multiple}_{jt_0}^{Stick}$. A commercial station's reservation value $\widehat{v}_{jt_0}$ is then the higher of the realized draws of its discounted broadcast cash flow value and its stick value as specified in equations 2.4.1-2.4.3; a non-commercial station's reservation value $\widehat{v}_{jt_0}$ is its stick value. Our estimates imply that the average TV station in our data has a cash flow value of \$42.2 million and a stick value of \$4.5 million. For 31.6% of TV stations, our estimates indicate that the reservation value is given by its stick value rather than its cash flow value.

Example. As an example, Figure 2.4.1 shows a sample draw from our estimated reservation values for auction-eligible UHF licenses in the Philadelphia, PA, DMA. The licenses are ordered by their reservation value, and we overlay each license's 2012 advertising revenues from the BIA dataset. In addition, we label each license by its network affiliation on the horizontal axis. It is immediately apparent that our estimated valuations correlate with advertising revenues and network affiliation. In addition, it is clear that reservation values can differ greatly across licenses.

Reservation values are not the same as naive bids in the auction, as pointed out in Section 2.2.1; since a license is shown a personalized price based on its broadcast volume, two licenses with the same reservation value may have very different clock prices at which they would withdraw from the reverse auction. Figure 2.4.2 plots, for the same draw of valuations as above, each license's broadcast volume against its reservation value. While there is some positive correlation, it is far from perfect, and the vertical cluster of licenses on the left indicates that a number of licenses with similar reservation values have broadcast volume levels that are multiples of one another.

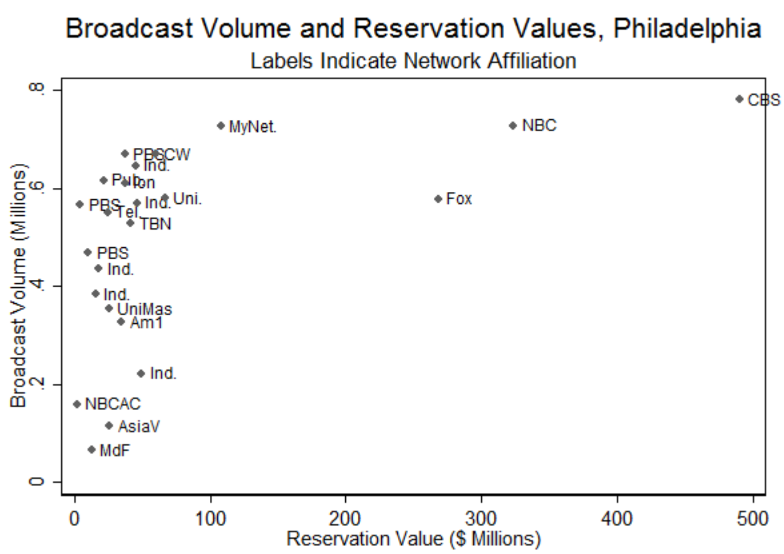
As a result of the variation in broadcast volumes, naive bids have a different distribution than reservation values. Figure 2.4.3 plots naive bids compared to reservation values.

Figure 2.4.1: Sample draw of valuations for Philadelphia licenses



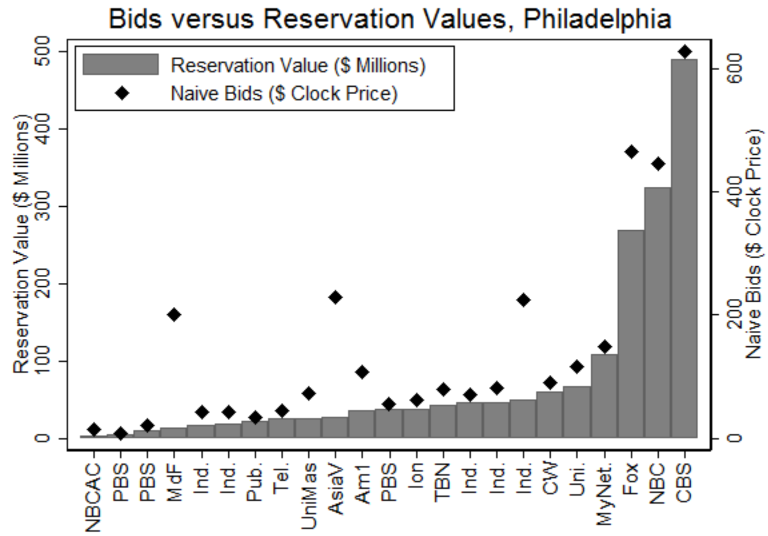
Notes: This chart shows reservation values (left axis), advertising revenues (right axis), and network affiliations (horizontal axis) for all auction-eligible UHF licenses in the Philadelphia DMA, for a single draw from our estimated distributions of valuations and multiples. The Philadelphia ABC affiliate broadcasts from the VHF spectrum and so is not included here.

Figure 2.4.2: Correlation of broadcast volume and reservation values



Notes: This scatterplot shows broadcast volume (left axis) against reservation values (horizontal axis) for all auction-eligible UHF licenses in the Philadelphia DMA, for a single draw from our estimated distributions of valuations and multiples.

Figure 2.4.3: Naive bids and reservation values for Philadelphia licenses



Notes: This chart shows reservation values (left axis), naive bids (right axis), and network affiliations (horizontal axis) for all auction-eligible UHF licenses in the Philadelphia DMA, for a single draw from our estimated distributions of valuations and multiples. ABC Philadelphia broadcasts from the VHF spectrum and is not included here.

Stations with a relatively low broadcast volume - that are shown a relatively low price in the auction - would withdraw from the auction at relatively higher clock prices. For example, the licenses affiliated with MdF and AsiaV have valuations similar to many other stations, but far lower broadcast volume, meaning that they are shown a relatively lower price than other stations for the same clock price in the auction. Consequently, they would withdraw from the auction at a higher clock price than a station with a similar valuation but higher broadcast volume, and so their naive bids in terms of the clock price are relatively high.

In other contexts, we would interpret the naive bids in Figure 2.4.3 as the elements of a supply curve. Here, however, that would ignore repacking constraints. Since the licenses are not perfectly interchangeable in repacking, the supply of licenses at a given point in the auction depends on what other licenses have already been surrendered. To illustrate, we return to the Philadelphia example in Section 2.5.2, where we show our simulation auction outcomes for this particular draw of reservation values.

While reservation values alone ignore repacking constraints that together make up supply, they are nevertheless useful in assessing the implications of the basic model in Section 2.2.2 descriptively, outside of full model simulations. We test whether pre-auction acquisitions are concentrated in vulnerable DMAs where the change in closing clock price is likely large due to supply reduction: we relate the propensity of a purchase by one of the three private equity firms in a probit regression to the increase in reservation values that results from removing from the auction either one or two licenses that would otherwise sell given the median number of licenses the FCC expects to be repurchased in the DMA (the term in parentheses on the left hand side of equation 2.2.2). Controlling for population and number of licenses, we find in unreported results that private equity firms were more likely to acquire licenses in DMAs where the distribution of reservation values is relatively steep around expected demand levels and strategic supply reduction is thus likely to be profitable.

2.4.2 Simulations

To quantify the impact of strategic supply reduction, we solve for all equilibria of a simplified localized version of the reverse auction. Since it is possible, albeit unlikely, that due to interference constraints, the withdrawal of a license in New York, NY, sets the price of a license in Los Angeles, CA, through a series of domino effects, the reverse auction is truly national. Checking the feasibility of repacking a particular station is an NP-complete computational problem that can easily take hours to run. Indeed, according to the FCC's own reports, Round 22 of the first stage of the reverse auction was delayed by one day since the FCC computing engine could not determine the necessary outcomes on time.[113] The computational challenge is further compounded here as we study the impact of strategic supply reduction by enumerating all auction equilibria and integrate over the distribution of estimated cash flows using Monte Carlo simulation.

As a step towards making the analysis computationally feasible, we reduce the size of the nationwide repacking problem by taking repacking constraints into account only at a regional level. Our approach is as follows: for a "focal" DMA $m \in \{1, 2, ..., 204\}$, we

[113] See `https://auctiondata.fcc.gov/public/projects/1000/reports/reverse_announcements`, accessed on December 9, 2016.

Table 2.3: Repacking "regions" of licenses

	Mean	Min	25%	Median	75%	Max
Number of Eligible UHF Licenses:						
By DMA	8.2	1	4	7	11	28
By Repacking Region	83.5	3	54	82.5	112.5	191
Distance Between Eligible UHF Licenses:						
By DMA	34.86	0	1.52	22.44	49.11	414.92
By Repacking Region	184.19	0	107.44	176.61	252.91	779.5

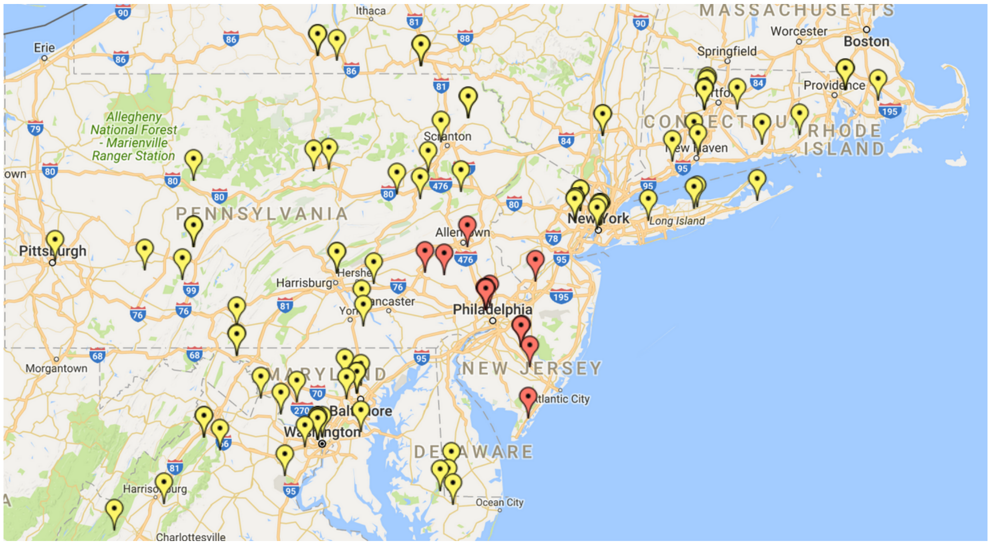
Notes: Statistics are for 204 DMAs with eligible licenses. Distances are in miles. All pairwise distances between stations' broadcast towers in the same DMA or Repacking Region are computed, and statistics are for the full sample of 17,438 pairwise distances within DMAs and 1,751,112 pairwise distances within Repacking Regions.

define the "region" of DMA m as the set of all DMAs in which at least one station has an interference constraint with at least one station in DMA m. We simulate the auction for a focal DMA m taking all stations in that DMA's region into account, even those stations that do not have any direct interference constraints with licenses in the focal DMA. The object of interest is the payouts in the focal DMA alone. Table 2.3 shows that a typical repacking region consists of a far larger set of licenses, and a far more distant set of licenses, than those in a DMA alone. Figure 2.4.4 shows the set of TV stations considered to be a part of the Philadelphia region for repacking purposes in our simulations.

We require a number of additional simplifications. First, to sidestep the forward auction and the multi-stage nature of the overall incentive auction, we fix the clearing target in the reverse auction to be the clearing target used by the FCC in the first round of the actual incentive auction, namely 126 MHz, the clearing target at the time we performed our simulations below. The assumption of a fixed clearing target does not imply that the FCC's demand for a given license is inelastic because there are many sets of licenses that can be repacked to meet the target.

Second, we assume full participation by auction-eligible stations. This is a conservative assumption: concerns have been raised about the possibility of some license holders, such as nonprofit or religious stations, being motivated by considerations beside profitability and choosing not to participate in the reverse auction, even though it would in likelihood be profitable for those stations to surrender their license. In the popular press, several commercial broadcasting chains have shown little interest in the auction, with the CEO of

Figure 2.4.4: The repacking "region" for Philadelphia, PA



Notes: Each pin represents the facility location of an auction-eligible UHF broadcast TV station. Stations in the Philadelphia DMA are denoted by red pins. Stations denoted by yellow pins are in other DMAs that have at least one station that has a repacking constraint with a Philadelphia station. A total of 126 broadcast stations are considered to be in the Philadelphia region for the purposes of our analysis.

Sinclair Broadcasting Group, which operates 74 stations, stating he "hasn't heard of any broadcaster who has said they have anything for sale."[114] We consider the effect of partial participation specifically by non-commercial station types in Section 2.5.5 and find that total payouts can easily double, even in the absence of strategic bidding, if participation is low among such stations.

Third, we focus exclusively on the full surrender of UHF licenses into the auction in line with the FCC's own simulation exercise. We thus set aside VHF licenses and, similarly, do not model a TV station's additional option of moving from a higher to a lower frequency band in order to free up more desirable parts of the spectrum. The price that a VHF station is offered for going off the air and the price that a UHF or a VHF station is offered to move channels are fixed fractions of the price that a UHF station is offered for going off the air. For this reason, the FCC's own auction simulations focus solely on the number of UHF licenses required to meet a given clearing target in its repacking simulations. We also do not consider the option of channel-sharing arrangements. Channel-sharing refers to a sit-

[114] "FCC can auction spectrum, but will broadcasters sell?", Joe Flint, The Los Angeles Times, February 17 2012.

uation where two TV stations enter into a private agreement to share a license to 6 MHz of spectrum and split the proceeds from selling the other license into the auction. It is unclear how attractive this option is, in part due to technological constraints.[115] Channel-sharing arrangements are likely to boost participation in the reverse auction, thereby effectively reducing the number of UHF licenses required from regular auction participants to meet a given clearing target.

Fourth, we assume complete information among the owners of TV stations in that they know the broadcast volume φ_{jt_0} and reservation value v_{jt_0} of every TV station. The assumption of complete information greatly simplifies the analysis relative to an asymmetric information formulation. The net effect of this assumption on prices and payouts is ambiguous as it makes it easier for firms to implement supply reduction strategies while also eliminating any possible ex-post regret for multi-license owners. Furthermore, while knowledge of reservation values is a strong assumption, many large broadcasters have engaged consultants to help them estimate valuations heading into the auction. In addition, industry groups of smaller broadcasters have helped their members to similarly estimate valuations in their DMAs. While there may be some residual uncertainty about reservation values, we conjecture that a model with incomplete information has similar but possibly less sharp implications as our current model. In particular, strategic supply reduction with incomplete information manifests itself by a multi-license owner raising her bid above a station's reservation value instead of outright withdrawing the station from the auction, which could have a smaller impact on closing prices.

Our baseline is the outcome of the reverse auction with naive bidding. Under naive bidding, we simply ignore the ownership pattern in the data and treat TV stations as independently owned. In line with the discussion in Section 2.2.2, TV license j remains in the auction until the base clock price drops below $\frac{v_j}{\varphi_j}$. To account for uncertainty in our estimates of reservation values, we construct reservation values by repeatedly drawing realizations of the cash flow error term ϵ_j and the multiples $Multiple_j^{CF}$ and $Multiple_j^{Stick}$.

[115] 6 MHz of spectrum is insufficient for two high-definition video streams. The FCC has piloted a channel-sharing arrangement in Los Angeles, CA, showing that it is technologically feasible for one high-definition video stream and one or more standard-definition video streams to share 6 MHz of spectrum. 6 MHz of spectrum may no longer suffice if a TV station eventually transitions from a high-definition to a ultra-high-definition (4K) video stream.

Given a particular set of implied reservation values, we proceed as follows to compute the naive bidding outcome (a formal coding of the algorithm is presented in Appendix B.2):

1. Any participating license in a focal DMA's repacking region whose reservation value is above its starting price is repacked into the available spectrum. If it is not possible to accommodate all of these licenses in the available channel space (UHF channels 14-30 for the 126 MHz clearing target), then the auction is considered to have failed.[116]

2. All remaining licenses are considered "active" in the auction. They are sorted in descending order by their reservation price and are denoted by $j \in \{1, ..., J\}$, where $j = 1$ is the station with the highest reservation value in terms of the base clock price. As the base clock price falls, licenses withdraw one by one. Each time a license j withdraws, we verify that each of the remaining "active" stations in $\{j + 1, ..., J\}$ could still feasibly be repacked if it were to withdraw in the next round of the auction. If station $k \in \{j + 1, ..., J\}$ can no longer be repacked due to j having withdrawn, it is no longer "active", and instead is "frozen" at the current base clock price. The payout to station k is therefore set by station j, who made it no longer feasibly to repack k.

3. The base clock price drops and the process continues until all licenses are either repacked or frozen.

Unless otherwise noted, we report average auction outcomes based on 100 draws from the distribution of reservation values below. With 204 DMAs, each with its own repacking region, and 100 simulation draws, we determine payouts for 20,400 simulated auctions.

We next account for the ownership pattern in the data. We assume that a multi-license owner engages in strategic supply reduction by withholding one or more of its TV stations from the reverse auction at the outset of the auction. Hence, an owner of n_o TV stations has $2^{n_o} - 1$ possible combinations of licenses to consider bidding into auction, and if there are N_o multi-license owners in the repacking region, there are $\prod_{o=1}^{N_o}(2^{n_o} - 1)$ strategy profiles

[116]In practice, it is very rare for the auction to fail. The fail rate for our main results is under 0.7% of simulations, and those cases involve many strategic withdrawals of licenses. It is unclear although perhaps unlikely that massive withdrawals of licenses could constitute equilibria.

across firms, assuming again full auction participation by single-license owners. Given computational constraints, this creates an infeasibly large strategy space for at least some repacking regions; there are 17 owners with 20 or more TV stations and one owner with 93 TV stations in the data. Hence, we impose a restriction on the strategy space, by limiting strategic bidding to licenses in the focal DMA, rather than all licenses owned by a given firm across the repacking region. This reduces both N_o to only multi-license owners and n_o to only licenses in the focal DMA. For example, a multi-license owner of a given license in the Philadelphia, PA, DMA considers strategic bidding only for that and any other licenses in the Philadelphia, PA, DMA, but not for those held in, say, the Harrisburg, PA, DMA, even though we consider such stations in repacking.

For each of the resulting strategy profiles, we re-run the above algorithm to determine the payouts associated with this strategy profile, assuming that firms bid their valuations for the set of licenses they consider bidding into the auction under the particular profile under consideration. We use these payouts to determine whether a particular strategy profile is an equilibrium outcome of each multi-license owner's strategic choice of licenses to bid into the auction in the focal DMA by verifying the absence of any unilaterally profitable deviations from the strategy. With no or one multi-license owner in a DMA, there is a unique pure-strategy equilibrium set of licenses to bid into the auction. With more than one multi-license owner, there may be multiple equilibria; we enumerate all of them.

To be able to accommodate the large dimensional strategy space under strategic bidding, we make one final simplification in estimating auction outcomes: we do not assert feasibility or compute payouts for stations outside the focal DMA; we instead assume they exit the auction at the same base clock price when they would have exited under naive bidding over the entire repacking region. We assert feasibility only for active licenses in the focal DMA when these licenses exit. Our main results compare naive and strategic outcomes under this simplification. In Section 2.5.6 we show that in select major markets that we tested, this simplification introduced an error of less than 0.20% in the strategic payouts we computed relative to asserting feasibility and computing payouts for all non-focal DMA stations. It had virtually no effect on naive payouts across all markets. At the same time, computational speed increased by a factor of 15 to 20. To further illustrate how we complete our simulations, Appendix Figure B.2.1 shows the reverse auction process in detail for two alternative strategy profiles of multi-license owners in Philadelphia, PA, for a given draw of simulated valuations. Appendix Figure B.2.2 graphically depicts the auction progress under the first strategy profile by plotting the locations of licenses that withdraw and/or are frozen.

Restricting strategic bidding to the focal DMA, the Pittsburgh, PA, DMA has 4,601 strategy profiles to consider, the largest number of any DMA in our data. The total sum of strategy profiles across the 204 DMAs is 17,316, implying solving a total number of 1,731,600 simulated auctions.[117] We recognize that restricting strategic bidding to the focal DMA is likely to create a lower bound on the effect of strategic bidding by multi-license owners. In Section 2.5.4 we return to the Philadelphia case study and explicitly allow a particular multi-license owner to strategically bid an additional license from a neighboring DMA; we demonstrate large effects.

Despite the simplifications, our analysis is near the bound of what can be computed in a reasonable amount of time. The over two million simulations in this paper were completed on a combination of the Wharton High-Performance Computing Cluster and the Amazon EC2 Cloud Computing Platform over a period of just under one month during the summer of 2016, typically utilizing over 500 dedicated cores (without hyperthreading) simultaneously.

2.5 Results

2.5.1 Naive versus strategic bidding

We compare the outcome of the reverse auction under naive bidding with the outcome under strategic bidding when we account for the ownership patterns in the data. Table 2.4 shows the main results: payouts to broadcast TV licenses holders under naive and strategic bidding, broken down into different subsets of DMAs. As there may be multiple equilibria when more than one firm in a DMA controls multiple licenses, we present moments of the

[117] In contrast, if we allowed multi-license owners to bid strategically within each repacking region, the number of strategy profiles would be a completely unmanageable $6.89 \cdot 10^{44}$.

resulting payout distributions across equilibria. In a given DMA, we record the minimum, mean, median, and maximum payout level across equilibria for strategic bidding for each simulated set of reservation values. We then average each of these four moments across simulation draws. Table 2.4 then reports the sum of the payouts across particular groupings of DMAs for each moment. In the following, we focus on comparing payouts under naive bidding to the average equilibrium payout under strategic bidding.[118]

The first thing to remark upon is that strategic bidding generally leads to higher payouts in the reverse auction. This need not be the case: there exist equilibria in DMAs where total payouts are lower due to strategic behavior, although higher for the firm engaged in supply reduction.[119] At the mean payout level across equilibria, strategic bidding is found to increase total payouts from \$16.999 billion to \$20.740 billion, an increase of 22%. Moving down the table, we see that nearly 99% of payout increases are concentrated in DMAs with two or more multi-license owners. Further, we see that DMAs in which private equity firms are active are an important source of increased payouts. Those 18 DMAs are large, accounting for 70.4% of payouts in the base case; however, they account for 95.7% of the total increase in payouts at the mean due to strategic bidding. In addition, we see that there is some skew to the distribution, with some exceptionally large payout increases at the high end of the distribution for certain strategic equilibria. For example, in the 18 private equity active DMAs, the average *maximum* strategic equilibrium payout across equilibria is 51% greater than the payout under the naive base case. Panel B shows the large spillover effect of strategic bidding: single-license owners, who have no incentive to reduce supply, see payout increases of 19.1% at the mean strategic bidding equilibrium. While multi-licenses owners benefit more in percentage terms, the level of payout increases is actually greater among single-license owners.

Table 2.4 masks significant heterogeneity in the impact of strategic supply reduction. Even under naive bidding, there are an average of just over 90 DMAs that see payouts of zero across simulations. Under strategic bidding, an average of 23 DMAs, or just 11% of

[118] In the rare cases where there is no pure strategy equilibrium, we assume firms revert to naive bidding.

[119] A firm can selfishly increase its own profits through strategic bidding, but by withdrawing a license from the auction, it also affects which other licenses are sold in. In some occasions, more expensive licenses are substituted with less expensive ones due to strategic bidding by others.

Table 2.4: Total payouts to broadcast TV license holders by DMA type

		Naive Bidding	Strategic Bidding				Payout Increase
Panel A: DMA Payouts ($B) In:	# DMAs	Base	Mean	Min	Median	Max	At Mean
All DMAs	204	16.999	20.740	18.519	20.554	23.301	22.0%
		(0.751)	(0.898)	(0.876)	(1.235)	(1.686)	
DMAs with Multi-License Owners	125	15.611	19.352	17.130	19.166	21.193	24.0%
		(0.718)	(0.875)	(0.864)	(1.220)	(1.666)	
DMAs with 2+ Multi-License Owners	72	14.410	18.108	15.892	17.922	20.663	25.7%
		(0.711)	(0.867)	(0.860)	(1.209)	(1.663)	
Private Equity Active DMAs	18	11.978	15.558	13.382	15.372	18.070	29.9%
		(0.704)	(0.853)	(0.861)	(1.203)	(1.649)	
Panel B: Owner Payouts ($B) Among:							
Single-License Owners (within DMA)	-	12.009	14.309	12.113	14.201	16.556	19.1%
		(0.505)	(0.679)	(1.564)	(0.776)	(1.487)	
Multi-License Owners (within DMA)	-	4.981	6.207	(5.496)	6.222	6.859	24.6%
		(0.363)	(0.367)	(0.584)	(0.427)	(0.440)	

Notes: In the case of naive bidding, payouts are averaged over 100 simulated sets of reservation values. In the case of strategic bidding, the moments of the payout distribution are calculated across all possible equilibria for a given simulated auction and then averaged over 100 auction simulations. Standard deviations of the moments across the 100 simulations are in parentheses. The mean, min, median, and max payouts are computed within DMA, and aggregated to the DMA grouping under consideration in Panel A. In Panel B, payouts are first aggregated among the sets of single-license and multi-licenses owners. The sum of payouts among the two sets of owners does not always equal the value for "All DMAs" in Panel A since statistics are taken among those two groups before averaging and ratios between the two groups differ across markets.

Table 2.5: Auction surplus to private equity firms

	# Stations	Total Purchase Price ($M)	Total Profits Naive ($M)	Strategic ($M)
NRJ	14	235.51	690.79	1,473.04
			(81.31)	(140.56)
OTA	20	77.05	446.79	1,237.10
			(44.94)	(144.61)
LocusPoint	9	54.75	200.65	400.44
			(24.64)	(62.22)

Notes: Profits are averages over 100 auction simulations, and in the case of strategic bidding, averages over equilibria, summed across all DMAs where the firms hold licenses. Total profits are defined as total proceeds from the auction for stations that sell, plus the reservation values of stations that do not sell, less the purchase prices paid by the firms for the stations.

DMAs, show payout increases from strategic bidding.[120]

Our simulations allow us to determine the profitability of strategic supply reduction for the three private equity firms. Table 2.5 presents the results. As discussed in Section 2.3.2 these firms have acquired TV stations with high broadcast volume but low valuations as going concerns. Even under naive bidding, the firms stand to profit as their payouts in the reverse auction plus the value of any unsold licenses substantially exceed their total acquisition costs.

The simulations bear out the implication of the model in Section 2.2.2 that a multi-license owner sells stations with higher broadcast volume into the auction but withholds stations with lower broadcast volume. Table 2.6 shows the average broadcast volume and reservation value of the licenses owners decide to keep and sell under naive and strategic bidding simulations. Comparing attributes of unsold stations under naive and strategic bidding, we see that owners on average keep stations of lower value and sell stations of higher value under strategic bidding. This is counter-intuitive, until one sees that the stations kept have lower broadcast volume, and so are less valuable in the auction, while those sold have higher broadcast volume, making them particularly attractive to sell into the auction. In general, we see that strategic behavior leads to a higher amount of broadcast volume being acquired to reach the clearing target, increasing the total payouts.

There are two potential efficiency losses from strategic bidding by multi-license own-

[120]On average, one DMA per simulation will see a decrease in payouts from strategic bidding.

Table 2.6: Station characteristics for auction licenses

	Bidding Strategy			
	Naive		Strategic	
License Averages	Unsold	Sold	Unsold	Sold
Broadcast Volume (000s)	156.893	214.051	100.554	343.060
	(2.920)	(8.853)	(4.544)	(18.041)
Reservation Value ($M)	44.959	8.845	32.742	20.072
	(1.254)	(0.827)	(3.750)	(3.929)

Notes: Statistics describe averages of station characteristics for selling and non-selling licenses across all 204 DMAs and 100 auction simulations per DMA, and in the case of strategic bidding, first averaged across equilibria for each auction simulation. Standard errors based on 100 simulation draws are in parentheses.

ers: first, such behavior changes the set and number of licenses surrendered in the auction; second, such behavior risks reducing the amount of spectrum that is repurposed in the auction.

The set of stations surrendered under strategic bidding differs from that under naive bidding, as implied by Table 2.6. Therefore, strategic bidding by multi-license owners distorts the set of licenses that are surrendered from a socially optimal set to a different set, allowing perhaps lower-value licenses to remain on-air while higher-value licenses are surrendered. Such efficiency losses are likely significant; our simulations indicate that the average license that remains unsold under naive bidding has a reservation value of less than half of the reservation value of an unsold license in the average strategic equilibrium.

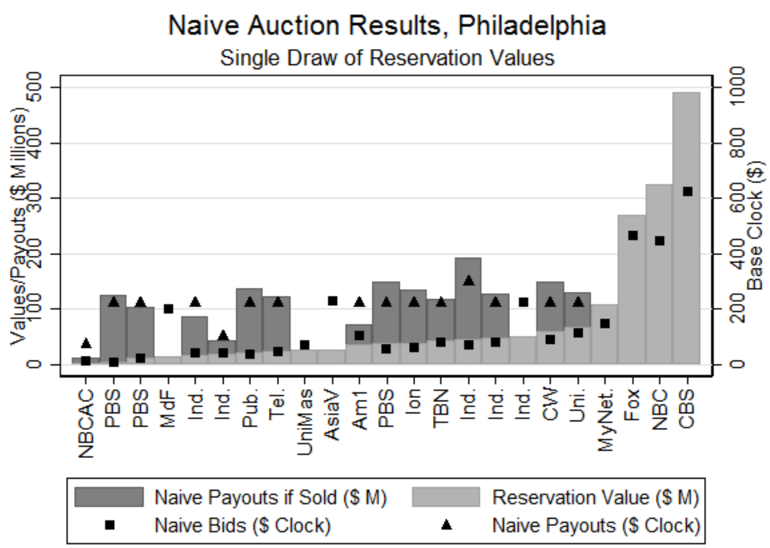
In addition, there is a risk that strategic bidding by multi-license owners could cause a stage of the auction to fail, leading to a reduction of the overall clearing target. As the forward auction is outside of the scope of this paper, we cannot address this prospect numerically, although we mention it as a possibility.

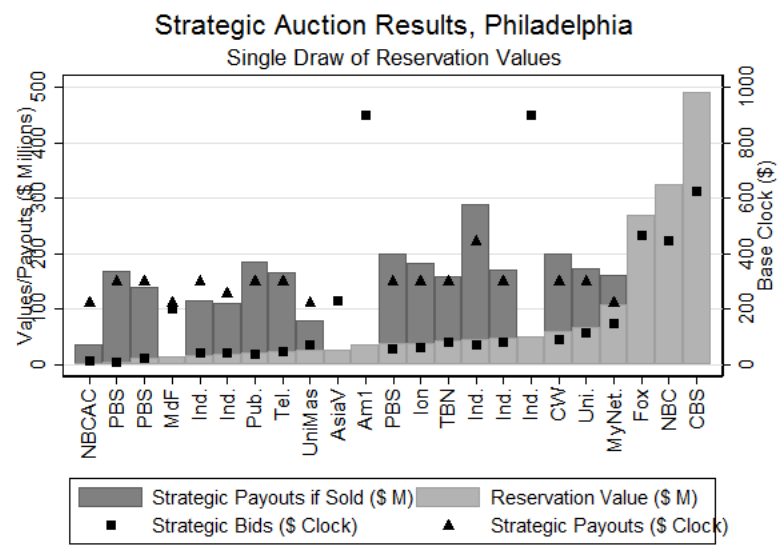
2.5.2 Case study: Philadelphia, PA

In this section we illustrate the impact of strategic bidding for a particular realization of valuations for stations in the Philadelphia, PA repacking region. Figure 2.5.1 shows graphs of outcomes first under naive, and then under strategic bidding. Both charts show all eligible UHF licenses in the focal Philadelphia, PA, DMA ordered by their simulated reservation values in light gray on the left y-axis. We further display on the second y-axis payouts in terms of base clock prices as triangles; recall that the payout in terms of base clock maps to license-specific payouts on the left y-axis through the broadcast volume of each license. Lastly, we display on the second y-axis the firms' bids in terms of base clock prices as squares. These correspond to the stations' transformed reservation prices and indicate the base clock price at which the station would drop out of the auction, if it is not frozen at a higher base clock price. The first image shows the outcome under naive bidding. In the second, we present a strategic equilibrium of this simulation where two multi-license owners each are able to increase their own auction surplus by withdrawing one of their licenses from the auction (these licenses are identified by strategic bids of $900, which is the starting base clock price of the auction). Note that the withdrawal of two licenses increases payouts for several other licenses, implying a positive spillover to those licensees, but the two firms find it individually rational to withdraw licenses solely based on their own profit motives. Total payouts in this DMA go from nearly $1.7 billion for 15 licenses under naive bidding to over $2.5 billion for 16 licenses in the strategic equilibrium presented here. This example thus clearly indicates the inefficiencies associated with strategic bidding: not only is it possible that the mix of stations that sells in the auction is distorted (e.g., UniMas does not sell under naive bidding but does under strategic bidding, while MyNetwork TV sells under strategic bidding but not under naive), but strategic bidding may also result in the government needing to purchase a larger number of licenses to reach its predetermined clearing target given the interference patterns between the stations that are now not selling in the auction.

