Commitment, Vertical Restraints, and Dynamic Pricing of Durable Goods

Øystein Daljord*

January 6, 2019

Abstract

The Norwegian book industry has historically been regulated by a fixed price agreement which lets the publisher fix the retail price for a limited time period. The agreement is enforced by the association of publishers and the association of booksellers and commits the retailers to not discount a title early in its lifecycle. A new competition act in 2005 forced a weakening of the agreement: the restraint period was shortened and the price restraint itself was softened. Two changes in the lifecycle sales followed: retailers started discounting titles earlier, and demand shifted towards the earlier discounts. In a market with forward-looking consumers, commitment to limited future discounts can be profitable by discouraging intertemporal substitution. I quantify the commitment value of the agreement relative to standard vertical restraints with and without commitment. The commitment value is evaluated using a dynamic market equilibrium model with forward-looking demand side and a forward-looking supply side at parameters estimated from the data. The profitability of a fixed price agreement with commitment is estimated to be less than the profitability from alternative vertical restraints without commitment.

*Booth School of Business, University of Chicago, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA. E-mail: oeystein.daljord@chicagobooth.edu. Web: faculty.chicagobooth.edu/oystein.daljord A special thanks to Harikesh Nair, Lanier Benkard and Wes Hartmann. Thanks also to Tim Bresnahan, Chris Conlon, Liran Einav, Jeremy Fox, Pedro Gardete, Jakub Kastl, James Lattin, Jon Levin, Erik Madsen, Josh Mollner, Sridhar Narayanan, Christian Riis, Ali Yurukoglu and Pavel Zruymov for comments. Thanks to Bokdatabasen for access to their database, to Henrik Lande and Gjermund Nese at the Norwegian Competition Authority for facilitating access to the sales data, to Retriever for access to their media archives, and to Mike Saunders and Marcus Edwall for help with software. Thanks to Shuo Xie for excellent research assistance.

1 Introduction

Fixed price agreements regulate the pricing of books in many countries.[1] The Norwegian fixed price agreement is a voluntary trade agreement between the association of publishers and the association of booksellers.[2] Under the agreement, the publishers decide on the retail prices of their titles for a limited time period, e.g. a year.

A contract term that allows the manufacturer to decide the retail price of its products is a vertical restraint known as Resale Price Maintenance (RPM). RPM is usually featured in bilateral contracts between a manufacturer and a retailer. Standard theories explain RPM as a channel strategy that helps the manufacturer coordinate the retail price level. RPM can set a retail price ceiling (Spengler (1950)), which prevents double marginalization, or it can set a retail price floor, which gives retailers incentives to offer complementary services (Telser (1960)). The evidence in Section 4 however shows that retail prices in the Norwegian book market hardly changed in the period of introduction though retailers could provide discounts under the new Agreement. Instead, the prices unravel later in the lifecycle. It appears the Agreement did not coordinate price levels, it coordinated price paths.

A fixed price agreement differs from standard bilateral RPM in two respects. Firstly, it is a multilateral agreement, between all publishers and all retailers, on using RPM. Though the Agreement regulates a term of sale between a publisher and a retailer, the Agreement also allows for enforcement of the restraint by rival publishers and retailers. A detailed discussion of the Agreement is given in Section 2. Secondly, the price restraints hold for a limited time period. The time limited price restraint, coupled with external enforcement, implies not just a price level, but it also commits the industry to particular dynamic price strategy: titles will not be discounted early in the lifecycle.

While a seller of a durable good, such as books, may want to price discriminate myopic consumers by gradually lowering the price over time, a commitment to limited future discounts may be more profitable if consumers are forward looking. Coase (1972) points out that forward looking consumers who anticipate future discounts may strategically delay their purchases. Now the seller is effectively competing against expectations of its own future prices. Expectations of future discounts can even become self-fulfilling as shrinking current demand forces the seller to discount to retain profits. In the limit, intertemporal substitution lead prices to immediately unravel to costs, and even a monopolist may fail to make a profit. There is a rich theory literature on dynamic pricing of durable goods with forward looking consumers since Coase (1972), see Waldman (2003) for a survey.

Stokey (1979) shows that faced with consumers as forward looking as itself, a monopolist seller of a durable good would in fact prefer to not price discriminate at all, but rather commit to charge a constant price. By giving up control over its own prices,

[1] Fixed price agreements regulate book markets in France, Germany, Austria, Netherlands, Lebanon, Italy, Spain, Portugal, Greece, Hungary, Israel, Slovenia, Argentina, South Korea, Japan and Mexico (International Publishers Association (2014)).

[2] In some countries, the fixed price agreements are encoded as government law.

the seller curbs competition from expectations of its own future discounts. A fixed price agreement effectively commits the industry collectively to a price strategy that its members may find hard to implement individually. In this paper, I quantify the commitment value of a fixed price agreement.

The question is motivated by a change in the Norwegian Competition Act in 2005 which forced a revision of the industry's fixed price agreement ("Agreement"). The change in the Agreement forced by the new Competition Act is interpreted as a natural experiment in vertical restraints. Following a political process, described in Section 2, the price restraint period in the new Agreement was shortened by eight months, and softened from a retail price point to a retail price interval. The change in the lifecycle price strategies is evidence of dynamic effects of fixed price agreements that to the best of my knowledge have not been documented before.

The commitment value of the Agreement depends on the consumers' price expectations and their willingness to delay their purchases. Whereas the shift in price strategies following the new Agreement are directly observable, consumers' price expectations and time preferences must be inferred from the sales data. I estimate a dynamic discrete choice model of demand of the Rust (1994) kind on market level data that allows for substitution between retailers and over time. Using the identification strategy in Abbring and Daljord (2018), the shift in lifecycle sales following the new Agreement is taken as informative of consumers time preferences. The strategy exploits that at virtually the same introductory prices, consumers purchased a smaller share of lifecycle demand in the introductory period under the new Agreement, when discounts came earlier, then under the old Agreement. The shift in demand towards earlier discounts is consistent with forward-looking consumers adapting the timing of their purchases to the new price strategies.

The commitment value is taken to be the profits that a fixed price agreement can generate beyond the profits that can be had form standard, alternative bilateral vertical contracts, such as dynamic two-part tariffs and linear contracts, which do not offer commitment. Since it is not publicly known which vertical contracts the industry switched to under the new Agreement, it is hard to interpret a direct comparison of profits under the old and the new Agreement. I therefore turn to an analysis of counterfactual vertical contracts in Section 9.

The counterfactual analysis compares the profitability of a fixed price agreement to standard alternative, vertical contracts. The vertical contracts are evaluated in a market equilibrium model, where the demand parameters are estimated from the data. In the model, a monopolist publisher supplies an oligopoly of retailers using vertical contracts. The counterfactual contracts vary with the level commitment and retail price coordination they offer. Assuming that the vertical contracts are common knowledge, in particular whether there is a fixed price agreement or not, the contracts affect both the consumers price expectations and the retailers' price incentives.

While the observed changes in the prices and sales following the new Agreement are

consistent with an industry that lost commitment power, the counterfactual analysis in Section 10 suggests that commitment value of the Agreement is modest. For instance, at an estimated discount factor of 0.780, the profitability of the Agreement relative to a dynamic two-part tariff, which coordinates intrabrand pricing without commitment, is -2.49% for a typical title and -9.79% for a bestseller. In comparison, no intrabrand price coordination reduces the industry profits by an estimated -3.79% for a typical title and -22.50% for a bestseller. The analysis therefore suggests that though the Agreement offers commitment, the industry can achieve higher profits by using more flexible vertical restraints that coordinates the retail pricing without commitment.

The analysis uses concepts and frameworks from the mostly distinct literatures on vertical contracts and dynamic pricing. Coasian dynamics have recently been studied empirically in a variety of markets, such as college textbooks (Chevalier and Goolsbee (2009)), consumer electronics (Conlon (2012)), video games (Nair (2007)), fashion goods (Krishnamurthi and Soysal (2016)), and sports event tickets (Sweeting (2012)). There is a rich theory literature on the effects of vertical contracts and channel coordination across the fields of operations research, economics, and marketing, see Cachon (2003) for a survey, but there are few studies of channel coordination in markets for durable goods. One exception is Desai et al. (2004), which develops a theory for a dynamic, channel-coordinating two-part tariff in a two-period durable goods market with forward-looking consumers. I use a similar concept as the basis of one of the counterfactual vertical contracts in the analysis.

Though the data are comprehensive market level sales data, the analysis relies on strong assumptions on both the demand and the supply side. Firstly, the natural experiment is of a before-and-after kind. The legislation change did not offer an obvious control group for a difference-in-difference design. There are no instruments for prices in the demand estimation, the counterfactual analysis is limited to competition between retailers for the same title, and the analysis abstracts from competition between titles within a retailer. For these reasons, the analysis should be considered more of an inquiry into dynamic effects of a particular kind of vertical restraints at empirically plausible values, than an evaluation of the Agreement in the Norwegian book market. A discussion of the limitations is given in Section 10.1.

This paper relates to a literature of empirical evidence on vertical contracts and channel coordination, e.g. Besanko et al. (2005) on retail pass-through, Villas-Boas (2007) on identification of unobserved vertical contracts, Ho et al. (2012) on full line forcing in the video rental industry, Mortimer (2008) studies revenue sharing in the same industry, Hristakeva (2017) on vendor contracts, and Asker and Ljungquist (2010) on the impact of vertical integration in investment banking. None of these studies however explicitly considers dynamic effects of vertical restraints.

2 Natural experiment

The old Agreement dated back to the 1960s.[3] The Agreement was a voluntary and legally binding contract between the Association of Booksellers and the Association of Publishers which specified the terms of sales in the industry, including price restraints.

The Agreement regulated the lifecycle pricing of books in two ways. It let the publisher fix the price for its titles for the year of publication plus the following year, and the price restraint period was followed by an industry coordinated clearance sale. The fixed price was often printed onto the cover of the book. There was little price promotion in the industry, except for the yearly clearance sale. Each publisher could fix the retail price at any level, but once set, the price could not be revised in the restraint period.[4] The clearance sale was trade marked ('Mammut'), heavily advertised, and offered discounts on the order of 40%.

Secondly, the Agreement's clause 5 specified arbitration clauses in case of non-compliance.

> *Violation of the provisions of this Agreement may be prosecuted and, if necessary, by any of the two associations, any publisher and any bookstore or combinations of these who through their union are affiliated by the Trade Agreement. Each association further commits to, within the framework of the individual association bylaws, to take appropriate measures against its own members who may be guilty of violations of this Trade Agreement.*

Though the price restraint itself is a bilateral agreement between a publisher and a retailer on the retail price of a title, the arbitration clause exposes the vertical unit, i.e. a publisher and a retailer, to potential legal action by rival firms if the vertical unit was to deviate from the fixed price in the restraint period. Beyond allowing for legal actions of rival firms, the Agreement also allowed either Association to enforce further sanctions within the confines of each association's bylaws.

The arbitration clauses allow for external enforcement of an otherwise bilateral agreement that commits the publisher and the retailer to the fixed price. The threat of being taken to court, or made subject to other punishments, counteracts incentives to deviate from the fixed price in the restraint period. The data in Figure 3 in Section 4 shows that retailers for the most part complied with the restraints in the old Agreement.

As part of European legislative integration, the Norwegian Competition Act was aligned with its European Union counterpart in 2004. Following the legislation change, the Norwegian Competition Authority deemed the Agreement unlawful and called for

[3] See "Bokavtalen" in Store Norske Leksikon (Norwegian Encyclopedia) .

[4] The Trade Agreement has been suspected of facilitating horizontal collusion among publishers. The argument is that with RPM, it is easier for publishers to detect deviations on observable retail prices than in alternative vertical contract terms, such as unobservable wholesale prices. See Jullien and Rey (2007) for one treatment of this argument.

abolishment of the fixed price agreement in any form. As European integration is a political process that evolves independently of developments in the Norwegian book industry, the legislation change can be considered an exogenous change to the vertical restraints.

The industry voiced strong and united opposition against the new legislation. The Association of Booksellers, the Association of Publishers and the Association of Authors rallied together against the new legislation and called for exemption from the competition law.[5] The industrywide support of the Agreement suggests that it solves a price coordination problem that standard bilateral vertical contracts can not. Following a public debate, a political compromise was reached.[6]

The new Agreement was effective from May 1st 2005. There were two main changes

- A *shortening* of the price restraint period by eight months, from the year of publication plus one year to the year of publication plus four months.
- A *softening* of the fixed price to a price band. Whereas the price restraint under the old Agreement was both a floor and a ceiling, the new Agreement gave retailers discretion to discount the fixed price by up to 12.5%.