Appendix Figure B.2.1 shows the precise details of these two simulations, with the naive bidding simulation in the left column and strategic on the right. The figure shows the order in which licenses withdraw from the auction and the base clock price at which they withdraw, labeling licenses by their FCC Facility ID number, and including their network affiliation in parentheses. Importantly, the figure emphasizes that we repack a large region around the DMA, by highlighting the few licenses in the Philadelphia, PA, DMA in bold. The left column of the figure shows how major network affiliates in large markets withdraw from the auction immediately, as starting base clock prices are too low. For example, License #9610 is CBS New York, and it withdraws immediately from the auction.

Figure 2.5.1: Sample outcomes under naive and strategic bidding in Philadelphia





Notes: Data are for a single simulation draw of reservation values. Under strategic bidding, two stations withdraw from the auction (effectively having naive bids above the opening price). This raises payouts for many other stations in this particular simulation. The strategic outcome shown here is an equilibrium for this simulation, as the two private equity firms that withdraw licenses are able to increase their profits from this behavior. From the left, the stations owned by NRJ are numbers 11 and 16 (Am1 and Ind affiliations), while the ones owned by LocusPoint are in positions 1 and 17 (NBCAC and Ind affiliations). The simulation takes into account the repacking of all 126 licenses in the Philadelphia repacking region.

Reading down the column, we see how licenses withdraw and how all remaining licenses are then checked for feasibility in repacking. If a license can no longer be feasibly repacked due to a withdrawal, its payout is frozen based on the current clock price. We only list freezes for licenses in the focal Philadelphia, PA, DMA.

The right column of the figure shows how the strategic withdrawal of two licenses changes the prices at which other licenses become frozen in the auction. For example, license #39884 becomes frozen by the withdrawal of license #73333 at a clock price of $298.15 in the naive bidding simulation, but in the strategic simulation, it is frozen earlier by the withdrawal of license #63153 at a clock price of $444.89, increasing its payout by 49% even though its owner is not the firm withdrawing licenses. The two multi-license owners who own pairs (#74464, #55305) and (#61111, #72278), find withdrawing licenses a profitable strategy individually.[121]

2.5.3 Partial remedy

We have so far shown that strategic supply reduction may lead to increased payouts and efficiency losses in the reverse auction. We next propose a change to the auction rules and show how it limits the potential for rent-seeking. The model in Section 2.2.2 shows that strategically reducing supply is more likely to be profitable if the increase in the closing base clock price from withholding a license can be leveraged by selling a license with high broadcast volume into the auction. Our proposal aims to weaken this mechanism by limiting the strategy space of multi-license owners. In particular, we stipulate that a multi-license owner must first withdraw her highest broadcast volume license. Once that has been withdrawn from the reverse auction, the owner may withdraw her second highest broadcast volume license, and so on.

Table 2.7 shows how the rule change affects our main results. The increase in payouts from strategic bidding are 80% less than in Table 2.4 at the mean. Interestingly, under

[121] NRJ who withdraws #74464 foregoes a surplus on that license of $37.5 million = ($221.79-$106.71)*326127.3, which is the naive clock payout less reserve value measured in terms of base clock price times broadcast volume, in exchange for increasing the payout in terms of clock price for license #55305 from $221.79 to $298.15, which when multiplied by a broadcast volume of 570169.3 is $43.5 million. LocusPoint that withdraws #72278 loses nothing on that license as it does not sell under naive bidding, but its withdrawal increases the payout to license #61111 by $23.8 million, by raising the freezing clock price from $72.66 to $221.79, given a broadcast volume of 159417.7.

this alternative policy, the outcome under strategic bidding presented in Figure 2.5.1 for the Philadelphia, PA, DMA would no longer involve feasible strategies since there, both firms withhold their lower broadcast volume stations. This would thus cease to be an equilibrium outcome.

Efficiency requires that, for otherwise identical licenses, the licenses with the lowest reservation values are sold into the auction. In the spirit of the literature on regulation where effort is not verifiable (Laffont and Tirole, 1986), the rule change leverages the fact that broadcast volume, unlike cash flows, is observed and contractible. Our estimates imply that broadcast volume is positively correlated with reservation value: averaged across simulation runs the correlation is 0.47 overall and 0.44 within DMA.[122] The rule change therefore mitigates efficiency losses by requiring that licenses with higher broadcast volumes, and likely higher reservation values, are withdrawn first from the reverse auction.

The rule change has two potential shortcomings, aside from legal considerations. First, a multi-license owner may be able to circumvent the rule change by selectively entering her licenses into the reverse auction in the first place. However, the rules of the auction may be further rewritten to compel a multi-license owner to either participate with all her licenses in the auction or not at all. Second, and perhaps more importantly, forcing lower broadcast volume licenses to sell before higher broadcast volume licenses may complicate the repacking process to the extent that licenses with higher broadcast volume and potentially also higher interference count may no longer sell and have to be repacked.

2.5.4 Multi-market strategies

Strategic bidding may extend beyond market borders if multi-license owners withhold a license in a DMA from the reverse auction to drive up the closing base clock price in a neighboring DMA where they also own a license. Here, we illustrate how such strategies may work continuing with the Philadelphia, PA, case study. As mentioned above, it is not computationally feasible to consider all multi-market strategies in a repacking region, given the prevalence of multi-license ownership and the need to simulate over uncertainty in reservation values.

[122] Within DMA correlations are averaged over 184 DMAs that have three or more stations.

Table 2.7: Total payments by DMA type under rule change

Payouts ($B) under:		Naive Bidding	Strategic Bidding				Payout Increase
	# DMAs	Base	Mean	Min	Median	Max	At Mean
All DMAs	204	16.999	17.735	17.097	17.826	18.343	4.3%
		(0.751)	(0.753)	(0.729)	(0.840)	(0.834)	
Multi-License Owners	125	15.611	16.347	15.709	16.437	16.955	4.7%
		(0.718)	(0.715)	(0.694)	(0.806)	(0.798)	
2+ Multi-License Owners	72	14.410	15.129	14.495	15.220	15.734	5.0%
		(0.711)	(0.711)	(0.690)	(0.797)	(0.799)	
Private Equity Active	18	11.978	12.656	12.034	12.747	13.249	5.7%
		(0.704)	(0.698)	(0.688)	(0.779)	(0.790)	

Notes: Data are averages over 100 auction simulations. Standard deviations across 100 simulations are in parentheses. Auction mechanism assumes that multi-license holders withdraw license with highest broadcast volume first. See Notes to Table 2.4.

In late 2012 NRJ purchased WGCB-TV in the Harrisburg, PA, DMA for $9 million. While NRJ owns no other TV stations in the Harrisburg, PA, DMA the firm had previously purchased WTVE and WPHY in the Philadelphia, PA, DMA in late 2011 and early 2012 for $30.4M and $3.5M respectively. WGCB-TV has a very high interference count and may interfere with 161 stations in the repacking process. A closer look shows that WGCB-TV is not actually located in Harrisburg, PA, but in Red Lion, PA, towards both the Philadelphia, PA, and Baltimore, MD, DMAs. Figure 2.5.2 shows how the broadcast contours of WGCB-TV and WTVE overlap. Hence, if NRJ withdraws WGCB-TV from the reverse auction, this may sufficiently complicate the repacking process to increase demand in the Philadelphia, PA, DMA, and potentially other DMAs.

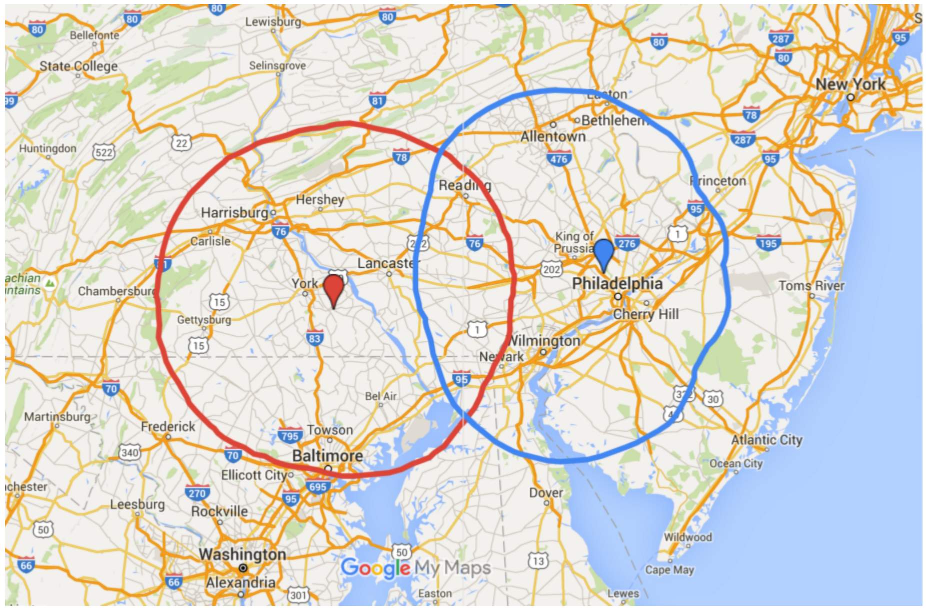
Table 2.8 shows the effect on payouts in the Philadelphia, PA, focal DMA of allowing NRJ to bid its license to WGCB-TV strategically in concert with its licenses in the Philadelphia, PA, DMA. We continue to assume that all remaining multi-license owners in the Philadelphia, PA, DMA, to the extent that they own licenses outside that focal market, consider strategic bidding for their Philadelphia stations only. From a practical perspective, this relocates the license of WGCB-TV from the Harrisburg DMA to the Philadelphia DMA for ownership purposes without actually relocating the broadcast tower. This change alone increases the number of strategy profiles to consider from 729 to 1701 for each simulation draw.

The first row in Table 2.8 shows Philadelphia's results from our main results in Table 2.4. The first and second rows show that the partial remedy proposed in Section 2.5.3 is very effective in this DMA. The third row shows that total payouts can increase dramatically in this DMA if NRJ bids its Harrisburg license strategically. The fourth row shows that the partial remedy is no longer particularly effective in this situation, suggesting that our above estimate is likely an upper bound of the true degree to which the remedy would restrain the extent to which strategic bidding can influence payouts.

Note that all payouts in the table exclude any payout to WGCB-TV, so they are directly comparable across scenarios. In particular, one can see that having an additional strategic lever is valuable in the tail end of equilibrium payouts in this case study.

More generally, cross-market ownership can be seen as positive for the auction, as sell-

Figure 2.5.2: Multi-market strategy in the Mid-Atlantic



Notes: Map plots reception contours from FCC TV Query database for WGCB-TV (Harrisburg) in red and WTVE (Philadelphia) in blue. Contour plots reflect reception of DTV signals from the broadcast towers. Image via Google Maps.

ing licenses should be complementary across markets if it allows a larger clearing target to be attained. However, in this context, it has the potential to be negative if withdrawing WGCB-TV from the auction sufficiently complicates repacking in Philadelphia: NRJ may find it worthwhile to withdraw WGCB-TV from the auction if either the proceeds from selling WGCB-TV are low, or if NRJ's increase in profits from its Philadelphia licenses is large.

2.5.5 Partial participation

So far, we have conservatively assumed full participation on behalf of all eligible UHF licensees. We now consider what payouts may be under reduced participation. In particular, the results in Table 2.9 show total payouts when we assume that across simulation draws, both a) none of the religious stations participate, and b) a random subset of 50% of non-commercial stations does not participate, for a total of 253 non-participating licenses in each simulation draw. The results highlight how important full participation is to the auction's success. First, limited participation makes the failure of the auction significantly

Table 2.8: Case study: Philadelphia PA DMA multi-market strategy

Payouts ($B) under:	Naive Bidding	Strategic Bidding				Payout Increase
	Base	Mean	Min	Median	Max	At Mean
Philadelphia, Base Case	2.172	2.481	2.320	2.499	2.614	14.2%
	(0.241)	(0.466)	(0.450)	(0.548)	(0.637)	
Philadelphia, Base Case, Remedy	2.172	2.215	2.189	2.223	2.231	2.0%
	(0.241)	(0.294)	(0.244)	(0.328)	(0.333)	
Philadelphia, NRJ Strategic with WGCB-TV	2.172	4.065	2.141	3.995	6.190	87.2%
	(0.241)	(0.201)	(0.322)	(0.258)	(0.525)	
Philadelphia, NRJ Strategic with WGCB-TV, Remedy	2.172	3.803	2.172	3.824	5.614	75.1%
	(0.241)	(0.291	(0.241)	(0.318)	(0.426)	

Notes: See notes to Table 2.4. We depict total payouts under NRJ bidding all of its licenses in the Philadelphia repacking region strategically in the third and fourth rows. Payouts in rows 3 and 4 do not include any payments to WGCB-TV.

Table 2.9: Naive outcomes under reduced participation

Payouts ($B) under:	# DMAs	Naive Bidding Full	Naive Bidding Limited	Payout Increase from Limited Participation
All DMAs, No-Fail Simulations	204	16.999	33.869	99.2%
		(0.751)	(2.005)	
No-Fail Positive Payout DMAs	45	5.545	10.341	86.5%
		(0.265)	(0.867)	

Notes: Data are averages over 100 auction simulations. Standard deviations across 100 simulations are in parentheses. Limited participation means that no religious stations participate (108 licenses) and that a random subset of 50% non-commercial stations do not participate (145 additional licenses). The "All DMAs" row is conditional on auction non-failure. "No-Fail Positive Payout DMAs" is a set of markets that always see positive payouts, yet never see the auction fail across 100 simulations.

more likely, as all of the licenses in the non-participating group, and any other stations that were identified as participating but that have reservation values in terms of the base clock price above the auction's starting point, need to be repacked for the auction to even be able to start. This means that in 4.2% of all 20,400 naive bidding simulations under reduced participation, the auction immediately concluded as a failure, which never occurred under full participation. It is immediately clear that low participation - before any strategic withdrawal of licenses - can have dramatic effects on payouts, nearly doubling them. The first row of Table 2.9 shows that, conditional on the auction not failing, payouts would effectively double due to non-participation by religious and some non-commercial licensees. To avoid comparing subsets of simulation draws from different markets, the second row limits the sample to DMAs that always see strictly positive payouts and yet never see auction failure under lowered participation; the effect is nearly as dramatic, with an 86.5% increase in payouts.

The fact that the mechanism is sensitive to participation leads us to a second recommendation. The likely reason many small broadcasters would choose to remain on-air relates to "must-carry" provisions in FCC regulations. While the regulations are fairly complex, they stipulate that a large cable operator must carry any and all local broadcast stations, unless such a station has opted-out and requested retransmission fees.[123] There-

[123] Any cable operator offering more than 12 channels must set aside one-third of their channel capacity for local commercial broadcasters. Any cable operator offering more than 36 channels must carry all non-commercial and educational broadcasters.

fore, continuing to broadcast guarantees that many small licenses will be carried on cable, which greatly broadens their reach or potential advertising audience. One simple measure to increase participation, therefore, would be to allow broadcasters to relinquish their spectrum licenses but retain their must-carry status, so that they can continue to operate as businesses and reach viewers through cable systems.

2.5.6 Robustness of the repacking approach

Here, we consider the effect of how our simulations limit repacking on outcomes. Our main results for both naive and strategic bidding rely on simulations that do not assert the feasibility of repacking of licenses outside the focal DMA, but instead assume that these licenses withdraw or are frozen at the same time as when we assert feasibility for all stations in the entire region under naive bidding. A more complete but computationally intensive analysis of strategic bidding would assert feasibility for all regional licenses to determine at which points licenses withdraw or are frozen in the auction under each strategy profile, in effect treating all licenses in the region as if they were in the focal DMA. We denote the full analysis of bidding by all stations in the repacking region as "robust repacking" in the following. To assess the implication of our simplification, we perform two exercises. First, we compare our naive repacking results to naive robust repacking results for all DMA markets; second, we compare our strategic repacking results to strategic robust repacking results for two important DMAs.

We first compare naive bidding outcomes when we robustly assert feasibility for all licenses in the region to the above outcomes when assuming that licenses outside the focal DMA are frozen or repacked at the same point in the auction as occurs under robust repacking, without explicitly asserting feasibility. The latter simplified repacking procedure greatly decreases the computational burden by more than an order of magnitude. Table 2.10 shows the results for all DMAs. The assumption has a very limited effect on payouts under naive bidding, with robust repacking reducing total payouts by just under 0.2%. In addition, the correlation between payouts in all 20,400 simulations is 0.9997.

We then consider how strategic bidding outcomes would change if we continued to

Table 2.10: Naive outcomes under robust repacking

Payouts ($B) under:	Naive Bidding
All DMAs	16.9994
	(0.7513)
All DMAs,	16.9655
Robust Repacking	(0.7330)

Notes: Data are averages over 100 auction simulations. Standard deviations across 100 simulations are in parentheses. Robust repacking implies that payouts and feasibility were computed and asserted for all licenses in the region, as opposed to only the focal DMA.

account for feasibility of all licenses in the region, instead of only those licenses in the focal DMA. To assess these issues, we simulated the robust regional repacking for all strategy profiles for the New York, NY, (729 strategy profiles for each simulation draw) and Washington, DC, (189 strategy profiles for each simulation draw) DMAs, as doing so for all DMAs would not be computationally feasible. Note that while we thus treat all licenses in the region as though they were part of the focal DMA for repacking purposes, we continue to assume that strategic bidding occurs only between licenses within the focal DMA, but not between all licenses a firm owns across the full repacking region. Table 2.11 shows the results of this exercise, and affirms that the impact of the repacking simplification on strategic outcomes is very small. Our intuition is that robust repacking is more flexible and so leads to lower payouts, although the results in these two markets suggest that the impact is negligible.

2.6 Conclusions

In this paper we explore ownership concentration as a means to seek rents in the context of the U.S. government's acquisition of broadcast TV licenses in the ongoing incentive auction. Ownership concentration is an important policy concern as the FCC has worried about encouraging a healthy supply of licenses in the reverse auction and has viewed outside investors as more likely to part with their licenses than potentially "sentimental" owners. Our prospective analysis shows that this is likely to give rise to strategic supply reduction and raise the cost of acquiring spectrum.

Table 2.11: Strategic outcomes under robust repacking

Payouts ($B) under:	Total	Naive Bidding	Strategic Bidding				Payout Increase
	Simulations	Base	Mean	Min	Median	Max	At Mean
New York, NY	72,900	2.363	3.086	2.363	3.096	3.893	30.6%
		(0.216)	(0.295)	(0.216)	(0.330)	(1.013)	
New York, NY	72,900	2.363	3.086	2.363	3.095	3.893	30.6%
Robust Repacking		(0.216)	(0.296)	(0.216)	(0.330)	(1.013)	
Washington, DC	18,900	0.290	0.327	0.317	0.328	0.338	12.9%
		(0.071)	(0.065)	(0.063	(0.067)	(0.085)	
Washington, DC	18,900	0.290	0.326	0.316	0.327	0.337	12.9%
Robust Repacking		(0.071)	(0.065)	(0.063)	(0.066)	(0.0084)	

Notes: Moments of the payout distribution are calculated across all possible equilibria for a given simulated auction and then averaged over 100 auction simulations. Standard deviations of the moments across the 100 simulations are in parentheses. Robust repacking implies that payouts and feasibility were computed and asserted for all licenses in the region, as opposed to only the focal DMA. Payout increase percentages are computed with full precision and so may not correspond to values computed from this table.

In particular, we argue that firms may engage in rent-seeking by attempting to reduce supply of broadcast TV licenses in the reverse auction. We conduct a large-scale valuation exercise for all UHF auction-eligible broadcast licenses in order to highlight the potential for strategic supply reduction and quantify the resulting increases in payouts and efficiency losses. The effect of ownership concentration can be substantial, and this paper is the first attempt to quantify this effect.

For the auction's initial clearing target of 126MHz, on which we base the simulations in this paper, the first stage of the reverse auction resulted in a spectrum acquisition cost of $86.4B, far exceeding the revenue the FCC was able to realize in the forward auction, a total of only $23.1B. Our base specification shows instead payouts of only roughly $17B in the reverse auction, which would have been low enough to end the incentive auction after the first stage.

While the goal of our paper is not to attempt to predict the exact reverse auction outcome, our broad findings indicate that it is likely that both participation below 100% and strategic bidding may have contributed to the failure to clear the opening target. Our results in Section 2.5.5 suggest that participation of less than 50% of eligible noncommercial stations nearly doubles spectrum acquisition costs, even in the absence of any strategic bidding. Strategic supply reduction, even when constrained to licenses in the same DMA market, increases broadcaster payouts by 22%, on average, as we show in Section 2.5.1. In a case study of cross-market strategic supply reduction, as would have likely taken place in the auction itself, payouts due to strategic bidding increased by a much more significant 87.2%. While computational constraints do not allow us to investigate whether the particular Philadelphia case study is representative of the effect of cross-market strategic supply reduction in other DMAs, the results in Section 2.5.4 suggest that our baseline strategic supply reduction effects are likely a lower bound on the true extent of rent seeking that could arise in the auction due to multi-license ownership.

The execution of the incentive auction, the most novel auction designed since the inception of spectrum auctions, is an incredibly difficult task that, based on current indications, has been very successfully tackled. We do not take a stand on whether any specific action would have altered the bidding results in the reverse auction. For example, it is impossible for us to assess whether full participation could ever have been achieved. We hope nevertheless that our work proves useful in designing future auctions geared at repurposing spectrum toward more efficient use.

Chapter 3

Do Concealed Gun Permits Deter Crime?[124]

3.1 Introduction

Shall-issue laws are state laws providing for the liberal issue of concealed gun permits analogous to getting a drivers license. Setting off a long controversy, Lott and Mustard (1997) (henceforth LM) reasoned that SILs increase the probability that a given would-be perpetrator's crime will fail because he can no longer tell which prospective victim may carry a gun and respond with threats or gun shots. In this controversy the weapon of choice has been the differences-in-differences (DD) estimator applied to state and county panel data spanning various intervals of time. Researchers have come to divergent conclusions spanning "more guns, less crime" to "more guns, more crime."

Elementary dynamic analysis highlights the possibility of three different effects of the introduction of SILs - one effect on those already vested in a life of violent crime, another effect on those teetering between entering such a life and the alternatives and, thereafter, a selection effect on the exit of those who chose to enter in the presence of SILs. With panel data on individual potential and actual violent criminals, an empirical specification to measure these effects would be straight forward. Unfortunately state (not individual) panels of crime rates for various types of violent crimes constitute the best available data.

[124] Joint with Marjorie B. McElroy, Duke University.

To date the research on the impact of SILs has ignored any forward-looking behaviors and insights from analysis of the dynamics - insights such as the contemporaneous responses of existing violent criminals may differ between those who were hit with SILs after they became violent criminals and those who selected into a life of violent crime despite the presence of SILs. Rather, variations on a static DD approach have been employed, typically estimating one effect of SILs for each type of violent crime. We argue that DD estimators can be viewed as weighted sums of three effects where the weights depend on the shares of three corresponding sub-populations (potential entrants, those who were hit with SILs after they became violent criminals, and those who selected into a life of violent crime despite the presence of SILs). As the sub-populations change systematically as more time elapses since the passages of SILs, so will the DD estimates. Thus suppose because the time series lengthens as the years roll by, an early investigator applies DD to a sample period including the immediate aftermath of SILs but not a longer run and a later investigator includes many time periods long after SILs passed. Then the DD estimate of the first will tend to estimate a surprise effect (muddied by a bit of a selection effect mixed in) and the DD estimate of the second investigator will weigh the selection effect more heavily. And since these effects bear different magnitudes, the DD estimate produced by the second investigator will tend to be different from the first investigator. This sensitivity of the DD estimate to the time span of the sample period provides a setup for a long controversy!

This situation likely arose because there seemed to be no way to incorporate the basic insights into panel data on crime aggregated to state (county, city) averages. In contrast, the CPDM proposed here, while using data aggregated to the state level can, nonetheless, tease out the three separate effects dictated by almost any dynamic model. We attack the problem indirectly - first by building a model of entry and exits from careers in violent crime and wrapping up all three effects in a net entry (= entry minus exits) equation. Under appropriate assumptions we link this to the observed changes in the number of crimes at the state level, a well-measured dependent variable. In addition, we develop appropriate proxies for the relevant sub-populations of violent criminals. With these in hand, we specify a Cohort Panel Data Model and provide maximum likelihood estimators of the three different effects of SILs on violent crime rates for all violent crimes as well as the four components.

Assuming that violent crime is a career, we provide a straightforward dynamic interpretation of what we term LM's deterrence hypothesis. Namely, SILs reduce the prospective value of a criminal career and also the continuation value for existing criminals. This is sufficient to sign the three effects and we strongly reject this hypothesis. We show how the CPDM nests the standard DD model thereby revealing exactly how the DD scuttles the basic implications from dynamics. Tests resoundingly reject the restrictions that reduce the CPDM to a DD model.

Our paper is related to recent work that closely examine the empirical specifications of DD. Bertrand, Duflo and Mullainathan (2004) (henceforth BDM) reviews a large set of DD papers and points out the underestimated standard errors due to serially correlated outcomes. In this paper, similar to Iyvarakul, McElroy and Staub (2011), we recognize that the point estimates are even biased in the DD specification in a large subset of the papers reviewed in BDM due to heterogeneous agents' dynamic decision making. By applying the more general CPDM to the crime setting in this paper, we show the wide application and robustness of CPDM in any setting that involves decision making of forward-looking agents.

This paper also sheds light on the controversial literature on concealed carry weapons, where almost all papers have employed variations of DD as their main statistical specification. LM was the first to use a large panel data set and essentially a DD specification, exploiting the different timing of state SIL passages, to rigorously study the effects of SILs on violent crimes. Since then, several papers have found the opposite, or facilitating effects of guns on crimes (Ayres and Donohue, 2003*b*,*a*) (henceforth AD); some have found no effects (Black and Nagin, 1998; Dezhbakhsh and Rubin, 1998); while some others have confirmed LM's findings (Plassmann and Tideman, 2001)[125]. While most of these studies make use of the same crime and law passage data set and a DD specification, they mainly differ in the lengths of their samples and various controls (time trends and demographics) used. We show that after accounting for serially correlated error terms as suggested by BDM, most of the results (those of both LM and AD) are rendered insignificant. Further-

[125] Moody and Marvell (2008) presents a more thorough literature review of the debate on SILs.

more, the estimates vary with the size of the sample, suggesting that the DD model is a misspecification. In contrast, the CPDM yields significant results that are invariant to the lengths of different sample periods (see Section 3.5.2).

This paper also fits in the broader literature on the economics of crimes. We construct a novel proxy for age-specific violent crime rates to study entry and exit behaviors of individual violent criminal cohorts. Similar to the economics of crimes and sociology literature (Hirschi and Gottfredson, 1983), we find consistent distributions of violent crimes across ages and further parameterized an exit function of violent criminals by age. Our results suggest that the recent liberalizations of gun laws, in addition to increasing overall violent crimes, also increased the turnover - both entry and exit - of violent criminals, effectively increasing the number of people with violent crime records, while reducing the duration of their violent criminal careers on average. Higher turnover of violent criminals has large social implications for criminal records, poverty, labor market outcomes, and etc. These results are consistent with and complement the recent work on the reasons and effects of the prison boom in the U.S. (Neal and Rick, 2014; Johnson and Raphael, 2012)[126].

Finally, our CPDM embeds a structural model of criminal discrete choices, extending Gary Becker's rational criminal framework (Becker, 1968) to the dynamic setting. Similar to the structural labor and crime literature (Wolpin, 1984; Imai and Krishna, 2004), we model individual criminals as forward-looking agents with heterogeneous propensity to commit crimes who dynamically optimize utility. However, while these papers estimate criminal behaviors with very special samples of micro data (e.g. the Philadelphia Birth Cohort Study), we believe that state panel data are more widely accessible to researchers and representative of general population and criminal population to study the overall crime patterns. Instead of solving individual-level Bellman equations, we are also able to aggregate to the cohort, state and year level for the simple estimation procedure that still captures average costs and benefits of entry and exit decisions.

The rest of the paper is organized as follows: Section 3.2 sets up the model, Section 3.3 introduces data and descriptive evidence, Section 3.4 describes the empirical specification

[126] Johnson and Raphael (2012) also exploits the dynamics as an instrument to identify the effects of changes in incarceration rates on changes in crime rates with state panel data. We explicitly address the dynamic adjustments of criminals as well as the heterogeneity among criminals with our CPDM.

in detail, Section 3.5 presents results and Section 3.6 concludes.

3.2 Model

This section presents a spare model that captures the essential consequences of forward looking behaviors on the part of potential and actual violent criminals in order to identify the differing effects of SILs across three sub-populations as well as the total effect. Treating violent crime as an occupation lets us capture the effects of SILs on entry into and exit from a career in violent crime in a familiar way. Potential entrants are all those who are capable but not yet criminals; potential exitors are all those who are currently violent criminals. To simplify the language, in this paper, we refer to careers in violent crimes as "careers" and use violent criminals and criminals interchangeably. We also refer to the potential entrants and exitors as the "entry cohort" and the "exit cohort" even though it is not, strictly speaking, a cohort but a stage of life.

Assume the choice governing entry is captured by a value function and those governing exit by a continuation function. The passage and presence of SILs affect both. Begin with the entry cohort. Let (s, t) denote state s in period t and let N^{En} = the number of potential entrants in (s, t).Then a familiar, straightforward reduced-form representation of decisions to enter careers in violent crime would be

$$Entry_{st} = (\alpha_0 + \alpha_1 I_{st}^{SIL} + \epsilon_{st}^{En}) N_{st}^{En} \tag{3.2.1}$$

where $I_{st}^{SIL} = 1$ if SILs are in effect, and ϵ_{st}^{En} is a well-behaved random error to be discussed. Parameters to be estimated are the base entry rate, α_0, and the impact of SILs on entry, α_1. Note that the dependent variable $Entry_{st}$ is unobserved.

With forward looking behaviors, the contemporaneous effects of SILs on exits from careers in violent crime depend on whether this career was chosen before or after the passage of SILs. For those whose entry was prior, the passage of a SIL induces a surprise change in the continuation value of this career and consequently exit rates change by the surprise effect, denoted by β_2. In the case that the advent of SILs causes continuation values to fall, the exit rates increase and $\beta_2 > 0$, and vice versa. Use $N_{st}^{Surprised}$ to denote the size of the surprised cohort.

In contrast to the surprised cohort, those who chose their careers in violent crime after the passage of SILs presumably capitalized the effect of SILs on the value of a career in violent crime when they selected into careers of violent crime. Use $N_{st}^{Selected}$ to denote the size of this selected cohort. For if the pool of potential entrants is heterogeneous in their "quality" (proclivity for violent crime) the change in the value of the violent career path induced by SILs will affect not just the quantity of entrants as in Equation 3.2.1 but also their quality and, in turn, change their exit rate down the road. This is captured by the selection effect β_1. In the case that the advent of SILs decreases continuation values, the marginal and average violent criminal will have a higher quality, be more buffered from negative career shocks, and thus have a lower probability of exiting or $\beta_1 < 0$, and vice versa. These effects are captured in the reduced form exit equations,

$$Exit_{st}^{Selected} = (\beta_0 + \beta_1 I_{st}^{SIL} + \epsilon_{st}^{Ex}) N_{st}^{Selected} \tag{3.2.2}$$

$$Exit_{st}^{Surprised} = (\beta_0 + \beta_2 I_{st}^{SIL} + \epsilon_{st}^{Ex}) N_{st}^{Surprised} \tag{3.2.3}$$

Thus, as shown below, in contrast to diff-in-diff specifications, this enables the CPDM to explain turning points in criminal activity and not just either upswings or downturns. Finally, subtracting exits from entrances gives the net increase in criminals,

$$\begin{aligned} NetEntry_{st} = & (\alpha_0 + \alpha_1 I_{st}^{SIL} + \epsilon_{st}^{En}) N_{st}^{En} \\ & - (\beta_0 + \beta_1 I_{st}^{SIL} + \epsilon_{st}^{Ex}) N_{st}^{Selected} \\ & - (\beta_0 + \beta_2 I_{st}^{SIL} + \epsilon_{st}^{Ex}) N_{st}^{Surprised} \\ = & (\alpha_0 + \alpha_1 I_{st}^{SIL}) N_{st}^{En} - (\beta_0 + \beta_1 I_{st}^{SIL}) N_{st}^{Selected} - (\beta_0 + \beta_2 I_{st}^{SIL}) N_{st}^{Surprised} + \epsilon_{st} \end{aligned} \tag{3.2.4}$$

where the error $\epsilon_{st} = \epsilon_{st}^{En} N_{st}^{En} - \epsilon_{st}^{Ex} N_{st}^{Selected} - \epsilon_{st}^{Ex} N_{st}^{Surprised}$ is mean zero, heteroskedastic, and can be written as $\sigma^2 = \left[(N_{st}^{En})^2\pi + (N_{st}^{Selected})^2 + (N_{st}^{Surprised})^2\right]\sigma_{Ex}^2$, where $\pi = \frac{\sigma_{En}^2}{\sigma_{Ex}^2}$ is a parameter to be estimated. Should $Var(\epsilon_{st}^{En}) = Var(\epsilon_{st}^{Ex})$, then $\sigma^2 = Var(\epsilon_{st}^{En})$ and the variance is homoskedastic.

Equation 3.2.4 is the basic model for the CPDM. Later in the empirical work, we investigate the effect of floodgate and aging effects to this model. Our approach highlights the importance of three separate effects of SILs: α_1 a direct effect on entry of youths into violent criminal careers and β_1 the subsequent selection effect on their exits; and β_2 the surprise effect on cohorts of older criminals who began their careers prior to SILs. Further, these three parameters capture the two fundamental implications of dynamic analysis. These are (i) the impact of SILs on behaviors are not symmetric between potential entrants and exitors (youths in their entry windows and violent criminals) - roughly, the α's are not equal to the corresponding β's; and (ii) the impact of SILs on exits from violent criminal careers differs between those who began their careers before the advent of SILs and those who began after - $\beta_1 \neq \beta_2$.

Given ideal panel data on individuals, we could observe entries and exits of potential and actual criminals and form subsamples of criminals according to whether their entry preceded or post-dated the advent of SILs. Then the strategy would be to estimate each of these three separate effects - using something like diff-in-diff - on the corresponding three sub-samples. In reality such data are not on the visible horizon. Unlike other occupations, the pool of criminals as well as their entries and exits go unobserved. The panel data we do have are aggregated to the state (or county or city) level and, of course, do not parse out the criminal population, much less record entry dates. Thus a three-separate-regression estimation strategy for state panel data that parallels that for micro panel data is precluded. In particular, this strategy is precluded because the crime rates (dependent variables) available are for the entire state population, not for the three key sub-populations. The point of using the cohort panel data model is that, despite observing only the impact of SILs on violent crimes aggregated to the state level, nonetheless the CPDM provides a way to identify the three fundamental dynamic effects of SILs - α_1, β_1, and β_2.

Table 3.1: Effects of SILs on criminal careers

Cohorts	Before SIL $I_{st}^{SIL}=0$	After SIL $I_{st}^{SIL}=1$
Entry N_{st}^{En}	α_0	$\alpha_0+\alpha_1$
Exit		
$N_{st}^{Selected}$	β_0	$\beta_0+\beta_1$
$N_{st}^{Surprised}$	β_0	$\beta_0+\beta_2$
Net Entry $N_{st}^{Ex}=N_{st}^{Selected}+N_{st}^{Surprised}$	$\alpha_0 N_{st}^{En}-\beta_0 N_{st}^{Ex}$	$(\alpha_0+\alpha_1)N_{st}^{En}-\beta_0 N_{st}^{Ex}$ $-\beta_1 N_{st}^{Selected}-\beta_2 N_{st}^{Surprised}$

$$\alpha_0 N_{st}^{En}+\alpha_1 I_{st}^{SIL} N_{st}^{En}-\beta_0 N_{st}^{Ex}-\beta_1 I_{st}^{SIL} N_{st}^{Selected}-\beta_2 I_{st}^{SIL} N_{st}^{Surprised}$$

Notes: Breakdown of the CPDM into entry and exit, before and after SIL. Multiplying cohort sizes in column 1 with average effects in columns 2 & 3 yields the respective contributions of each cohort to the total effect of SILs on criminal careers. Summing across rows then gives the total effect, or equivalently, our CPDM.

3.2.1 Implications

It is worth pausing to create a sketch of the model as contained in Table 3.1. The first two blocks in Table 3.1 show the contribution of each cohort (entry, selected and surprised) to the aggregate net entry rate with the second and last columns giving these contributions before and after SILs, respectively. In the third block of rows, weighting each row by its share and then subtracting exits from entries gives the net entry rate before and after SILs. N_{st}^{Ex} is the number of all potential exitors. Finally weighting the second and last share-weighted column total net entries by $(1-I_{st}^{SIL})$ and I_{st}^{SIL} gives the desired net entry rate for each (s,t) in the last block. Note that the expression in the last block is the same with Equation 3.2.4.

We use this table to lay out, in turn, the evolution of the crime rate over time, the implications of the deterrence hypothesis, the nesting and testing diff-in-diff specifications as special cases of the CPDM.

Table 3.2: Evolutions of criminal cohorts

Impacts on Criminal Cohorts	Before Passage Old Eqm. $t < t^*$	At Passage $t = t^*$	After Passage Transition Years $t^* < t < t^{**}$	New Eqm. $t \geq t^{**}$
$s_{st^*}^{*Surprised}\beta_2$	0	β_2	$0 < s_{st}^{*Surprised}\beta_2 < \beta_2$	0
$s_{st^*}^{*Selected}\beta_1$	0	0	$0 < s_{st}^{*Selected}\beta_1 < \beta_1$	β_1

Notes: Exit cohort sizes and contributions to the total effect over time. Cohorts are normalized by the total exit cohort size N_{st}^{Ex}.

3.2.1.1 Evolutions of criminal cohorts

Under the CPDM, how would passages of SILs affect crime rates? As Equation 3.2.4 and Table 3.1 show, the obvious effects are captured by α_1, $\beta_1$and β_2 that affect entry and exit of the corresponding sub-populations. We turn to how the size and share of each sub-population evolve over time.

First set aside the entry cohort and presume it is exogenous (i.e., fertility is independent of SILs). Divide the selected ($N_{st}^{Selected}$) and surprised ($N_{st}^{Surprised}$) cohorts by the total exit cohort (N_{st}^{Ex}) so they sum to one, $s_{st}^{*Selected} + s_{st}^{*Surprised} = 1$. Prior to SILs, crime evolves according the pre-SIL entry and exit rates as they hit the associated entry and exit cohorts. Further, note that as of the period when SILs become effective (t^*), essentially all criminals would have entered before this. Thus in t^* none of the stock of criminals were selected into crime under SILs so that $s_{st^*}^{*Selected} = 0$ and thus $s_{st^*}^{*Surprised} = 1$. This contrasts with the long run here defined as beginning when the last survivor in the surprised cohort retires or exits (t^{**}) . By then the cohort shares have reversed: $s_{st^*}^{*Selected} = 1$ and $s_{st^*}^{*Surprised} = 0$ and they remain there going forward. Most importantly, for t in between t^* and t^{**}, the shares evolve systematically with $s_{st^*}^{*Selected}$ growing (approaching 1) at the expense of $s_{st^*}^{*Surprised}$ (approaching 0). These shares are the weights on the selection and surprise effects. Hence, the impact of these effects on crime rates go from the surprise effect (β_2) dominating in the immediate aftermath of the passage of SILs, then fading as these older criminals exit and the fraction selected into crime grows until, in the long run, only the selection effect of SILs remains. These trends are summarized in Table 3.2.

3.2.1.2 The deterrence hypothesis

LM's deterrence hypothesis has a natural interpretation in terms of the CPDM. Recall the channel they envisioned was that in the presence of concealed guns born by law-abiding citizens, violent criminals faced lower payoffs in the form of increased risk from their intended victim because they can no longer tell which victims are unarmed and which not. This translated into our CPDM model as lowering the value of entering a career in violent crime and also lowering the continuation value for those who are already criminals. Consequently, we interpret the deterrence hypothesis as implying that SILs reduce entry via lowering the career value, i.e., $\alpha_1 < 0$. Also, thereafter, those who select into crime are fewer in number but more hardcore than otherwise, i.e., $\beta_1 < 0$. Finally, and this likely gets closest to what LM had in mind: the advent of SILs is a negative surprise for the continuation value for current criminals and they exit at higher rates than otherwise, i.e., $\beta_2 > 0$.

3.2.1.3 Nesting DD in CPDM

To show that the CPDM nests the basic DD we return to the two basic insights from a dynamic model of entry and exit into crime. These are (i) differential impacts of SILs between potential entrants and exitors (youths in their entry windows and violent criminals) - roughly, the α's are not equal to the corresponding β's; and (ii) the impact of SILs on criminals' exits by those who began their careers before and after the advent of SIL are not equal, i.e., $\beta_1 \neq \beta_2$. It is exactly the denial of these insights that reduces the CPDM to the DD estimators.

Let us impose these in turn on the specification of the CPDM in Equation 3.2.4. First deny insight (ii) by imposing the restriction that those who became criminals before and after the advent of SIL exhibit the same contemporaneous responses to the presence of SILs, or $\beta_1 = \beta_2 = \beta_*$, a common value. In that case Equation 3.2.4 becomes

$$NetEntry_{st} = (\alpha_0 + \alpha_1 I_{st}^{SIL})N_{st}^{En} - (\beta_0 + \beta_* I_{st}^{SIL})N_{st}^{Ex} + \epsilon_{st} \tag{3.2.5}$$

Then further deny insight (i) by imposing that the contemporaneous impact of SILs on the crime rate is the same for potential entrants as for criminals, or $\alpha_0 = -\beta_0$ and $\alpha_1 = -\beta_*$. Equation 3.2.5 is reduced to

$$NetEntry_{st} = \alpha_0 N_{st} + \alpha_1 I_{st}^{SIL} N_{st} + \epsilon_{st} \tag{3.2.6}$$

where $N_{st} = N_{st}^{En} + N_{st}^{Ex}$ is the total relevant population at risk to contribute to the net change in the number of criminals. Equation 3.2.6 is then the familiar DD form and is, as everyone knows, completely static.

3.3 Data and descriptive evidence

We draw from several sources of data in this paper in order to build up the cohorts in the CPDM and to overcome data difficulties in traditional studies of crimes.

To construct the basic dependent variables (violent crimes), we follow the literature and obtain data from the Uniform Crime Report (UCR) maintained by the Federal Bureau of Investigation (FBI). The UCR data starts from 1977, as used in LM, but we focus on the period 1980-2011 due to other data constraints (BJS, see below). UCR reports violent crime and arrest rates at the state-year level in five categories: (1) murder and nonnegligent manslaughter, (2) forcible rape, (3) robbery, (4) aggravated assault, and (5) total violent crimes. Crime rates are used to construct dependent variables in our empirical specification, while state-level arrest rates are proxies for state police enforcement intensities, as is often used in the literature. Demographic control variables are obtained through the Regional Economic Information System (REIS) of the Bureau of Economic Analysis (BEA). These variables include real per capita personal income, income maintenance, unemployment insurance, and retirement payment for people older than 65 on the state-year level and are again broadly used in this literature to control for state-level income and welfare conditions over time. Table 3.3 summarizes these crime and control variables.

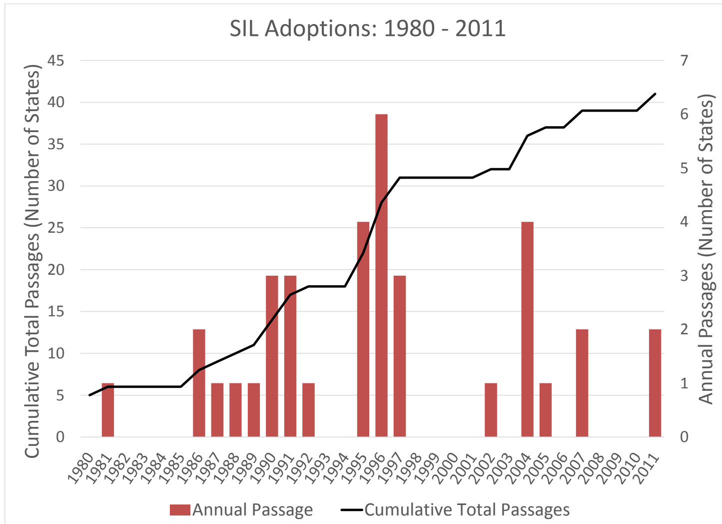
We obtain single-age population estimates from the Census on the state-year level to

Table 3.3: Main sample summary statistics: all states, 1980-2011

	Mean	SD	Min	Max	N
Crime Rates					
(Crimes/100,000 pop.)					
Violent	480.21	308.17	47.01	2921.80	1632
Murder	6.69	6.95	0.16	80.60	1632
Rape	35.08	13.42	7.30	102.18	1632
Robbery	145.30	151.26	6.40	1635.06	1632
Agg. Assault	293.13	171.62	31.32	1557.61	1632
Arrest Rates					
(Arrests/100,000 pop.)					
Violent	167.40	109.95	3.13	1313.82	1600
Murder	5.14	4.97	0	52.00	1599
Rape	10.41	6.48	0	92.49	1598
Robbery	35.98	44.75	0.16	1251.85	1597
Agg. Assault	116.06	74.70	2.78	656.23	1600
Control Variables					
State Pop. (M)	5.24	5.83	0.41	37.69	1632
Pop. Density ($pp/mile^2$)	313.25	1191.56	0.62	9306.41	1632
Inc. Mainten. ($)	404.46	179.72	104.26	1282.19	1632
Income ($000s)	28.87	7.01	15.01	64.88	1632
Unemploy. Insur. ($)	142.92	103.09	18.86	780.47	1632
Retire. Pay. ($000s)	3.53	1.13	1.18	7.00	1632

Notes: Crime type definitions - murder and nonnegligent manslaughter is defined as the willful (nonnegligent) killing of one human being by another; rape is defined as the carnal knowledge of a female forcibly and against her will; robbery is defined as the taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or by putting the victim in fear; aggravated assault is defined as an unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury.

Figure 3.3.1: SIL adoptions trend



Notes: Bars indicate the number of SIL passages in each year (right axis) and the line shows the total number of SIL states so far (left axis).

construct age-specific entry cohorts in our model. For more homogeneous effects, we focus only on the male population in this paper[127].

There has also been controversy over the exact years of passage of SILs in several states in the literature. We conduct our independent research in the SIL passage years in all states and show them in Appendix C.2.1. Our coding of the passage years is aligned with AD and extends it 2011. We plot in Figure 3.3.1 these SIL passages over time. The upward trended line over the three decades suggests explosive increases in the number of SIL states from 5 to 41. By 2011, 41 states have SILs in place and 36 of these were passed during our sample period 1980-2011. Many states have been persuaded to adopt SILs by political lobbyists as well as strong academic influence (e.g. LM), corroborating the importance to understand effects of SILs. We also identify the causal effects of SILs by exploiting the variations in the timing of state adoptions.

It is well known that U.S. crime rates peaked shortly after 1990 and have been falling

[127]Violent crimes reported to be committed by females are far less than those by males and are likely to be different in nature.

rather smoothly ever since. Also, our CPDM with $\beta_2 < 0$ and $\beta_1 > 0$ can explain an upswing followed by a downturn in the crime rate. This does not, nonetheless, make the CPDM a good candidate for explaining the national peak in crimes in the early 1990's. This can be seen in Figure 3.3.2. There states are partitioned into five groups with the states within a group all adopting SILs about the same time[128]. The first group of states adopted SILs prior to 1985 or have always had equivalent laws as SIL and the last group includes states that adopted SILs in 2011 or never adopted SIL by 2011. If the swings were all explained by the CPDM model, the peak crime rates for each group would all occur some years after that group adopted SILs and Figure 3.3.2 would have a series of humps whose max moves to the right as adoption years become more recent. But that is not the case. Instead, Figure 3.3.2 shows that for all groups, crime rates peak around 1990. Thus the CPDM for SILs could explain deviations from the overwhelming national peak in the early 1990's. But it is an unlikely candidate for explaining the huge national swing. On the other hand, it is important to control for non-linear time trends in the empirical specification.

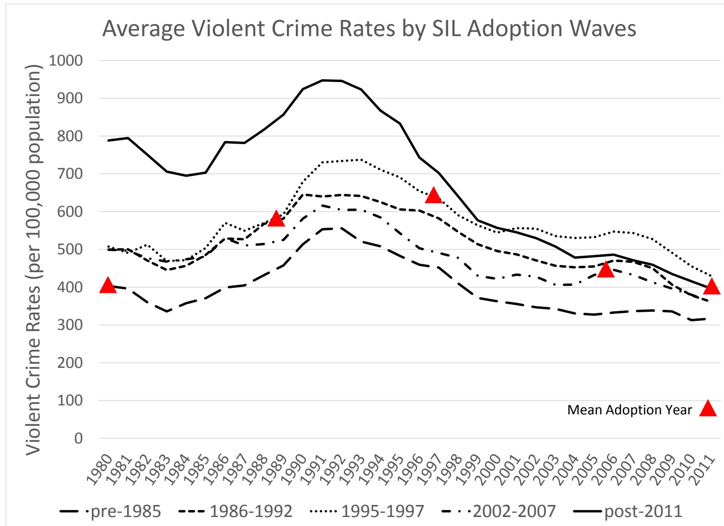
Importantly, the patterns in Figure 3.3.2 argue against the endogeneity of SILs. For example, the group of states with the second lowest crime rate was the last group to pass SILs while the group with the lowest crime rate was the earliest. In short Figure 3.3.2 gives no reason to suspect that high (or low) crime rates cause states to pass SILs.

To visualize the effects of SILs on violent crimes estimated from a typical DD specification, we compare average crime rates of the treated states vs. the non-treated states. The multiple treatment dates (16 unique years for the 36 states that adopted SILs within our sample) make it difficult to present the treatment and control groups graphically using the standard multiple-event DD as in Equation 3.2.6. We follow Gormley and Matsa (2011) here[129] - define a 20-year window around each treatment date t_* (normalizing t_* to zero), use all states who never adopted SILs within the window as the control group and states that exactly adopted SIL in t_* as the treated group, and call the two groups to-

[128]See Figure 1 of Ayres and Donohue (2003*b*) for comparison. We follow them for this categorization but extend it into a longer panel and finer groupings.

[129]In the rest of the paper, we use the standard multiple-event DD as our DD specification for estimations but only use the Gormley and Matsa (2011) procedure here for graphically comparing the treated and the control.

Figure 3.3.2: Violent crime rates by SIL passage years

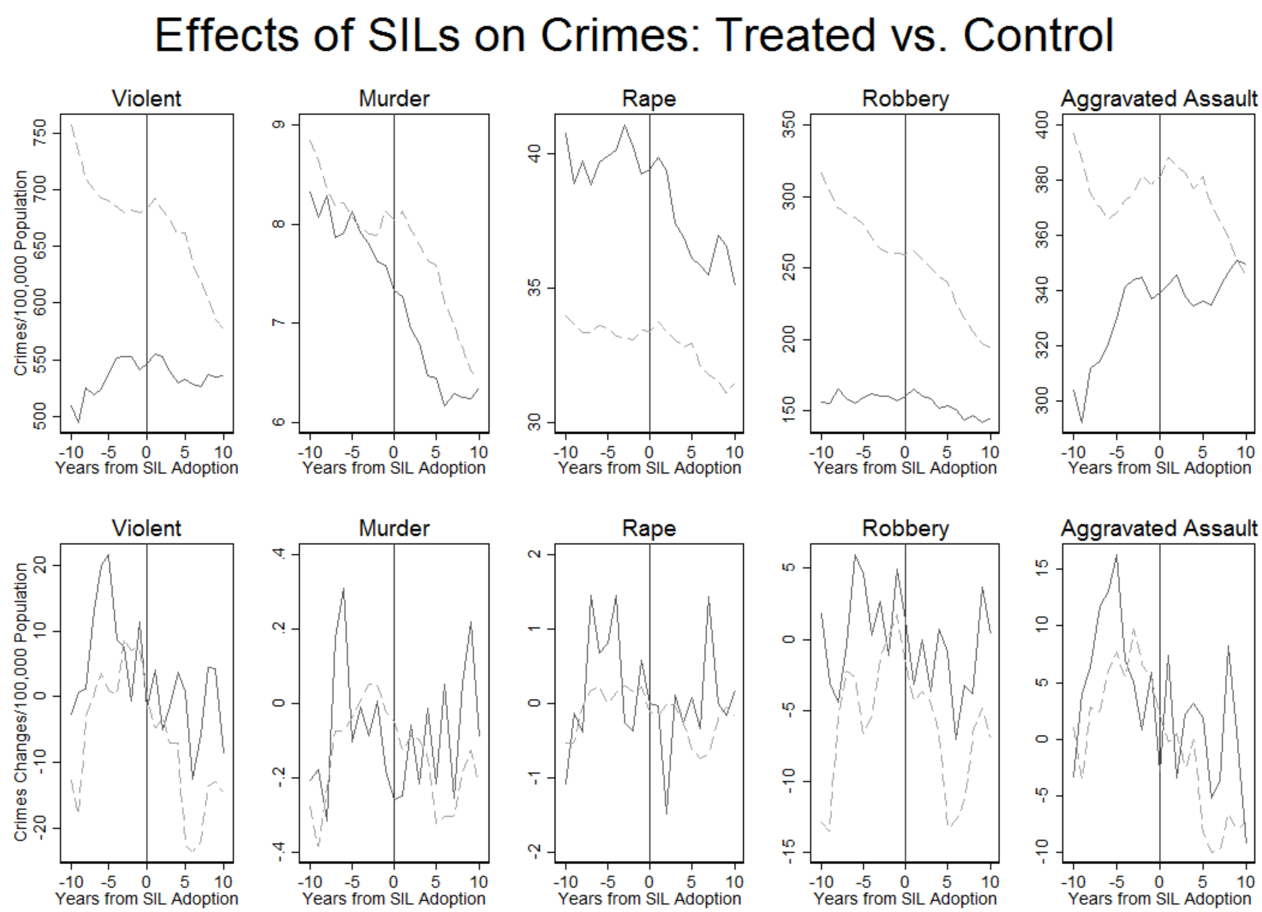


Notes: Total violent crime rates are averaged across states (weighted by state population) within same waves and plotted over time. The solid triangles indicate the average SIL passage year of the group.

gether a cohort. Then we average across cohorts for the overall treated and control groups. The results are shown in Figure 3.3.3. In the top row, we plot the levels of crime rates for each crime category, where solid lines are for the treated group and dashed lines for the control group. We find only ambiguous evidence of the effect of SILs - already showing evidence against LM and AD. In particular, the declining crime rates (or in some cases, the "inverted-V" shape) of the treated group cannot be used as evidence for the deterrence hypothesis in comparison to the control group. In the second row, we plot the same for the changes (or net entry) in each crime category. Visually, a small positive effect of SILs can be detected in the treated group compared with the control - the DD is able to capture the more nuanced effect when specified on the changes, while still leaving much dynamics to be explained.

The final data set we use is the national arrests by age groups data from the Bureau of Justice Statistics (BJS). The BJS arrests data differs from the UCR arrests data in that it reports arrest rates on the age group-year level for each crime category. It covers the period

Figure 3.3.3: Treatment effects of SILs



Notes: Solid lines represent crime rates for the treated group. Dash lines represent crime rates for the control group. Treated and control groups are constructed similar to Gormley and Matsa (2011). X-axis indicates years from SIL adoption, centered at the adoption year. Y-axis indicates the number of (changes of) crimes per 100,000 population.

1980-2011 and reports in 17 age groups: 9 or younger, 10-12, 13-14, 15, 16, 17, 18-20, 21-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, and 65 or older. Together with the UCR data, we then impute age-specific arrest and crime rates, the lack of which is a traditional data problem in studies of crimes, due to the nature of crime reporting (see Appendix C.2.2 for the imputation procedure).

3.4 Empirical specification

In this section we turn back to the CPDM and bring it to the data. We lay out an empirical strategy here to construct the relevant cohorts and to estimate the CPDM parameters.

Recall our CPDM in Equation 3.2.4 and Table 3.1 - unfortunately we do not observe the dependent variable in Equation 3.2.4. The link between the unobserved number of new criminals and the observed net increase in crimes is $\kappa_{ast} = \frac{crimes}{criminals}$. Multiplying Equation 3.2.4 through by κ_{ast} converts the criminals dependent variable to the change in crime rate. Ideally, we would know both components of the change in crime rates, κ_{ast} and $\Delta criminals_{st}$. But given the impracticality of a large representative panel on the number of criminals, this seems, at best, beyond the visible horizon. The simplest practical assumption is that κ is constant across all criminals. In that case, multiplying through Equation 3.2.4 by κ converts the dependent variable to the observe change in the crime rate and changes the interpretation of the coefficients. Thus, the parameters to be estimated become $\alpha_i' = \kappa\alpha_i$ in place of α_i and $\beta_i' = \kappa\beta_i$ in place of β_i. Thus, α_0' the baseline new crimes/year attributed to the entry cohort, α_1' the change in these crimes due to SILs, and so forth. Note that the percent increase in entry rate due to SILs is identified as $\frac{\alpha_1-\alpha_0}{\alpha_0} = \frac{\alpha_1'-\alpha_0'}{\alpha_0'}$ because the κ's cancel and the analogous result holds for β_1 and β_2.

In an abuse of notation we re-use the α's and β's and write the basic CPDM for crime rates as

$$NetEntry_{st} = \alpha_0 N_{st}^{En} + \alpha_1 I_{st}^{SIL} N_{st}^{En} - \beta_0 N_{st}^{Ex} - \beta_1 I_{st}^{SIL} N_{st}^{Selected} - \beta_2 I_{st}^{SIL} N_{st}^{Surprised} \tag{3.4.1}$$

The model above assumes that $\kappa_{ast} = \frac{crimes}{criminals}$ is constant for all criminals at all ages and in every (s, t). We have nothing to add to the usual discussions of holding such parameters constant every s, but need to deal with the obvious fact that intensity κ varies across ages and in response to SILs. Our dependent variable is the time-differenced rate, $\frac{NumberOfCrimes}{Population}$. In turn the number of crimes is the number of criminals times the average $\frac{crimes}{criminals}$. Data limitations preclude parsing out changes in this intensity between changes in the components. If data permitted, we could pursue a more complex model that distinguished, for example, the effects of SILs on the numbers of entrants and their average intensity κ. But, it is not hard to see that such a dynamic model would predict that either both effects are positive or both effects are negative and our entry parameter α_1 measures the combination of these two. Hence, although we refer to "entries and exits of criminals," a more accurate descriptor would be "increases and decreases in the crime level." We prefer, however, "entries and exits" because it constantly reminds us of the dynamic decision making underpinning our model.

In constructing the cohorts, the entry cohorts are ideally composed of all capable (reasonable ages, discussed below) males that are not violent criminals[130] already. Since the number of violent criminals at a time in a state is unobservable to us and is relatively small compared to the total population (violent crimes / total population are 0.48% on average), we simply use the male populations as the entry cohorts. Exit cohorts, on the other hand, are even harder to construct. Criminal populations are obviously unobservable. Much of the crime literature suffer from this unavoidable data difficulty and in this paper we try to remedy it using proxies. Older males' population is a potential candidate to proxy for the exit cohorts but it lacks correlation with the actual criminal cohorts and variation from the entry cohorts (perfectly collinear when weighted by total male population). Crime rates are better proxies for the exit cohorts if we believe that criminals across different states, years, and ages commit similar number of crimes. The only remaining issue is that the UCR crimes data only vary at the state-year level and we need age-specific exit cohorts to identify the selection and surprise effects. We thus supplement the UCR crime rates data

[130] We loosely define violent criminals as anyone who has committed at least one of the four types of violent crimes in a year.

with the BJS age-specific arrests data to impute the age-specific crimes[131] (Appendix C.2.2).

Now before we can specify the entry and exit cohorts, we need one more piece of information (assumption in this case). Remember that we needed the age-specific crimes data to construct the variables $N_{st}^{Selected}$ and $N_{st}^{Surprised}$ for the identification of the selection and surprise effects. The reason is that we need to know which age groups entered when to categorize them into young and old cohorts (see below for specific procedures). To do so, we opt for a parsimonious specification[132] in which we define entry and exit windows. Figure C.2.1 (Appendix C.2.2) suggests that violent crimes peak around age 20, across types of crime and time. Classic sociology theory, discussed in Hirschi and Gottfredson (1983), also confirms that the age distribution of criminals does not vary across times, places, or types of crime[133]. We thus define our entry-only window to be age 13-21, and exit-only window 22-64. The cutoffs of these windows are also empirically informed, beyond what the theory suggests. The age range 13-64 covers, on average, 98% of the crimes committed in a given state-year and allows for easier parametrization (constant entry rate and quadratic exit rates, see below)[134]. The age 21 that divides our entry and exit windows is picked out by maximizing the log-likelihood of the estimated baseline CPDM (see below).

Part of the main contribution of this paper is to capture the heterogeneous treatment effects of SILs due to the dynamically optimizing behaviors of different cohorts. The selected and surprised cohorts in the model thus tease these effects (selection and surprise) apart from the base exit rate. We define the selected and surprised cohorts as follows. With age-specific crimes (or criminals, as proxied for), an age cohort belongs to the selected cohort if the entirety of its entry window (13-21) is spent after the SIL passage in that state. Similarly, an age cohort is part of the surprised cohort if the entirety of its entry window is

[131] The imputation procedure and the use of proxy variables will likely introduce measurement errors, which we assume to be uncorrelated with the regressors, as typically done in this literature.

[132] An alternative is to specify a nonlinear probability model to figure out the proportions of people of different entry dates within age groups.

[133] This suggests that any legislation differences and changes would impact the whole distribution of ages similarly. In Figure C.2.1, we observe that the far tails (beyond age 40) of the distributions become fatter over time (from left to right), suggesting an aging criminal population. However, the distributions still peak around age 20 and thus do not affect our choice of entry and exit windows.

[134] This is also the reason why we do not allow overlapping entry and exit windows. The relatively narrow entry window of 13-21 allows for a plausibly constant entry rate but the variations in young male population do not pick up all entry variations. Thus allowing exit in the same region would severely bias the exit parametrization.

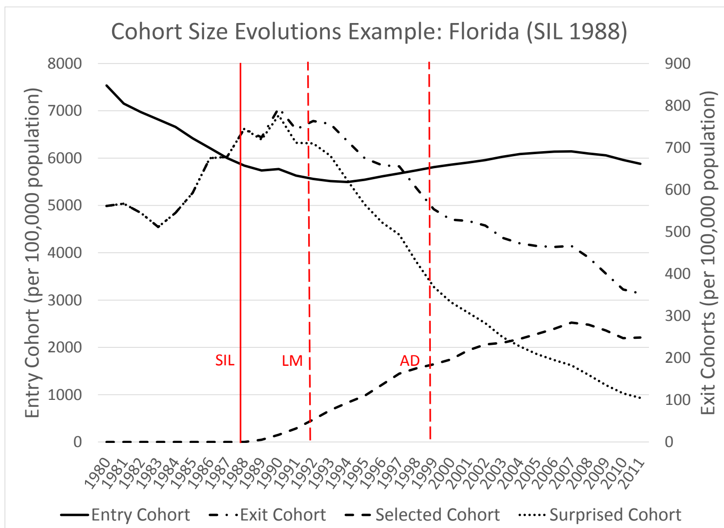
spent before the SIL passage in that state. For age cohorts that experience SIL passage during their entry windows, we divide the cohort by weights corresponding to the number of years within their entry windows before and after SIL passage[135].

Key to our identification of CPDM is the difference in the evolution of different cohorts over time after the passage of SILs. Expanding on AD's case studies on the populous state Florida, where SIL went into effect in 1988, we illustrate these evolutions in Figure 3.4.1. The entry cohort measures male population between 13 and 21 and is relatively stable and exogenous to the SIL passage. The total exit cohort (of violent criminals) measures the current stock of violent criminals and thus fluctuates with violent crime rates and exhibits the "inverted-V" shape following the national pattern. The exit cohort is further divided into the surprised and the selected cohorts after the adoption of SIL. As time goes by, the selected cohort converges again to the total exit cohort while the surprised cohort disappears as the violent criminal stock is replaced with entrants from the post-SIL era. A new equilibrium establishes as the selected cohort coincides with the total exit cohort. In Figure 3.4.1 we also show the lengths of samples used in LM and AD. In examples like Florida, where SIL is adopted before 1992, LM's sample weighs more on the surprise effect in a DD model while AD's sample weighs more on the selection effect. We show in Section 3.5.2 that in the full national sample, given gradual passages of SILs among different states, DD is biased by the changing weights of surprise and selection effects while CPDM tease them apart consistently.

Figure 3.4.2, on the other hand, shows the evolutions of the average ages of the different cohorts. While the overall entry and exit cohorts stay relatively constant in age, the surprised cohort on average grows in age over time due to the lack of replenishment of new entries and will eventually all reach retirement age. The selected cohort also on average grows in age due to the initial aging of its constituents but will be balanced out by new entries and converge to the total exit cohort around age 34 when the surprised cohort dies out. The differences and changes in average ages across cohorts and time pose an challenge to the identification of the selection and surprise effects in our model, which we

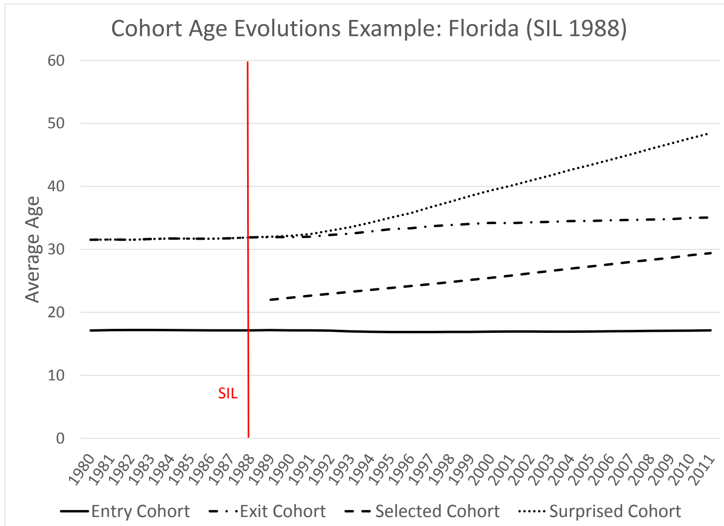
[135]This is internally consistent in the model when we estimate a constant entry rate. See more discussion below on aging effects and non-constant entry rates.