The changes are illustrated in Figure 1.[7]

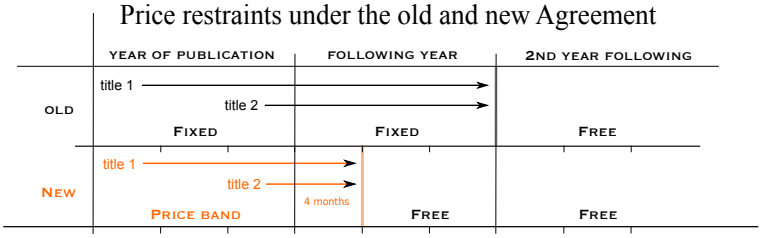

Figure 1: The price restraints followed calendar time. Some title a released early in the year and some other title b released later in the year would both have their price restraints lifted at the end of the following year under the old Agreement. Following the expiration of the price restraint period, the titles went to the clearance sale with discounts on average in the range of 40% to 50%. The end of the clearance sale marks the end of the lifecycle for the typical title. Under the new Agreement, the same titles a and b would again have their restraints lifted at the same calendar time, but now May 1st the year after publication rather than December 31st. The clearance sale continued to be held in spring, but was no longer part of the Agreement itself.

The softening of the price restrictions in the Agreement is interpreted as a natural experiment in vertical restraints. The experiment is of the before-and-after kind.

[5]Exemptions from the European competition law can be given for industries that make goods considered to be of particular importance to national identity and is widely allowed for cultural goods, see Canoy et al. (2005).

[6]To give some context of the media attention devoted to the new Agreement, a search on the keywords 'Book Trade Agreement' in Retriever, a comprehensive Scandinavian media archive, over the period of public debate gives about half the search hits that the contemporaneous 'Salt Lake City Olympics' generates in a comparable period. The numbers offer some perspective of the media interest that the new Agreement attracted in a nation where winter sports are very popular.

[7]The new Agreement also implied changes for *Bokklubben*, a mail-order retailer which was allowed to discount a limited set of titles by up to 25% under the old Agreement, but which was subject to the same terms as all other retailers under the new Agreement. As *Bokklubben* had a negligible share of the book market, I abstract away from *Bokklubben*'s role in the industry.

To interpret the shift in price strategies as caused by the change in the Agreements, the titles published under the old Agreement and the new Agreement are assumed to drawn from a stationary distribution which itself was not affected by the changes to the Agreement. For instance, there is a number of crime titles released each year, some are popular, others less so, but the assumption requires that there are no systematic changes over time to the quality or range of the released crime titles that affect the price incentives. Secondly, the preferences are assumed stationary as well. If that is true, we can interpret the difference in distributions under the new Agreement and the old Agreement as caused by the new Agreement.

Changes to market structure in the window of the data are possible confounders. The Norwegian book market had a stable market structure at the time with five main publishers and four main book retail chains. In the last tertile of 2007, two of the main publishers merged.

3 Data

Scanner data on sales were collected from the four largest book retail chains. The data covers the legislation change effective in May 2005. The data make up around 40% of total national sales over the period. The data are aggregated over four months (tertiles), and across stores within each chain. Observations are on title level identified by an Electronic Article Number (EAN). The data contain observations about 27000 titles. The sales are long tailed, with about 2% of titles making up about half of the total sales.

The data are merged on the EAN identifier with a comprehensive catalogue of title characteristics provided by Bokdatabasen, an industry logistics company. The catalogue contains data on the fixed price, genres, and various other characteristics such as page counts, edition etc and is used by retailers for logistical purposes and ordering. Prices are calculated as revenue divided by quantity sold in each period for each chain. Interviews with industry representatives indicate that pricing policies were mostly uniform across retailers within the chains. The summary statistics of the scanner data are given in Table 1.

	Old		New	
	Mean	Std dev	Mean	Std dev
Price	193.47	122.16	180.17	122.59
Fixed price	222.46	126.02	218.36	128.01
Quantity	32.77	166.07	41.42	227.64
Year of sale	2004.00	0.00	2006.05	0.82
Observations	67655		245804	

Table 1: Summary statistics. Data on title-tertile-retailer level. Old and New sampled at shares 0.36 and 0.47 of the population, respectively.

4 Impact of the new Agreement

Figure 2 shows the data on sales under the old and the new Agreement. The lifecycle of a title is taken to be about three years. After three years in the market, the sales of most titles are exhausted. By then, all price restraints are lifted, under both the old and the new Agreement. There are two full life cycles in the data. These are the titles published in 2004, and the titles published in 2005. Ideally, we would compare the price and demand paths for a typical title over the lifecycle before and after the legislation change. The data however cover only one year of sales prior to the legislation change.

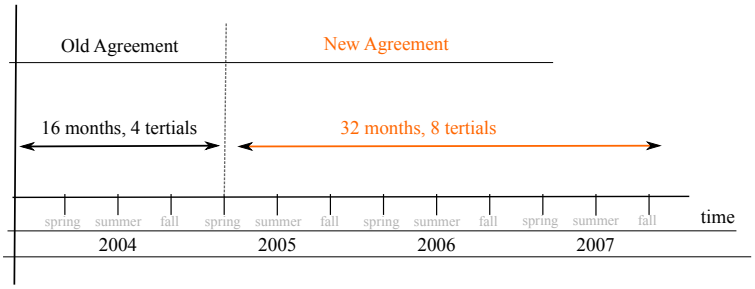

Figure 2: Agreements and data time line

To establish the lifecycle sales patterns under the old Agreement, the prices and sales patterns are assumed to be comparable across cohorts within a given year, where a cohort is taken to be all titled released in a given year. To compare sales across years, the titles each year are thought of as drawn from a stable distribution. Though the sales of individual titles can not be directly compared, for instance each year has its own bestsellers, the assumption allows for comparisons of moments of the sales distribution across years. A lifecycle price path is constructed under the old Agreement by calculating the average price path of new titles in 2004, splice it with the mean price path of titles that were one year old in 2004 and lastly, with titles that were two years old in 2004.[8] The resulting price path serves as a representative price path before the legislation change.

The tertiles are labelled spring, summer and fall. The average representative prices under the old and the new Agreement are graphed in Figure 3. The prices are normalized to the fixed price. A price of 1 implies that a title in that tertile on average retailed at the fixed price while a price of say 0.5 means a title retailed at 50% discount. The normalization allows comparison of price paths across different price points and controls for seasonal variation in the composition of demand. The prices are plotted against time and contains a total of nine periods. Confidence intervals of the means are interpolated between the data points to display the variance. Standard errors are reported at non-standard levels to display visible variation in dispersion across time. A plot of the nominal prices is given in Figure 12.

[8] Note that as titles are released over the year of publication, the set of titles in the sample is growing.

The retailers are seen to have largely respected the fixed price policy under the old Agreement: titles were retailing close to the fixed price in the price restraint period. Towards the end of the restraint period, there are some signs of retailers allowing discounts on the fixed price, on average about 5%. The deviations show that the price restraints were effectively a price floor and not a price ceiling.

The restraint period was followed by the industry coordinated clearance sale where titles were discounted by about 45% on average. The clearance sale marks the end of the lifecycle for most titles. The average prices are seen to increase some after the clearance sale. The bounce back may have a variety of causes. Firstly, there is a selection of particularly popular titles that still sell after the sale, and these titles may command higher prices.

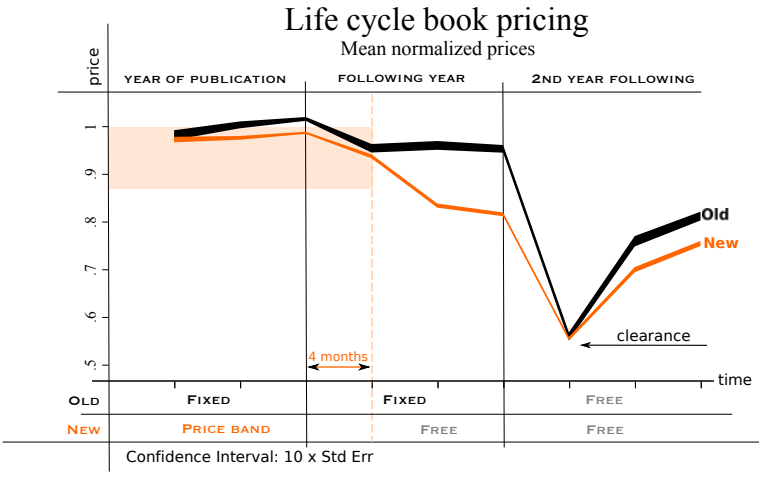

Figure 3: Mean of retail prices normalized to the fixed price.

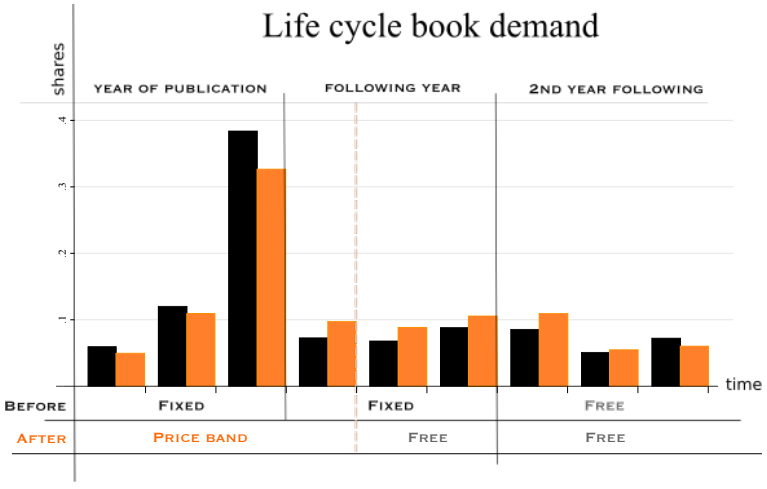

Figure 4: Lifecycle demand shares.

With the new Agreement, the retailers could discount the fixed price by up to 12.5% at discretion and the restraint period was shortened by eight months. The resulting price band is illustrated by the shaded area in Figure 3. Three changes stand out.

First, the retailers' discretion to discount the fixed price by up to 12.5% under the new Agreement went largely unused. The prices are comparable before and after in the introductory period showing that the restraints were largely non-binding in the early period under the old Agreement. Second, the changes in pricing are seen towards the end of the restraint period in the following year. Retailers started discounting prices in the summer of the following year under the new Agreement, when restraints were completely lifted. The change in the pricing shows that the old Agreement was mainly effective in keeping prices high towards the end of the restraint period. It is also evidence of dynamics in the pricing incentives over the course of the lifecycle. Third, prices fall to about the same level at the clearance sale. The industry is seemingly not serving new groups of lower valuation consumers under the new Agreement by dropping prices deeper than before. On the contrary, it seems that the same consumer groups are served, but at different prices over the course of the lifecycle.

The changes in the lifecycle pricing strategies are not well explained by the standard theories of RPM. For instance, if double marginalization had been a first order issue, the price levels in the introductory period would expectedly increase under the new Agreement, yet they largely stay put. If the main effect of the old Agreement was to coordinate price competition between retailers, the introductory period price levels would expectedly drop. The fact that the introductory price levels hardly change between the two Agreements suggests that competition between retailers is modest.

Figure 4 plots the corresponding shares of sales over the lifecycle. Note that the prices in the introductory period were comparable before and after the legislation change. The changes to prices came towards the end of the following year when the restraints were completely lifted. Demand is however seen to shift from comparable introductory prices, maintained through the early periods, towards earlier discounts under the new Agreement. The new Agreement presumably affects the incentives for timing the release of the titles. It may be that when the restraint period shortened under the new Agreement, then publishers prefer to release titles earlier in the year to increase the time a title is price restrained. Figure 5 however shows that there are relatively modest changes in the distribution of the release dates following the new Agreement, and not unequivocally to earlier in the year.

The assumptions of a stationary quality distribution and stationary preferences are strong. With a substantial shortening of the restraint period, the publishers could plausibly respond by changing their release data strategies. For instance, publishers might prefer to release titles earlier in the year be protected by the restraints for a longer period of time. Einav (2007) finds evidence of strategic timing of releases in the U.S. motion picture industry. The data carry information on release dates. Figure 5 shows there a change in the distribution of release dates towards releasing titles earlier in the year. On average, the change is relatively small, about half a month, and the shift is not unequivocally to the left. The muted supply side response may be related to the seasonality of demand. It could be more important to release a title in the market at peak demand for that genre than it is to lengthen the price restraint period.

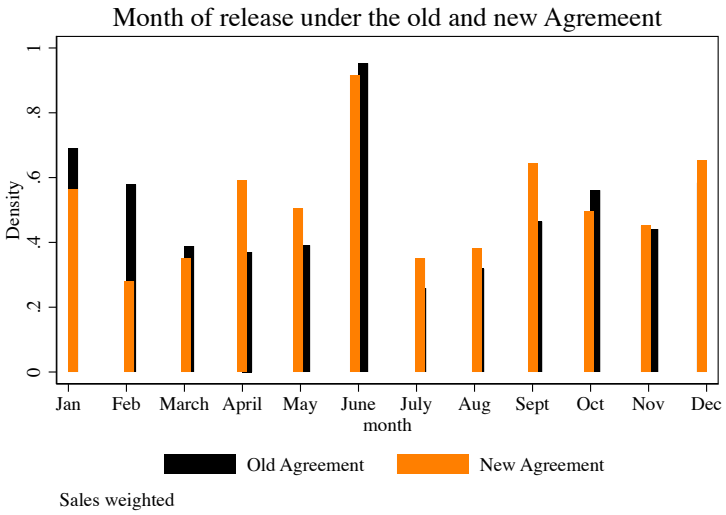

Figure 5

The evidence is consistent with a change in the consumers price expectations under the new Agreement. Expecting earlier discounts, consumers may be more willing to wait for a future discount at otherwise comparable prices. More demand is served at lower prices over the course of the lifecycle. These patterns in both price and demand patterns are qualitatively stable across years and across genres, see Figures 13 and 14 in the Appendix for the sales broken down over years. Whereas the pricing strategies are similar across titles, there is more variation in the demand responses across years.