Figure 3.4.1: Illustration of cohort size evolutions



Notes: Entry cohort is measured in 100,000 population on the left axis. Exit cohorts are measured in 100,000 population on the right axis. The solid vertical line indicates SIL passage in Florida in 1988. The vertical dashed lines indicate where LM and AD's samples end, respectively.

Figure 3.4.2: Illustration of average cohort age evolutions



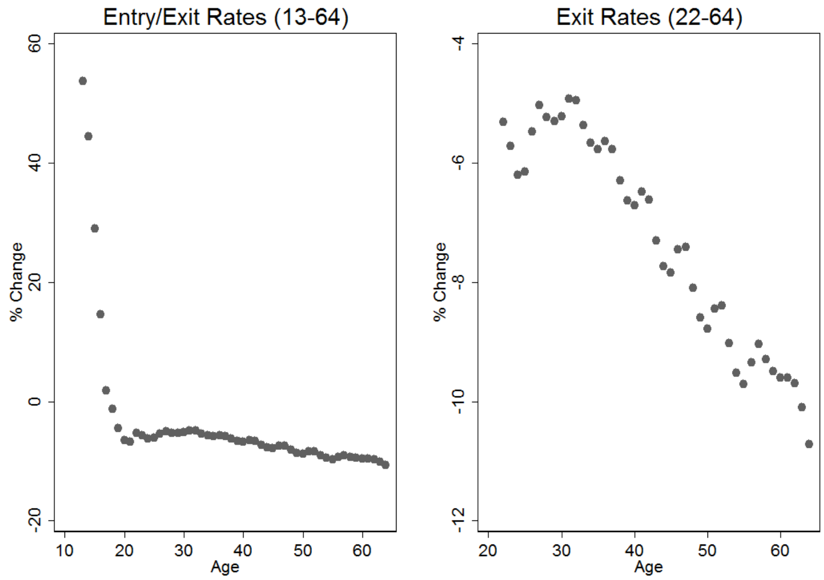
Notes: Evolutions of the average age in each cohort. The solid vertical line indicates SIL passage in Florida in 1988.

now turn to address.

Much of this paper is concerned with capturing the heterogeneity in cohorts as defined by the timing of their entries into crimes and the passage of SILs. However, there is another dimension of heterogeneity, intertwined with our cohorts definition, which we have so far ignored - the heterogeneity in ages. People of different ages have different physical conditions (important for committing violent crimes), have accumulated different levels of human capital (either human capital in the crime career that results in different skills or human capital outside crimes that results in different values of outside options), have different lengths of potential career left until retirement in crimes (important if we think that people dynamically optimize in choosing their careers), and etc.

In theory, these competing forces over the life cycle likely result in a non-linear base exit probability (irrelevant to the passage of SILs) that bottoms out in male criminals' 30's or 40's. Ignoring this crucial fact (and only estimating a constant exit rate) will bias estimates for selection and surprise effects in our model due to their differences and evolutions in ages. To fit the exit rate over the life cycle empirically, in Figure 3.4.3, we plot an empirical

Figure 3.4.3: Aging effects on violent crime entry and exit



Notes: Fraction changes of the imputed crime rates against ages, averaged across all state-year observations. The left panel shows the entire criminal career span defined in this paper (13-64, entry and exit). The right panel zooms in on the 22-64 age range.

distribution of exit rates derived from the imputed age-specific crime rates. Specifically, we compute the fraction changes from crime age cohort a in year t to crime age cohort $a+1$ in year $t+1$ in total violent crimes averaged over all state-years and plot them against age. The positive region in the left panel indicates net entry and confirms our choice of entry window again[136]. The right panel suggests that the exit rate for total violent crimes averages about 8% (without controlling for anything), bottoms out in the early 30's, and increases until retirement. Therefore, Figure 3.4.3 presents empirical evidence for not only our choice of the entry window but also the functional form we use to parametrize the aging effects on base exit rates.

We thus parametrize the average base exit rate β_0 as a quadratic function in age as follows[137],

[136]These fraction changes do not reflect entry probability since the denominators are current criminals but not potential entrants. The graph, however, does suggest aging effects on entry as well but specifying a non-constant entry rate will result in non-linearity of the model in differentiating selected from surprised cohorts. Since the aging effects on entry do not interfere with the identification of other coe¢cients in the model, we only estimate the average entry rate using a constant term.

[137]Derivation of Equation 3.4.2: $\beta_0 N^{Ex}_{st} = \beta_0 \sum_{a=22}^{64} N^{Ex}_{ast} = \sum_{a=22}^{64} \gamma^a N^{Ex}_{ast} = \sum_{a=22}^{64} (\gamma_0 + \gamma_1 a + \gamma_2 a^2) N^{Ex}_{ast} =$

$$\beta_0 N_{st}^{Ex} = \gamma_0 \sum_{a=22}^{64} N_{ast}^{Ex} + \gamma_1 \sum_{a=22}^{64} a N_{ast}^{Ex} + \gamma_2 \sum_{a=22}^{64} a^2 N_{ast}^{Ex} \tag{3.4.2}$$

We then estimate (γ_0, γ_1, γ_2) in place of β_0 in the CPDM with aging effects.

In a dynamic model, the surprise effect only measures the average change in exit rates among the surprised cohort. However, with heterogeneity in proclivity for crime, remaining careers till retirement, etc., the marginal criminals are to be surprised first, with less incumbent criminals to be surprised as time passes after the SIL passage. We therefore expect the surprise effect to be most salient in the immediate years following passages of SILs and to gradually taper off over time. We thus non-parametrically decompose the surprise effects into several floodgate effects over the years succeeding the passage of SILs. Specifically, we let $\beta_2 = \sum_{j=0}^{9+} \lambda_j I_{st}^j$, where I_{st}^j are dummies indicating the j^{th} year after SIL passage. λ_j's then represent the evolution of the surprise effects after the initial passages of SILs.

Combining everything discussed above and building upon the baseline CPDM equation, we arrive at the following estimating equation for CPDM with both aging and floodgate effects.

$$\begin{aligned}\Delta_{st}^C(NetEntry) =& \alpha_0 \sum_{a=13}^{21} N_{ast}^{En} + \alpha_1 \sum_{a=13}^{21} N_{ast}^{En} I_{st}^{SIL} - \gamma_0 \sum_{a=22}^{64} N_{ast}^{Ex} - \gamma_1 \sum_{a=22}^{64} a N_{ast}^{Ex} - \gamma_2 \sum_{a=22}^{64} a^2 N_{ast}^{Ex} \\ & - \beta_1 I_{st}^{SIL} \sum_{a=22}^{64} N_{ast}^{Selected} - \sum_{j=0}^{9+} \lambda_j I_{st}^j \sum_{a=22}^{64} N_{ast}^{Surprised} + \Gamma X_{st} + \epsilon_{st} \end{aligned} \tag{3.4.3}$$

We estimate this equation separately for each crime type as well as the total violent crimes. The dependent variable, net entry, is constructed as the difference between the number of crimes in state s in year $t+1$ and year t weighted by state population in year t, i.e. $\Delta_{st}^C = (C_s^{t+1} - C_s^t)/Pop_s^t$. All cohort variables on the right-hand side are also weighted by state population in year t for consistency. X_{st} include all control variables (state population, population density, real per capita personal income, income maintenance, unemployment insurance, and retirement payment for people older than 65), state and year fixed

$\gamma_0 \sum_{a=22}^{64} N_{ast}^{Ex} + \gamma_1 \sum_{a=22}^{64} a N_{ast}^{Ex} + \gamma_2 \sum_{a=22}^{64} a^2 N_{ast}^{Ex}$

effects, and state-specific linear and quadratic time trends. ϵ_{st} is assumed to be autocorrelated over time within each state.

We also follow the dynamic panel data literature (e.g. Anderson and Hsiao (1982)) and use the lagged variables $N^{Ex}_{as,t+1}$, $N^{Selected}_{as,t+1}$ and $N^{Surprised}_{as,t+1}$ as instruments for all exit cohorts in the model[138]. The additional identifying assumption being made is that crime rates C_{st} follow an $AR(1)$ process over time within each state. We then use two stage least squares to estimate the CPDM.

3.5 Results

In this section we present the estimates of our CPDM, test for the deterrence hypothesis as well as the model specification, further compare DD to the CPDM, and finally decompose the three effects from CPDM in a counterfactual example.

3.5.1 CPDM estimates

Table 3.4 presents estimates of 4 different specifications of the CPDM, with and without the aging and the floodgate effects, on the total violent crimes only.

The baseline model estimates only the three basic effects (direct, selection, and surprise) on top of the base entry and exit rates. Only the CPDM parameters are reported and signed under the deterrence hypothesis in parentheses. The signs of the precisely estimated direct effect α_1 and selection effect β_1 contradict those predicted under the deterrence hypothesis, which we thus strongly reject. The two signs are, however, internally consistent within the model - more entry into violent crimes after SIL passages will lead to higher rate out of the criminal force when it comes to exit - a labor force shakeout. The surprise effect, on the other hand, is estimated to increase exit rates post-SIL for cohorts who became criminals before SIL passages. The older incumbent cohort is still shocked negatively despite the positive reactions of the potential entrants. We thus only find evidence on partial deterrence of SILs on the incumbent criminals.

To interpret the magnitudes of our estimates, we note again that the model is estimated

[138]This is because the exit cohorts are imputed partially with the crime rates.

Table 3.4: CPDM specifications

	Baseline	w/ Aging	w/ Floodgate	w/ Both
Entry α_0	0.0188***	0.0111†	0.0210***	0.0148*
	(0.0145)	(0.2742)	(0.0049)	(0.1418)
SIL Entry α_1 (-)	0.0042***	0.0055***	0.0037***	0.0046**
	(0.0049)	(0.0023)	(0.0054)	(0.0208)
Exit β_0	0.5310***		0.5269***	
	(0.0000)		(0.0000)	
Exit γ_0		51.7108***		51.8624***
		(0.0032)		(0.0027)
Exit γ_1		-3.2453***		-3.2551***
		(0.0024)		(0.0020)
Exit γ_2		0.0483***		0.0484***
		(0.0016)		(0.0013)
Selection β_1 (-)	0.2628**	-0.1023	0.3759***	0.1161*
	(0.0429)	(0.3787)	(0.0179)	(0.1697)
Surprise β_2 (+)	0.1129***	0.1086***		
	(0.0006)	(0.0195)		
Floodgate λ_0			0.1071***	0.0890**
			(0.0002)	(0.0537)
Floodgate λ_1			0.0962***	0.0822**
			(0.0120)	(0.0669)
Floodgate λ_2			0.1091***	0.0886*
			(0.0007)	(0.1531)
Floodgate λ_3			0.0960***	0.0706†
			(0.0020)	(0.2672)
Floodgate λ_4			0.0770***	0.0555
			(0.0132)	(0.4294)
Floodgate λ_5			0.0558*	0.0301
			(0.1262)	(0.7143)
Floodgate λ_6			0.0630*	0.0355
			(0.1309)	(0.7152)
Floodgate λ_7			0.0351	-0.0199
			(0.3907)	(0.8304)
Floodgate λ_8			0.0008	-0.0609
			(0.9880)	(0.5588)
Floodgate λ_{9+}			-0.0050	-0.1172
			(0.9487)	(0.3825)
Log-likelihood	-7694	-7696	-7688	-7682
F-statistics	175.5	138.0	111.6	95.4
Nb. Obs.	1549	1549	1549	1549

Notes: All regressions are run on the total violent crimes. Arrest rates of violent crimes, demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). λ_j's measure the surprise effect in the j^{th} year after SIL passage. Key coefficients relevant for testing the deterrence hypothesis are signed in parentheses. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

with crime rates as proxies instead of actual criminal populations. The dependent variable as well as all the exit cohorts are measured in the number of crimes (all variables are then weighted by every 100,000 state population), while the entry cohorts are measured in the number of potential entrants. Therefore, we have actually estimated the following equation,

$$\begin{aligned}\Delta_{st}^{C}(NetEntry) =&\hat{\alpha}_0\kappa \sum_{a=13}^{21} N_{ast}^{En} + \hat{\alpha}_1\kappa \sum_{a=13}^{21} N_{ast}^{En} I_{st}^{SIL} - \beta_0 \sum_{a=22}^{64} N_{ast}^{Ex} \\ &- \beta_1 I_{st}^{SIL} \sum_{a=22}^{64} N_{ast}^{Selected} - \beta_2 I_{st}^{SIL} \sum_{a=22}^{64} N_{ast}^{Surprised} + \Gamma X_{st} + \epsilon_{st}\end{aligned}$$

where κ is the number of crimes committed by a career violent criminal in a year and assumed to be constant across age, state, and time. Now the $\hat{\alpha}$'s and β's measure the corresponding entry and exit probabilities into and out of the criminal force (since we can divide the equation through by κ). Since we can not separately identify the $\hat{\alpha}$'s from κ due to data limitations, we only roughly interpret the magnitudes of the α's. The estimated α_1 suggests that, if a criminal commits 10 violent crimes a year, we estimate a 0.19% entry probability into violent criminals from the pool of all males between 13-21 in the absence of SIL. On the other hand, without knowing κ, we estimate a 22.3% (= 0.0042/0.0188) increase in this entry probability due to the direct effect of SIL. For exits, in the absence of SIL, violent criminals are estimated to exit with 53.1% probability annually. The estimated selection and surprise effects suggest that the criminal cohort that entered after SIL passages experience an additional 26.3 percentage points in exit probability with SIL due to the dilution in criminal quality from the higher entry rate, while the cohort that entered before SIL passages is surprised and exits with a probability increase of 11.3 percentage points.

Building upon the baseline model, we first introduce the aging effects that parametrize the base exit rate. We find very strong empirical evidence supporting the aging effects on base exit rates for violent criminals. All aging parameters are strongly significant. The resulting parabola of exit rates constructed from these estimates suggests the lowest exit rate around age 34, evidence for the peak of violent criminals' careers as a consequence of aging and human capital accumulations. All CPDM coefficients stay unchanged from the baseline model except for the selection effect. Aging effects take away the significance of the selection effect coefficient due to the differences in average ages across these different cohorts. The previously estimated selection effect is thus an artifact of the fact that the selected cohort is on average much younger, which is now absorbed away by the aging effects.

On the other hand, if we just relax the surprise effect to be flexible over time with the floodgate effects, the estimated CPDM parameters (except the surprise effect) stay almost unchanged from the baseline specification, while the surprise effect gets less precisely estimated over time as cohorts drop out of our sample. We refuse the temptation of re-running the regressions with ex-post cutoffs but only report them in Table C.3 for robustness. The estimated magnitudes of the surprise effect also confirm the theory and taper off over time, capturing the reactions of the older cohort. Combining all of above, we arrive at our preferred specification with both aging and floodgate effects, as stated in Equation 3.4.3 and shown in the last column of Table 3.4.

We maximize the log-likelihood of the total violent crime regression to arrive at the entry window cutoff at age 21. F-statistics of the full model strongly reject null hypotheses that all coefficients of the model (except state and year fixed effects) are zeros and provide measures of the fit of the model.

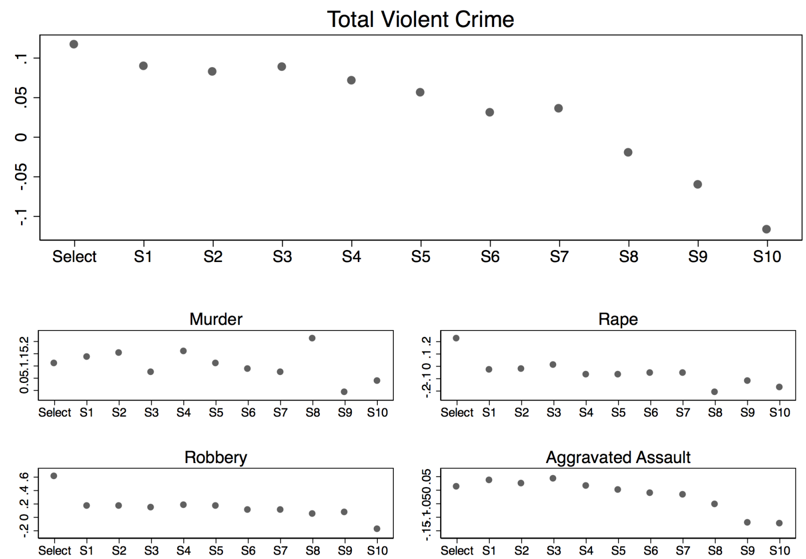
We conduct hypothesis and specification tests in Table 3.5. Here we first formally test that the parabola of exit rates bottom out around age 33.6, statistically significant from zero. We also show that the aging effects and floodgate effects are both jointly significant where applicable. Although it is obvious from the point estimates in Table 3.4 that the deterrence hypothesis ($\alpha_1 > 0$, $\beta_1 < 0$, and $\beta_2 > 0$) will be rejected, we present the formal one-sided hypothesis tests in Table 3.5. Finally, we turn to the specification tests of DD. Specifically, the null hypotheses are the two restrictions in Section 3.2.1.3 that reduce the CPDM to Equation 3.2.5 and Equation 3.2.6. Namely, (1) $\beta_1 = \beta_2 = \beta_*$ and ((2) $\alpha_0 = -\beta_0$, (3) $\alpha_1 = -\beta_*$). Note that when we specify the non-parametric floodgate effects, (1) and (3) require all the floodgate effects to be the same with the selection effect to reduce to DD.

Table 3.5: Hypothesis and specification tests: CPDM specifications

	Baseline	w/ Aging	w/ Floodgate	w/ Both
Turning point		33.6***		33.6***
		(0.0000)		(0.0000)
Joint significance tests				
Aging (γ's)		17.21***		16.90***
		(0.0002)		(0.0002)
Floodgates (λ's)			32.27***	34.72***
			(0.0004)	(0.0001)
Deterrence hypothesis tests				
$\alpha_1 < 0$	0.0042***	0.0055***	0.0037***	0.0046**
(one-sided)	(0.0025)	(0.0012)	(0.0027)	(0.0104)
$\beta_1 < 0$	0.2628**	-0.1023	0.3759***	0.1161*
(one-sided)	(0.0215)	(0.8107)	(0.0090)	(0.0849)
$\beta_2 > 0$	0.1129	0.1086	0.0972	0.0773
(one-sided)	(0.9997)	(0.9903)	(0.9997)	(0.9185)
Diff-in-diff nested specification tests				
(1) $\beta_1 = \lambda_j, \forall j$	1.25	3.03*	15.22†	23.96***
	(0.2628)	(0.0815)	(0.1241)	(0.0077)
(2) $\alpha_0 = -\beta_0$	162.31***	296.04***	175.48***	323.69***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(3) $-\alpha_1 = \beta_1 = \lambda_j, \forall j$	16.02***	6.80**	32.18***	44.15***
	(0.0003)	(0.0333)	(0.0000)	(0.0000)
(2) & (3)	171.75***	313.08***	475.39***	588.90***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Notes: Age of the turning point, F-statistics for the joint significance tests, point estimates of the CPDM parameters, and F-statistics for the DD tests are shown. For specifications with floodgate effects, we replace β_2 with the weighted cumulative surprise effect of the first five floodgate effects, i.e. $\sum_{j=0}^{4} \lambda_j \frac{\sum_i I_{st}^{ij}}{\sum_{i,j} I_{st}^{ij}} > 0$. One-sided p values are in parentheses for the deterrence hypothesis tests. Two-sided p values are in parentheses for the rest. †, *, **, and *** indicate one-sided (deterrence hypothesis tests) and two-sided (rest) statistical significance at the 15, 10, 5, and 1 percent level.

Figure 3.5.1: Estimated selection and floodgate surprise effects



Notes: Point estimates of selection (leftmost on each plot) and floodgate effects from CPDM on violent crimes and sub-categories (Table 3.6).

Bottom of Table 3.5 then shows results that strongly reject the DD specification across all four specifications of the CPDM.

We further estimate our preferred specification on the four sub-categories of violent crimes. Table 3.6 shows the results. We first note that most of the estimated CPDM parameters (with exception of surprise effects in later years) are significant and very consistent across crime types, suggesting much stronger results compared to the existing literature on SILs.

The floodgate surprise effects are again higher and more precisely estimated at the beginning and taper off nicely in later years. The pattern persists across all crime types as well. Figure 3.5.1 plots the floodgate surprise effects against the estimated selection effect (leftmost). The selection effects are generally higher than or equal to the surprise effects, suggesting again against the deterrence hypothesis. The differences between the two effects (particularly in rape and robbery) also imply the misspecification of a DD.

In the same vein of Table 3.5, we show results of formal hypothesis and specification tests on Table 3.4 in Table 3.7. We find the turning points to be statistically significant and

Table 3.6: CPDM with aging and nonparametric floodgate effects: Crime types

	Violent	Murder	Rape	Robbery	Agg. Ast.
Entry α_0	0.0148*	0.0003**	0.0009†	0.0127**	0.0087*
	(0.1418)	(0.0906)	(0.2303)	(0.0379)	(0.1455)
SIL Entry α_1 (-)	0.0046**	0.0001**	-0.0001	0.0012**	0.0015*
	(0.0208)	(0.0931)	(0.5455)	(0.0571)	(0.1625)
Exit γ_0	51.8624***	27.6572***	15.8851**	127.8388***	15.3134***
	(0.0027)	(0.0000)	(0.0626)	(0.0001)	(0.0017)
Exit γ_1	-3.2551***	-1.5197***	-1.1057**	-7.9613***	-1.0533***
	(0.0020)	(0.0000)	(0.0296)	(0.0001)	(0.0002)
Exit γ_2	0.0484***	0.0193***	0.0183***	0.1182***	0.0173***
	(0.0013)	(0.0000)	(0.0122)	(0.0000)	(0.0000)
Selection β_1 (-)	0.1161*	0.1101	0.2250	0.6073**	0.0133
	(0.1697)	(0.8271)	(0.3451)	(0.0978)	(0.8487)
Floodgate λ_0	0.0890**	0.1354*	-0.0322	0.1698**	0.0348
	(0.0537)	(0.1987)	(0.5159)	(0.0880)	(0.3547)
Floodgate λ_1	0.0822**	0.1521*	-0.0224	0.1725***	0.0242
	(0.0669)	(0.1805)	(0.7011)	(0.0163)	(0.5000)
Floodgate λ_2	0.0886*	0.0749	0.0075	0.1423*	0.0407
	(0.1531)	(0.4367)	(0.9043)	(0.1748)	(0.3396)
Floodgate λ_3	0.0706†	0.1601†	-0.0659	0.1748†	0.0158
	(0.2672)	(0.2101)	(0.4377)	(0.2149)	(0.6921)
Floodgate λ_4	0.0555	0.1107*	-0.0671	0.1667†	0.0010
	(0.4294)	(0.1630)	(0.3416)	(0.2820)	(0.9826)
Floodgate λ_5	0.0301	0.0862	-0.0550	0.1114	-0.0115
	(0.7143)	(0.4756)	(0.4754)	(0.5664)	(0.8261)
Floodgate λ_6	0.0355	0.0735	-0.0552	0.1097	-0.0187
	(0.7152)	(0.4470)	(0.4965)	(0.6520)	(0.7663)
Floodgate λ_7	-0.0199	0.2120**	-0.2131**	0.0434	-0.0535
	(0.8304)	(0.1474)	(0.0782)	(0.8584)	(0.3525)
Floodgate λ_8	-0.0609	-0.0091	-0.1217	0.0761	-0.1211**
	(0.5588)	(0.9305)	(0.3146)	(0.7633)	(0.0908)
Floodgate λ_{9+}	-0.1172	0.0378	-0.1722*	-0.1743	-0.1241*
	(0.3825)	(0.7961)	(0.1564)	(0.6609)	(0.1631)
Log-likelihood	-7682	-2286	-4012	-6827	-7088
F-statistics	95.4	672.5	62.3	171.8	122.2
Nb. Obs.	1549	1548	1547	1546	1549

Notes: Arrest rates (of corresponding crime categories), demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). λ_j's measure the surprise effect in the j^{th} year after SIL passage. The F-statistics test for the joint significance of all estimated coefficients and reject the null (all coefficients are equal to zero) in all specifications. Key coefficients relevant for testing the deterrence hypothesis are signed in parentheses. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

consistent across crime types, with murder being slightly higher (39) and rape and aggravated assault lower (30), reflecting the peak of the combination of male physical conditions and criminal skill accumulations. All other tests show similar results across crime types as the total violent crime as shown in Table 3.5.

3.5.2 Comparing DD to CPDM

We further compare DD to our CPDM in this section, in relation to the evolutions of cohorts. Expanding on the Florida example depicted in Figure 3.4.1, Figure 3.5.2 shows the evolutions of average cohort sizes (across states) over time. Again, the total entry cohort measures male population between 13-21 and is stable over time (exogenous to SIL passages). Interacting the entry cohort with SIL passages, the double-solid line exhibits the growth of SIL states as shown in Figure 3.3.1. The total exit cohort again follows the national trend of violent crimes. However, note that the total exit cohort is not the sum of the selected and surprised cohorts nationally as states adopt SILs at different times and some states never do so. The surprised cohort first increases as more states adopt SILs and then starts decreasing in late 1990's as the old criminal cohorts exit without being replenished. Finally, the selected cohort keeps gradually increasing as more states adopt SILs and more new criminals having entered under SILs.

Top of Table 3.8 presents the evolutions of the shares of these cohorts for different sample lengths (LM, AD, and this paper). s^{En} is the share of the entry cohort as a fraction of the total population at risk (the sum of entry and exit cohorts). $s^{*Selected}$ and $s^{*Surprised}$ are defined similarly as in Section 3.2.1.1, as a fraction of the total exit cohort. Note that the share of the surprise cohort is highest in the middle sample due to the dynamics.

Given these evolutions, we then compare the corresponding DD estimates in these different samples with our CPDM. We estimate two standard DD models as follows,

$$C_{st} = \alpha + \beta I_{st}^{SIL} + \Gamma X_{st} + \varepsilon_{st} \tag{3.5.1}$$

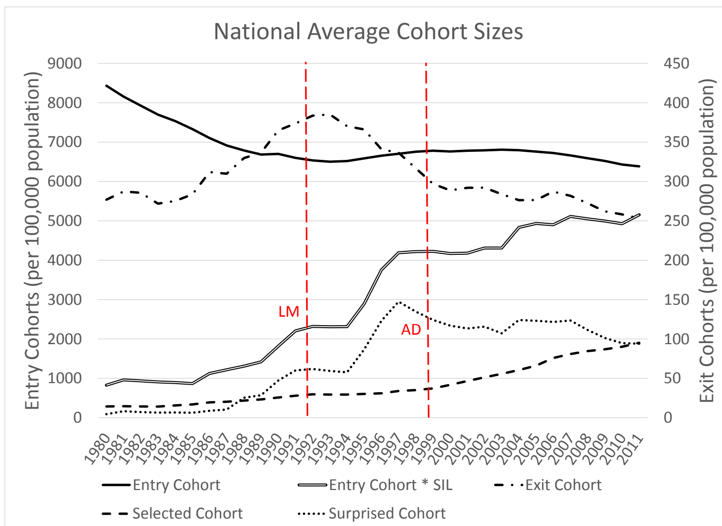
$$\Delta_{st}^{C} = \alpha' + \beta' I_{st}^{SIL} + \Gamma' X_{st} + \varepsilon_{st} \tag{3.5.2}$$

Table 3.7: Hypothesis and specification tests: Crime types

	Violent	Murder	Rape	Robbery	Agg. Ast.
Turning point	33.6***	39.3***	30.2***	33.7***	30.5***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Joint significance tests					
Aging (γ's)	16.90***	51.10***	8.24**	58.91***	25.52***
	(0.0002)	(0.0000)	(0.0162)	(0.0000)	(0.0000)
Floodgates (λ's)	34.72***	7.79	32.43***	32.60***	18.30*
	(0.0001)	(0.6498)	(0.0003)	(0.0003)	(0.0501)
Deterrence hypothesis tests					
$\alpha_1 < 0$	0.0046**	0.0001**	-0.0001	0.0012**	0.0015*
(one-sided)	(0.0104)	(0.0466)	(0.7273)	(0.0286)	(0.0813)
$\beta_1 < 0$	0.1161*	0.1101	0.2250	0.6073**	0.0133
(one-sided)	(0.0849)	(0.4136)	(0.1726)	(0.0489)	(0.4244)
$\beta_2 > 0$	0.0773	0.1268	-0.0360	0.1653	0.0234
(one-sided)	(0.9185)	(0.9052)	(0.2782)	(0.9346)	(0.7363)
Diff-in-diff nested specification tests					
(1) $\beta_1 = \lambda_j, \forall j$	23.96***	15.59†	32.60***	21.51**	17.52*
	(0.0077)	(0.1120)	(0.0003)	(0.0178)	(0.0636)
(2) $\alpha_0 = -\beta_0$	323.69***	216.66***	38.21***	430.67***	74.06***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(3) $-\alpha_1 = \beta_1 = \lambda_j, \forall j$	44.15***	25.25***	32.65***	35.45***	18.47*
	(0.0000)	(0.0084)	(0.0006)	(0.0002)	(0.0712)
(2) & (3)	588.90***	364.62***	82.23***	742.73***	154.53***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Notes: Age of the turning point, F-statistics for the joint significance tests, point estimates of the CPDM parameters, and F-statistics for the DD tests are shown. We replace β_2 with the weighted cumulative surprise effect of the first five floodgate effects, i.e. $\sum_{j=0}^{4} \lambda_j \frac{\sum_i I_{st}^{ij}}{\sum_{i,j} I_{st}^{ij}} > 0$. One-sided p values are in parentheses for the deterrence hypothesis tests. Two-sided p values are in parentheses for the rest. †, *, **, and *** indicate one-sided (deterrence hypothesis tests) and two-sided (rest) statistical significance at the 15, 10, 5, and 1 percent level.

Figure 3.5.2: Evolutions of national average cohort sizes



Notes: Entry cohorts (solid lines) are measured in 100,000 population on the left axis. Exit cohorts (dashed/dotted lines) are measured in 100,000 population on the right axis. Cohorts are averaged across all states. The vertical dashed lines indicate where LM and AD's samples end, respectively.

where Equation 3.5.1 is estimated with levels of violent crimes, while Equation 3.5.2 uses year-to-year changes of violent crimes. I_{st}^{SIL} is the standard multiple-event DD dummy that equals one if state s has SIL in place at time t. X_{st} includes the same set of controls as in our CPDM in Equation 3.4.3. Note that Equation 3.5.2 is the same with Equation 3.2.6 after weighting by the total population. The first two samples highly resemble the data used in LM and AD[139]. The DD specifications in Equations 3.5.1 and 3.5.2 are more general and robust than the "dummy variable model" of LM and the "hybrid model" of AD[140]. We also account for auto-correlated errors by clustering at the state level.

The estimates are shown in the middle panel of Table 3.8. Similar to BDM, we find that most of the effects are essentially zero (with no consistency in signs) after controlling for trends and auto-correlations of errors. We only find significant effects (about 7% reduction in crimes following passages of SILs) with the 1980-1999 sample on the levels of crime rates[141]. The DD estimates reflect the evolutions and offsetting effects of the different cohorts. We have found that the surprise effect increases exit rates and thus decreases net entry rates and levels of crimes - the effect of SIL on crimes is thus dominantly negative when the surprised cohort dominates in the 1980-1999 sample. The reversed trends of the entry and exit cohorts, together with the positive entry and selection effects, also contribute to the negative DD estimate in the 1990's sample. In the full sample, as the exit cohort shrinks with the national trend, the surprised cohort decreases, and the tapering off of the surprise effect over time, we see very weak evidence of positive effects estimated by DD. The DD estimates are also largely insignificant as the entry effects offset the surprise and selection effects.

We then turn to the CPDM estimates of the varying sample lengths in the bottom panel. For comparison, we only show estimates from the baseline CPDM using ordinary least squares[142]. We find strongly significant results with consistency in the estimated signs

[139]We differ with them in data in two ways. Both of their data begin with 1977 while ours is cut off at 1980 due to the availability of the cohort population data. We also estimate a DD from 1977 but only report results from 1980 (which are similar) for comparison with the CPDM. While AD also use state-level panel data, LM uses county-level crime data. We also use state-level data for comparison with CPDM but the DD estimates are similar on the county level as well.

[140]We defer further discussions on the literature to Appendix C.1.2. See Table C.7 and Table C.8 for details.

[141]These results largely contradict with findings of LM and AD. See Appendix C.1.2 for replications of LM and AD, and comparisons of different DD specifications.

[142]For robustness, see Appendix C.1.1.4 for OLS estimates of the full model.

Table 3.8: DD vs. CPDM for different sample lengths

	1980-1992	1980-1999	1980-2011
Average Cohort Sizes			
s^{En}	0.9585	0.9559	0.9574
$s^{En}I^{SIL}$	0.1707	0.2762	0.4230
$s^{Ex} = 1 - s^{En}$	0.0415	0.0441	0.0426
$s^{*Selected}$	0.4727	0.3235	0.3585
$s^{*Surprised}$	0.5273	0.6765	0.6415
Diff-in-Diff in Levels			
SIL Dummy	-0.8789	-38.0306*	0.8612
	(0.9550)	(0.0582)	(0.9580)
Diff-in-Diff in Changes			
SIL Dummy	-1.6291	-0.2657	2.6447
	(0.8881)	(0.9758)	(0.6256)
Baseline CPDM			
Entry	0.1626*	0.0091	0.0190**
	(0.0949)	(0.7368)	(0.0107)
SIL Entry	0.0017	0.0028*	0.0041***
	(0.3214)	(0.0538)	(0.0015)
Exit	1.0768***	0.3946***	0.3232***
	(0.0000)	(0.0003)	(0.0000)
Selection	0.2604	0.3174*	0.1943†
	(0.6146)	(0.0595)	(0.1342)
Surprise	0.0381	0.1054***	0.1067***
	(0.5829)	(0.0000)	(0.0001)
Nb. Obs.	657	994	1549

Notes: All regressions are run on the total violent crimes. All regressions are run using OLS. Arrest rates of violent crimes, demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate two-sided statistical significance at the 15, 10, 5, and 1 percent level.

across different sample lengths. In the shorter samples, the CPDM also struggles to precisely estimate base exit rates (column 1) and base entry rates (column 2), which may bias the dynamic selection and surprise effects slightly upwards due to the aging effects of exit. Despite of this, the CPDM also consistently estimates the direct entry effect across all samples, which, together with the consistently estimated signs of other parameters, yields the most important policy implications.

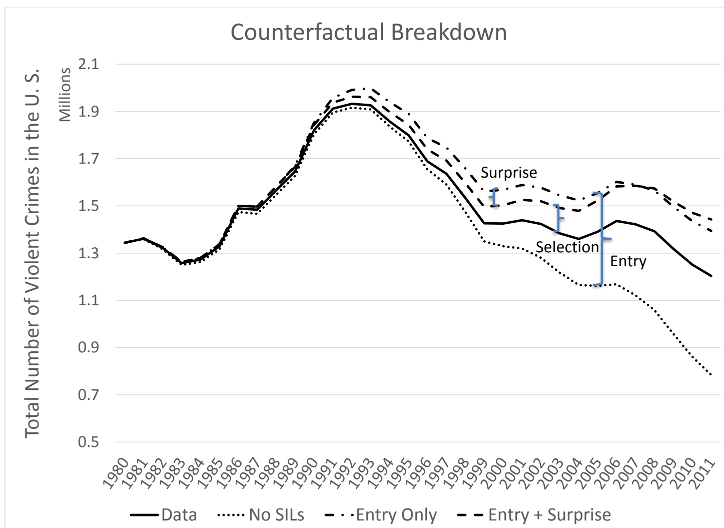
3.5.3 Counterfactual example

In order to make direct policy evaluations with the CPDM, accounting for the entry, selection and surprise effects, we consider the following counterfactual example. In this example, we eliminate SILs from all states and compute the counterfactual crime levels in the U.S. had we never adopted SILs. To do so, we start with crime rates in 1980 at the beginning of our sample, let the CPDM predict changes in crimes from year to year for all states while shutting down all post-SIL effects (entry, selection, and surprise), and then simulate crime levels for all states in all following years. The result is shown in Figure 3.5.3. The actual data (solid line) shows that violent crimes totaled at 1.3 million in the U.S. in 1980, peaked at 1.9 million in 1992, and settled at 1.2 million in 2011. When we take away the effects of SILs (dotted line), we find a drop in violent crimes that shows the dynamic properties that the CPDM captures. After eliminating SILs, the counterfactually predicted crime rates track the actual crime rates very closely for 2/3 of the sample and only diverge in the last 1/3, although by year 2000, 3/4 of the states have already adopted SIL. For example, in 1995, the counterfactual prediction only shows a 1.4% (about 26000 crimes annually) reduction in crime levels. By 2011, there is a large reduction of 34.8% (or about 419000 crimes) in total violent crimes[143].

We then further decompose this gap between the levels of crimes into the three effects captured by CPDM. From the dotted line where there are no post-SIL effects, we first add back only the direct entry effect (dash-dotted line). Graphically, the entry effect is positive and significant, driving up the total violent crime level to about 1.4 million in 2011. Adding on top of that the surprise effects (dashed line), which increase exit rates in the first few years following SIL passages and taper off after, shifts down the overall curve but dissipates at the end of the sample. Finally, the remaining gap between the dashed line and the solid line represents the selection effect, which captures the increased exit rates from the lesser criminals who entered post-SIL. As expected, this gap keeps widening over time as

[143] We interpret the large drop as an upper bound for the amount of crime reductions if SILs were eliminated. The reason is that, although we have eliminated all post-SIL effects in the counterfactual simulation, we keep the stock of criminals (i.e. base exit cohorts) constant. With lower entry rate absent SILs, we should see a smaller stock of criminals and consequently less exits as well, which would shift up the dotted line. We ignore this second-order effect in this exercise.

Figure 3.5.3: Decomposition of entry, selection, and surprise effects



Notes: Counterfactual national violent crime levels if no SILs were ever adopted. Brackets indicate magnitudes of individual entry, selection, and surprise effects, decomposed in contribution to the total effect of SILs on crimes. Estimates obtained from OLS regressions.

the younger cohorts replace their older counterparts.

3.6 Conclusions

In this paper, we use a more general cohort panel data model to bring a consistent and unified answer to the debate of the effects of shall-issue laws on violent crimes. The CPDM incorporates dynamic decision-making by forward-looking agents through the estimation of (i) a direct effect of SIL passages on entry (into violent crime careers), (ii) a selection effect on exit for those who entered the violent crime under SIL, and (iii) a surprise effect on exit for those who entered prior to the advent of SIL. We find all three effects to be positive - suggesting that in addition to the deterrence effect on existing criminals (who entered before SIL), the passages of SIL also substantially lower the barrier of entry for new potential criminals. The combined effect is large - eliminating all passed SILs from the beginning would reduce total violent crimes by about one third by 2011.

We further show that in contexts where heterogeneous agents make forward-looking decisions the standard DD is a model misspecification due to the lack of dynamic considerations. Our CPDM reduces to the standard DD with restrictions that shut down the three effects. The estimated coefficients strongly reject such restrictions and thus rule the DD as misspecified. We then compare the CPDM and DD estimates on samples with varying lengths corresponding to the literature (LM and AD). We find that the DD estimates fluctuate systematically based on the evolutions of cohort shares - leading to the heated debate in the literature. The CPDM, on the other hand, yields consistent and highly significant results across different sample lengths.

Appendix A

Appendix to Chapter 1

A.1 Stylized model of product life cycle formation

In this section, I present a stylized model to show that bell-shaped product life cycles endogenously arise in technologically progressive markets under mild conditions. The model is based on a standard Logit demand system and only assumes that (1) the quality frontier of the market expands, and (2) the production cost of a new product decays over time.

I first set up the model. The industry evolves over continuous time t, although all agents behave myopically. There is a one-dimensional continuous product space J (quality spectrum). The product space is exogenously filled with single-product firms j (generalized to multi-product firms below) at all times, occupying the $[q_j - \frac{1}{2}dq, q_j + \frac{1}{2}dq]$ portion of the product space.

Two technologies are developed exogenously outside this industry. One is the frontier of the quality spectrum, below which the continuum of products (or firms) lives, represented by $Q(t)$, i.e., $\forall j \in J$, $q_j \in [\underline{Q}, Q(t)]$. I assume that the quality frontier increases over time due to upstream product innovations, i.e., $\frac{dQ}{dt} > 0$. The other technology is the cost of production, $C(t, q)$, which is assumed to fall over time given cheaper component costs, i.e., $\frac{\partial C(t,q)}{\partial t} < 0$.

The demand system follows a standard Logit model. Consumer i's utility in purchasing product j at time t is given by

$$u_{ijt} = \beta q_j - \alpha p_j(t; q_j) + \epsilon_{ijt},$$

where $\beta > 0$ is consumers' marginal utility for the quality of the product, $\alpha > 0$ is the price sensitivity, and ϵ_{ijt} is distributed *i.i.d.* Type I Extreme Value. Prices are endogenously set to maximize profits in this market,

$$\overrightarrow{p} \in arg \max_p (p - C(t,q))s(p,q).$$

Now, to show the bell shape of product life cycles, I will solve for the market share of each product as a function of time and show that under very mild conditions, this function is increasing in the interval after release and decreasing after. Integrating out the Logit preference shocks, the market share of any product j at time t is expressed as

$$s_j(t; q_j) = \frac{exp(\beta q_j - \alpha p(t; q_j))dq}{1 + \int_{\underline{Q}}^{Q(t)} exp(\beta\omega - \alpha p(t; \omega))d\omega}.$$

Since prices are endogenously determined in a Logit model with continuous product space, as $dq \to 0$, markups become constant from the first-order conditions of the pricing game[144]:

$$s_j + (p_j - C_j)\frac{\partial s_j}{\partial p_j} = s_j - (p_j - C_j)\alpha s_j = 0 \Rightarrow p_j = C_j + \frac{1}{\alpha},$$

and we can thus express the share function in terms of production costs over time,

$$s(t,q) = \frac{exp(\beta q - \alpha C(t,q) - 1)dq}{1 + \int_{\underline{Q}}^{Q(t)} exp(\beta\omega - \alpha C(t,\omega) - 1)d\omega}.$$

To arrive at closed-form solutions, I first make a few simplifying functional form assumptions, which will subsequently be relaxed in the numerical simulations. I first assume that the production cost function is independent of the quality of the products, i.e., $C(t,q) = C(t), \forall q$ (Assumption I). This assumption significantly simplifies the share func-

[144]Therefore, the following analysis also directly applies to multi-product firms as cross-price elasticities approach zero and product ownership becomes irrelevant.

tion to the following:

$$s = \frac{\beta e^{\beta q} dq}{\beta e^{\alpha C+1} + e^{\beta Q} - e^{\beta \underline{Q}}},$$

and its derivative with respect to t,

$$\frac{\partial s}{\partial t} = \frac{e^{\beta q} dq}{(\beta e^{\alpha C+1} + e^{\beta Q} - e^{\beta \underline{Q}})^2} \left(-\alpha \beta e^{\alpha C+1} \frac{\partial C}{\partial t} - \beta e^{\beta Q} \frac{\partial Q}{\partial t} \right),$$

and therefore

$$sign\left(\frac{\partial s}{\partial t}\right) = sign\left(-\alpha e^{\alpha C+1} \frac{\partial C}{\partial t} - e^{\beta Q} \frac{\partial Q}{\partial t}\right).$$

Now, under Assumption I, it is clear that the countervailing forces of the decreasing production costs and the expanding quality frontiers could drive the bell shape of product life cycles. Formally, if I additionally assume that the decay of production costs eventually has to taper off, i.e., $\lim_{t\to\infty} C = \underline{C}$ and $\lim_{t\to\infty} \frac{\partial C}{\partial t} = 0$, then it is always the case that $\exists \bar{t}$ such that $\forall t > \bar{t}$, $\frac{\partial s}{\partial t} < 0$, or, in other words, the market share of the product eventually also has to decrease. Now for the time path of sales to be bell-shaped or nonmonotonic, we simply need $-\alpha e^{\alpha C+1} \frac{\partial C}{\partial t} > e^{\beta Q} \frac{\partial Q}{\partial t}$ for some t. If, for example, let $C(t) = C_0 - At$ and $Q(t) = Q_0 + Bt$, then we have a unique peak of the product life cycles at $t^* = \frac{ln(\frac{\alpha A}{B}) - (\beta Q_0 - \alpha C_0 - 1)}{\alpha A + \beta B}$.

I now slightly relax Assumption I. Let's instead assume that the production cost is linear in quality at any point in time, i.e., $C(t,q) = c(t)q$ (Assumption II). Without loss of generality, I assume that $\underline{Q} = 0$ and let $\tilde{c} = \beta - \alpha c$. Then I can rewrite the market share as

$$s = \frac{\widetilde{c} e^{\widetilde{c} q} dq}{e\widetilde{c} + e^{\widetilde{c} Q}},$$

and its derivative with respect to t,

$$\frac{\partial s}{\partial t} = \frac{e^{\widetilde{c} q} dq}{(e\widetilde{c} + e^{\widetilde{c} Q} - 1)^2} \left([(e\widetilde{c}^2 + \widetilde{c} e^{\widetilde{c} Q} - \widetilde{c})q - \widetilde{c} Q e^{\widetilde{c} Q} + e^{\widetilde{c} Q} - 1] \frac{\partial \widetilde{c}}{\partial t} - \widetilde{c}^2 e^{\widetilde{c} Q} \frac{\partial Q}{\partial t} \right),$$

and therefore,

$$sign\left(\frac{\partial s}{\partial t}\right) = sign\left([(e\widetilde{c}^2 + \widetilde{c} e^{\widetilde{c} Q} - \widetilde{c})q - \widetilde{c} Q e^{\widetilde{c} Q} + e^{\widetilde{c} Q} - 1] \frac{\partial \widetilde{c}}{\partial t} - \widetilde{c}^2 e^{\widetilde{c} Q} \frac{\partial Q}{\partial t} \right).$$

It is now harder to draw conclusions about the shape of product life cycles based on the primitives of the market, even under the functional form Assumption II. I therefore resort to numerical simulations to further explore the shapes of and variations in product life cycles generated by this model. In the numerical simulations below, I will also be able to relax the continuous time and product space assumptions.

Numerical simulations and comparative statics

In my empirical model of product portfolio choices with product life cycles in Section 1.3, managers form beliefs about the total lifetime profitability of a new product, based on how patterns of product life cycles correlate with characteristics of the product and the market at the time of its release. These profit-maximizing product managers, however, need not to care about the actual shape of product life cycles. This section provides intuition for the formation of product life cycles in high-tech markets, and why the eventual lifetime profitability of a product systematically correlates with its quality, and the competitiveness of the market at the time of its release, via numerical simulations.

I first set up the numerical setting. The product space is discretized with $dq = 0.2$, or 26 starting products within the initial quality spectrum of $[0, 5]$. The industry is simulated forward for 50 periods. The quality frontier grows linearly ($Q(t) = 5 + 0.2t$) and the per-quality production cost exponentially decays, so that it halves every 18 periods ($c(t) = 1.3 \times \left(\frac{1}{2}\right)^{\frac{t}{18}}$). Production (marginal) cost is linear in the quality of products ($C(t, q) = c(t)q$). Demand parameters: $\beta = 1$ and $\alpha = 1$. Equilibrium prices are numerically solved (instead of constant markups).

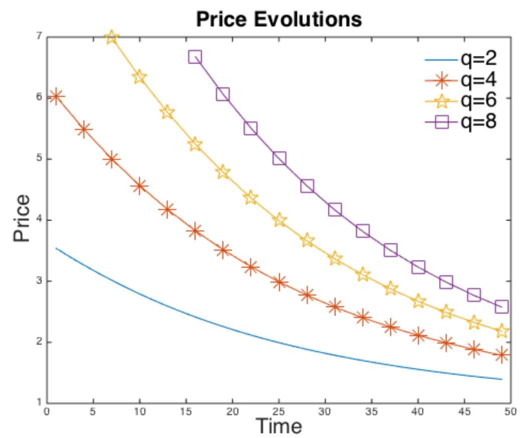
Figure A.1.1 illustrates the intuition behind product life cycle formation in technologically progressive markets, as discussed in Section 1.2. In Figure A.1.1a, production costs for different product-quality levels fall at different speeds over time, which are translated into speeds at which prices fall for different-quality products (while markups are endogenously determined in this model, differential cost reductions still significantly shift the prices—this is especially true if the number of products is large, and markups do not vary much).

Figure A.1.1b then shows that differential price paths are mapped onto consumers' mean utility. As the mean utility of a product increases more quickly than most products in the market at the beginning of its release (because that is when its cost falls most quickly), its market share rises accordingly. As the speeds at which products' costs decline converge, their market shares also converge to their respective levels. When the quality frontier expands, and more and better products are introduced, the market share of existing products will eventually vanish. Figure A.1.1c shows the resulting bell-shaped product life cycles.

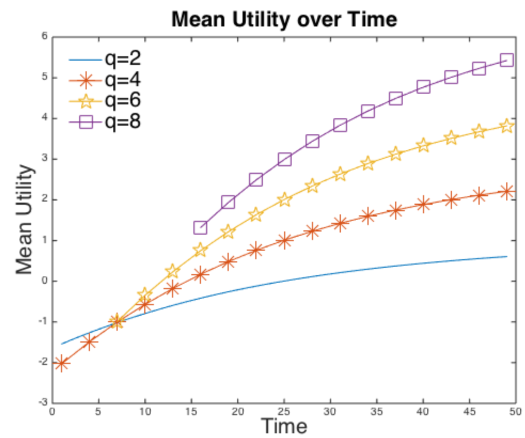
I now turn to comparative statics of product life cycles with respect to characteristics of the product and the market. As evidenced in Figure 1.2.8, lifetime sales of a product accumulated over its product life cycle systematically correlate with its quality and market competitiveness at release time. With the stylized model in hand, I illustrate these comparative statics with simulated product life cycles in Figure A.1.2. To isolate the effects of product quality and market competition on product life cycles, I fix the quality frontier, i.e., $Q(t) = 5$, so that no new products are introduced to change the market structure. Figure A.1.2a shows the product life cycles for products of different quality in the same market over time: While low-quality products might have higher short-run sales when first-released, given their low prices, their product life cycles taper off very quickly; on the other hand, higher-quality products have much higher potential in their future sales, with higher peaks and slower tapering off in their product life cycles. Figure A.1.2b shows how product life cycles vary with the level of market competition (number of equally spaced products in simulation) for the same product in different markets: Increased competition not only reduces the product's short-run sales, but also its lifetime sales—to a much larger extent—by shrinking its product life cycle (i.e., taper off faster).

Figure A.1.1: Product life cycle formation: Intuition

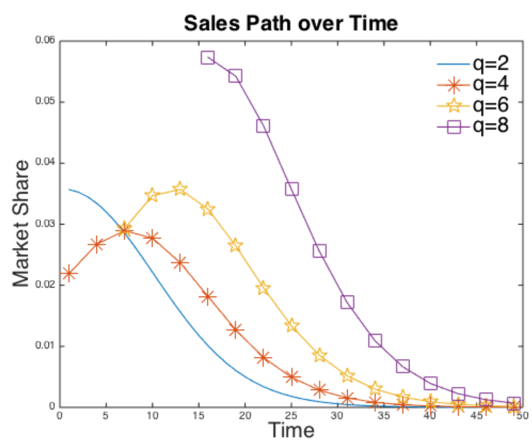
(a) Prices



(b) Mean utility



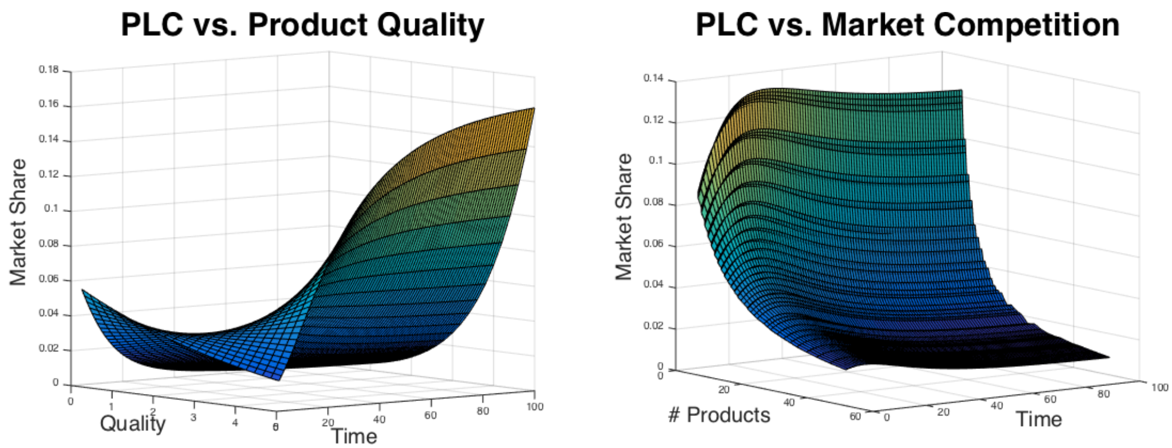
(c) Market shares



Notes: This set of figures shows the intuition behind product life cycle formation in technologically progressive markets. Roughly, differences in cost reduction between different product-quality levels drive differential price paths shown in Figure A.1.1a, which are then mapped onto mean utility in Figure A.1.1b. These differential trends in mean utility cause newly released products' market share to rise, as their prices fall faster, and eventually vanish as they become obsolete, resulting in bell-shaped product life cycles shown in Figure A.1.1c. Plotted are numerical simulation results.

Figure A.1.2: Product life cycle: Comparative statics

(a) Product quality **(b)** Market competition



Notes: This figure shows how product life cycles correlate with product quality and market competition. In particular, Figure A.1.2a shows that, in the same market, a higher-quality product might have lower short-run sales, given its high release-price, but eventually much larger lifetime sales, because its product life cycle peaks higher, and tapers off much more slowly. On the other hand, Figure A.1.2b shows that, for the same product, increased market competition not only reduces the product's short-run sales, but, to a much larger extent, its lifetime sales, by shrinking its product life cycle. Plotted are numerical simulation results.

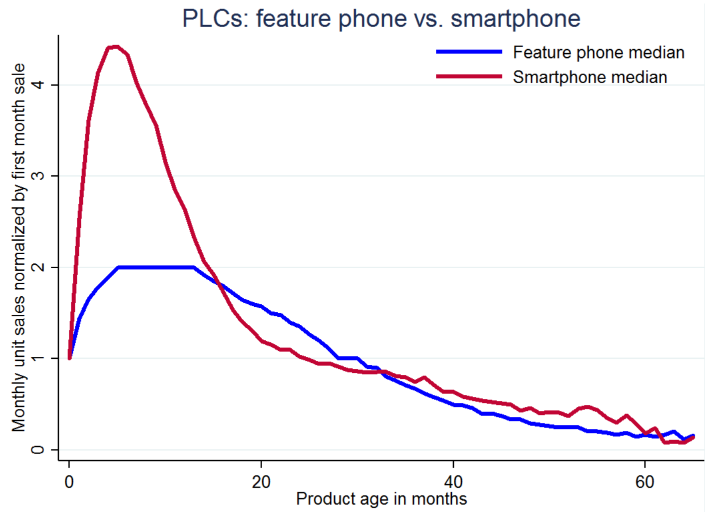
Evidence: Comparing mature vs high-tech industry

The model of product life cycle formation presented in this section relies on the assumption of decreasing production costs and expanding quality frontiers. In other words, according to the model, product life cycles should be more salient in technologically progressive markets and much flatter in more mature industries, in which the extent to which production cost falls and quality frontier expands is small. This also justifies managers' use of fixed hurdle rates in approximating dynamic payoffs in those industries, as pointed out in the literature (see Section 1.1). While a formal test of this theory must be left for future research that uses cross-industry data on product life cycles[145], this section takes a first step in that direction by presenting evidence with one example.

Specifically, I make use of the separate feature phone handset data from GfK, which offers the same level of detail as the smartphone data used in this paper. I compare the

[145]One that is similar to Polli and Cook (1969) in spirit.

Figure A.1.3: PLCs in mature vs. high-tech industry



Notes: This figure shows that, between two otherwise similar industries, the more mature (feature phone) industry exhibits much flatter product life cycles, compared to the high-tech (smartphone) industry. Plotted from monthly mobile phone handset sales from China between Jan. 2009 and Nov. 2014; pooled across provinces and products; aligned by the release month of each product in each province. Monthly unit sales are normalized by release-month sales; the blue (red) line shows median feature phone (smartphone) sales within age (in months) cohorts.

observed product life cycles of feature phone handsets with those of smartphones. The advantage of this comparison is that across industries, feature phones are very close to smartphones in most aspects, except for the maturity of its technology. Figure A.1.3 shows this comparison of the median product life cycle of smartphone handsets (as seen in Figure 1.2.7) and that of the feature phones; both are from the Chinese market between 2009 and 2014, and both are normalized by first-month sales. The peak of the median smartphone PLC is more than twice as large as that of feature phones, which only reaches two times initial sales. Figure A.1.3 thus lends some support to the theory presented in this section and the different uses for product life cycles and hurdle rates across different industries.

A.2 Appendix: Data

A.2.1 Other players and related industries

There are several other players in this market that, although important in certain aspects[146], are not crucial to my study of product choice, and thus I abstract away from them. Most of the hardware innovations are driven by the chipsets used in smartphones. Chipmakers[147] thus often establish exclusive contracts with major smartphone manufacturers for their top-of-the-line chipsets. However, this is not an issue, given my focus on non-flagship products, which typically feature older models of chipsets that are available to all firms. I also do not explicitly model the behaviors of the carriers and retailers. According to industry reports[148], channels on average require 25% margins from handset manufacturers. In my empirical work, I deduct channel margins from the profits of the manufacturers accordingly. Finally, I also ignore other more inconsequential players, such as contract factories (original equipment manufacturers), design houses (original design manufacturers)[149], and other operating system and software developers. The existence of these parties, to some extent, justifies my assumptions of no capacity constraints or economies of scope in the estimation.

A.2.2 Product definition

Defining what constitutes a product is often not easy[150]. This is especially the case for consumer electronics, where some specifications can be tweaked/customized across different models without much costs to manufacturers. In this paper, I first collapse handset models with different customization, such as storage space, and compatibility with different carrier networks[151]. Moreover, smartphone manufacturers sometimes release a re-

[146] Sinkinson (2014) and Yang (2016) explore these aspects in the US smartphone market.

[147] The three most dominant chipmakers in this market are MediaTek, Qualcomm, and Spreadtrum.

[148] Nomura Global Markets Research, "China Smartphone chips: LTE changes the balance," https://www.nomura.com/events/china-investor-forum/resources/upload/China_Smartphone_chips.pdf, accessed March 27, 2016.

[149] These firms help smartphone manufacturers design their bill of materials (BOM). In other words, they come up with the list of components for each handset, given the needs of the smartphone manufacturer.

[150] See Kaplan and Menzio (2015), where price dispersions vary significantly depending on how broadly products are defined.

[151] GfK data are on the model level, and therefore do not differentiate on industrial design, such as the color of the phone.

branded older model in the same product line. There are many reasons firms do this, from a product-line management perspective; I do not focus on this aspect in this paper, and collapse these models based on their characteristics. I also winsorize products by dropping extreme specialty phones[152]. The result is a sample of 691 major smartphone products, down from 1,782 models in the original sample.

Before I detail the procedures used to identify different versions of the same product in this section, it is worth first discussing why they exist. Firms in this market introduce multiple versions of the same handset for several potential reasons. The first is technical. Due to differences in the networks of the three carriers as well as upgrades from the second to the third and then to the fourth generation of networks, smartphone manufacturers often have to introduce multiple versions of the same handset to capture consumers of each carrier. The timing of these releases is sometimes coordinated, but often different. The second reason is product customization. For example, once the first model is introduced, additional set-up costs for variations in the model with different storage space is usually minimal. On the benefit side, these variations of the original model help firms further segment the market or price discriminate. The third reason relates to product life cycle management. As pointed out by Enis, La Garce and Prell (1977), once a product reaches the plateau of its life cycle, additional introduction of model tweaks helps extend the time the product stays on the plateau before dropping off. While I acknowledge the importance of all of the above potentially strategic firm behaviors, this paper does not focus on these aspects, and will thus simply collapse different versions of the same product.

The procedure is as follows. I define a product B to be a "copycat" of a product A if the following criteria are met: 1) they are in the same product line; 2) B is released after A; 3) among all the products that satisfy 1) and 2), the characteristics of A are closest (metrics defined below) to those of B, with distance D; and 4) D is less than some threshold T. This problem resembles a clustering algorithm, as copycats defined here are simply clusters of products with very similar characteristics. A standard issue in clustering is the chains

[152]For example, I drop ultra-luxury phones targeted as high-end gifts, such as the Samsung Clamshell series, which sell for more than 8000 RMB (compared to initial release prices of iPhones at around 6000). I also drop phones cheaper than 600 RMB (release price). Finally, I drop specialty models, such as the Nokia 808 Pureview, which features a 41 mega-pixel camera.

of neighbors. The algorithm used in this paper is similar to the Jarvis-Patrick algorithm (Jarvis and Patrick, 1973) with one shared neighbor threshold. Finally, the distance metric is a standard fraction Euclidean distance with the threshold chosen at 0.2 empirically, i.e.,

$$D_{AB} = \sqrt{\frac{1}{6} \sum_{k=1}^{6} \left(\frac{log(x_B^k) - log(x_A^k)}{SD(log(x^k))} \right)^2}$$

where the six characteristics include camera resolution, display size, display resolution, thickness of the phone, CPU clock speed, and battery capacity; each characteristic is weighted by its own standard deviation across all models before the summation is taken.

A.2.3 List of potential products

Table A.1 presents the list of all potential products defined in Section 1.5 in the counterfactual analysis.

A.3 Computational details

A.3.1 Demand

To evaluate the integral in equation (1.3.2), I first construct the empirical distributions of income in each market. I observe the average income of each quintile of the income distribution among urban and rural residents in each province/year. I assume that income is distributed log-normally in each province/year (urban and rural separately), and estimate the mean and standard deviation of each distribution using Simulated Method of Moments.

Then, for each market (province/month), I assume that income distributions remain constant within the calendar year. I then draw 40 consumers from each market's estimated income distribution, 20 urban and 20 rural. The integral in equation (1.3.2) is then evaluated with 20 Gauss-Hermite quadrature points each for urban and rural residents in a market. In particular, the corresponding quadrature weights are weighted by the ratio of urban/rural population in the market.

Table A.1: Jan. 2013 potential products, major characteristics

Brand	Product	CPU clock speed (GHz)	Display size (inch)	Camera resolution (MP)
Samsung	Galaxy Grand	1.2	7.99	5
	Galaxy Style	1.2	5.04	4.3
	Galaxy Infinite	1.2	3.15	4
	Google Nexus	1.2	4.92	4.6
Huawei	Ascend W1	1.2	5.04	4
	Ascend Y310	1	3.15	4
	C8813	1.2	5.04	4.5
Lenovo	A590	1	3.15	5
	LePhone 2802	1	3.15	4
	LePhone 2908	1.2	3.15	4.5
Coolpad	5876	1	3.15	4.5
	7251	1.2	5.04	5
	8070	1	3.15	4
ZTE	Grand S	2.3	12.78	5
	U816	1.2	3.15	4.5
Oppo	R2007	1.3	5.04	4.7
Vivo	S11	1	5.04	4.3
Market (Jan. 2013)	Median	0.9	5.04	3.5
	Max	2	12.83	5.5

Notes: This table shows that potential products defined in Section 1.5 are of significant quality compared to existing products in the market at the time.

The same procedure follows in evaluating the integral in equation (1.3.4).

A.3.2 Point estimates of maintenance and introduction costs

To arrive at point estimates for maintenance and sunk introduction costs, I minimize the objective function in equation (1.4.4). The upper (UB_{jm}) and lower (LB_{jm}) bounds are pre-computed from equation (1.4.1), equation (1.4.2), equation (1.4.7), and equation (1.4.8), and provide identification for the point estimates. I will use maintenance costs for illustration. Procedurally, I pre-draw $s = 100$ draws of cost shocks η^s_{jmt} from an *i.i.d.* $N(0, \sigma_\eta)$ distribution, where σ_η remains to be estimated. In practice, I construct the objective function on the log-scale, i.e., $F(\theta, X_{jm}|\eta^s_{jmt}) = \theta^F X_{jm} + \eta^s_{jmt}$, where $\theta = (\theta^F, \sigma_\eta)$. Similarly, both UB_{jm} and LB_{jm} are converted to log-scale, as in Figure 1.4.2. I then obtain the point estimates shown in Table 1.7 by solving $(\theta^F, \sigma_\eta) \in argminQ(\theta)$.