5 Model setup

The goal of the empirical analysis is to quantify the commitment value of the old Agreement. While the old Agreement could provide commitment to future prices, it also precluded retail price competition. To separate the effect of commitment from the effect of restricting retailer price competition, I turn to counterfactual exercises.

The modeling approach follows two steps. In the first step, substitution patterns along the horizontal dimension and the time dimension are estimated from the sales data. Consumers with unit demand make purchase decisions considering the current states and prices and their expectations about future prices. These beliefs are recovered from the observed state transitions under the assumption of rational expectations. Beyond that, no assumptions are made on the pricing games in the demand estimation.

In the second step, a dynamic oligopoly model is used to evaluate the profitability of counterfactual vertical contracts using the demand parameters estimated in the first step. The contracts vary with the level of commitment and price coordination they offer. The oligopoly model is evaluated at the parameters estimated in the first step. The oligopoly model has a forward-looking demand side and a forward looking supply side. On the supply side, a publisher offers the retailers various vertical con-

tracts. Within the limits of contract terms, the retailers then set prices taking into account the impact of their prices on both the current and the future demand.

5.1 Market

Demand is represented by a finite horizon, discrete choice, adoption model with discrete unobserved type heterogeneity. The finite horizon model allows for non-stationary pricing strategies, such as the a time limited fixed price agreement.

Each title is assumed to be an independent market. The assumption restricts substitution to be between retailers and over time for a given title, which is the central focus of this paper, but assumes away substitution between titles. The independent market assumption is strong, but hard to relax. Firstly, with many titles and relatively few periods of observations, there is limited variation in the data to estimate patterns of substitution between titles. Secondly, substitution between titles leads to a high dimensional state space that complicates the computation of value functions. Though the inclusive value sufficiency assumption of Melnikov (2013) can collapse the state space to a manageable dimensionality, it implies a set of strong assumptions on substitution patterns and choice sets that seem implausible in this setting.

The market for each title consists of one publisher, J retailers, and a large, but finite number of M heterogenous and forward looking consumers. Each title is introduced prior to period 1. The periodization of the model follows the four month periodization in the data. A market lasts for T periods, where T is between 7 and 9, i.e. the maximum lifecycle is 3 years. After the T'th period, the market ends and no further transactions are made.

5.2 Consumers

At the start of the first period, there are M consumers in the market. Consumers, indexed by i, either purchase from one of the J retailers, or wait until the next period. The choice set is denoted $\mathcal{D} = \{0, 1, \ldots, J\}$, where $\{0\}$ is the choice to not purchase from any of the retailers. If a purchase is made, the consumer leaves the market never to return. A consumer that chooses to wait faces the same choice set in the next period.

The product characteristics that are observable to the econometrician are prices at all retailers $p_t = [p_{1,t}, \ldots, p_{J,t}] \in \mathbb{R}^J_+$ and product and market characteristics $x_{j,t} \in \mathcal{X} \subseteq \mathbb{R}^K$ for all $j \in \mathcal{D}\backslash\{0\}$. The product characteristics include deterministic functions of time, such as seasonality and tastes for novelty, and time invariant characteristics, such as retailer fixed effects.

One state variable is which type of fixed price agreement is in the market. By assumption, this state variable enters the information set, but not the utility, and is assumed to be commonly known in the market. The assumption implies that the fixed price agreement affects the consumer's expectations over the future prices of the title, but does not directly affect the value of a given title. A change in expectations over future prices caused by the change in the agreements is central to the identification of

forward-looking behaviour, which is discussed in Section 6.

All product characteristics, including prices, are assumed commonly observed in every period. A full description of the product characteristics is given in Section 9.2. Each consumer privately observes a vector of utility shocks $\epsilon_{i,t} \in \mathbb{R}^{J+1}$ in each period prior to making the choice. The private utility shocks are drawn from an absolutely continuous distribution, independently of of the observable states. There are two types of consumers, indexed by $l \in \{1, 2\}$ with probability mass m and $1 - m$, respectively. The current period utility of a consumer of type l choosing option j in period t is $u^l_{j,t}(p_{j,t}, x_{j,t}, \epsilon_{i,j,t}) = u^l_j(p_{j,t}, x_{j,t}) + \epsilon_{i,j,t}$, additively separable in the observable states and the private shocks.

A consumer who purchases a title at any of the J retailers collects utility $u^l_j(p_{j,t}, x_{j,t}) + \epsilon_{i,j,t}$ and leaves the market never to return. A consumer of type l that chooses to wait collects a current period utility normalized to zero, so

$$u^l_{j,t}(p_{j,t}, x_{j,t}, \epsilon_{i,j,t}) = \begin{cases} u^l_{j,t}(p_{j,t}, x_{j,t}) + \epsilon_{ijt} & \text{if } j \geq 1 \\ \epsilon_{i,0,t} & \text{if } j = 0. \end{cases}$$

The normalization of the utilities is standard, but restricts counterfactual behaviour in dynamic models. The implications of the normalization for counterfactual choice probabilities are discussed in Section 6 on identification.

The expected future utility streams are discounted by a constant factor β, which is assumed equal across consumers. In each period, a consumer makes a choice $d \in \mathcal{D}$ to maximize the expected lifetime utility.

$$V^l_t(p_t, x_t, \epsilon_{i,t}) = \max_{j \in \mathcal{D}} \{u^l_{j,t}(p_{j,t}, x_{j,t}) + \epsilon_{i,j,t} + \beta E\left[V^l_{t+1}(p_{t+1}, x_{t+1}, \epsilon_{i,t+1}) | I_t\right]\} \quad (1)$$

The *choice specific value function* $v^l_{j,t} : \mathcal{X} \to \mathbb{R}$ gives the expected discounted lifetime utility, prior to learning $\epsilon_{i,t}$, of making choice j in period t, and then choose optimally in the remaining periods.

$$v^l_{j,t}(p_t, x_t) = \begin{cases} u^l_{j,t}(p_{j,t}, x_{j,t}) & \text{if } j \geq 1 \\ \beta E\left[V^l_{t+1}(p_{t+1}, x_{t+1}, \epsilon_{t+1}) | I_t\right] & \text{if } j = 0 \end{cases}$$

The choice probabilities are

$$Pr[d^l_{i,t} = j | p_t, x_t] = Pr\left[\epsilon_{i,t} : v^l_{j,t}(p_t, x_t) + \epsilon_{i,j,t} \geq \max_{k \in \mathcal{D}} \{v^l_{k,t}(p_t, x_t) + \epsilon_{i,k,t}\}\right]$$

The *residual demand* $R^l_t \in [0, 1]$ for $l = 1, 2$, records the share of consumers of type 1 who are still in the market at time t. Assuming no sampling error, the type specific aggregate demand is

$$D^l_{j,t} = R^l_t Pr\left[d^l_{i,t}(p_t, x_t, \epsilon_{i,t}) = j | p_t, x_t\right]$$

The residual demand of type l is defined recursively

$$\begin{aligned} R^l_{t+1}(p_t, x_t, R^l_t) &= R^l_t Pr\left[d^l_{i,t}(p_t, s_t, \epsilon_{i,t}) = 0 | p_t, x_t\right] \\ &= D^l_{0,t}(p_t, x_t, R_t) \end{aligned} \tag{2}$$

with initial condition $R^1_1 = R^2_1 = 1$. The aggregate demand in period t is the sum of the demand of the two types is

$$D_{j,t}(p_t, x_t, R_t) = mD^1_t(p_t, x_t, R^1_t) + (1-m)D^2_t(p_t, x_t, R^2_t).$$

In the demand estimation, consumers are assumed to form rational expectations over the state transitions given their information sets I_t. The information set includes all variables that affect the consumers utility and expectations in period t, including which Agreement is in place. Rational expectations imply that the consumers' expectations coincide with the observed distribution of state transitions. The counterfactuals uses the same demand model, but there the expectations rational expectations will be determined in equilibrium rather than estimated from the data, see Section 9.2 for further discussion.

6 Identification

It is well known that the standard dynamic discrete choice models of Rust (1994) is non-parametrically underidentified. Magnac and Thesmar (2002) shows that without further assumptions, a static model can rationalize all choice data as well as any dynamic model. Conditional on the discount factor, Magnac and Thesmar's Proposition 2 shows that a normalized utility function is identified. It is therefore common to fix the discount factor at an a priori plausible value. Since the consumers' discount factor determines the profitability of the fixed price agreement in the model, assuming the discount factor assumes the conclusion, which seems unsatisfactory.

Common intuition suggests that if there is variation in the data that holds the current payoffs fixed, but changes the continuation values, then the current period choice response to the change in continuation values is informative about time preferences. I first argue that the clear shift in the price paths that followed with the new Agreement may have changed consumers expectations of future prices and consequently their expected value of waiting for future discounts.

The introductory prices were approximately equal under the old and the new Agreement. The clear shift in demand in the year of publication is consistent with a change in the consumers' beliefs about the future prices. Consumers who at the time of publication expected lower future prices under the new Agreement, may be more willing to postpone their consumption towards future discounts.

We can think of the transition from the old and the new Agreement as shifting price expectations, and hence the continuation value, without changing the consumption value of a purchase itself. Abbring and Daljord shows that this intuition can be for-

malized as exclusion restrictions that set identifies the discount factor. Specifically, if there exists a pair of states x_{old} and x_{new} such that the exclusion restriction

$$u_{i,t}(x_1) = u_{j,t'}(x_2) \tag{3}$$

for some pair $i \in \mathcal{D}, j \in \mathcal{D}\backslash\{0\}$, $x_1, x_2 \in \mathcal{X}$ and $t \in 1, \ldots, t', t' \in 1, \ldots, T$, with either $i \neq j$, or $x_1 \neq x_2$, or $t \neq t'$, then the identified set is finite. The intersection of the sets identified from each moment condition is assumed to be a point.

The assumption that the new Agreement in May 2005 came unexpected to consumers is more questionable. There was a lively debate in the media over the future of the old Agreement in the fall of 2004. To the extent consumers factored in a likely change to a new pricing regime with lower prices, it will bias the discount factor towards zero by attenuating the contrast between demand under the old and the new Agreement.

Following Abbring and Daljord, the auxiliary moment conditions derived under the exclusion restrictions in (3) are

$$\ln\left(\frac{s_{j,t}^{old,l}(p_t, x_t)}{s_{0,t}^{old,l}(p_t, x_t)}\right) - \ln\left(\frac{s_{j,t}^{new,l}(p_t, x_t)}{s_{0,t}^{new,l}(p_t, x_t)}\right) = \sum_{\tau=t+1}^{T} \beta^{\tau-t}\left(E[\ln(s_{0,\tau}^{new,l}(p_\tau, x_\tau))|I_t^{new}] - E[\ln(s_{0,\tau}^{old,l}(p_\tau, x_\tau))|I_t^{old}]\right)$$

for $t = 1, \ldots, 3$ of the lifecycle in 2004 (old) and 2005 (new), for otherwise equal states x_t and p_t, and for both consumer types, and where j is taken to be a purchase of an inside good. Since there are no complete title level time series under the old Agreement, the moment conditions are constructed using the aggregate data in Section 4. From Magnac and Thesmar Proposition 2, a non-stationary utility function is identified conditional on the discount factor being known. This implies that a taste for novelty is separately, but not independently, identified from the discount factor.

The expectations in the single agent demand model are recovered as the observed distribution of state transitions under the assumption of rational expectations, following Rust (1994). The demand is estimated independently of the pricing games in the counterfactuals. The counterfactual analysis in Section 9 will explicitly model the vertical contracts as pricing strategies that follow as solutions to particular dynamic games. These counterfactual pricing strategies will depend on the consumers' expectations. Unlike in the demand estimation, the expectations in the counterfactual analysis are derived as equilibrium objects using Stokey (1981)'s concept of a rational expectations equilibrium, see further discussion in Section 9. Computing counterfactuals where rational expectations are endogenous and determined jointly with prices and demand in equilibrium, given utility parameters recovered from a single agent demand model approach, is standard in the literature, e.g. Nair (2007) and Lee (2013).

The price expectations are estimated separately under the old and the new Agreement. Under the old Agreement, there is little variation in prices, so price expectations are

set equal to the fixed price in the restraint period and the average sales price at the expiration of the restraint period. For the title level data under the new Agreement, the price expectations are estimated from the observed price transitions under the new Agreement. We can think of the discount factor as informed by data on changes in demand between the old and the new Agreement holding price fixed, while the utility function is estimated from the lifecycle demand under the new Agreement. The auxiliary moment conditions are added to the criterion function.[9]

Kasahara and Shimotsu (2009) shows that two latent consumer types are non-parametrically identified for $T \geq 3$ with type independent transitions and some mild regularity conditions on the covariates which are plausibly met in the data. Market prices are commonly thought of as being determined simultaneously with demand in observational data. The most common solution to simultaneity of prices and demand is to find instruments that are correlated with the retailers price incentives, but do not directly cause demand itself. Ideally, I would have instruments with variation across titles, across retailers, and over time. That is a tall order. Standard sources of instruments include marginal cost shifters (Working (1927)), variation in the density of the product space (Berry et al. (1995)), and common components in geographical price variation (Hausman (1996), Nevo (2001)).