A.3.3 Counterfactual

I first describe the procedure used in computing the equilibrium of my model with PLCs (last column in Table 1.13). PLC beliefs are first computed according to equation (1.3.8) and based on new market characteristics after removing the fringe. The portfolio game in equation (1.3.9) is solved using a Gauss-Seidel iterated best-response algorithm. I solve the game for 100 simulations for each of the 31 markets. In each simulation, I first randomize the move orders of the 7 firms in consideration and draw the entry cost shocks for each of the 17 potential products from the estimated distribution of the shock. I then iterate through the firms based on the random order drawn in that simulation. In each iteration, one firm computes its payoffs for each of its 2^n possible strategies (where n is the number of its potential products) according to equation (1.3.9). Note that each of its strategies defines a product market configuration (given the other firms' current portfolio strategies, which are initiated at all 17 products being introduced, as well as all incumbent products). For the firm to compute its payoffs under that strategy, a new vector of equilibrium prices for all products (including incumbents) has to be re-computed according to equation (1.3.4). The firm in consideration in this iteration then chooses the strategy that yields the highest payoff and updates the strategy vector (of length 17, filled with zeros and ones). In the next iteration, the next firm observes the updated strategy vector and makes its decisions accordingly. The algorithm keeps iterating over firms and terminates when the strategy vector ceases to change. I record the final strategy vector and report the average in the first row of Table 1.13. Other equilibrium outcomes can then be computed based on the product market configuration defined by the final strategy vector.

To solve the model without PLCs, the game solution algorithm is identical, except that when firms compute equation (1.3.9) for each of its strategies, I keep PLC beliefs fixed at the level at which fringe firms were present. For the column with product introductions fixed at the observed level, I do not solve the portfolio game, but simply use observed product configurations to compute the other equilibrium outcomes.

Appendix B

Appendix to Chapter 2

B.1 Appendix: Data

B.1.1 Sample construction and primary variables

In this appendix, we describe how we construct the sample of DMAs and TV stations used and discuss several details of the data sources we rely on. Our objective is to infer a TV station's reservation value going into the auction from its cash flow or population coverage scaled the appropriate multiple. While the auction is scheduled for March 2016, we infer a TV station's reservation value as of 2012 as the latest year of availability for both the BIA and NAB data. Our analysis is further made difficult by the fact that different data sources cover different TV stations.

B.1.1.1 DMAs

The U.S. is divided into 210 DMAs. DMAs are ranked annually according to market size as measured by the total number of homes with at least one television (henceforth, TV households, measured in thousand). Table B.1 lists the top ten DMAs in 2012 along with some characteristics from the BIA data.

Table B.1: Top ten DMAs (2012)

Rank	DMA	TV Households	Station Count	Income ($)
1	New York, NY	7,388	19	49,518
2	Los Angeles, CA	5,570	24	36,972
3	Chicago, IL	3,493	18	40,500
4	Philadelphia, PA	2,993	19	42,034
5	Dallas-Ft. Worth, TX	2,571	17	37,215
6	SF-Oakland-San Jose, CA	2,507	18	53,448
7	Boston, MA	2,380	16	48,294
8	Washington, DC	2,360	11	49,495
9	Atlanta, GA	2,293	13	33,726
10	Houston, TX	2,185	14	40,704

Notes: Includes all auction-eligible commercial full-power and low-power (class-A) TV stations. Income is average per capita disposable personal income. Source: BIA.

Table B.2: TV station counts by power output and type of use and service (2012)

	Type of Use and Service						
	Commercial		Non-commercial				
	UHF	VHF	UHF	VHF	UHF	VHF	Total
Full-power							
Primary	950	292	281	104	1,231	396	1,627
Satellite	57	55	0	0	57	55	112
Low-power (Class-A)	376	42	8	1	384	43	427
Total	1,772		394		1,672	494	2,166

Notes: Only stations that are eligible for participation in the incentive auction included. Primary stations denote the owner's main station in the DMA. Satellite stations are full-power relay stations re-broadcasting for the primary stations. Non-commercial stations carry educational or public broadcast programming.

B.1.1.2 TV stations

Table B.2 shows counts of auction-eligible TV stations as of 2012, broken down by power output, type of use, and type of service. There are a total of 2,166 auction-eligible TV stations. We focus on the 1,672 UHF stations that the FCC includes in its repacking simulations.

B.1.1.3 BIA data

After restricting to full-power (primary and satellite) and low-power (class-A) stations, the BIA data provides us with 24,341 station-year observations from 2003 to 2013. Commercial stations make up 19,595 observations and non-commercial stations, including dark stations, for 4,746 observations.

For commercial stations, advertising revenue is missing for 6,058, or 30.9%, station-year observations. Table B.3 shows the share of station-year observations with missing advertising revenue for commercial stations. Advertising revenue is missing for almost all satellite stations because BIA subsumes a satellite's advertising revenue into that of its parent primary station.[153] Missing values are further concentrated among low-power (Class-A) stations, among stations affiliated with Spanish-language networks (Azteca America, Independent Spanish, Telemundo, Unimas, and Univision) and other minor networks, and among independent stations. There are no discernible patterns in missing values along other dimensions of the data such as the market size.

We impute advertising revenue for commercial stations where it is missing by regressing the log of advertising revenue (in \$ thousand) $\ln AD_{jt}$ on station, owner, and market characteristics X_{jt}. We run this regression separately for each year from 2003 to 2013 and use it to predict advertising revenue AD_{jt}. We include in X_{jt} the log of the station's population coverage (in thousand), an indicator for whether the TV station has multicast sub-channels, power output fixed effects (primary and class-A), fixed effects for the eleven affiliations in Table B.3, fixed effects for the interaction of affiliation groups (see Section B.1.2.1) with U.S. states, an indicator for whether the owner owns more than one TV station in the same DMA, ownership category fixed effects (whether the owner owns one, between two and ten, or more than ten TV stations across DMAs), the number of TV stations in the DMA, the number of major network affiliates in the DMA, the wealth and competitiveness indices for the DMA (see Section B.1.2.1), and the log of the number of TV households (in thousand) in the DMA. Finally, we account for the contribution of any

[153] We enforce this convention for the 36 station-year observations where a satellite has non-missing advertising revenue. We manually link satellite stations to their parent primary stations because BIA does not provide this information. The 114 satellite stations in Table B.2 belong to 80 primary stations.

satellite stations to advertising revenue by including in X_{jt} the number of satellite stations that belong to the primary station N_{jt}^{SAT}.The adjusted R^2 is 0.99 in all years in logs and 0.75 on average in levels, suggesting that we capture most of the variation in advertising revenue across stations and years.

With the estimates in hand, we predict advertising revenue AD_{jt} as $\widehat{AD}_{jt} = e^{\widehat{\ln AD}_{jt}+\frac{\hat{\sigma}^2}{2}}$ to account for the non-zero mean of the log-normally distributed error term with estimated variance $\hat{\sigma}^2$. We proceed as follows: First, for a primary station we impute advertising revenue AD_{jt} where missing as $\widehat{AD}_{jt}$. Second, we compute the contribution of satellite stations (if any) to the advertising revenue of their parent primary station as $AD_{jt} - AD_{jt}/e^{\hat{\beta}_{SAT}N_{jt}^{SAT}}$, where $\hat{\beta}_{SAT}$ is the estimated coefficient for the number of satellite stations. For a primary station we net out the contribution of satellite stations by replacing advertising revenue AD_{jt} with $AD_{jt}/e^{\hat{\beta}_{SAT}N_{jt}^{SAT}}$. Third, for a satellite station we impute advertising revenue AD_{jt} by allocating the contribution of satellite stations to the advertising revenue of their parent primary station in proportion to the population coverage of the satellite stations.

B.1.1.4 NAB data

NAB collects detailed financial information for commercial full-power stations. In 2012, NAB received 785 responses on 1,288 originated questionnaires, corresponding to a response rate of 60.9%.

NAB reports the data at various levels of aggregation. Table B.4 shows the resulting 66 tables for 2012.[154] The number of tables fluctuates slightly year-by-year because NAB imposes a minimum of ten TV stations per aggregation category to ensure confidentiality.[155] Note that a TV station may feature in more than one table. For example, WABC-TV is the ABC affiliate in New York, NY. Its data is used in calculating statistics for (1) markets of rank 1 to 10; (2) major network affiliates; (3) all ABC affiliates; and (4) ABC affiliates in markets with rank 1 to 25.

[154]We exclude 15 aggregation categories that are defined by total revenue from each year's NAB report because the BIA data is restricted to advertising revenue.

[155]Some years, in particular, break out United Paramount and Spanish-language networks but not other minor networks. We conclude that the response rate of other minor networks is very low and thus exclude other minor networks from most of the subsequent analysis.

Table B.3: Missing advertising revenue for commercial stations

	Station-year count	Missing advertising revenue	
		Station-year count	%
Full-power			
Primary	13,490	937	6.95
Satellite	1,252	1,216	97.12
Low-power (Class-A)	4,853	3,905	80.47
Major networks			
ABC	2,497	420	16.82
CBS	2,423	314	12.96
Fox	2,272	318	14.00
NBC	2,445	376	15.38
Minor networks			
CW	850	99	11.65
MyNetwork TV	745	133	17.85
United Paramount	269	37	13.75
Warner Bros	267	24	8.99
Spanish-language networks	1,747	563	32.23
Other	3,159	1,781	56.38
Independent	2,921	1,993	68.23
Total	19,595	6,058	30.92

Notes: United Paramount and Warner Bros merged in 2006 to form CW. Spanish-language networks include Azteca America, Telemundo, Univision, UniMas, and Independent Spanish stations.

For each aggregation category, NAB reports the mean, 1^{st}, 2^{nd}, and 3^{rd} quartiles for cash flow and detailed revenue source categories. We define non-broadcast revenue as the sum of total trade-outs and barter, multicast revenue, and other broadcast related revenue. We further define advertising revenue as the sum of local, regional, national, and political advertising revenues, commissions, and network compensations. Because we do not observe correlations between the detailed revenue source categories, we can construct the mean of non-broadcast revenue and advertising revenue but not the quartiles. We present sample moments of cash flow and non-broadcast revenue for select aggregation categories in Table B.5.[156]

[156] To validate the data, we compare the mean of advertising revenue from the NAB data to suitably averaged advertising revenue from the BIA data. The resulting 662 pairs of means from the two data sources exhibit a correlation of 0.92.

Table B.4: NAB tables (2012)

Table	Description	Table	Description
1	All Stations, All Markets	34	ABC, CBS, FOX, NBC, Markets 176+
2	All Stations, Markets 1-10	35	ABC, All Markets
3	All Stations, Markets 11-20	36	ABC, Markets 1-25
4	All Stations, Markets 21-30	37	ABC, Markets 26-50
5	All Stations, Markets 31-40	38	ABC, Markets 51-75
6	All Stations, Markets 41-50	39	ABC, Markets 76-100
7	All Stations, Markets 51-60	40	ABC, Markets 101+
8	All Stations, Markets 61-70	41	CBS, All Markets
9	All Stations, Markets 71-80	42	CBS, Markets 1-25
10	All Stations, Markets 81-90	43	CBS, Markets 26-50
11	All Stations, Markets 91-100	44	CBS, Markets 51-75
12	All Stations, Markets 101-110	45	CBS, Markets 76-100
13	All Stations, Markets 111-120	46	CBS, Markets 101+
14	All Stations, Markets 121-130	47	FOX, All Markets
15	All Stations, Markets 131-150	48	FOX, Markets 1-50
16	All Stations, Markets 151-175	49	FOX, Markets 51-75
17	All Stations, Markets 176+	50	FOX, Markets 76-100
18	ABC, CBS, FOX, NBC, All Markets	51	FOX, Markets 101+
19	ABC, CBS, FOX, NBC, Markets 1-10	52	NBC, All Markets
20	ABC, CBS, FOX, NBC, Markets 11-20	53	NBC, Markets 1-25
21	ABC, CBS, FOX, NBC, Markets 21-30	54	NBC, Markets 26-50
22	ABC, CBS, FOX, NBC, Markets 31-40	55	NBC, Markets 51-75
23	ABC, CBS, FOX, NBC, Markets 41-50	56	NBC, Markets 76-100
24	ABC, CBS, FOX, NBC, Markets 51-60	57	NBC, Markets 101+
25	ABC, CBS, FOX, NBC, Markets 61-70	58	CW, All Markets
26	ABC, CBS, FOX, NBC, Markets 71-80	59	CW, Markets 1-25
27	ABC, CBS, FOX, NBC, Markets 81-90	60	CW, Markets 26-50
28	ABC, CBS, FOX, NBC, Markets 91-100	61	CW, Markets 51-75
29	ABC, CBS, FOX, NBC, Markets 101-110	62	MNTV, All Markets
30	ABC, CBS, FOX, NBC, Markets 111-120	63	MNTV, Markets 1-50
31	ABC, CBS, FOX, NBC, Markets 121-130	64	MNTV, Markets 51+
32	ABC, CBS, FOX, NBC, Markets 131-150	65	Independent, All markets
33	ABC, CBS, FOX, NBC, Markets 151-175	66	Independent, Markets 1-25

Notes: Data comes from NAB annual directory for 2012. Market numbers refer to a market's rank in terms of size. NAB rules prohibit aggregation when there are too few respondents in a particular grouping, which determines the market size ranges. Tables with total revenue breakouts are excluded.

Table B.5: Sample moments for select aggregation categories (2012)

	Cash Flow ($ million)				Non-broadcast Revenue ($ million)
		Percentile			
	Mean	25th	50th	75th	Mean
All Stations	7.798	1.243	3.752	9.178	2.977
All Stations, Markets 101-110	4.120	1.704	3.619	6.444	2.102
All Major Affiliates	9.244	1.936	4.929	10.90	3.326
ABC Affiliates, Markets 1-25	32.40	15.09	27.15	42.46	7.596
NBC Affiliates, Markets 101+	3.652	1.293	3.283	5.901	1.883
All CW Affiliates	3.929	0.355	1.798	3.224	2.884
MyNetwork TV, Markets 1-50	3.124	1.270	1.799	3.215	2.507
All Independent Stations	2.786	-0.020	1.288	4.327	2.195

Notes: Data comes from NAB annual directory for 2012. A select few categories are reported (see Table B.4 for all categories). Non-broadcast revenues are constructed as the sum of total trade-outs and barter, multicast revenues, and other broadcast related revenues. We thus only obtain the mean as we lack information on the correlations of the respective distributions.

B.1.2 Cash flows

B.1.2.1 Functional forms

We parameterize $\alpha\left(X_{jt};\beta\right)$, $RT\left(X_{jt};\gamma\right)$, and $F\left(X_{jt};\delta\right)$ as a function of station and market characteristics X_{jt} as

$$\begin{aligned}
\alpha\left(X_{jt};\beta\right) &= \sum_{a=1}^{9}\beta_0^a I(Affiliation_{jt}=a) + \sum_{s=2003}^{2012}\beta_0^s I(t=s) + \beta_1 Fox_{jt}\cdot t + \beta_2 CompIndex_{jt}, \\
RT\left(X_{jt};\gamma\right) &= \exp\left(\gamma_0+\gamma_1 t+\gamma_2\ln(MktSize_{jt})\right), \\
F\left(X_{jt};\delta\right) &= \delta_0 + \delta_1 WealthIndex_{jt} + \sum_{h=1}^{3} I(Group_{jt}=h)\cdot\left(\delta_2^h\ln(MktSize_{jt})+\delta_3^h\ln(MktSize_{jt})^2\right),
\end{aligned}$$

where $I(\cdot)$ is the indicator function. $Affiliation_{jt}$ refers to nine of the eleven affiliations[157] in Table B.3 and $Group_{jt}$ to groupings of affiliations (detailed below). $MktSize_{jt}$ is the number of TV households in the DMA and $WealthIndex_{jt}$ and $CompIndex_{jt}$ are the wealth and competitiveness indices for the DMA.[158]

[157] We normalize the parameter on the indicator for Spanish-language networks to zero. We exclude any TV station affiliated with other minor networks from the estimation, see footnote 155. To predict the cash flow for such a TV station from our parameter estimates, we use its station and owner characteristics X_{jt} and the parameter on the indicator for Independent.

[158] To parsimoniously capture market characteristics, we conduct a principal component analysis of the market-level variables prime-age (18-54) population, average disposable income, retail expenditures, adver-

We allow the share $\alpha\left(X_{jt};\beta\right)$ of advertising revenue retained as cash flow to vary flexibly by year and network affiliation. We allow for a separate time trend for Fox affiliates as their profitability grew substantially over time. The competitiveness index $CompIndex_{jt}$ accounts for differences in the competitive environment across DMAs.

We specify $RT\left(X_{jt};\gamma\right)$ as an exponential function of a time trend and market size in light of the rapid growth of retransmission fees. We make no attempt to separately estimate an error term for non-broadcast revenue and assume it is one part of ϵ_{jt} in equation 2.4.4 due to additivity.[159]

Lastly, we let fixed cost $F\left(X_{jt};\delta\right)$ vary flexibly with market size and the network affiliation. To streamline the specification, we subsume the affiliations in Table B.3 into three groups with similar cost structures: (1) ABC, CBS, and NBC; (2) Fox, CW, and Warner Bros; (3) My Network TV, United Paramount, Spanish-language networks, and Independents. We include the wealth index $WealthIndex_{jt}$ in the fixed cost to reflect the differential cost of operating in different DMAs.

B.1.2.2 Data

We combine the station-level data on advertising revenue from BIA with the aggregated data from NAB. The NAB data yields 3,313 moments across aggregation categories and the years from 2003 to 2012 as shown in Table B.7.[160] There are a total of 11,801 station-year observations from the BIA data that meet NAB's data collection and reporting procedure and therefore map into a table of a NAB report.

B.1.2.3 Estimation

We use a simulated minimum distance estimator. We draw $S = 100$ vectors of cash flow error terms $\epsilon^s = \left(\epsilon^s_{jt}\right)$, where ϵ^s_{jt} is the cash flow error term of TV station j in year t in

tising revenues, number of primary TV stations, and number of major network affiliates. The first principal component, denoted as $CompIndex_{jt}$, loads primarily on to prime-age population, advertising revenues, number of primary TV stations, and number of major network affiliates. The second principal component, denoted as $WealthIndex_{jt}$ loads primarily on to average disposable income and retail expenditures.

[159] We obtain very similar estimates when we separately estimate such an error term.

[160] We drop the year 2013 from the BIA data as 2012 is the latest year of availability for the NAB data. We further drop TV stations affiliated with other minor networks from the BIA data, see footnotes 155 and 157.

draw s. Denote by $\overline{CF}_{gt}$, CF^1_{gt}, CF^2_{gt}, and CF^3_{gt} the mean, 1^{st}, 2^{nd}, and 3^{rd} quartiles of the cash flow distribution reported by NAB in year t for aggregation category $g = 1, \ldots, G_t$, where G_t is the number of aggregation categories in year t. Similarly, denote by $\widehat{\overline{CF}}_{gt}(\theta; \epsilon^s)$, $\widehat{CF}^1_{gt}(\theta; \epsilon^s)$, $\widehat{CF}^2_{gt}(\theta; \epsilon^s)$, and $\widehat{CF}^3_{gt}(\theta; \epsilon^s)$ the analogous moments of the predicted cash flow distribution for the TV stations that feature in aggregation category g in year t. Our notation emphasizes that the latter depend on the parameters $\theta = (\beta, \gamma, \delta, \sigma)$ and the vector of cash flow error terms ϵ^s in draw s. We use similar notation, replacing $\overline{CF}$ with $\overline{RT}$, for the mean of the non-broadcast revenue distributions. To estimate θ, we match the moments of the predicted and actual distributions across aggregation categories and years. Formally,

$$\hat{\theta} = \arg\min_{\theta} \sum_{t=2003}^{2012} \sum_{g=1}^{G_t} \left(\overline{CF}_{gt} - \frac{1}{S} \sum_{s=1}^{S} \widehat{\overline{CF}}_{gt}(\theta; \epsilon^s) \right)^2 + \sum_{q=1}^{3} \left(CF^q_{gt} - \frac{1}{S} \sum_{s=1}^{S} \widehat{CF}^q_{gt}(\theta; \epsilon^s) \right)^2 + \left(\overline{RT}_{gt} - \widehat{\overline{RT}}_{gt}(\theta) \right)^2 .$$

We constrain the standard deviation of the error term to be positive. Our interior-point minimization algorithm terminates with a search step less than the specified tolerance of 10^{-12}. We use multi starts to guard against local minima. Our estimates are robust to different starting values.

B.1.2.4 Results

Table B.6 reports the parameter estimates. The estimates are in line with our expectations: major network affiliates retain a higher share of advertising revenue than minor networks, with Fox having a positive trend; independent and WB stations retain the highest share of advertising revenues, albeit with the smallest revenue base; the retained share falls over time, bottoming out in 2009 before bouncing back in recent years; the retained share is lower in more competitive markets. Finally, non-broadcast revenue has grown significantly in recent years and there are economies of scale in fixed cost.

Figure B.1.1 plots the distributions of the estimated retained share $\alpha(X_{jt}; \beta)$, non-broadcast revenue $RT(X_{jt}; \delta)$, and fixed cost $F(X_{jt}; \delta)$. Reassuringly, without imposing restrictions,

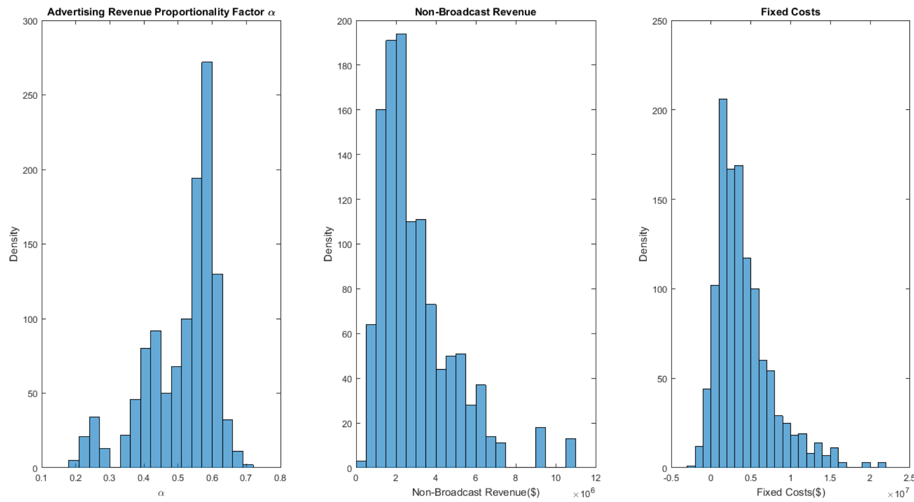
Table B.6: Cash flow parameters estimates

	Estimates
Retained share $\alpha(X_{jt};\beta)$ of advertising revenue	
ABC	-0.035
CBS	-0.062
Fox	-0.382
NBC	-0.054
CW	-0.113
MyNetwork TV	-0.356
United Paramount	-0.364
Warner Bros	0.013
Spanish-language networks (normalized)	0
Independent	-0.210
Fox $\times$ Trend	0.018
2003	0.692
2004	0.666
2005	0.642
2006	0.630
2007	0.599
2008	0.567
2009	0.529
2010	0.600
2011	0.619
2012	0.636
$CompIndex$	-0.021
Non-broadcast revenue $RT(X_{jt},\gamma)$ (log $)	
Intercept	6.513
ln(MktSize)	0.527
Trend	0.135
Fixed cost $F(X_{jt};\delta)$ ($ million)	
Intercept	63.941
$WealthIndex$	1.052
Group 1 $\times$ ln(MktSize)	-12.851
Group 2 $\times$ ln(MktSize)	-11.696
Group 3 $\times$ ln(MktSize)	-10.210
Group 1 $\times$ ln(MktSize)^2	0.643
Group 2 $\times$ ln(MktSize)^2	0.535
Group 3 $\times$ ln(MktSize)^2	0.419
σ ($ million)	1.030

Notes: Group 1 is ABC, CBS, and NBC; group 2 is Fox, CW, and Warner Bros; and group 3 is My Network TV, United Paramount, Spanish-language networks, and Independents.

we estimate α to be between 0.2 and 0.7 in 2012, with an average of 0.51; non-broadcast revenue is estimated to be between $0.21 million and $10.9 million; fixed cost is estimated

Figure B.1.1: Estimated retained share of advertising revenue, non-broadcast revenue, and fixed cost (2012)



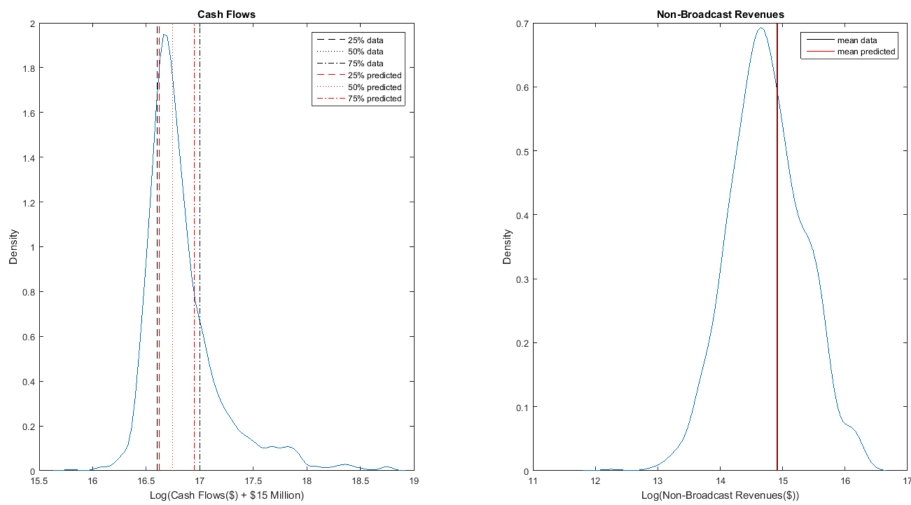
Notes: Plots are distributions of estimated retained share of advertising revenue α (left), non-broadcast revenue in dollars (middle), and fixed cost in dollars (right) in 2012. Includes 1,172 commercial full-power stations in 2012 used in the cash flow estimation.

to be contributing negatively to cash flow in 95% of cases, averaging $4.12 million, with the highest fixed cost estimated to be up to $21.9 million in 2012.

The cash flow model fits the data well. Figure B.1.2 plots the predicted distributions of cash flow and non-broadcast revenue, superimposed with the corresponding moments from the NAB data for all TV stations in 2012. Cash flow is estimated to be between -$6.4 million and $127 million across TV stations in 2012, with an average of $7.2 million (compared to $7.8 million reported by NAB). The 25^{th} ($1.6 million), 50^{th} ($3.6 million), and 75^{th} ($7.8 million) percentiles of the predicted distribution are overlaid in red lines (dashed, dotted, and dash-dotted, respectively). The black lines of the same patterns refer to the corresponding moments in the NAB data. Non-broadcast revenue is estimated to average $3.0 million (compared to $3.0 million reported by NAB).

To further assess the fit of the cash flow model, Table B.7 compares the cash flow and non-broadcast revenue moments as reported in NAB to the corresponding predicted moments, broken down type of moment, affiliation, year, and market rank. It provides three different measures of fit, namely the correlation between actual and predicted moments

Figure B.1.2: Estimated cash flow and non-broadcast revenue with moments (2012)



Notes: Plots are kernel densities of estimated cash flows (left) and non-broadcast revenues (right) in 2012 in log terms. Cash flows (in dollars) are shifted by $15 million to avoid negative numbers. Black lines indicate data moments and red lines indicate model-predicted moments. Data moments are from the all-station category in NAB (year 2012, table 1). Includes 1,172 commercial full power stations in 2012 used in the cash flow estimation procedure. The cash flow distribution is plotted from one simulation with station-specific errors.

Table B.7: Cash flow and non-broadcast revenue moments and fit measures

		Number of Moments	Correlation	Mean Abs Deviation $ million	%
All		3313	0.98	0.91	18.11
Type	Cash flow, mean	663	0.99	0.89	13.16
	1^{st} quartile	662	0.97	0.90	35.24
	2^{nd} quartile	663	0.98	0.90	17.73
	3^{rd} quartile	663	0.98	1.36	15.16
	Non-broadcast revenue, mean	662	0.84	0.49	28.65
Affiliation	Major network	1995	0.98	1.00	16.13
	Minor network	350	0.93	1.00	38.55
	Independent	110	0.73	0.95	62.31
Year	2003	329	0.98	1.03	19.52
	2004	325	0.99	0.87	15.05
	2005	330	0.98	0.95	19.17
	2006	310	0.99	0.96	16.26
	2007	344	0.98	0.90	19.68
	2008	350	0.98	0.86	20.43
	2009	330	0.98	0.66	23.08
	2010	330	0.98	0.84	16.30
	2011	335	0.97	0.90	18.68
	2012	330	0.98	1.11	16.51
Market Rank	1-25	460	0.98	2.37	14.74
	26-50	385	0.96	0.94	15.97
	51-100	930	0.93	0.63	19.98
	101+	799	0.87	0.47	32.21

Notes: Correlations refer to correlations between predicted and observed dollar magnitudes for a particular subsample of distribution moments. Percent mean deviations measured as a share of observed dollar magnitudes.

as well as the absolute deviation in millions of dollars and in percent magnitudes, and percent of absolute deviations. Overall, our cash flow model predicts the 3,313 moments with a 0.98 correlation. Of the 330 moments from 2012, our predicted moments have a 0.98 correlation with the actual moments reported by NAB; on average, our predicted moments miss the actual moments by $1.11 million, or 16.5%.

B.1.3 Multiples

B.1.3.1 Priors

Industry analysts give a range of \$0.15 to \$0.40 per MHz-pop for the stick multiple and a range of 10 to 12 for the cash flow multiple.[161] Therefore, our prior is that the stick multiple is distributed log-normally with mean $\mu^{Stick}_{prior} = -1.4$ and standard deviation $\sigma^{Stick}_{prior} = 0.5$ (corresponding to a mean of \$0.25 per MHz-pop and a standard deviation of \$1 per MHz-pop, thereby covering \$0.15 to \$0.40 per MHz-pop with probability 0.68). According to industry analysts, while the stick multiple is believed to be much larger for larger markets, the cash flow multiple is believed to be symmetrically distributed. Our prior is therefore that the cash flow multiple is distributed normally with mean $\mu^{CF}_{prior} = 11$ and standard deviation $\sigma^{CF}_{prior} = 1$.

B.1.3.2 Data

As discussed in Section 2.4.1, our data consists of 136 transactions between 2003 and 2012 based on cash flow and 201 transactions between 2003 and 2013 based on stick value. For cash flow transactions, we infer the cash flow multiple from the transaction price and the estimated cash flow $\widehat{CF}_{jt}$ using equation 2.4.2. For stick value transactions, we infer the stick multiple from the transaction price using equation 2.4.3.

B.1.3.3 Estimation

For cash flow transactions, we estimate the following model for the multiple to construct its conditional likelihood function:

$$Multiple^{CF}_{jt} = \beta X_{jt} + \epsilon_{jt}, \tag{B.1.1}$$

where X_{jt} includes owner, station, and market characteristics. Specifically, we include in X_{jt} an indicator of whether a station has multicast sub-channels, the station's population

[161] See "Opportunities and Pitfalls on the Road to the Television Spectrum Auction," Bond & Pecaro white paper, December 12, 2013, available at http://www.bondpecaro.com/images/Bond_Pecaro_Spectrum_White_Paper_12122013.pdf, accessed on November 15, 2015.

coverage (in thousand), the wealth and competitiveness indices for the DMA, power output fixed effects (primary, satellite, and class-A), ownership category fixed effects (whether the owner owns one, between two and ten, or more than ten stations across markets), fixed effects for the eleven affiliations in Table B.3, and a full set of year fixed effects. The adjusted R^2 is 0.68 and we take $\hat{\sigma}^{CF}_{likelihood} = 4.52$ to be the standard deviation of the 136 estimated residuals.

For stick value transactions, we estimate the following model:

$$\ln Multiple^{Stick}_{jt} = \beta X_{jt} + \epsilon_{jt}, \tag{B.1.2}$$

where we include in X_{jt} the log of the station's output power, the log of the station's population coverage, the wealth and competitiveness indices for the DMA, power output fixed effects, ownership category fixed effects, affiliation fixed effects, and year fixed effects. The adjusted R^2 is 0.67 and we take $\hat{\sigma}^{Stick}_{likelihood} = 0.97$.

B.1.3.4 Posteriors

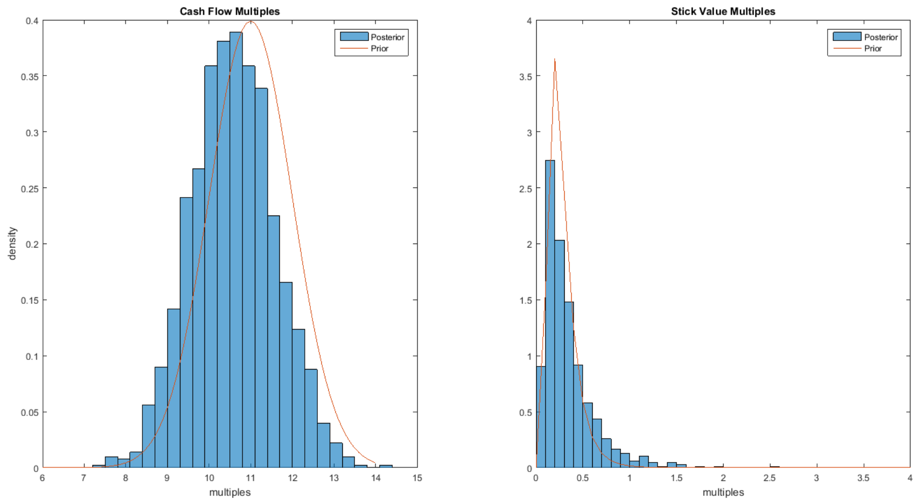
With the estimates in hand, we can predict multiples for any TV stations. To obtain the posterior for the cash flow, respectively, stick multiple, we update our prior with the conditional likelihood function using Bayes rule as

$$\mu_{posterior} = \frac{\mu_{prior}\sigma^2_{likelihood} + \mu_{likelihood}\sigma^2_{prior}}{\sigma^2_{prior} + \sigma^2_{likelihood}},$$

$$\sigma^2_{posterior} = \frac{\sigma^2_{likelihood}\sigma^2_{prior}}{\sigma^2_{prior} + \sigma^2_{likelihood}},$$

where $\mu^{CF}_{likelihood} = \widehat{Multiple}^{CF}_{jt_0}$ for cash flow transactions and $\mu^{Stick}_{likelihood} = \ln \widehat{Multiple}^{Stick}_{jt_0}$ for stick value transactions and we set $t_0 = 2012$. The posterior standard deviation of the cash flow multiple is 0.98 and that of the stick multiple is 0.44. Because the posterior mean depends on X_{jt}, Figure B.1.3 illustrates the estimated posterior distribution for the cash flow, respectively, stick multiple in one particular simulation run for the 1,672 UHF licenses that the FCC includes in its repacking simulations. The prior distributions are overlaid in red dashed lines.

Figure B.1.3: Prior and posterior distributions of cash flow and stick multiples



Notes: Probability density function for the prior distributions and estimated posterior distributions. Plotted from one simulation (one station-specific draw for each station).

B.2 Algorithm details

There are N TV stations in the focal DMA and the DMAs in its repacking region. As in section 2.2.2, denote as v_j the reservation value of TV station j and as φ_j station j's broadcast volume. We order TV stations such that $\frac{v_1}{\varphi_1} \leq \frac{v_2}{\varphi_2} \leq \ldots \leq \frac{v_N}{\varphi_N}$. Breaking ties in favor of selling, we have that if the base clock price P exceeds $\frac{v_N}{\varphi_N}$, then all stations $1, \ldots, N$ would be willing to relinquish their license. If the base clock price is $\frac{v_N}{\varphi_N} > P \geq \frac{v_{N-1}}{\varphi_{N-1}}$, then only stations $1, \ldots, N-1$ relinquish their license, while station N prefers to continue operating and drops out of the auction, at which point it has to be repacked, and so on.

A given clearing target of spectrum maps into a certain number of TV station channels to be cleared for wireless service. For example, the auction's initial 126 MHz clearing target would have corresponded to clearing 21 channels out of a total of 37 non-dedicated UHF channels in each DMA. In some DMAs, this would be possible without purchasing spectrum from TV stations since not all channels are allocated. Denote the channels available for repacking TV stations that choose to continue operating after the reverse auction by R. For simplicity, we suppress the dependency of R on the spectrum clearing target.

The repacking feasibility checker SATFC takes as inputs the remaining available channels R and a set of TV stations X, together with their interference profile and the channels to which they could be repacked.[162] Denote this as $SATFC(X,R)$. It then returns one of three possible outcomes:SAT, $UNSAT$, or $TIMEOUT$. SAT denotes that the set of stations X could be satisfactorily repacked. $UNSAT$ indicates that is is not possible, leading to a station being frozen, while $TIMEOUT$ indicates that within the maximum amount of time allotted to one run of the feasibility checker, it was not possible to ascertain whether repacking would be feasible. Based on testing, the FCC has set set the TIMEOUT parameter to one minute, which we also use. Their testing found that this time limit offered an efficient trade off between total computational burden and accuracy, and interpret TIMEOUT as an UNSAT.

We partition the set of TV stations $\{1,\ldots,N\}$ into a set of "active" stations A, a set of "frozen" (or "conditionally winning") stations F, and a set of "inactive"stations I that, given the initial clock price $P = 900$ immediate drop out of the auction (see Competitive Bidding Procedures for Broadcast Incentive Auction 1000, Including Auctions 1001 and 1002, Public Notice FCC 14-191, p. 105, available at `https://www.fcc.gov/document/broadcast-incentive-auction-comment-pn`).

Finally, denote as PO_j the payout to TV station j from the reverse auction in terms of the base clock price; the station's ultimate selling price would be $\varphi_j PO_j$. We iteratively solve the reverse auction in Matlab interfacing with SATFC via a Java bridge. To initialize, we set $P = 900$, and define the set of active stations $A = \left\{s \in \{1,\ldots,N\} | \frac{v_s}{\varphi_s} \leq 900\right\}$, the set of frozen stations $F = \emptyset$, and the set of inactive stations $I = \left\{s \in \{1,\ldots,N\} | \frac{v_s}{\varphi_s} > 900\right\}$. That is, all TV stations with valuation less than or equal to 900 participate in the reverse auction. If $SATFC(I,R) \neq SAT$, then the remaining stations cannot be repacked. In this case, we declare the auction as failed and set $PO_s = 0$ for all $s \in \{1,\ldots,N\}$.

Otherwise, we proceed as follows:

1. REPEAT

[162] We use Perl scripts to create repacking region-specific domain and interference files to use with the SATFC feasibility checker to simulate the auction. This speeds up computation and decreases the amount of memory overhead required for large-scale parallel computing.

(a) For each $s \in A$ run $SATFC(I \cup \{s\}, R)$.

i. If $SATFC(I \cup \{s\}, R) = UNSAT$, station s cannot be repacked. Change its status to frozen: set $A \leftarrow A \setminus \{s\}$, $F \leftarrow F \cup \{s\}$. Its payout is the current base clock price: $PO_s = P$.

(b) If $A \neq \emptyset$ then set $s = \max_{s \in A}(\frac{v_s}{\varphi_s})$, $P = \frac{v_s}{\varphi_s}$, $A \leftarrow A \setminus \{s\}$, $I \leftarrow I \cup \{s\}$, and $PO_s = 0$.

2. UNTIL $A = \emptyset$

Step (i) changes the status of any currently active TV station that cannot be repacked in addition to the currently inactive TV stations to frozen (p. 108 and pp. 112–113, FCC 14-191). If a TV station is frozen, it receives a payout equal to the current base clock price P. P, in turn, is determined by the TV station most recently marked as inactive (or, possibly, the opening base clock price of 900).

At the end of steps (a)–(b) we are guaranteed that changing the status of any remaining active TV station to inactive preserves feasibility. Step (c) then finds the remaining active TV station with the highest value and changes its status from active to inactive. This TV station receives a payout of zero.

B.2.1 Illustration

Figure B.2.1: Simulation algorithm examples

DMA: 4, Simulation: 7, Strategy: 1
126 Total Licenses Considered, 23 in Focal DMA

Processing Withdrawls.

1: #9610 (CBS) withdraws…✓
2: #47535 (NBC) withdraws…✓
3: #22206 (FOX) withdraws…✓
4: #47904 (NBC) withdraws…✓
5: #15569 (NBC) withdraws…✓
6: #70158 (IND) withdraws…✓
7: #68136 (IND) withdraws…✓

Auction begins.

8: #68135 (IND) withdraws at $873.34…✓
9: #22207 (FOX) withdraws at $748.20…✓
10: #52077 (FOX) withdraws at $667.83…✓
11: #53115 (CBS) withdraws at $639.13…✓
12: #25453 (CBS) withdraws at $626.80…✓
13: #4688 (ABC) withdraws at $476.46…✓
14: #51568 (FOX) withdraws at $463.92…✓
15: #63153 (NBC) withdraws at $444.89…✓
16: #23341 (CBS) withdraws at $443.64…✓
17: #60653 (NBC) withdraws at $427.76…✓
18: #74215 (UNI) withdraws at $422.03…✓
19: #71218 (CBS) withdraws at $400.52…✓
20: #73318 (ABC) withdraws at $365.24…✓
21: #50780 (NBC) withdraws at $357.84…✓
22: #74170 (NBC) withdraws at $323.50…✓
23: #73333 (TEL) withdraws at $298.15…×
➢ **#39884 (IND) frozen at $298.15**
24: #10213 (FOX) withdraws at $254.81…✓
25: #147 (FOX) withdraws at $242.21…✓
26: #62219 (FOX) withdraws at $241.57…✓
27: #13929 (REL) withdraws at $235.37…✓
28: #191822 (OTH) withdraws at $229.57…✓
29: #73120 (NBC) withdraws at $228.28…✓
30: #72278 (IND) withdraws at $224.06…✓
31: #74197 (MY) withdraws at $221.79…×
➢ **#60560 (UNI) frozen at $221.79**
➢ **#74464 (Am1) frozen at $221.79**
➢ **#12499 (CW) frozen at $221.79**
➢ **#55305 (IND) frozen at $221.79**
➢ **#7623 (TBN) frozen at $221.79**
➢ **#51984 (ION) frozen at $221.79**
➢ **#48465 (PBS) frozen at $221.79**
➢ **#23142 (TEL) frozen at $221.79**
➢ **#191340 (OTH) frozen at $221.79**
➢ **#28480 (PUB) frozen at $221.79**
➢ **#36989 (PBS) frozen at $221.79**
➢ **#48481 (PBS) frozen at $221.79**
32: #10758 (FOX) withdraws at $208.75…✓
33: #52075 (MY) withdraws at $206.82…✓
34: #167543 (MdF) withdraws at $201.53…✓
35: #50357 (OTH) withdraws at $175.93…✓
36: #30576 (CW) withdraws at $172.90…✓
37: #11260 (ABC) withdraws at $170.23…✓
38: #190915 (FOX) withdraws at $151.07…✓
39: #72313 (CBS) withdraws at $149.28…✓
40: #73879 (MY) withdraws at $148.81…✓
41: #51567 (MY) withdraws at $144.80…✓
42: #20287 (ABC) withdraws at $137.30…✓
43: #71508 (ABC) withdraws at $127.85…✓
44: #73374 (CW) withdraws at $124.65…✓
45: #70493 (ME) withdraws at $100.53…×
➢ **#9739 (IND) frozen at $100.53**
46: #16455 (ABC) withdraws at $98.56…✓
47: #72623 (OTH) withdraws at $93.75…✓
48: #50063 (ION) withdraws at $83.70…✓
49: #74216 (UnM) withdraws at $72.66…×
➢ **#61111 (NBC) frozen at $72.66**

Simulation concludes as all licenses in the focal DMA have either been withdrawn or frozen.

DMA: 4, Simulation: 7, Strategy: 325
126 Total Licenses Considered, 23 in Focal DMA

Processing Withdrawls.

1: #9610 (CBS) withdraws…✓
2: #47535 (NBC) withdraws…✓
3: #22206 (FOX) withdraws…✓
4: #47904 (NBC) withdraws…✓
5: #15569 (NBC) withdraws…✓
6: #70158 (IND) withdraws…✓
7: #74464 (Am1) withdraws…✓
8: #68136 (IND) withdraws…✓
9: #72278 (IND) withdraws…✓

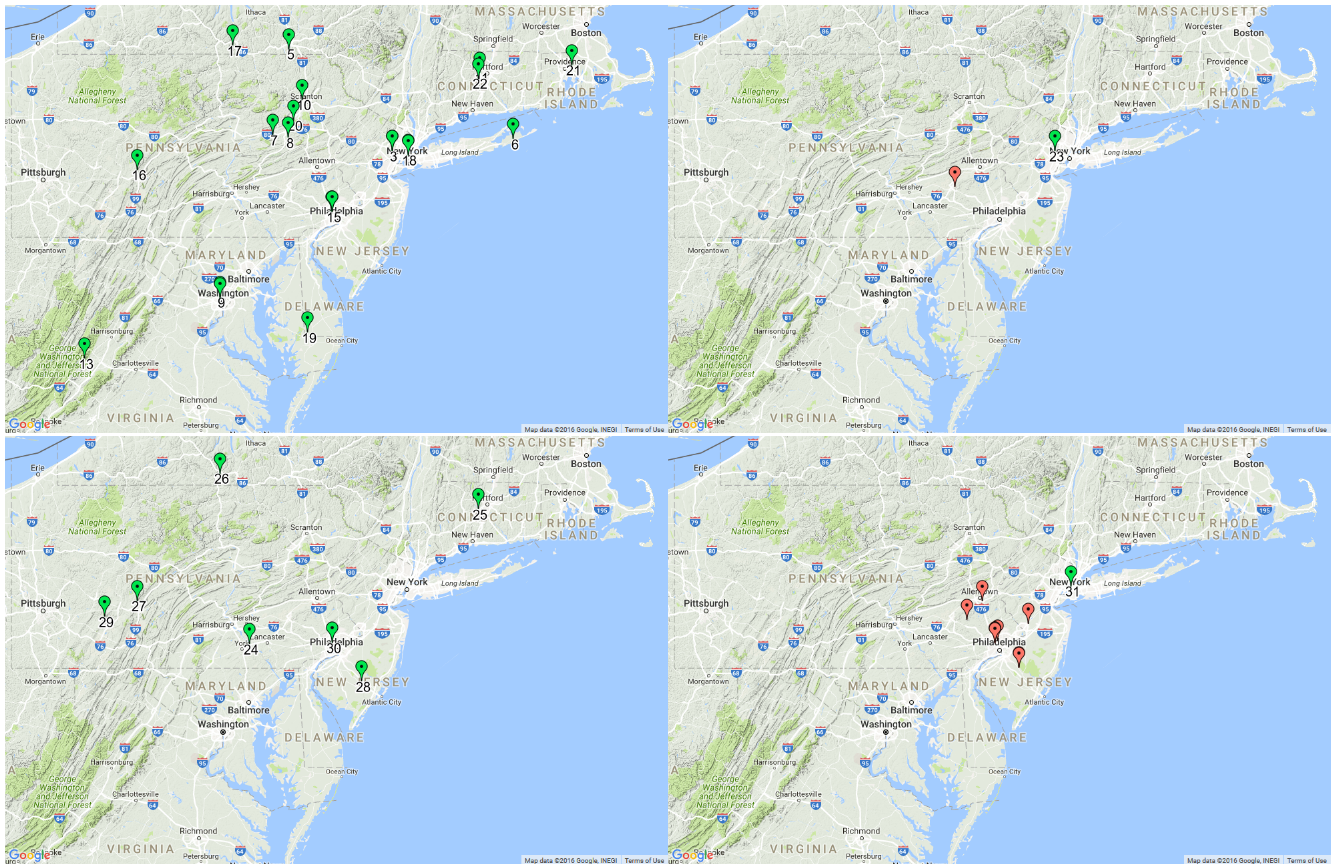
Auction begins.

10: #68135 (IND) withdraws at $873.34…✓
11: #22207 (FOX) withdraws at $748.20…✓
12: #52077 (FOX) withdraws at $667.83…✓
13: #53115 (CBS) withdraws at $639.13…✓
14: #25453 (CBS) withdraws at $626.80…✓
15: #4688 (ABC) withdraws at $476.46…✓
16: #51568 (FOX) withdraws at $463.92…✓
17: #63153 (NBC) withdraws at $444.89…×
➢ **#39884 (IND) frozen at $444.89**
18: #23341 (CBS) withdraws at $443.64…✓
19: #60653 (NBC) withdraws at $427.76…✓
20: #74215 (UNI) withdraws at $422.03…✓
21: #71218 (CBS) withdraws at $400.52…✓
22: #73318 (ABC) withdraws at $365.24…✓
23: #50780 (NBC) withdraws at $357.84…✓
24: #74170 (NBC) withdraws at $323.50…✓
25: #73333 (TEL) withdraws at $298.15…×
➢ **#60560 (UNI) frozen at $298.15**
➢ **#12499 (CW) frozen at $298.15**
➢ **#55305 (IND) frozen at $298.15**
➢ **#7623 (TBN) frozen at $298.15**
➢ **#51984 (ION) frozen at $298.15**
➢ **#48465 (PBS) frozen at $298.15**
➢ **#23142 (TEL) frozen at $298.15**
➢ **#191340 (OTH) frozen at $298.15**
➢ **#28480 (PUB) frozen at $298.15**
➢ **#36989 (PBS) frozen at $298.15**
➢ **#48481 (PBS) frozen at $298.15**
26: #10213 (FOX) withdraws at $254.81…×
➢ **#9739 (IND) frozen at $254.81**
27: #147 (FOX) withdraws at $242.21…✓
28: #62219 (FOX) withdraws at $241.57…✓
29: #13929 (REL) withdraws at $235.37…✓
30: #191822 (OTH) withdraws at $229.57…✓
31: #73120 (NBC) withdraws at $228.28…✓
32: #74197 (MY) withdraws at $221.79…×
➢ **#167543 (MdF) frozen at $221.79**
➢ **#73879 (MY) frozen at $221.79**
➢ **#74216 (UnM) frozen at $221.79**
➢ **#61111 (NBC) frozen at $221.79**

Simulation concludes as all licenses in the focal DMA have either been withdrawn or frozen.

Notes: Checkmarks (✓) indicate that all remaining active stations can feasibly be repacked when the listed station withdraws. Crosses () indicate that at least one active license can no longer feasibly be repacked when the listed station withdraws. Events concerning licenses in the focal DMA are in bold. Licenses are numbered by FCC ID numbers. Prices are in terms of the base clock price.

Figure B.2.2: Simulation algorithm example maps



Notes: Maps show condensed simulation process from Figure B.2.1. Green pins denote stations that withdraw from the auction and are labeled with their rank in the order of station withdrawals. Red pins denote stations whose price is frozen because they can no longer be repacked given the green stations' withdrawals in prior rounds. The top left map shows that the first 22 withdrawals in the region are processed with no constraints violated. The 23rd leads to a station being frozen in the top right map. The bottom left map shows that withdrawals 24 through 30 are then processed without incident, but the bottom right shows that the 31st station withdrawing and being repacked leads to the price being frozen for a handful of other stations.

Appendix C

Appendix to Chapter 3

C.1 Appendix: Specifications

C.1.1 Robustness

C.1.1.1 Entry windows and retirement ages

In our main specification, we choose the starting age of the entry window and the retirement age based on the empirical distribution of arrests over ages. We then choose the cutoff between the entry window and the exit window by maximizing the log-likelihood of the baseline CPDM estimation[163]. In this section, we arbitrarily vary these three cutoffs and show that our results are robust. Table C.1 presents the results estimated on our preferred specification.

C.1.1.2 Aging effects

In this section, we explore different functional forms of the aging effects on base exit rates and the robustness of the CPDM to the different parametrizations. Table C.2 presents the results. We note that, although in the last column the cubic term is statistically significant, we believe that the more parsimonious quadratic polynomial is sufficiently flexible. On the other hand, we have robust estimates across all specifications except the selection effect in the last column, which is imprecisely estimated.

[163]The reported standard errors do not take into account the uncertainty of the cutoffs.

Table C.1: Entry window and retirement cutoffs (entry start-entry end-retirement)

	(11-21-64)	(15-21-64)	(13-19-64)	(13-23-64)	(13-21-54)	(13-21-74)
Entry α_0	0.0168**	0.0119†	0.0137†	0.0106†	0.0160**	0.0128*
	(0.0899)	(0.2806)	(0.2570)	(0.2469)	(0.0855)	(0.1795)
SIL Entry α_1	0.0040***	0.0055**	0.0051**	0.0040**	0.0034**	0.0050***
	(0.0136)	(0.0342)	(0.0258)	(0.0229)	(0.0320)	(0.0130)
Exit γ_0	52.5361***	51.1861***	34.6856***	70.3408***	47.2253***	54.1816***
	(0.0028)	(0.0030)	(0.0000)	(0.0128)	(0.0013)	(0.0015)
Exit γ_1	-3.2961***	-3.2166***	-2.2753***	-4.2844***	-2.9579***	-3.3876***
	(0.0021)	(0.0022)	(0.0000)	(0.0103)	(0.0008)	(0.0011)
Exit γ_2	0.0490***	0.0479***	0.0350***	0.0623***	0.0445***	0.0498***
	(0.0014)	(0.0014)	(0.0000)	(0.0076)	(0.0004)	(0.0007)
Selection β_1	0.1158*	0.1141*	0.1352**	0.1024	0.2259**	0.1029†
	(0.1702)	(0.1871)	(0.0960)	(0.3266)	(0.0454)	(0.2056)
Floodgate λ_0	0.0927**	0.0842**	0.0666**	0.1056**	0.0738**	0.0994**
	(0.0447)	(0.0689)	(0.0691)	(0.0568)	(0.0601)	(0.0303)
Floodgate λ_1	0.0851**	0.0779**	0.0598**	0.0979**	0.0649**	0.0943**
	(0.0568)	(0.0860)	(0.0919)	(0.0709)	(0.0903)	(0.0372)
Floodgate λ_2	0.0911*	0.0844*	0.0644*	0.1070*	0.0715*	0.1020*
	(0.1390)	(0.1806)	(0.1912)	(0.1453)	(0.1716)	(0.1008)
Floodgate λ_3	0.0734†	0.0660	0.0477	0.0888†	0.0505	0.0817†
	(0.2431)	(0.3115)	(0.3411)	(0.2447)	(0.3337)	(0.2004)
Floodgate λ_4	0.0583	0.0506	0.0344	0.0747	0.0304	0.0664
	(0.4004)	(0.4838)	(0.5436)	(0.3674)	(0.6134)	(0.3233)
Floodgate λ_5	0.0329	0.0246	0.0131	0.0474	0.0041	0.0441
	(0.6843)	(0.7731)	(0.8452)	(0.6196)	(0.9544)	(0.5765)
Floodgate λ_6	0.0386	0.0295	0.0168	0.0555	0.0062	0.0554
	(0.6852)	(0.7724)	(0.8342)	(0.6184)	(0.9427)	(0.5559)
Floodgate λ_7	-0.0145	-0.0303	-0.0323	-0.0098	-0.0461	0.0053
	(0.8729)	(0.7607)	(0.6692)	(0.9273)	(0.5650)	(0.9526)
Floodgate λ_8	-0.0521	-0.0752	-0.0706	-0.0544	-0.0913	-0.0341
	(0.6030)	(0.5033)	(0.4057)	(0.6452)	(0.3309)	(0.7285)
Floodgate λ_{9+}	-0.1025	-0.1346	-0.1171†	-0.1216	-0.1396†	-0.0837
	(0.4200)	(0.3496)	(0.2866)	(0.4208)	(0.2589)	(0.5118)
Nb. Obs.	1549	1549	1549	1549	1549	1549

Notes: All regressions are run on the total violent crimes. Arrest rates of violent crimes, demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1 and γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). λ_j's measure the surprise effect in the j^{th} year after SIL passage. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

Table C.2: Parametrizations of aging effects

	Constant	Linear	Quadratic	Cubic	Translog	Quad. Exp.
Entry α_0	0.0210***	0.0082†	0.0148*	0.0108†	0.0143*	0.0148*
	(0.0049)	(0.2139)	(0.1418)	(0.2354)	(0.1319)	(0.1418)
SIL Entry α_1	0.0037***	0.0025**	0.0046**	0.0063***	0.0039**	0.0046**
	(0.0054)	(0.0311)	(0.0208)	(0.0049)	(0.0320)	(0.0208)
Exit γ_0	0.5269***	-5.6213***	51.8624***	-153.8857***	626.4235***	4.8548***
	(0.0000)	(0.0000)	(0.0027)	(0.0001)	(0.0004)	(0.0077)
Exit γ_1		0.1913***	-3.2551***	14.5884***	-363.0427***	-1.2217***
		(0.0000)	(0.0020)	(0.0002)	(0.0003)	(0.0038)
Exit γ_2			0.0484***	-0.4523***	52.3943***	0.0484***
			(0.0013)	(0.0004)	(0.0003)	(0.0013)
Exit γ_3				0.0045***		
				(0.0006)		
Selection β_1	0.3759***	0.2822**	0.1161*	0.0081	0.1701**	0.1161*
	(0.0179)	(0.0221)	(0.1697)	(0.9323)	(0.0816)	(0.1696)
Floodgate λ_0	0.1071***	0.0529**	0.0890**	0.1186***	0.0781**	0.0890**
	(0.0002)	(0.0970)	(0.0537)	(0.0157)	(0.0692)	(0.0537)
Floodgate λ_1	0.0962***	0.0454*	0.0822**	0.1137**	0.0711**	0.0822**
	(0.0120)	(0.1841)	(0.0669)	(0.0209)	(0.0875)	(0.0669)
Floodgate λ_2	0.1091***	0.0596*	0.0886*	0.1284**	0.0769*	0.0886*
	(0.0007)	(0.1572)	(0.1531)	(0.0473)	(0.1838)	(0.1531)
Floodgate λ_3	0.0960***	0.0476†	0.0706†	0.1213**	0.0577	0.0706†
	(0.0020)	(0.2518)	(0.2672)	(0.0768)	(0.3244)	(0.2672)
Floodgate λ_4	0.0770***	0.0296	0.0555	0.1171**	0.0397	0.0555
	(0.0132)	(0.5070)	(0.4294)	(0.0996)	(0.5456)	(0.4294)
Floodgate λ_5	0.0558*	0.0089	0.0301	0.0957†	0.0145	0.0301
	(0.1262)	(0.8647)	(0.7143)	(0.2320)	(0.8517)	(0.7143)
Floodgate λ_6	0.0630*	0.0113	0.0355	0.1098†	0.0178	0.0355
	(0.1309)	(0.8438)	(0.7152)	(0.2345)	(0.8463)	(0.7152)
Floodgate λ_7	0.0351	-0.0282	-0.0199	0.0527	-0.0334	-0.0199
	(0.3907)	(0.5733)	(0.8304)	(0.5417)	(0.6999)	(0.8304)
Floodgate λ_8	0.0008	-0.0607	-0.0609	0.0210	-0.0744	-0.0609
	(0.9880)	(0.3248)	(0.5588)	(0.8298)	(0.4480)	(0.5587)
Floodgate λ_{9+}	-0.0050	-0.0835	-0.1172	-0.0466	-0.1206	-0.1172
	(0.9487)	(0.3108)	(0.3825)	0.7168	(0.3368)	(0.3825)
Nb. Obs.	1549	1549	1549	1549	1549	1549

Notes: All regressions are run on the total violent crimes. Arrest rates of violent crimes, demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 and γ_3 are coefficients of the constant, linear, quadratic and cubic terms of the exit function (of age). For the translog function, we replace age with $log(age)$; for the quadratic experience column, we replace age with $age - 21$. λ_j's measure the surprise effect in the j^{th} year after SIL passage. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

Table C.3: CPDM with grouped floodgate effects

	Violent	Murder	Rape	Robbery	Agg. Ast.
Entry α_0	0.0130*	0.0003*	0.0010*	0.0121**	0.0077*
	(0.1989)	(0.1041)	(0.1839)	(0.0506)	(0.1946)
SIL Entry α_1 (-)	0.0051***	0.0001**	-0.0001	0.0013**	0.0017*
	(0.0082)	(0.0877)	(0.6576)	(0.0234)	(0.1082)
Exit γ_0	51.5256***	27.6088***	15.5613**	127.8716***	13.9920***
	(0.0034)	(0.0000)	(0.0554)	(0.0001)	(0.0038)
Exit γ_1	-3.2329***	-1.5160***	-1.0864**	-7.9605***	-0.9719***
	(0.0025)	(0.0000)	(0.0236)	(0.0001)	(0.0006)
Exit γ_2	0.0481***	0.0193***	0.0180***	0.1182***	0.0161***
	(0.0017)	(0.0000)	(0.0085)	(0.0000)	(0.0001)
Selection β_1 (-)	-0.0228	-0.2386	0.0909	0.4953*	-0.0708†
	(0.7713)	(0.6279)	(0.6114)	(0.1217)	(0.2219)
Floodgate λ_{0-1}	0.1036***	0.1471*	-0.0122	0.1914***	0.0438*
	(0.0122)	(0.1707)	(0.8044)	(0.0084)	(0.1923)
Floodgate λ_{2-4}	0.1012**	0.1330*	-0.0149	0.1904**	0.0441†
	(0.0879)	(0.1807)	(0.8109)	(0.0875)	(0.2535)
Floodgate λ_{5-9}	0.0445	0.1518†	-0.0625	0.1266	-0.0077
	(0.6051)	(0.2261)	(0.4402)	(0.5348)	(0.8893)
Floodgate λ_{10+}	0.0228	0.3195*	-0.0594	-0.0052	-0.0059
	(0.8338)	(0.1476)	(0.5055)	(0.9868)	(0.9378)
Log-likelihood	-7689	-2297	-4025	-6830	-7092
F-statistics	123.6	778.0	61.7	175.0	105.6
Nb. Obs.	1549	1548	1547	1546	1549

Notes: Arrest rates (of corresponding crime categories), demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). The F-statistics test for the joint significance of all estimated coefficients and reject the null (all coefficients are equal to zero) in all specifications. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

C.1.1.3 Floodgate effects

In our preferred specification, we adopt a non-parametric specification of the floodgate effects. In this section, we show that our estimates for all crime types are robust to more parametric specifications. Table C.3 presents the results when we group individual year fixed effects and Table C.4 shows the linear trend estimates.

Table C.4: CPDM with linear floodgate trend

	Violent	Murder	Rape	Robbery	Agg. Ast.
Entry α_0	0.0138*	0.0003**	0.0009†	0.0120**	0.0083*
	(0.1721)	(0.0912)	(0.2096)	(0.0516)	(0.1629)
SIL Entry α_1	0.0050***	0.0001**	-0.0001	0.0014***	0.0016*
	(0.0164)	(0.0962)	(0.6136)	(0.0168)	(0.1433)
Exit γ_0	51.6829***	27.7278***	16.3377**	127.9731***	14.8706***
	(0.0030)	(0.0000)	(0.0453)	(0.0001)	(0.0015)
Exit γ_1	-3.2418***	-1.5233***	-1.1334***	-7.9673***	-1.0244***
	(0.0022)	(0.0000)	(0.0198)	(0.0001)	(0.0002)
Exit γ_2	0.0482***	0.0194***	0.0187***	0.1183***	0.0168***
	(0.0015)	(0.0000)	(0.0076)	(0.0000)	(0.0000)
Selection β_1	0.0070	0.0271	0.1633	0.4449*	-0.0349
	(0.9364)	(0.9473)	(0.4631)	(0.1393)	(0.5816)
Floodgate cons.	0.1363***	0.1445*	0.0153	0.2343***	0.0735**
	(0.0005)	(0.1867)	(0.7482)	(0.0000)	(0.0220)
Floodgate slope	-0.0154**	-0.0058	-0.0150*	-0.0112	-0.0147**
	(0.0932)	(0.5364)	(0.1688)	(0.6826)	(0.0233)
Log-likelihood	-7685	-2291	-4022	-6832	-7090
F-statistics	134.7	834.3	38.5	152.9	98.1
Nb. Obs.	1549	1548	1547	1546	1549

Notes: Arrest rates (of corresponding crime categories), demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). The F-statistics test for the joint significance of all estimated coefficients and reject the null (all coefficients are equal to zero) in all specifications. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

C.1.1.4 OLS estimates

Table C.5 presents estimates from OLS without the dynamic panel instruments. We find similar results compared with Table 3.6 using IVs.

C.1.1.5 CPDM on levels

C.1.2 Literature replications

In this section, we review and test the robustness of model specifications in LM and AD. We use state-level panel data from 1980 onwards and only present results on the total violent crimes. LM adopts a simple "dummy variable model," where they only control for state and year fixed effects (but not trends). We first try to replicate their results with our data and then test its robustness with variations of the specification, controls, and sample lengths. Table C.7 shows the results. Column (1) resembles the most of their main specification. Specifically, the dependent variable is the log of crime rates and the demographic controls include arrest rates, state population, population density, real per capita personal income, income maintenance, unemployment insurance, and retirement payment for people older than 65. In particular, LM also control for a large set of race and age group variables (18 groups divided into three races - black, white, and others - and six age groups - 10-19, 20-29, 30-39, 40-49, 50-59, and 65+). We include the same controls in column 1 for comparison but later exclude them in our preferred DD specification. Similar to LM, we find a roughly 8.8% (vs. 5-10% in LM) reduction in violent crimes following SIL passages. In columns (2) and (3), we keep the same specification but expand the sample to 1999 and 2011, respectively. Despite having more observations in the sample, we find gradually smaller and less precisely estimated effects. With this specification and the full sample in (3), we find essentially zero effect of SILs on violent crimes. We then compare column (4) with (1) by dropping the controversial race and age controls. We also find small and almost insignificant effects. The last two columns are our preferred specifications[164], where we exclude the race and age controls but instead control for state-specific linear and quadratic time trends and account for serially correlated errors by clustering on the state level. We

[164] Column (6) corresponds to estimates reported in Table 3.8.