These instruments are unfortunately either weak, questionable, or unavailable in this application. For books, marginal costs are likely close to constant over the lifecycle of a title, non-price product characteristics are mostly time-invariant, and there is no geographical variation in the data. As pointed out in Rossi (2014), instruments may under these circumstances cause more problems than they solve. The prices are therefore left to instrument for themselves and are effectively assumed exogenous. This is a strong assumption that is likely violated.

There may still be an important component of exogenous variation in the prices. Firstly, the fixed prices are set by the publisher prior to publication. Predicting the sales potential of a title prior to release is notoriously hard, see e.g. Caves (2003). In Section B, a regression of total lifecycle sales of titles on the fixed price and other characteristics known at the time of introduction shows no correlation between the fixed price and the sales, which is consistent with the fixed prices being close to randomly assigned. Secondly, from interviews with management at the retail chains at the time of the data collection, individual titles were predominantly priced using heuristics, such as preset discount schedules for groups of titles. Randomly assigned fixed prices and preset discount schedules does not verify the assumption of exogenous prices, but implies the assumption may be close to hold for many titles.

Discounts for some titles were however offered in conjunction with advertising, which clearly violates exogeneity. Unfortunately, I do not know how many titles that affects.

[9] Abbring and Daljord shows that in a single agent model without latent types, the discount factor can be estimated independently of the utility function from the auxiliary moment condition. With latent types, as in the current model, the discount factor however must be estimated jointly with the full parameter vector.

To alleviate a possible downward bias in the price sensitivity estimates in the absence of instruments, the price elasticities are restricted to be larger than one in absolute value. The restriction imposes a weak form of profit maximization in the demand estimation, which for most observations is not binding .

The choice specific value functions do not contain a market level, product specific shock, which is often denoted $\xi_{j,t}$. Such shocks are often used to implement instruments in discrete choice models using a procedure like in Berry et al. (1995). The ξs allow the the choice specific value functions to be inverted from the choice data by exactly equating the predicted market shares to the observed market shares and standard IV methods can then be applied to the recovered choice specific value functions. Since in this application, prices are assumed exogenous, market level shocks are not required to implement instruments. Market level shocks, one for each retailer, however increases the state space by J dimensions. In both demand estimation and in the computation of the dynamic games, these additional states complicates computations, but without offering material economic insight into the research question. Therefore, as these market level shocks serve little economic purpose in this application, these shocks were dropped from to keep the model as simple as possible, while capturing the main economic concepts.[10]

From Magnac and Thesmar's Proposition 2, the utility function is non-parametrically identified only up to a normalization of an arbitrary reference choice. Kalouptsidi et al. (2016) shows that a large class of counterfactuals are not identified in dynamic discrete choice models when the utility of the reference choice is normalized, but not identified, i.e. the standard case. The counterfactuals I consider changes the transitions and belong to the non-parametrically underidentified class.

Kalouptsidi et al. considers identification of counterfactuals under parametric assumptions on the utility function. Its Corollary 10 shows that when the counterfactual changes the transition process for state variables that are part of the identified component of the utility function, counterfactual choice probabilities are identified. In our case, the counterfactuals change the price transition and the price effects are parametrically identified. The counterfactuals are hence identified if we are willing to assume that the parametric specification in the demand model is the true model. Hence, the counterfactuals are identified under stronger assumptions than the discount factor and the utility function.

7 Specification and data selection

The data for the demand estimation are selected as all new fiction and non-fiction titles with strictly positive sales at all retailers for consecutive periods. Gandhi et al. (2017) shows that dropping products with observed zero market shares introduces a bias and proposes a bias-corrected estimator. The estimator however assumes that

[10] A market level shock would make the dynamic games stochastic. It would also create a trade-off between commitment and flexibility to set prices in response to demand shocks. Though this trade-off is of economic interest, it is not ideally studied with an *iid* process for the same title.

there are no true zero shares. That assumption is violated in the data. Many of the observed zero shares are specialty titles that are not sold by all retailers. If I had inventory data across retailers, I could distinguish between true and false zeros. Then the choice sets could be adapted accordingly which would also introduce useful variation. Unfortunately, I have no inventory data. It is beyond the scope of this paper to adapt Gandhi et al.'s estimator to cases with unidentified true zeros.

There is substantial heterogeneity in sales between titles, both in volume and lifecycle demand distribution, which suggests using fixed effects at the title level. Title level fixed effects however come with two problems. Firstly, fixed effects imply thousands of free parameters in this application, which is computationally prohibitive. Secondly, with at most nine periods observed per title, fixed effects introduce incidental parameter bias.

It is common to discretize heterogeneity as a finite sample approximation to unrestricted heterogeneity. Such grouped fixed effects generally suffer from approximation bias. I use the two-step estimator Bonhomme et al. (2018) which trades off the incidental parameters bias of title level fixed effects against approximation bias of group fixed effects using a data driven, two-step, bias reduction approach. Bonhomme et al. shows that the estimator has desirable asymptotic properties when the underlying dimensionality of the heterogeneity is low, which is the case in this applications.

The estimator classifies the titles into groups using a k-means classifier in a first step and uses a bias reducing estimation procedure in a second step. The first step classifier is based on title specific moments of the data, h_i, that are informative about the heterogeneity. Let $\varphi(\alpha_{i,0})$ be the population value of h at the true individual specific vector $\alpha_{i,0}$, where φ is an unknown, injective function.

Given a number of K groups, the groups are determined by minimizing a mean square criterion $\hat{Q}(K)$ as

$$\hat{Q}(K) = \min_{h^K, \{k_i^K\}} \frac{1}{N} \sum_{i=1}^{N} (h_i - h^K(k_i^K))^2$$

where k_i^K is a group membership indicator and h^K is the group moment.

The goal of the classifier is to group titles that respond similarly to price changes. One challenge is that, even within genres, observationally equivalent titles display large variation in sales. In this application, the underlying heterogeneity the estimator approximates is a title specific, time invariant fixed effect and a time trend. The underlying heterogeneity is therefore two-dimensional by assumption. The fixed effect represents how popular a title is, perhaps its inherent quality, and the time trend a taste for novelty.

The key property of the moments h that are used in the classification is that they are uniquely informative about individual heterogeneity. As long as the population

moments uniquely characterize the units, which here are titles, they can be anything, including functions of endogenous objects such as sales. The moments I use are functions of shares of sales of the individual titles, of time, and of the fixed price

$$h_i = \frac{1}{T}\sum_{t=1}^{T} s_{i,j,t} - \gamma_t t - \gamma_{fp} fp_i + \upsilon_{i,j,t}$$

where t is a time trend, fp is the fixed price, and $\upsilon_{i,j,t}$ is a residual. The fixed price carries information about both the genre of the title the publisher's ex ante expectations of its total sales. As the demand for titles are exhausted at different rates, some faster than others possibly due to novelty, I include a time trend. I use a simple linear regression to fit the sample moments.

The number of groups K is chosen using the classifier

$$\hat{K} = \min_{K \in \{1,\dots,N\}} K \text{ s.t. } \hat{Q}(K) \leq \xi \frac{\hat{\bar{V}}_h}{T}$$

where $\hat{\bar{V}}_h$ is a consistent estimator of the asymptotic variance

$$\text{plim}_{N,T\to\infty} \frac{T}{N}\sum_{i=1}^{N}(h_i - \varphi(\alpha_{i,0}))^2$$

Here, ξ is a tuning parameter that controls how much signal the classifier extracts about the types from the moments. A lower value extracts more signal about the types from the moments, which reduces approximation bias at the cost of increasing variance and increasing incidental parameter bias. There are two considerations that guides the choice of ξ. The first is the dimensionality of the underlying heterogeneity. When the underlying dimensionality is low, which is the case in this application, fewer types are needed to achieve a good approximation to the underlying continuous heterogeneity. The second is how informative the moments h_i are about the types.

Since the moments I chose are rather informative, and there is a penalty for too many types, I decided on $\xi = 3$. The choice may be on the conservative side in terms of reducing bias from approximating the unobserved heterogeneity. The choice is somewhat arbitrary, but so is a choice like $\xi = 1$ used in Bonhomme et al. (2018). There is no obvious data driven method without a ground truth which is not available. As long as ξ is fixed as the number of observations (hypothetically) grows, the asymptotic properties are however guaranteed for either value it takes.

The observable states are $\{p_{i,j,t}, r_j, t, ssn_t\}$, where i indexes the title, j the retailer, k the group, and t indexes time. The utility specification is

$$u^l_{i,j,k,t} = \alpha^l + \alpha^l_p p_{j,k,t} + \alpha^l_{t,k} t + \gamma^k + \gamma^k_n t + \gamma_j r_j + ssn_t \gamma_{ssn} + \epsilon_{i,j,t}. \tag{4}$$

The parameter γ^k is a group fixed effect and r_j is a retailer fixed effect. Following Einav (2007) and Ho et al. (2012), the utility has a component $\gamma^k_n t$ that represents a group specific taste for novelty, where t is a linear time trend. Retailer fixed effects

are $\gamma_j r_j$, and seasonal fixed effects are $\gamma_{ssn} ssn_t$. The type specific parameters are superscripted by l. The utility parameters α, γ and β are estimated with full solution methods using a minimum distance criterion.

The bias reduction step in Bonhomme et al. is implemented by first estimating the parameters $\hat{\theta}$ with group fixed effects on the full sample. Then $\hat{\theta}_1, \hat{\theta}_2$ are estimated on randomly selected halves of the sample, respectively, where the groups are held fixed in each sub-sample. The bias-reduced estimate is $\hat{\theta}^{BR} = 2\hat{\theta} - \frac{\hat{\theta}_1 + \hat{\theta}_2}{2}$. Standard errors are calculated by block bootstrapping, where the classifications are again held fixed in the bootstrap samples.

Consumers are assumed to have rational expectations such that their expectations coincide with the true distribution of transitions. In the demand estimation, the consumer's information set is $I_t = \{p_t, x_t, \epsilon_t\}$. All observable, non-price characteristics x_t are deterministic with trivial transition distributions. The relevant transition probability distribution can therefore be written $F_t(p_{t+1}, \epsilon_{t+1}|I_t) = F_t(p_{t+1}|p_t, x_t)G(\epsilon)$, where G is identified *iid* EV1 by assumption. The conditional distribution $F_t(p_{t+1}|p_t, x_t)$ is estimated from the observed state transitions by FGLS first step. The expectations are estimated separately under the old and the new Agreement as described in Appendix D. The model fits fairly well with $R^2 = 0.97$ for the expectations under the new Agreement, which implies that the state transitions are not far from deterministic. Following the arguments in Skrainka and Judd (2011), I use the SparseGrid quadrature package of Heiss and Winschel (2008) to numerically integrate out the expectations in the demand functions.

The demand estimation imposes no equilibrium constraints from the supply side. In the counterfactuals in Section 9, the consumers expectations will be determined in equilibrium, rather than estimated from the data.

8 Estimation results

The parameter estimates are given in Table (2). The consumers discount factor β comes out at 0.780, substantially less than the real interest rate which is used as the retailers discount factor. The estimated discount factor is lower than implied by the real interest rate, but higher than e.g. Dubé et al. (2014). There is some variation in retailer fixed effects, reflecting that the retailers are differentiated. While fall is the prime season where most titles are introduced, the fall fixed effect is negative. Though aggregate sales is higher in fall, the average title sells fewer copies than in spring and summer, where the sales is concentrated on fewer titles.

Figure 6 plots the own price elasticities, averaged over retailers and titles, against time.

Table 2: Parameter estimates

Utility parameters

Parameter	**Coeff**	**Std Err**
β	0.780	0.025
summer	−0.166	0.129
fall	−0.300	0.138
retailer B	−1.039	0.254
retailer C	−0.110	0.093
retailer D	−0.826	0.103
m	0.215	0.178
Retailer A is the base level	Group fixed effects: yes	Nr of groups: 5

Type specific parameters

	Type 1		Type 2	
Parameter	**Coeff**	**Std Err**	**Coeff**	**Std Err**
α	1.333	3.694	−0.132	2.684
α_p	−1.112	0.913	−3.780	1.448
N	40388			

9 The commitment value of the old Agreement

The goal of the counterfactual analysis is to quantify the commitment value of the fixed price agreement. I take the commitment value to be the profits an agreement generates to the industry beyond what the industry can achieve with standard vertical contracts, without commitment. While the terms of the old Agreement are publicly known, it is not known which vertical contracts were used under the new Agreement. Therefore, a direct comparison of the industry performance under the old and the new Agreement is hard to interpret. Instead, I evaluate the performance of four counterfactual vertical contracts at primitives estimated in Section 5.1.

The supply side has one publisher that supplies J retailers, where $J = 4$. The terms of sale, including retail price restraints, are regulated by contracts between the publisher and the retailers. The retailers individually have less market power than the publisher. The publisher can extend its market power to the retail market using vertical restraints that prevent retailers from competing away the industry profits. Intertemporal price discrimination also requires market power. The publisher's market power also allows it coordinate prices over time to an extent that individual retailers can not. However, without commitment, even a monopolist publisher faces competition from expectations of its own future prices when consumers are forward-looking.

A fixed price agreement can however commit an integrated vertical unit to a price strategy it may not credibly sustain on its own. The FPA contract fixes the retail price for a limited period. Since the commitment by assumption is credible, the FPA contract also fixes consumers' expectations. We can ask how much the commitment of the FPA contract is worth relative to a vertical contract that lets the publisher optimally revise the price restraint in every period, a contract we call HC. The HC price restraints are set to coordinate the intrabrand competition, i.e. horizontally, but since it is revised in every period, it has no commitment.

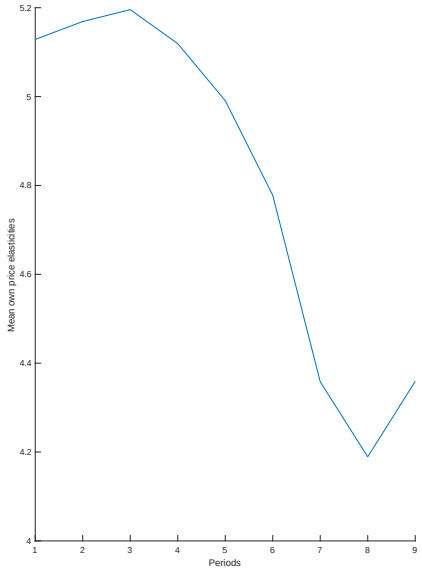

Figure 6: Own price elasticities

We may ask whether commitment is more important than integrating the vertical unit, without commitment. To construct a measure, I compute a standard linear contract NC with a time invariant wholesale price that offers no price coordination. The NC contract is suboptimal: the publisher does not coordinate the competition between neither retailers nor over time. This contract serves as a lower bound to the returns vertical integration. The measure of the returns to integrating the vertical unit, without commitment, is then the difference in profitability between the HC contract and the NC contract.

Though a fixed price agreement offers commitment by locking in the retail price for a given number of periods, it may not commit the industry to the optimal commitment strategy. I therefore compare the FPA contract to a contract that holds with commitment for all periods, FC, and which allows the price restraint to change over the lifecycle. The FC contract serves as an upper bound to the returns to both commitment and intrabrand price coordination. The four contracts are defined formally below.

The contracts FC, HC, and NC can be ranked in terms of profitability a priori. Since the publisher under FC can commit to any price strategy, and there is no uncertainty in the model in equilibrium, FC it is at least profitable as HC, and if there is any value to commitment, we get $FC > HC$. Since the seller can not make less profits by coordinating the retailers' pricing, $HC \geq NC$, which holds with strict inequality if there is any substitution between the retailers. The ranking of FPA can not be determined a priori, but is an empirical question. On the one hand, FPA offers commitment for some periods, which can profitably curb intertemporal substitution,

so FPA may therefore be more profitable then HC. The price restraint in the FPA is however constant over the restraint period. As discussed in Section , if $\beta < \rho$, the optimal commitment path allows some intertemporal price discrimination. The FPA also offers no coordination of prices after the restraint period expired. If β is sufficiently small, the flexibility of the HC restraints allow for intertemporal price discrimination which outweigh the commitment value of the FPA, and we get $HC > FPA$.

The outcomes of interest are the lifecycle price paths and profitability across the contract types. The counterfactual contracts are evaluated in a dynamic equilibrium model with a forward-looking supply side and a forward-looking demand side, where rational price expectations are formed endogenously on both sides of the market. The demand side is represented by the estimated demand from the previous section, with the important difference that price expectations are now determined within the model, while the expectations were treated as primitives in the estimation. The outcomes of interest are the lifecycle price paths and profitability across the contract types. The equilibria are calculated for two representative titles, a typical title and for a bestseller.

Whereas the vertical contracts employed in the industry emerge from a complex and unobserved bargaining process between retailers and publishers, I make several simplifying assumptions to abstract from the bargaining process. Following Matthewson and Winter (1984), the HC contract is assumed to include a sufficient number of instruments, such as quantity independent transfers, to support the price strategy that maximizes the profits to the vertical unit.[11] The publisher and the retailers are assumed to agree on the contract terms, which include price restraints and possibly transfers, that maximize the vertical units profits. Both the bargaining over the contract terms and the non-price terms themselves are however left unmodeled. No assumption is made on the bargaining, e.g. there is no assumption of Nash bargaining, except that the publisher and retailer can find some division of profits that support the strategies. A contract will therefore for our purposes be summarized by a set of retail price restraints which affect the retailers price strategies. The commitment in the FC and FPA contracts is by assumption, and not explicitly modeled as e.g. incentive compatible.

9.1 Publisher and retailers

The equilibrium price restraints and retailer pricing strategies are solutions to dynamic games. The publisher chooses price restraints p_t^F that are uniform across retailers. The price restraints dictate the retail prices, i.e. they are both ceilings and floors.[12]

Since all the observable, exogenous states x_t are deterministic, the dependence of $D_{j,t}$ on x_t is suppressed in the following. Retailer j's per-period profits depend on both its own prices $p_{j,t}$ and its rivals prices $p_{-j,t}$. A retailer's per-period profit function

[11] One example of such transfers is known as 'marketing support', where a publisher pays the retailer a fixed fee for what may or may not be related to marketing activities.

[12] Though the optimal price restraints vary across retailers, I focus on price restraints that are uniform across retailers as these are empirically relevant.

is

$$\pi_{j,t}(p_{j,t}, p_{-j,t}, w, R_t) = (p_{j,t} - w)D_{j,t}(p_t, R_t)$$

where w is the per unit wholesale price. The expected present value of profits to retailer j in period t is

$$\mathbb{E}_t \left[\sum_{\tau=t}^{T} \rho^{\tau-1} \pi_{j,t}(p_{j\tau}, p_{-j,\tau}, R_\tau) | p_t, R_t \right].$$

where ρ is the discount factor of the firms, which is assumed equal across retailers.

A *price strategy profile* is a set of of retail price strategies $\sigma_t = [\sigma_{1,t}, \ldots, \sigma_{J,t}]$. A *sequence* of strategy profiles is denoted $\boldsymbol{\sigma}_\tau = \{\sigma_{1,t}, \ldots, \sigma_{J,t}\}_{t=\tau}^{T}$, for $\tau = 1, \ldots, T$. Both retailers and the publisher form rational expectations $F_t(R_{t+1}|\sigma_t, R_t)$ over the evolution of the pay-off relevant states, which is the vector of residual demands R_t.

9.2 Consumer demand and equilibrium expectations

The consumer demand $D_t(p_t, R_t)$ uses the utility function and the discount factor estimated in Section . While the consumers expectations are estimated from the observed state transitions under the assumption of rational expectations in that section, these estimated expectations are in general not rational in the counterfactuals. Instead, both the sellers and the consumers are in the counterfactuals formed in equilibrium. Unlike in the demand estimation, which makes no assumptions about the particular game being played in the data, the expectations in the counterfactuals depend crucially on assumptions of the particular game being played.

For all contracts, I use Stokey (1981)'s concept of a rational expectations equilibrium (REE). An REE consists of two sets of functions. One is a pair of functions describing how the consumers' expectations are formed, one for each consumer type, and one is a function describing the seller's sales strategy.[13] These functions jointly satisfy the following conditions

1. The seller's price strategy maximizes the net present value of profits, given the expectation function of consumers
2. The consumers' expectations of the future prices are fulfilled along the realized pricing path

We distinguish here between the expectation function, which returns a consumers's beliefs given the observed states, and the expectations, which are the values the expectation function take along the equilibrium path. The equilibria are constructed by considering what would happen off the equilibrium path. If consumers faced prices off the equilibrium path, their beliefs change, their demand change, and subsequently the seller's payoff changes. Once the seller's payoff changes, the first criterion above may not be met, even if the second is met. Therefore, though the games move deterministically along the equilibrium path since there is no stochasticity in the game, the

[13] For the *NC* contract, there are multiple sellers. The concept generalizes straightforwardly to multiple sellers by allowing the supply side one function describing each retailer's price strategy.

rational expectations equilibrium concept is a non-trivial restriction.

The primitives that enter the counterfactuals are the utilities $\boldsymbol{u}$, the consumers' discount factor β and the firms' discount factor ρ, and the cost structure, summarized by the constant marginal cost c. The marginal cost c is taken to be known and set to 10% of the introductory retail price, in line with industry estimates. While β is estimated from the demand data, ρ is set equal to $\frac{1}{1+r}$, where the interest rate $r = 2\%$, approximately the risk free interest at the time. The endogenous objects are the price strategies $\boldsymbol{\sigma}$, demand $D_t(p_t, R_t)$, and price expectations $F(R_{t+1}|\sigma_t, R_t, t)$. Note that unlike in the demand estimation, the demand now explicitly conditions on the residual demand as a carrier of information on the seller's pricing incentives. This implies stronger assumptions on how expectations are formed than in the demand estimation, where expectations are estimated from the observed transitions, which are required to determine the expectations jointly with prices and demand in the dynamic games.

9.3 No Coordination

The publisher does not coordinate prices in the NC contract, i.e. there are no price restraints. In each period t, the retailers set prices simultaneously after observing R_t, the state of demand left in the market. The wholesale price w is set equal to the marginal cost c.[14] The timing is illustrated in Figure 7. The NC contract is not a plausible contract. It is instead a counterfactual contract that serves as a benchmark lower bound to the returns to integrating the vertical unit.

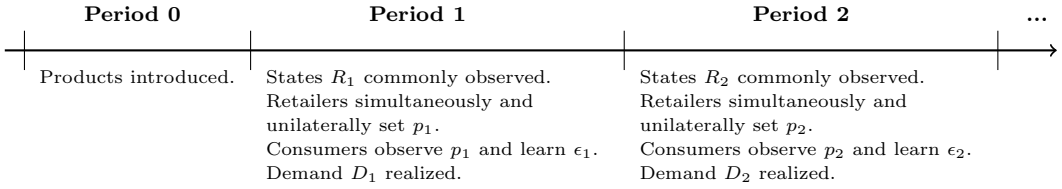

Figure 7: Timing No Coordination

The price strategies σ_t are Markovian, i.e. they map the current, pay-off relevant states to prices. A sequence of strategies for retailer j is written $\boldsymbol{\sigma}_{j,\tau} = \{\sigma_{j,\tau}\}_{\tau=t}^{T}$. The value function $\Pi_{j,t}$ gives each retailer's net present value of profits as of time t, conditional on the strategy profile $\boldsymbol{\sigma}_t = [\boldsymbol{\sigma}_{1,t}, \ldots, \boldsymbol{\sigma}_{J,t}]$, and is given as

$$\begin{aligned}\Pi_{j,t}(R_t; \boldsymbol{\sigma}_{j,t}, \boldsymbol{\sigma}_{-j,t}) = \pi_{j,t}(\sigma_{j,t}, \sigma_{-j,t}, R_t) + \\ \rho \int \Pi_{t+1}(R_{t+1}; \boldsymbol{\sigma}_{j,t+1}, \boldsymbol{\sigma}_{-jt+1}) \times \\ dF(R_{t+1}|\sigma_{j,t}, \sigma_{-j,t}, R_t, t) \\ \text{for all } \; j \in 1, \ldots, J \text{ and } \; t \in 1, \ldots, T.\end{aligned}$$

[14] There are many ways a contract can be suboptimal. One alternative suboptimal contract has the publisher charge a monopolist wholesale price in every period. This contract would introduce (dynamic) double marginalization. Though double marginalization is interesting in itself, there is no evidence of double marginalization in the data.

The solution concept is Markov Perfect equilibrium. A solution is a sequence of strategy profiles $\boldsymbol{\sigma}_1^{NC} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_J)$ such that for all retailers, in every period, the equilibrium condition

$$\Pi_{j,t}(R_t; \boldsymbol{\sigma}_{j,t}^{NC}, \boldsymbol{\sigma}_{-j,t}^{NC}) \geq \Pi_{j,t}(R_t; \boldsymbol{\sigma}_{j,t}', \boldsymbol{\sigma}_{-j,t}^{NC}) \tag{5}$$

holds. The condition implies that each retailer j prefers $\boldsymbol{\sigma}_{j,t}^{NC}$ to any alternative sequence of Markov strategies $\boldsymbol{\sigma}_{j,t}'$, conditional on its rivals pricing according to $\boldsymbol{\sigma}_{-j,t}^{NC}$. The resulting price strategies are hence time-consistent for each retailer.

An NC equilibrium is a fixed point where prices are set according to $\boldsymbol{\sigma}_1^{NC}$ in every period, and the price expectations of retailers and consumers are consistent with both the pricing and the demand. Note that there is no uncertainty about prices in equilibrium.