Table C.5: OLS estimates: CPDM preferred specification

	Violent	Murder	Rape	Robbery	Agg. Ast.
Entry α_0	0.0125*	0.0002	0.0010†	0.0090**	0.0083*
	(0.1258)	(0.3141)	(0.2291)	(0.0951)	(0.1153)
SIL Entry α_1 (-)	0.0032**	0.0002***	-0.0001	0.0010**	0.0011†
	(0.0410)	(0.0001)	(0.5929)	(0.0539)	(0.2377)
Exit γ_0	18.8892**	35.5137***	13.9282†	55.3285***	6.5723**
	(0.0328)	(0.0000)	(0.2331)	(0.0090)	(0.0828)
Exit γ_1	-1.2704**	-1.9969***	-0.9334*	-3.5189***	-0.5185***
	(0.0206)	(0.0000)	(0.1705)	(0.0059)	(0.0183)
Exit γ_2	0.0202***	0.0261***	0.0151*	0.0536***	0.0094***
	(0.0114)	(0.0000)	(0.1053)	(0.0031)	(0.0022)
Selection β_1 (-)	0.0727	0.9867***	0.2705*	0.2807**	-0.0814
	(0.3850)	(0.0039)	(0.1781)	(0.0857)	(0.3092)
Floodgate λ_0	0.0636**	0.3219***	-0.0182	0.1179*	0.0289
	(0.0894)	(0.0003)	(0.7077)	(0.1230)	(0.4013)
Floodgate λ_1	0.0529*	0.3383***	-0.0129	0.1339***	0.0120
	(0.1372)	(0.0001)	(0.8265)	(0.0104)	(0.7102)
Floodgate λ_2	0.0651†	0.2632***	0.0170	0.1002	0.0348
	(0.2309)	(0.0131)	(0.7802)	(0.3097)	(0.3887)
Floodgate λ_3	0.0468	0.3483***	-0.0589	0.1417†	0.0067
	(0.3519)	(0.0001)	(0.5021)	(0.2389)	(0.8494)
Floodgate λ_4	0.0330	0.3017***	-0.0612	0.1021	-0.0023
	(0.5540)	(0.0012)	(0.3384)	(0.3958)	(0.9537)
Floodgate λ_5	0.0121	0.2719***	-0.0455	0.0436	-0.0109
	(0.8558)	(0.0081)	(0.5212)	(0.7758)	(0.8212)
Floodgate λ_6	0.0230	0.2339***	-0.0526	0.0760	-0.0111
	(0.7584)	(0.0082)	(0.4802)	(0.6686)	(0.8440)
Floodgate λ_7	-0.0202	0.3339***	-0.2080**	-0.0059	-0.0369
	(0.7525)	(0.0008)	(0.0705)	(0.9727)	(0.4523)
Floodgate λ_8	-0.0581	0.1372*	-0.1092†	0.0400	-0.1056*
	(0.4782)	(0.1893)	(0.2737)	(0.8225)	(0.1237)
Floodgate λ_{9+}	-0.0752	0.0717	-0.1617*	-0.1756	-0.0710
	(0.4697)	(0.5762)	(0.1341)	(0.5674)	(0.3539)
Nb. Obs.	1549	1548	1547	1546	1549

Notes: Arrest rates (of corresponding crime categories), demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). λ_j's measure the surprise effect in the j^{th} year after SIL passage. Key coefficients relevant for testing the deterrence hypothesis are signed in parentheses. Standard errors are clustered at the state level. Two-sided p values are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

Table C.6: CPDM dependent variable: Changes vs. levels

	Change Δ_{st}^{C}	Level $C_{s,t+1}$
Lag C_{st}		1.1261***
(S.E.)		(0.4573)
Entry α_0	0.0148*	0.0135*
	(0.1418)	(0.1743)
SIL Entry α_1 (-)	0.0046**	0.0037**
	(0.0208)	(0.0394)
Exit γ_0	51.8624***	25.7653***
	(0.0027)	(0.0005)
Exit γ_1	-3.2551***	-1.7063***
	(0.0020)	(0.0001)
Exit γ_2	0.0484***	0.0270***
	(0.0013)	(3E-5)
Selection β_1 (-)	0.1161*	0.1987**
	(0.1697)	(0.0242)
Floodgate λ_0	0.0890**	0.0703*
	(0.0537)	(0.1158)
Floodgate λ_1	0.0822**	0.0640*
	(0.0669)	(0.1428)
Floodgate λ_2	0.0886*	0.0757†
	(0.1531)	(0.2219)
Floodgate λ_3	0.0706†	0.0603
	(0.2672)	(0.3314)
Floodgate λ_4	0.0555	0.0427
	(0.4294)	(0.5287)
Floodgate λ_5	0.0301	0.0169
	(0.7143)	(0.8330)
Floodgate λ_6	0.0355	0.0217
	(0.7152)	(0.8122)
Floodgate λ_7	-0.0199	-0.0288
	(0.8304)	(0.7343)
Floodgate λ_8	-0.0609	-0.0681
	(0.5588)	(0.4868)
Floodgate λ_{9+}	-0.1172	-0.1099
	(0.3825)	(0.3815)
Nb. Obs.	1549	1498

Notes: Arrest rates (of corresponding crime categories), demographic and welfare controls, state and year fixed effects and state-specific linear and quadratic time trends are controlled for but not reported. γ_0, γ_1, γ_2 are coefficients of the constant, linear and quadratic terms of the exit function (of age). λ_j's measure the surprise effect in the j^{th} year after SIL passage. Key coefficients relevant for testing the deterrence hypothesis are signed in parentheses. Standard errors are clustered at the state level. Two-sided p values (except for the lag variable, which shows the standard error) are in parentheses. †, *, **, and *** indicate one-sided statistical significance at the 15, 10, 5, and 1 percent level.

find no effects on both the log and the level of crimes. Overall, we find that the original LM specification is sensitive to controls, sample lengths, and assumptions on error structures.

On the other end of the debate, AD study the effects of SILs up to 1999 and employ a "hybrid model." In addition to the level shift in a standard DD, they include a trend-break (overall trend interacted with the SIL dummy) term post-SIL to capture the slope change. They find overall positive effects of SILs on violent crimes and positive "long run" effects of SILs suggested by their trend-break term. We argue that, however, in a DD specification, if our state-specific trends are flexible enough, we should not need the trend-break term. Therefore, in our preferred DD specification, we include state-specific quadratic time trends that will capture the "inverted-V" shape argued in this literature. Table C.8 presents the results. In column (1), we follow AD and drop the race and age controls. We find an overall increase of about 7.4% in crimes following SIL adoptions. We add the trend-break term in column (2) and find similar results to AD. In (3) and (4), we simply vary the sample length and again find inconsistent results over time. In (5), we add back the race and age controls for comparison. Finally, (6) and (7) are our preferred specifications[165]. We find the opposite effects compared to (1), after controlling for state-specific linear and quadratic time trends and clustering standard errors.

[165]Column (7) corresponds to estimates reported in Table 3.8.

Table C.7: Replication and variations of Lott and Mustard (1997)

	(1)	(2)	(3)	(4)	(5)	(6)
SIL Dummy	-0.0881***	-0.0276*	-0.0063	-0.0260†	-0.0015	-0.8789
	(0.0000)	(0.0716)	(0.5885)	(0.1298)	(0.9615)	(0.9550)
Dep. Var.	log	log	log	log	log	level
Sample	1980-1992	1980-1999	1980-2011	1980-1992	1980-1992	1980-1992
Race & age controls	Y	Y	Y	N	N	N
State trends	N	N	N	N	Y	Y
Clustering	N	N	N	N	Y	Y
Demographic controls	Y	Y	Y	Y	Y	Y

Notes: All regressions are run on the total violent crimes. Two-sided p values are in parentheses. †, *, **, and *** indicate two-sided statistical significance at the 15, 10, 5, and 1 percent level.

Table C.8: Replication and variations of Ayres and Donohue (2003*b*)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SIL Dummy	0.0737***	-0.0832**	-0.0260†	0.1088***	-0.0276*	-0.0385*	-38.0306*
	(0.0000)	(0.0135)	(0.1298)	(0.0000)	(0.0716)	(0.0665)	(0.0582)
SIL Trend		0.0107***					
		(0.0000)					
Dep. Var.	log	log	log	log	log	log	level
Sample	1980-1999	1980-1999	1980-1992	1980-2011	1980-1999	1980-1999	1980-1999
Race & age controls	N	N	N	N	Y	N	N
State trends	N	N	N	N	N	Y	Y
Clustering	N	N	N	N	N	Y	Y
Demographic controls	Y	Y	Y	Y	Y	Y	Y

Notes: All regressions are run on the total violent crimes. Two-sided p values are in parentheses. †, *, **, and *** indicate two-sided statistical significance at the 15, 10, 5, and 1 percent level.

C.2 Appendix: Data

C.2.1 State SIL passage years

C.2.2 Age-specific arrest and crime rates

We first use the BJS national arrests by age groups and the shape-preserving piecewise cubic hermite interpolating polynomials to impute age-specific arrests[166]. Figure C.2.1 presents the fit results for 1980 and 2010 in four crime categories.

To impute age-specific crime rates, let p_{st} be the probability of arrest for criminals in state s and year t, assuming it does not vary across ages. We also assume that every criminal commits κ crimes each year across states, years and ages. Let C be the number of crimes, A the number of arrests, and then we have, by definition, $\frac{C_{ast}}{\kappa} \cdot p_{st} = A_{ast}$, where the subscript a indicates age. Summing over ages and dividing the two equations, we get $\frac{\frac{C_{ast}}{\kappa} p_{st}}{\sum_a \frac{C_{ast}}{\kappa} p_{st}} = \frac{A_{ast}}{\sum_a A_{ast}}$, and after manipulations, $C_{ast} = \frac{A_{ast}}{\sum_a A_{ast}} C_{st}$. We, however, do not observe arrests on the age-state-year level and have to rely on an additional assumption that the arrests for each age group as a fraction of the total arrests do not vary across states, i.e. $\frac{A_{at}}{\sum_a A_{at}} = \frac{A_{ast}}{\sum_a A_{ast}}$. It is plausible that criminals of age 20 in Pennsylvania do no better or worse than those in North Carolina compared to other age groups in the same state. Then we arrive at the desired variable, age-specific crime rates, $C_{ast} = \frac{A_{at}}{\sum_a A_{at}} C_{st}$, where A_{at} are the age-specific national arrests imputed from BJS and C_{st} are the state-year level crime rates data from UCR. We then let the exit cohort $N_{ast}^{Ex} = C_{ast}$.

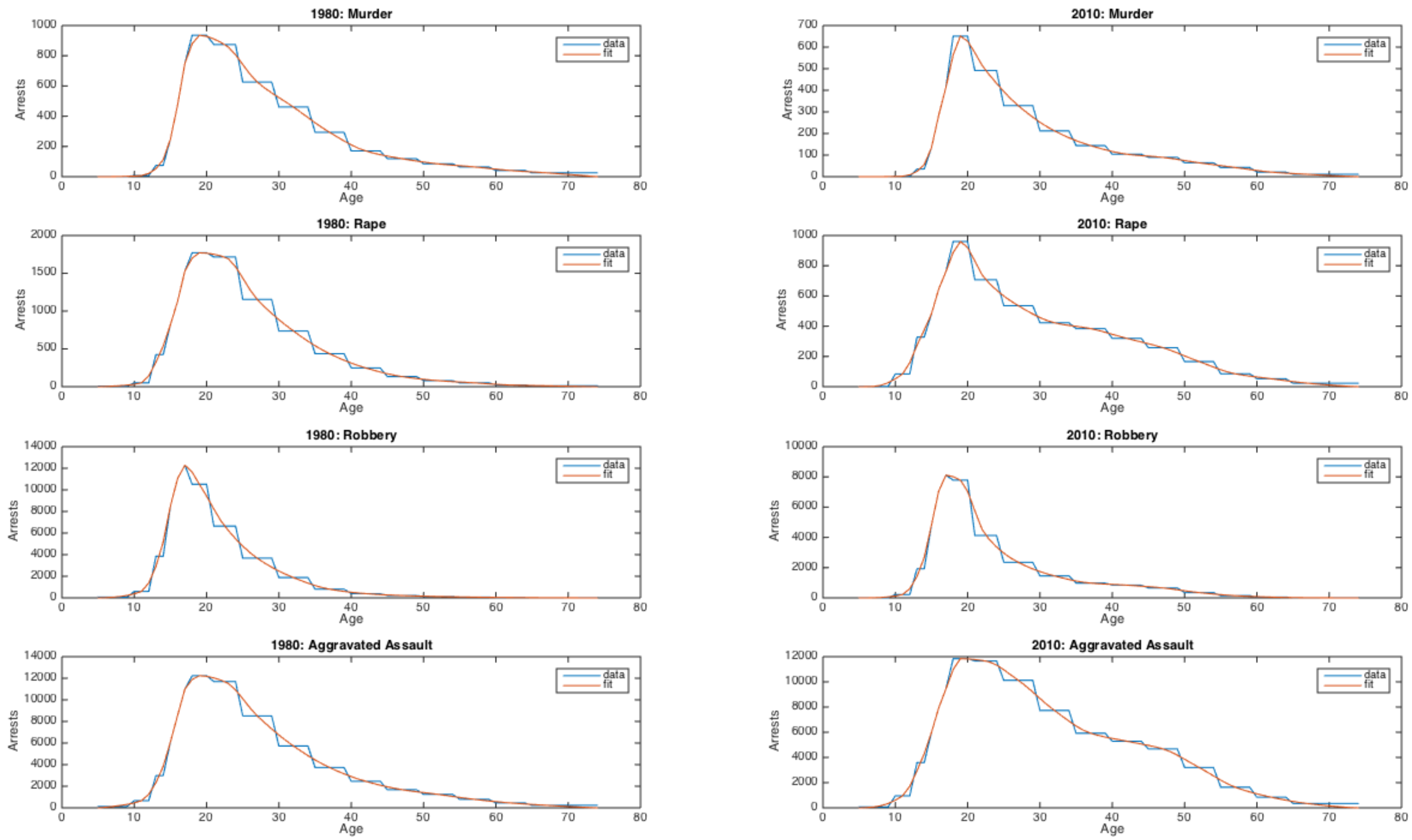
[166] Specifically, we assume there are no violent crimes committed by people younger than 5 or older than 74. We then assume that the mean age point in an age group has the average arrests in the age group. For example, there are 21 murders for age group 10-12 in 1987 and thus we let the 11-year olds have 7 murders in order to construct our data points. Then we interpolate over these data points using cubic hermite polynomials to impute arrests for each specific age.

Table C.9: State SIL passages

Pre-1985	**1986-1992**	**1995-1997**	**2002-2007**	**Post-2011**
Alabama (*always*)	Maine (*1986*)	Alaska (*1995*)	Michigan (*2002*)	Iowa (*2011*)
Vermont (*always*)	North Dakota (*1986*)	Arizona (*1995*)	Missouri (*2002*)	Wisconsin (*2011*)
New Hampshire (*1960*)	South Dakota (*1987*)	Tennessee (*1995*)	Colorado (*2004*)	California (*never*)
Washington (*1962*)	Florida (*1988*)	Wyoming (*1995*)	Minnesota (*2004*)	Delaware (*never*)
Connecticut (*1970*)	Virginia (*1989*)	Arkansas (*1996*)	New Mexico (*2004*)	Hawaii (*never*)
Indiana (*1981*)	Georgia (*1990*)	Nevada (*1996*)	Ohio (*2005*)	Illinois (*never*)
	Pennsylvania (*1990*)	North Carolina (*1996*)	Kansas (*2007*)	Maryland (*never*)
	West Virginia (*1991*)	Oklahoma (*1996*)	Nebraska (*2007*)	Massachusetts (*never*)
	Idaho (*1991*)	Texas (*1996*)		New Jersey (*never*)
	Mississippi (*1991*)	Utah (*1996*)		New York (*never*)
	Oregon (*1991*)	Kentucky (*1997*)		Rhode Island (*never*)
	Montana (*1992*)	Louisiana (*1997*)		District of Columbia (*never*)
		South Carolina (*1997*)		

Notes: State SIL passage years coding by adoption waves. Years in parentheses indicate the year that state SIL went into effect, typically the following year of law passage. We categorize any concealed carry laws that are equivalent to or more restrictive than may-issue laws (MILs) as non-SILs and laws that are more lenient than SILs as SILs.

Figure C.2.1: Arrests by age in 1980 and 2010 (# of arrests)



Notes: Rugged blue lines represent the data as we simply divide arrests evenly in an age group. Smooth orange lines represent the well-fitted single-age arrests.

Bibliography

Aghion, Philippe, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison and Patrick Legros. 2015. "Industrial Policy and Competition." *American Economic Journal: Macroeconomics* 7(4):1–32.

Anderson, Theodore W. and Cheng Hsiao. 1982. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18:47–82.

Asker, John. 2010. "A Study of the Internal Organisation of a Bidding Cartel." *American Economic Review* 100(3):724–762.

Ausubel, Lawrence, Peter Cramton, Marek Pycia, Marzena Rostek and Marek Weretka. 2014. "Demand reduction and inefficiency in multi-unit auctions." Working Paper, University of Maryland, College Park, MD.

Ayres, Ian and John J. Donohue. 2003*a*. "The Latest Misfires in Support of the More Guns, Less Crime Hypothesis." *Stanford Law Review* 55(4):1371–1398.

Ayres, Ian and John J. Donohue. 2003*b*. "Shooting Down the More Guns, Less Crime Hypothesis." *Stanford Law Review* 55(4):1193–1312.

Back, Kerry and Jaime Zender. 1993. "Auctions of divisible goods: on the rationale for the Treasury experiment." *Review of Financial Studies* 6(4):733–764.

Bajari, Patrick, C. Lanier Benkard and Jonathan Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75(5):1331–1370.

Bass, Frank M. 1969. "A New Product Growth Model for Consumer Durables." *Management science* 15(5):215–227.

Becker, Gary. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76(2):169–217.

Berry, Steven. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* 25(2):242–262.

Berry, Steven, Alon Eizenberg and Joel Waldfogel. 2016. "Optimal Product Variety in Radio Markets." *The RAND Journal of Economics* 47(3):463–497.

Berry, Steven, James Levinsohn and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4):841–890.

Berry, Steven and Joel Waldfogel. 1999. "Free Entry and Social Inefficiency in Radio Broadcasting." *The RAND Journal of Economics* 30(3):397–420.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1):249–275.

Black, Dan A. and Daniel S. Nagin. 1998. "Do Right-to-Carry Laws Deter Violent Crimes." *Journal of Legal Studies* 27:209–219.

Blonigen, Bruce A., Christopher R. Knittel and Anson Soderbery. 2013. "Keeping it Fresh: Strategic Product Redesigns and Welfare." NBER Working Paper.

Borenstein, Severin, James Bushnell and Frank Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92(5):1376–1405.

Bushnell, James and Catherine Wolfram. 2012. "Enforcement of Vintage Differentiated Regulations: The Case of New Source Review." *Journal of Environmental Economics and Management* 64:137–152.

Buzzell, Robert D. 1972. *Marketing: A Contemporary Analysis*. McGraw-Hill.

Cantillon, Estelle and Martin Pesendorfer. 2007. "Combination Bidding in Multi-Unit Auctions." CEPR Working Paper DP6083.

Chambers, John C., Satinder K. Mullick and Donald D. Smith. 1971. "How to Choose the Right Forecasting Technique." *Harvard Business Review* 49(4):45.

Chu, Chenghuan Sean. 2010. "The Effect of Satellite Entry on Cable Television Prices and Product Quality." *The RAND Journal of Economics* 41(4):730–764.

Coey, Dominic, Bradley Larsen and Kane Sweeney. 2015. "The Bidder Exclusion Effect." NBER Working Paper 20523.

Collard-Wexler, Allan. 2013. "Demand Fluctuations in the Ready-Mix Concrete Industry." *Econometrica* 81(3):1003–1037.

Conley, Timothy and Francesco Decarolis. 2016. "Detecting Bidders Groups in Collusive Auctions." *American Economic Journal: Microeconomics* forthcoming.

Cox, William E. 1967. "Product Life Cycles as Marketing Models." *The Journal of Business* 40(4):375–384.

Cramton, Peter and Jesse A. Schwartz. 2002. "Collusive Bidding in the FCC Spectrum Auctions." *Contributions in Economic Analysis & Policy* 1(1).

Crawford, Gregory S. and Matthew Shum. 2007. "Monopoly Quality Degradation and Regulation in Cable Television." *Journal of Law and Economics* 50(1):181–219.

Day, George S. 1981. "The Product Life Cycle: Analysis and Applications Issues." *Journal of Marketing* 45(4):60–67.

Dezhbakhsh, Hashem and Paul H. Rubin. 1998. "Lives Saved or Lives Lost? The Effects of Concealed-Handgun Laws on Crime." *American Economic Review* 88(2):468–474.

Dhalla, Nariman K. and Sonia Yuspeh. 1976. "Forget the Product Life Cycle Concept." *Harvard Business Review* 54(1):102–112.

Duggan, Mark and Fiona Scott Morton. 2006. "The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing." *Quarterly Journal of Economics* 121(1):1–30.

Eizenberg, Alon. 2014. "Upstream Innovation and Product Variety in the US Home PC Market." *The Review of Economic Studies* 81(3):1003–1045.

Engelbrecht-Wiggans, Richard and Charles Kahn. 1998. "Multi-unit auctions with uniform prices." *Economic Theory* 12(2):227–258.

Engelmann, Dirk and Veronika Grimm. 2009. "Bidding behaviour in multi-unit auctions - an experimental investigation." *Economic Journal* 119(537):855–882.

Enis, Ben M., Raymond La Garce and Arthur E. Prell. 1977. "Extending the Product Life Cycle." *Business Horizons* 20(3):46–56.

Fan, Ying. 2013. "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market." *The American Economic Review* 103(5):1598–1628.

Fan, Ying and Chenyu Yang. 2016. "Competition, Product Proliferation and Welfare: A Study of the U.S. Smartphone Market." Working paper.

Fowlie, Meredith. 2009. "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage." *American Economic Journal: Economic Policy* 1(2):72–112.

Fox, Jeremy and Patrick Bajari. 2013. "Measuring the Efficiency of an FCC Spectrum Auction." *American Economic Journal: Microeconomics* 5(1):100–146.

Gandhi, Amit and Jean-François Houde. 2015. "Measuring Substitution Patterns in Differentiated Products Industries." Working paper.

Goeree, Jacob, Theo Offerman and Randolph Sloof. 2013. "Demand reduction and preemptive bidding in multi-unit license auctions." *Experimental Economics* 16(1):52–87.

Goeree, Michelle Sovinsky. 2008. "Limited Information and Advertising in the US Personal Computer Industry." *Econometrica* 76(5):1017–1074.

Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy* 108(2):352–378.

Gormley, Todd A. and David A. Matsa. 2011. "Growing out of Trouble: Corporate Responses to Liability Risk." *Review of Financial Studies* 24(8):2781–2821.

Gowrisankaran, Gautam and Marc Rysman. 2012. "Dynamics of Consumer Demand for New Durable Goods." *Journal of Political Economy* 120(6):1173–1219.

Greenstein, Shane M. and James B. Wade. 1998. "The Product Life Cycle in the Commercial Mainframe Computer Market, 1968-1982." *The RAND Journal of Economics* 29(4):772–789.

Greenwald, Bruce and Joseph E. Stiglitz. 2006. "Helping Infant Economies Grow: Foundations of Trade Policies for Developing Countries." *The American Economic Review* 96(2):141–146.

Grimm, Veronika, Frank Riedel and Elmar Wolfstetter. 2003. "Low price equilibrium in multi-unit auctions: the GSM spectrum auction in Germany." *International Journal of Industrial Organization* 21(10):1557–1569.

Harrell, Stephen G. and Elmer D. Taylor. 1981. "Modeling the Product Life Cycle for Consumer Durables." *Journal of Marketing* 45(4):68–75.

Hirschi, Travis and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *The American Journal of Sociology* 89(3):552–584.

Hortacsu, Ali and Steven Puller. 2008. "Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market." *The RAND Journal of Economics* 39(1):86–114.

Hotz, V. Joseph, Robert A. Miller, Seth Sanders and Jeffrey Smith. 1994. "A Simulation Estimator for Dynamic Models of Discrete Choice." *The Review of Economic Studies* 61(2):265–289.

Hughes, G. David. 1990. "Managing High-Tech Product Cycles." *The Executive* 4(2):44–55.

Imai, Susumu and Kala Krishna. 2004. "Employment, Deterrence, and Crime in a Dynamic Model." *International Economic Review* 45:845–872.

Iyvarakul, Tongyai, Marjorie B. McElroy and Kalina Staub. 2011. "Dynamic Optimization in Models for Panel Data: A Cohort Panel Data Model of the Effects of Divorce Laws on Divorce Rates." Working paper.

Jarvis, Raymond A. and Edward A. Patrick. 1973. "Clustering Using a Similarity Measure Based on Shared Near Neighbors." *IEEE Transactions on Computers* 100(11):1025–1034.

Johnson, Justin P. and David P. Myatt. 2003. "Multiproduct Quality Competition: Fighting Brands and Product Line Pruning." *The American Economic Review* 93(3):748–774.

Johnson, Rucker and Steven Raphael. 2012. "How Much Crime Reduction Does the Marginal Prisoner Buy?" *Journal of Law and Economics* 55(2):275–310.

Jovanovic, Boyan. 1994. "The Life Cycle of a Competitive lndustry." *Journal of Political Economy* 102(2):322–347.

Jun, Byoung and Elmar Wolfstetter. 2004. "Signaling equilibria in a multi-unit English clock auction." Working Paper, Korea University, Seoul.

Kagel, John and Dan Levin. 2001. "Behavior in multi-unit demand auctions: experiments with uniform price and dynamic Vickrey auctions." *Econometrica* 69(2):413–454.

Kaplan, Greg and Guido Menzio. 2015. "The Morphology of Price Dispersion." *International Economic Review* 56(4):1165–1206.

Kawai, Kei and Jun Nakabayashi. 2015. "Detecting Large-Scale Collusion in Procurement Auctions." Working paper, UC Berkeley, Berkeley, CA.

Klepper, Steven. 1996. "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *The American Economic Review* 86(3):562–583.

Kotler, Philip and Gary Armstrong. 2010. *Principles of Marketing*. Pearson Education.

Kwoka, John E. 1996. "Altering the Product Life Cycle of Consumer Durables: The Case of Minivans." *Managerial and Decision Economics* 17(1):17–25.

Laffont, Jean-Jacques and Jean Tirole. 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94(3):614–641.

Lee, Robin S. and Ariel Pakes. 2009. "Multiple Equilibria and Selection by Learning in an Applied Setting." *Economics Letters* 104(1):13–16.

Levitt, Theodore. 1965. "Exploit the Product Life Cycle." *Harvard Business Review* 18:81–94.

Li, Ying, Joe Mazur, James Roberts and Andrew Sweeting. 2016. "Airline Mergers and the Potential Entry Defense." Working Paper.

List, John and David Lucking-Reiley. 2000. "Demand reduction in multiunit auctions: evidence from a sportscard field experiment." *American Economic Review* 90(4):961–972.

Lott, Jr, John R. and David B. Mustard. 1997. "Crime, Deterrence and Right-to-Carry Concealed Handguns." *Journal of Legal Studies* 26(1):1–68.

Mazzeo, Michael J. 2002. "Product Choice and Oligopoly Market Structure." *The RAND Journal of Economics* 33(2):221–242.

McFadden, Daniel. 2012. "Computing Willingness-to-Pay in Random Utility Models." *Trade, Theory and Econometrics* .

Menezes, Flavio. 1996. "Multiple-unit English auctions." *European Journal of Political Economy* 12(4):671–684.

Merryman, Scott. 2008. "CHINA_MAP: Stata module to provide map of China's provinces." *Boston College Department of Economics* .

Michelle Grantham, Lisa. 1997. "The Validity of the Product Life Cycle in the High-Tech Industry." *Marketing Intelligence & Planning* 15(1):4–10.

Milgrom, Paul and Ilya Segal. 2014. "Deferred-acceptance auctions and radio spectrum reallocation." Working Paper, Stanford University, Stanford, CA.

Milgrom, Paul, Lawrence Ausubel, Jon Levin and Ilya Segal. 2012. "Incentive auction rules option and discussion." Working Paper, Stanford University, Stanford, CA.

Moody, Carlisle E. and Thomas B. Marvell. 2008. "The Debate on Shall-Issue Laws." *Econ Journal Watch* 5(3):269–293.

Moore, Gordon E. 1965. "Cramming More Components onto Integrated Circuits." *Electronics* 38(8).

Moral, María José and Jordi Jaumandreu. 2007. "Automobile Demand, Model Cycle and Age Effects." *Spanish Economic Review* 9(3):193–218.

Neal, Derek and Armin Rick. 2014. "The Prison Boom and the Lack of Black Progress after Smith and Welch." NBER working paper.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69(2):307–342.

Ngwe, Donald. 2016. "Why Outlet Stores Exist: Averting Cannibalization in Product Line Extensions." *Marketing Science* (forthcoming).

Nosko, Chris. 2014. "Competition and Quality Choice in the CPU Market." Working Paper.

Oyer, Paul. 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113(1):149–185.

Pakes, Ariel, Jack Porter, Kate Ho and Joy Ishii. 2015. "Moment inequalities and their application." *Econometrica* 83(1):315–334.

Petrin, Amil. 2002. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy* 110(4):705–729.

Plassmann, Florenz and T. Nicolaus Tideman. 2001. "Does the Right to Carry Concealed Handguns Deter Countable Crimes? Only a Count Analysis Can Say." *Journal of Law and Economics* 44:771–798.

Polli, Rolando and Victor Cook. 1969. "Validity of the Product Life Cycle." *The Journal of Business* 42(4):385–400.

Porter, Robert and Douglas Zona. 1993. "Detection of Bid Rigging in Procurement Auctions." *Journal of Political Economy* 101(3):518–538.

Riedel, Frank and Elmar Wolfstetter. 2006. "Immediate demand reduction in simultaneous ascending-bid auctions: a uniqueness result." *Economic Theory* 29:721–726.

Ryan, Stephen P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica* 80(3):1019–1061.

Sampere, Juan Pablo Vazquez. 2014. "Xiaomi, Not Apple, Is Changing the Smartphone Industry." *Harvard Business Review* .

Schumpeter, Joseph Alois. 1934. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Vol. 55 Transaction Publishers.

Segerstrom, Paul S., Thirumalai C. A. Anant and Elias Dinopoulos. 1990. "A Schumpeterian Model of the Product Life Cycle." *The American Economic Review* pp. 1077–1091.

Seim, Katja. 2006. "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *The RAND Journal of Economics* 37(3):619–640.

Sinkinson, Michael. 2014. "Pricing and Entry Incentives with Exclusive Contracts: Evidence from Smartphones." Working Paper.

Spence, Michael. 1976. "Product Differentiation and Welfare." *The American Economic Review* 66(2):407–414.

Stavins, Joanna. 1995. "Model Entry and Exit in a Differentiated-Product Industry: The Personal Computer Market." *The Review of Economics and Statistics* 77(4):571–584.

Sweeting, Andrew. 2013. "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry." *Econometrica* 81(5):1763–1803.

Tellis, Gerard J. and C. Merle Crawford. 1981. "An Evolutionary Approach to Product Growth Theory." *Journal of Marketing* 45(4):125–132.

Terzi, Sergio, Abdelaziz Bouras, Debashi Dutta, Marco Garetti and Dimitris Kiritsis. 2010. "Product Lifecycle Management - From Its History to Its New Role." *International Journal of Product Lifecycle Management* 4(4):360–389.

Vernon, Raymond. 1966. "International Investment and International Trade in the Product Cycle." *The Quarterly Journal of Economics* 80(2):190–207.

Weber, Robert. 1997. "Making more from less: strategic demand reduction in the FCC spectrum auctions." *Journal of Economics and Management Strategy* 6(3):529–548.

White, Lawrence J. 1977. "Market Structure and Product Varieties." *The American Economic Review* 80(2):179–182.

Wilson, Robert. 1979. "Auctions of shares." *Quarterly Journal of Economics* 93(47):675–689.

Wolfram, Catherine. 1998. "Strategic Bidding in a Multiunit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales." *The RAND Journal of Economics* 29(4):703–725.

Wollmann, Thomas. 2016. "Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles." Working Paper.

Wolpin, Kenneth I. 1984. "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality." *Journal of Political Economy* 92(5):852–874.

Wood, Laurie. 1990. "The End of the Product Life Cycle? Education Says Goodbye to an Old Friend." *Journal of Marketing Management* 6(2):145–155.

Yang, Chenyu. 2016. "Vertical Integration and Innovation in the Chipset and Smartphone Industry." Working Paper.