9.4 Horizontal Coordination

Under the HC contract, the publisher coordinates the intrabrand retail pricing. in each period, the publisher observes the residual demand R_t before it sets a price restraint p_t^F which is uniform across retailers. The restraint is revised in every period to optimize the expected present value of joint profits to the vertical unit. The timing is illustrated in Figure 8.

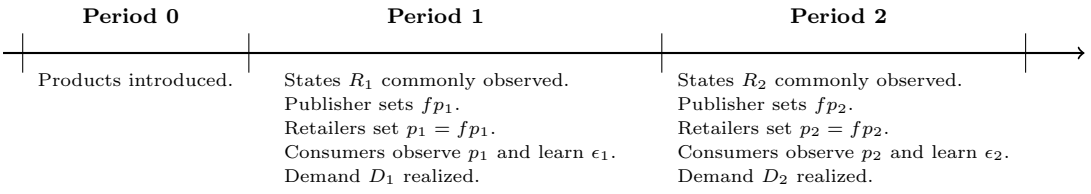

Figure 8: Timing Horizontal Coordination

The equilibrium vertical restraints $\boldsymbol{\sigma}_t$ are the price strategies that maximize the net present value of the joint profits to the vertical unit, taking into account that in the next period, the restraints will again be set with the same objective, i.e. the strategies are time consistent. The relevant marginal cost for the vertical unit is the marginal cost of production c.[15] The HC is a fixed point such that

$$\begin{aligned}\sigma_t^{HC} = &\arg\max_{p \in \mathbb{R}_+} \sum_{j=1}^{J} \pi_{j,t}(p, p, R_t, c) + \\ &\rho \sum_{j=1}^{J} \int \Pi_{j,t+1}(R_{t+1}; \boldsymbol{\sigma}_{t+1}^{HC}) dF(R_{t+1}|R_t, t)\end{aligned}$$

holds for all $t \in 1, \ldots, T$. The HC need not be implemented by a price restraint, but could be implemented by a dynamic two-part tariff where the publisher in each period

[15] The marginal cost of retail is assumed equal to zero.

sets a wholesale price $w_{j,\tau}$ such that no retailer unilaterally has an incentive to deviate from $\boldsymbol{\sigma}_{\tau}^{HC}$. These implementations are equivalent under the assumptions.

9.5 FPA

In the first period, the publisher sets the retail price p^F which is uniform across retailers and constant within the restraint period. The restraint period is set to four periods, corresponding to the modal title in the data being released in fall. The fixed price p^F is set to maximize the expected profits of the vertical unit given the information known in that periods, taking into account that following the restraint period, the retailers will unilaterally set prices in every period, without commitment, i.e. the retailers use the price strategies σ_t^{NC} for $t > 4$. In the non-restraint periods, the retailers price incentives depend on the wholesale price w.[16] The timing is illustrated in Figure 9.

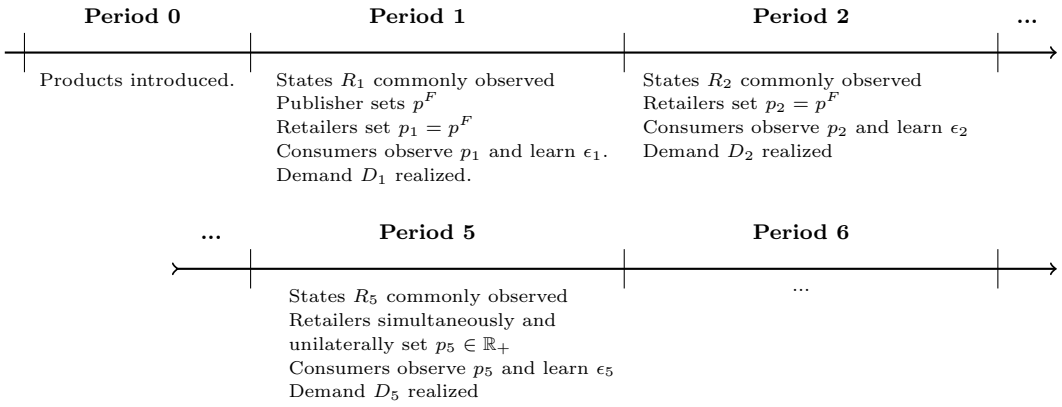

Figure 9: Timing FPA

The fixed price restraint solves

$$p^F = \arg\max_{p \in \mathbb{R}_+} \left\{ \mathbb{E}_1 \left[\sum_{j=1}^{J} \sum_{t=1}^{4} \rho^{t-1} \pi_{j,t}(p, p, R_t) | R_1 \right] + \rho^4 \mathbb{E}_1 \left[\sum_{j=1}^{J} \Pi_{j,5}(R_5; \boldsymbol{\sigma}_5^{NC}) | R_1, p \right] \right\}$$

for all $j = 1, \ldots, J$, and where σ_{jt}^{NC} satisfies (5) for each retailer in the periods after the restraint period expired, i.e. $t = 5, \ldots, 7$. The FPA strategies are hence

$$\sigma_{j,t}^{FPA} = \begin{cases} p^F & \text{if } t \leq 4 \\ \sigma_{j,t}^{NC} & \text{if } t > 4 \end{cases}$$

9.6 Full Commitment

Though the FPA strategy offers commitment, it is suboptimal for three reasons. It is constant over the course of the restraint period, it only offers commitment early in the lifecycle, and it does not offer intrabrand price coordination after the restraint period expired. To evaluate whether the FPA strategy is close to optimals, I calculate

[16] Rather than explore a more complex contract with a time varying wholesale price in the non-restraint period, I fix the wholesale price to c in all periods.

a strategy σ^{FC} where the publisher sets the price restraints with commitment in the first period and with full compliance by retailers. Consumers learn the price path at the announcement in the first period which removes any uncertainty over future prices. The timing is illustrated in Figure 10.

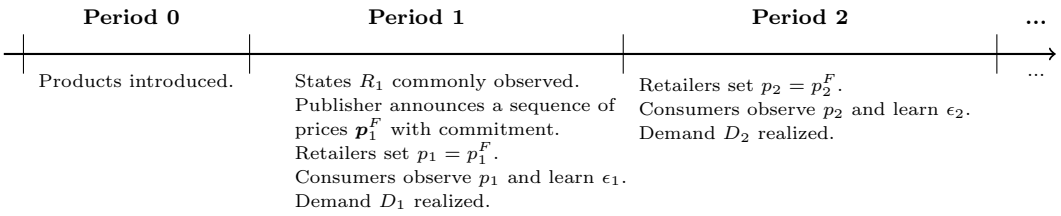

Figure 10: Timing Full Commitment

The commitment strategy maximizes the expected present value of the joint profits of the vertical unit as of the first period

$$\boldsymbol{\sigma}^{FC} = \arg\max_{\boldsymbol{p}\in\mathbb{R}_+^7} \sum_{j=1}^{J}\sum_{t=1}^{T} \rho^{t-1}\mathbb{E}\left[\pi_{j,t}(p_{j,t}, p_{-j,t}, R_t)|R_1\right]$$

The strategy is time-inconsistent, i.e. in any period t, the vertical unit, as well as each individual retailer, has an incentive to deviate from the strategy.

The algorithms used to find the equilibria are described in Section E.

10 Counterfactual results

For each contract type, the price strategies described in Section 9 are calculated along with the corresponding demand and profits. To maintain anonymity of the retailers, the reported prices and demand are averaged over the retailers. I calculate the contracts for two types of titles: a typical title and a bestseller. The title types correspond to the group fixed effects. A typical title is calculated at the average fixed effects, while the bestseller counterfactual is calculated at the fixed effects for the bestsellers.

The total profits of the three contracts measured relative to the HC contract and to the NC contract are given in Table 3 for both the typical title and the best seller. Two results stand out: there is most to gain from price coordination of bestsellers, and the FPA contract is less profitable than HC for both type of titles. In particular for a bestseller, the flexibility to revise the price in every period is more profitable than committing to a fixed price for the introductory period. It is mainly the low discount factor that makes the commitment offered by the FPA contract is relatively unprofitable.

We next turn to the shape of the price paths and the distribution of demand over time. The equilibrium prices and demand paths for both types of titles are plotted in Figure **??**. Starting with the bestseller, we note that the full commitment FC price

Table 3: Relative counterfactual profits

Type	**Contract**	**Change relative to HC**	**Change relative to NC**
Typical title	HC	0	3.94
	NC	−3.79	0
	FC	1.72	5.73
	FPA	−2.49	1.35
Bestseller	HC	0	29.03
	NC	−22.50	0
	FC	3.55	33.61
	FPA	−9.79	16.40

path declines over the lifecycle. The declining price path seems to contradict Stokey (1979)'s well-known result that the monopolist's commitment path is constant and equal to the static monopoly price. The difference follows from different assumptions. In Stokey, the consumers and the monopolist have the same discount factor, whereas here, the consumers' discount factor is smaller than the vertical unit's discount factor. The relative impatience of consumers in this application allows the vertical unit some room to commit to price discriminate over time without seeing too much demand substitute to future prices.[17]

The HC price path displays two important features. First the introductory price is about the same as in the full commitment strategy FC. Second, without commitment, the HC prices unravel. Having the discretion to revise the price in every period, the publisher acts on the incentive to lower the price as higher valuation consumers clear out of the market. Consumers who expect these discounts are more reluctant to purchase early than they would be at the same price under the commitment strategy FC, when they correctly believe that future prices are higher.

The terminal FC price is higher than the terminal HC price to persuade consumers to purchase early at high prices. Persuading consumers to purchase early under the commitment strategy is that it requires rationing consumers with willingness to pay in excess of the costs in the terminal period. The cost of rationing is borne later in the lifecycle, which creates the commitment problem.

In the NC equilibrium, retailers set prices unilaterally without commitment. Two features of the NC strategy stand out. Firstly, due to price competition between retailers, the introductory price level drops relative to the introductory prices under FC, HC, and FPA. Secondly, the NC price path flattens out relative to HC, reflecting that retailers have less market power individually than the publisher, leaving the retailer little scope for intertemporal price discrimination absent price coordination.

The introductory price under the FPA contract is similar in level to the FC contract

[17] If consumers and publishers have equal discount factors in the model, the FC price path is indeed constant.

for the four restraint periods. Following the restraint period, the retailers set prices unilaterally for the remainder of the lifecycle, i.e. they follow the NC strategies. Late in the lifecycle, the prices fall to a level marginally higher than under the NC contract.

The equilibrium paths are plotted in Figure **??** and the profits are reported in

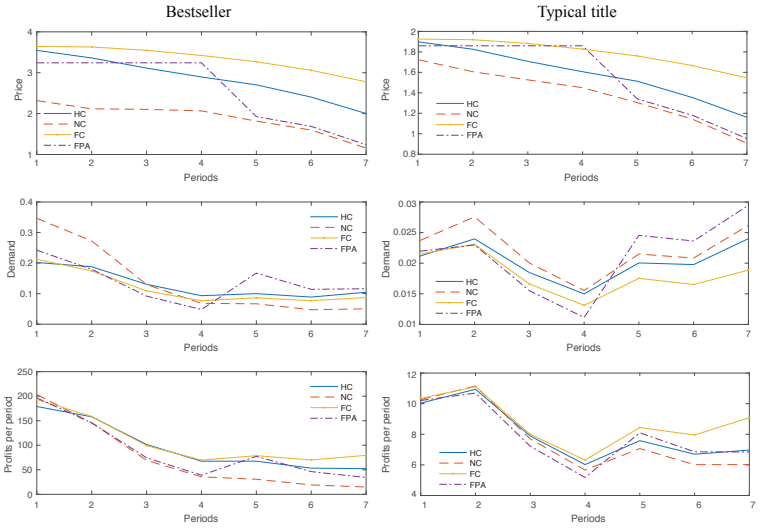

Figure 11: Equilibrium price, demand, and profits for the four contract types for a typical title

Table 3. The NC price paths show that without retail price coordination, the price paths would have flattened out and the introductory price levels would have dropped. Under HC, the publisher can still coordinate the prices between retailers with a bilateral dynamic two-part tariff, but without commitment power, the prices unravel as consumers expectations of future discounts become self-fulfilling. Moreover, the HC contract in fact performs better than the FPA contract, which offers commitment. The consumers are not patient enough for the returns to commitment under FPA to exceed the returns to price discrimination under HC.

For a typical title, the price paths of the four strategies follow the same qualitative pattern as for a bestseller, but at lower price levels. The introductory prices under the counterfactual contract HC and FPA exceed the introductory prices in the data by about 10%, but generate approximately the same discount towards the end of the lifecycle (35%). Again, going from FPA to HC in fact increases the introductory price marginally, unlike in the data, but then, in lack of commitment, the HC prices unravel.

We may compare the FPA equilibrium bestseller prices and demand to the aggregate data in Figure 4. Under the FPA strategy, the model predicts a demand increase at the expiration of the restraint period following from pent-up demand in the restraint period, which is not in the data. The price discount at the expiration of the restraint period under FPA is smaller than the price discount in the aggregate data, about 30% vs 40%, respectively. These results show that the counterfactual model does not match the empirical patterns in Section 4. Note however that the counterfactual model only uses information from the demand estimation, and does not impose any restrictions from the dynamic games.

The observed intertemporal price discrimination under the new Agreement in Figure 3 is however more similar to the HC price path than the NC price path, which suggests that the industry in fact shifted to some price coordinating vertical contracts. The results suggest that the flexibility of the pricing under the new Agreement may in fact be more profitable to the industry than the commitment offered by the old Agreement.

10.1 Caveats

The supply side model makes a number of restrictive assumptions. One is that the value of commitment in models with discrete time depends crucially on the periodization. The periodization of the model follows the data sampling intervals, which are four months. The periodization does not follow from evidence on how retailers actually set prices, but is an artifact of the data sampling.

Under a counterfactual contract like NC, the periodization of the model however implies a piecewise commitment: once the retail price is set at the start of the period, the retailers are by assumption committed to that price throughout the four month period. The model therefore likely overstates the profits of the NC contract. Without data on the actual frequency of price changes, there is a limit to how much can be done to ameliorate the problem. One robustness test is to carve up each four month period in shorter sub-periods, and allow the retailers to change prices in every sub-period. Such a robustness test is however hard to implement since the estimated demand parameters are not consistent with a finer periodization and more granular demand data are not available.

Though FPA is dominated by HC in the counterfactuals, there may be other reasons why the industry prefers a fixed price agreement. The simplicity of the price rules of the old Agreement, constant for the publication year plus the following year and

uniform across retailers, was easy for consumers to understand. That may have helped the consumers to form fairly precise price expectations. A consumer needed only see the current price and the year of introduction of a given title to accurately predict its price path. In contrast, title specific, dynamic price paths, like the FC path, which vary by title and over time, require formation of much more complex expectations.

The rigidity of the FPA strategy makes sense in a more realistic world where the retailers know more about local demand conditions than publishers, where there is uncertainty about the sales potential of a title, and where consumers have limited resources to form precise and accurate expectations over the lifecycle pricing. By fixing the price early in the lifecycle to prevent unravelling, and then leave it to better informed retailers to clear out the inventory later in the lifecycle, the Agreement may have struck a balance between commitment and pricing flexibility. The trade-off between commitment and flexibility to adapt the pricing to demand uncertainty is not considered in the analysis, but is likely non-trivial.

The demand model makes strong assumptions about both the retailers and the consumers information sets. Consumers have information on the prices at all stores, at all times, and form expectations based on their current information set, their perceptions of the residual demand in the market, and the state of competition in the market. One implication of the full information assumption is that the model leaves no room for a trade-off between the simplicity of the FPA pricing rule and the returns to intertemporal price discrimination. That level of information and rationality puts the FPA at a disadvantage relative to the optimal FC price strategy, that allows for some price discrimination over time.

Moreover, the model predicts fairly large differences in introductory prices between bestsellers and a typical title, assuming that the publisher knows the sales potential of a title prior to introduction. Most titles within a genre are however introduced at the same price point, reflecting the unpredictable sales potential of individual titles.

The analysis abstracted from competition between publishers. Competition between publishers restricts the publishers ability to implement a price skimming strategy by vertical restraints. The effect of increased competition between publishers is therefore to reduce the value of commitment. Competition between publishers is particularly relevant in a market with vertical restraints where the market performance is determined largely by the competition in the upstream market, and not among retailers. A full analysis of competition between publishers would however require an analysis of competition between titles, which has not been explicitly addressed here. However, beyond changing the form of the vertical restraints, the new Agreement did not have much impact on the structure of the upstream publishing market. The lack of impact on the competition between publishers may explain why the changes to the introductory prices were so limited.

The analysis allowed for heterogenous utility function parameters, but assumed homogenous time preferences. This assumption is standard, but not grounded in em-

pirical evidence. We may ask how the results would change if we instead allowed for heterogeneity in time preferences, but homogenous utility. Then declining price paths would price discriminate consumers differently. Consumers would self select on willingness to substitute over time, and not on willingness to pay, as in the current specification. Later in the lifecycle, more patient consumers would remain in the market and the competition from future prices would increase. We would therefore expect further unraveling of prices later in the lifecycle. It therefore seems plausible that a supply side model with heterogeneity in time preferences could rationalize the data as well.

Can these models be tested against each other without making assumptions on the pricing game? As shown in Abbring and Daljord, the model under the identifying exclusion restrictions has empirical content: there is a non-trivial region of the data space that rejects the model.[18] If the model with heterogeneity in the time preferences has different empirical content than the model with heterogeneity in the utility function, then the two models can be tested without making assumptions on the supply side. Data may fall in a region that rejects one model, but not the other. Different exclusion restrictions lead to different empirical content and seems to be a necessary condition to test the models against each other. This is an interesting task, but deriving such a test is beyond the scope of this paper.

11 Discussion

The book industry's united support for the old Agreement suggests that it solved a price coordination problem that its members could not solve by standard, bilateral vertical contracts like RPM and multi-part tariffs. The main difference between standard vertical restraints and fixed price agreement is that the Agreement commits the members of the industry to a particular price path.[19] Still, the results in this paper still show that at the discount factor recovered from the data, the returns to more flexible price skimming offered by a dynamic two-part tariff, without commitment, exceed the returns to the commitment price strategy implied by the old Agreement. Such dynamic two-part tariffs are used in other markets for durable goods. Nair (2007) notes that dynamic two-part tariffs are used the video game industry to control the price paths in a competitive retail market.

The Norwegian book market is only one of many markets that has seen fixed price agreements abolished. The evidence from evaluations of abolishment is mixed, see Canoy et al. (2005), though to the best of my knowledge, none has studied the impact on of fixed price agreements on lifecycle pricing. The evidence from the change in the Norwegian Agreement stands in contrast to the evidence from the US e-book market which effectively adopted a fixed price agreement in the spring of 2010 when Apple and six of the largest publishers imposed the agency pricing model. Under the agency

[18] In contrast, a dynamic discrete choice model in which the discount factor is assumed has no empirical content, see Rust.

[19] According to standard theories, both RPM and multi-part tariffs can achieve the same retail price targets, see e.g. Tirole (1988).

pricing model, the publishers set the retail price and leaves the retailer with a fixed percentage of the revenue, in this case 30%. Once the agency pricing model was instated in the US e-book market, the Department of Justice documented that the the e-book prices levels immediately increased by about 20% in the spring of 2010.[20]. The mixed evidence across countries suggests that fixed price agreements serve a variety of purposes, where the dominant effects may vary between countries.

The prevalence of fixed price agreements in book markets is partly related to the relative ease with which the industry is allowed exemptions in jurisdictions where RPM is illegal, such as in the European Union. Markets for books are however not the only ones to adopt fixed price agreements. Industrywide adoption of RPM has emerged in other markets for intellectual property in jurisdictions where RPM is legal, for instance in music (iTunes store) and software (AppStore). A variety of the fixed price agreement recently surfaced in the US e-book market. Apple entered an agreement with five of the top six publishing houses in the US to create an e-store that used the agency model pricing model. The agency pricing model is RPM by another name, and collective adoption of RPM in the industry is strikingly similar to a fixed price agreement. Under the agency price model, Apple received a fixed share of the retail price set by the publisher in exchange for selling the title through its web store. Coupled with a Most Favoured Nation clause, the agency pricing model however saw the Department of Justice charging Apple and the publishers with price collusion.[21] The government was concerned about interbrand price coordination. However, the results in this paper points to commitment, coordination of intrabrand prices over time, as a potentially relevant effect in markets for durable goods with collective adoption of the agency pricing model. The antitrust implications may be different for intrabrand price coordination than for interbrand price coordination.

12 Summary

This paper studies the dynamic effects of a fixed price agreement in a market for durable goods. When the price restraints of the old Agreement were weakened, the prices appeared to unravel. While vertical price restraints like RPM are usually thought of as restricting the retail price level in a static market, the changes in the sales show that the Agreement had an impact on the lifecycle price strategies. It is however not clear from the evidence whether the Agreement restricted the industry from implementing more profitable price skimming strategies, or if the prices and profits unraveled when commitment was lost with the new Agreement.

Using parameters estimated from scanner level data covering about half of the national market around the time of the new Agreement, the commitment effect of a fixed price agreement is quantified in a dynamic equilibrium model with forward looking demand and a forward looking supply side with a vertical structure. Given the modest size of the consumer discount factor that is recovered from the sales data, the estimated value of the commitment of the fixed price agreement is dominated by al-

[20] See https://www.justice.gov/atr/case/us-v-apple-inc-et-al
[21] The dispute was eventually settled outside of court.

ternative and more flexible vertical restraints that intertemporally price discriminates consumers, but without commitment.

References

Abbring, J. and O. . Daljord (2018). Identifying the discount factor in dynamic discrete choice models. Mimeo, University of Chicago. 3, 15, 16, 33

Asker, J. and A. Ljungquist (2010). Competition and the structure of vertical relationships in capital markets. *Journal of Political Economy 118*(3), 599–647. 4

Berry, S., J. Levinsohn, and A. Pakes (1995, July). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890. 16, 17

Besanko, D., J. P. Dub'e, and S. Gupta (2005). Own-brand and cross-brand retail pass-through. *Marketing Science 24*, 123–137. 4

Bonhomme, s., T. Lamadon, and E. Manresa (2018). Discretizing unobserved heterogeneity. Mimeo, University of Chicago. 18, 19, 20

Cachon, G. (2003). Supply chain coordination with contracts. In S. Graves and T. deKok (Eds.), *Handbooks in Operations Research and Management Science: Supply Chain Management*. Amsterdam: North-Holland. 4

Canoy, M., J. vanOurs, and F. van der Ploeg (2005). The economics of books. Working paper, CESifo. 6, 34

Caves, R. E. (2003). Contracts between art and commerce. *The Journal of Economic Perspectives 17*(2), 73–84. 16, 38

Chevalier, J. and A. Goolsbee (2009). Are durable goods consumers forward-looking? evidence from college textbooks. *Quarterly Journal of Economics 124*(4), 1853–1884. 4

Coase, R. (1972). Durability and monopoly. *The Journal of Law and Economics 15*, 143–149. 2

Conlon, C. (2012). A dynamic model of prices and margins in the lcd tv industry. Working paper. 4

Desai, P., O. Koenigsberg, and D. Purohit (2004). Strategic decentralization and channel coordination. *Quantitative Marketing and Economics 2*, 5–22. 4

Dubé, J.-P., G. J. Hitsch, and P. Jindal (2014). The joint identification of utility and discount functions from stated choice data: An application to durable goods adoption. *Quantitative Marketing and Economics 12*(4), 331–377. 20

Einav, L. (2007). Seasonality in the u.s. motion picture industry. *RAND Journal of Economics 38*(1), 127–145. 10, 20

Gandhi, A., Z. Lu, and X. Shi (2017). Estimating demand for differentiated products with zeroes in market share data. Working paper, UW Madison. 18

Hausman, J. (1996). Valuation of new goods under perfect and imperfect competition. In T. Bresnahan and R. Gordon (Eds.), *The Economics of New Goods*. University of Chicago Press. 16

Heiss, F. and V. Winschel (2008). Likelihood approximation by numerical integration on sparse grids. *Journal of Econometrics 144*(1), 62–80. 20

Ho, K., J. Ho, and J. Mortimer (2012). The use of full-line forcing contracts in the video rental industry. *American Economic Review 102*(2), 686–719. 4, 20

Hristakeva, s. (2017). Vertical contracts with endogenous product selections: An empirical analysis of vendor allowance contracts. Technical report, UCLA. 4

Jullien, B. and P. Rey (2007). Resale price maintenance and collusion. *RAND Journal of Economics 38*, 983–1001. 5

Kalouptsidi, M., R. Souza-Rodriguez, and P. Scott (2016, June). Identification of counterfactuals in dynamic discrete choice models. Working paper, NBER. 17

Kasahara, H. and K. Shimotsu (2009). Nonparametric identification of finite mixture modes of dynamic discrete choices. *Econometrica 77*, 135–175. 16

Krishnamurthi, L. and G. Soysal (2016). How does adoption of the outlet channel impact retail stores: Conflict or synergy? *Management Science 62*, 2692–2704. 4

Lee, R. (2013). Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review 103*(7), 2960–3000. 15

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete choice processes. *Econometrica 70*, 801–816. 14, 15, 17

Matthewson, G. and R. Winter (1984). An economic theory of vertical restraints. *RAND Journal of Economics 15*, 27–38. 23

Melnikov, O. (2013). Demand for differentiated durable products:the case of the U.S. computer printer market. *Economic Inquiry 51*, 1277–1298. 12

Mortimer, J. (2008). Vertical contracts in the video rental industry. *Review of Economic Studies 75*(1), 165–199. 4

Nair, H. (2007). Intertemporal price discrimination with forward-looking consumers: Application to the us market for console video-games. *Quantitative Marketing and Economics 5*(3), 239–292. 4, 15, 33

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*, 307–342. 16

Rossi, P. (2014). Even the rich can make themselves poor: A critical examination of iv methods in marketing applications. *Marketing Science 33*, 655–672. 16

Rust, J. (1994). Structural estimation of Markov decision processes. In R. Engle

and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, pp. 3081–3143. Amsterdam: North-Holland. 3, 14, 15, 33

Skrainka, B. and K. Judd (2011). High performance quadrature rules: How numerical integration affects a popular model of product differentiation. Working paper, Stanford. 20

Spengler, J. (1950). Vertical integration and antitrust policy. *Journal of Political Economy 58*(4), 347–352. 2

Stokey, N. (1979). Intertemporal price discrimination. *Quarterly Journal of Economics 93*(3), 355–371. 2, 29

Stokey, N. (1981). Rational expectations and durable goods pricing. *Bell Journal of Economics 12*(1), 112–128. 15, 24

Sweeting, A. (2012). Price dynamics in perishable goods markets: The case of secondary markets for major league baseball tickets. *Journal of Political Economy 120*, 1133–1172. 4

Telser, L. (1960). Why should manufacturers want fair trade? *The Journal of Law and Economics 3*, 86–105. 2

Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge: MIT Press. 33

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *Review of Economic Studies 74*, 625–652. 4

Waldman, M. (2003). Durable goods theory for real world markets. *Journal of Economic Perspectives 17*(1), 131–154. 2

Working, E. J. (1927). What do statistical "demand curves" show? *Quarterly Journal of Economics 41*, 212–235. 16

A Calculating market shares

The market size must be estimated to calculate the market shares. The market size is estimated proportional to the maximum observed sales for a title in its genre. The proportionality factor is set to 1.5. Experiments with different constants show that the results are not particularly sensitive to varying factors in a range from 1 to 2.

B Regression of total sales on fixed price

The identifying assumption of exogenous prices in the demand estimation is strong. Though this assumption is questionable, the fixed prices may be set close to randomly. These are set by the publisher prior to the introduction of a title. The conventional wisdom is that it is hard to predict the sales potential of individual titles prior to release, e.g. Caves (2003) "Nobody knows" principle. To check if the fixed prices are close to randomly assigned, I run a cross sectional regression of the total lifecycle sales on the fixed price and genre fixed effects for the titles released under the new Agreement that are used in the demand estimation. The results are given in Table 4. The coefficient on the fixed price is small and statistically insignificant. The results are consistent with the publishers having limited pricing relevant information at the time of a title's release. The fixed prices may plausibly be close to random, conditional on information known to the publisher at the time of release. The fixed price strongly affects the lifecycle pricing of titles on average and suggests that a significant share of the variation in prices may be considered exogenous. This argument falls short of verifying the assumption of exogenous prices.

.

Table 4: Regression of total quantity on fixed price

	(1)
	.
Fixed price	-0.337
	(-0.24)
Non-fiction	11.16
	(0.03)
Fiction	-230.9
	(-0.59)
Constant	4163.6***
	(13.45)
r2	0.00
N	4109

Paperback used as base

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C Nominal prices over the lifecycle

The average nominal transaction prices are plotted over the lifecycle in Figure 12. The apparent increase in the prices in the first three tertiles follow from seasonal changes in the composition of demand, and not from an increase in the titles themselves. Consumers buy different and more expensive titles in fall than in the low volume first tertile. The average prices are however seen to be surprisingly close in the second and third tertile.

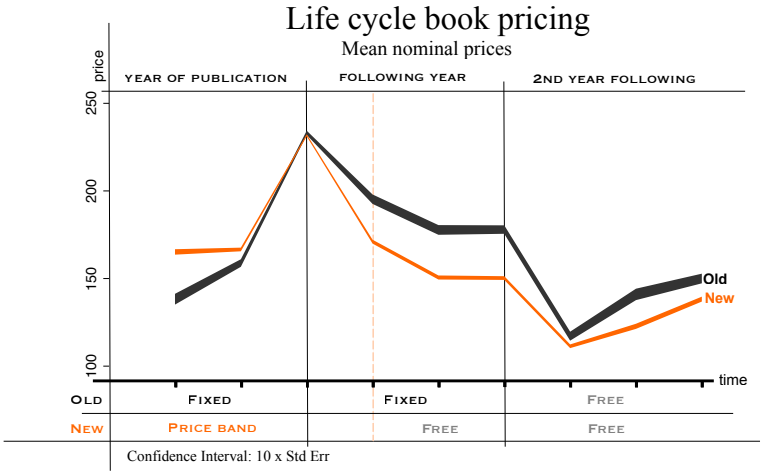

Figure 12

D Estimation of the state transitions

Partition the observable, payoff relevant states in the stochastic and deterministic processes $I_t = (p_t, x_t)'$, respectively. The state transitions are assumed to follow the first order Markov process

$$I_{t+1} = \Theta I_t + \eta_t$$

for $t = 1, \ldots, T-1$ where Θ includes retailer specific price coefficients. Partition the parameters Θ

$$\Theta = \begin{bmatrix} \theta_{pp} & \theta_{px} \\ 0 & \theta_{xx} \end{bmatrix}$$

where the zero block in the lower left corner is a restriction on the parameters that follows from the exogeneity of x. By forward iteration, write as

$$I_{t+r} = \Theta^{\cdot r} I_t + \sum_{\tau=1}^{r} \Theta^{r-\tau} \eta$$

The shocks η are assumed mean zero, multivariate normal, and serially uncorrelated. The moment conditions are

$$E[\eta_t|I_t] = 0 \quad \text{for all } t. \tag{6}$$

$$E\left[\eta_t\eta_t'|I_t\right] = \Sigma \text{ for all } t. \tag{7}$$

$$E\left[\eta_t\eta_{t+r}|I_t\right] = 0 \quad \text{for all } t \neq r. \tag{8}$$

The assumptions on η jointly define a martingale difference sequence adapted to the observable information set I_t. The first two moments of the expectations conditional on I_t are hence tractable functions of I_t itself.

$$\begin{aligned} E\left[I_{t+r}|I_t\right] &= \Theta^r I_t + \sum_{\tau=1}^{r} \Theta^{r-\tau} E\left[\eta_{t+\tau}|I_t\right] \\ &= \Theta^r I_t \end{aligned}$$

since $E\left[\eta_{t+\tau}|I_t\right] = 0$ for $\tau \geq 1$. Under assumptions (6)-(8), the second moment is

$$\begin{aligned} V\left[I_{t+r}|I_t\right] &= V\left[\sum_{\tau=1}^{r} \Theta^{r-\tau}\eta_{t+\tau}|I_t\right] \\ &= \sum_{\tau=1}^{r} \Theta^{r-\tau}\Sigma\Theta'^{r-\tau} \end{aligned}$$

Normality along with rational expectations implies that the first two moments of the transition process completely describe the expectations. The distributions are assumed stationary, so $F_t(p_{t+1}|p_t, x_t) = F(p_{t+1}|p_t, x_t)$.

The transition process is estimated from the observed price series using a Seemingly Unrelated Regression with lagged prices. Stacking the price equations for each title l, we get

$$\begin{pmatrix} \mathbf{p_1} \\ \vdots \\ \mathbf{p_L} \end{pmatrix} = \begin{pmatrix} \theta_{pp}\mathbf{lp_1} + \theta_{px}\mathbf{x_1} \\ \vdots \\ \theta_{pp}\mathbf{lp_L} + \theta_{px}\mathbf{x_L} \end{pmatrix} + \begin{pmatrix} \eta_\mathbf{1} \\ \vdots \\ \eta_\mathbf{L} \end{pmatrix} \tag{9}$$

where $\mathbf{lp}$ are laged prices. We now have standard linear system of equations that by reorganizing the parameters Θ can be written as

$$\mathbf{p} = [\mathbf{lp}, \mathbf{x}]\Theta_{FGLS} + \eta \tag{10}$$

The estimation follows a two-step procedure. In the first stage, OLS is run on (10). The estimated covariance $\hat{\mathbf{\Omega}}$ is

$$\hat{\mathbf{\Omega}} = I_{LT} \otimes \hat{\mathbf{\Sigma}}$$

where $\hat{\mathbf{\Sigma}} = \frac{1}{LT}\hat{\eta}\hat{\eta}'$ is a consistent estimator of covariance. The FGLS estimator is now the familiar

$$\Theta_{FGLS} = \left([\mathbf{lp}, \mathbf{x}]'\hat{\mathbf{\Omega}}^{-1}[\mathbf{lp}, \mathbf{x}]\right)^{-1} [\mathbf{lp}, \mathbf{x}]'\hat{\mathbf{\Omega}}^{-1}\mathbf{p}$$

By standard arguments, the errors $\hat{\eta}$ are asymptotically normal with covariance $\hat{\mathbf{\Sigma}}$.

The process is stable at $\hat{\Theta}$ with the characteristic roots of the price parameters all being of modulus less than one.

D.1 Descriptive evidence over years

Figure (13) shows that the price paths are fairly similar in all years following the deregulation. Prices start declining earlier under the new Agreement, but fall to about the same level. Similarly, the shift of demand from early at high prices to later at lower

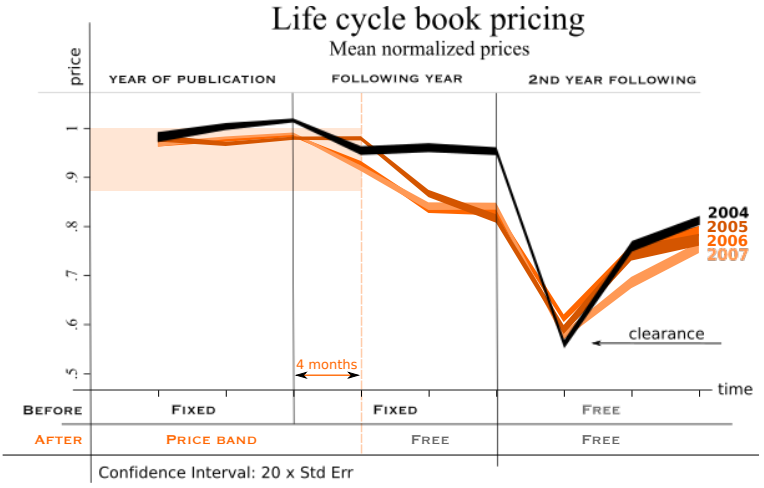

Figure 13: Lifecycle price paths year by year.

prices is also seen to be stable across the three years under the new Agreement, with some variation.

E Algorithms

The exogenous state variables x_t are assumed commonly observed by both consumers, firms, and the analyst. For notational convenience, the dependence of the value functions on x_t, which all evolve deterministically, is suppressed and written as $v_{j,t}(p_t)$. The endogenous objects are the prices strategies $\sigma_t \in \mathbb{R}_+^J$, the demand of both types $D_t^1 \in [0,1]^{J+1}$, $D_t^2 \in [0,1]^{J+1}$, and the price expectations F_t. The games are all solved by backwards induction. The algorithm finds exact solutions on a grid of the state space and interpolates between.

The dynamics are driven by the option value of no-purchase, $v_{0,t}$. Since we consider Markov price strategies, the conditional expectations over future states are written as

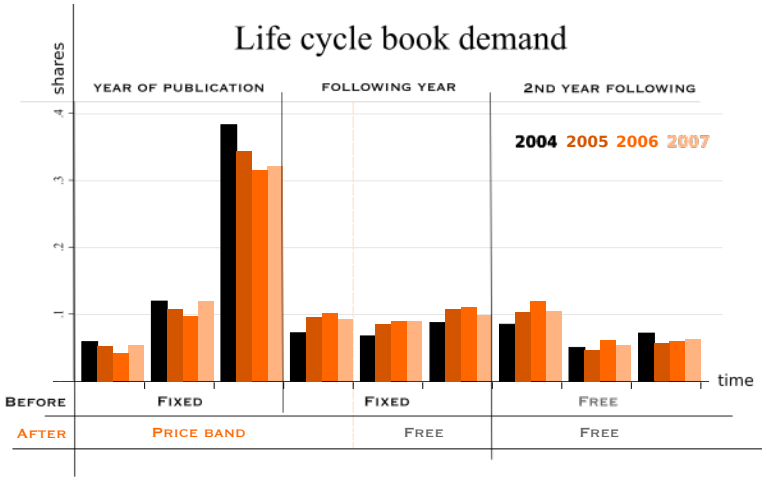

Figure 14: Lifecycle demand year by year.

expectations over the pay-off relevant states. The pay-off relevant states, beyond the deterministic states x_t, are the residual demands $R^e = [R_1^e, R_2^e] \in [0,1]^2$. Integrating out ϵ in closed form, the choice specific value function of the no-purchase option is

$$v_{0,t}(R_t) = \beta \int \ln \left(\sum_{j=0}^{J} \exp(v_{jt}(R_{t+1})) \right) dF_t(R_{t+1}|R_t).$$

Once the conditional expectations $F_t(R_{t+1}|R_t)$ are set, the demand D_t and the residual demands $R_{t+1} = D_{0,t}$ are deterministically rolled over to the next period and hence F_t is degenerate.

It is convenient to condition $v_{0,t}$ on an arbitrary, expected next period residual demand vector $R^e = [\tilde{R}_1^e, \tilde{R}_2^e] \in [0,1]^2$.

$$v_{0,t}(\tilde{R}^e) = \beta \ln \left(\sum_{j=0}^{J} \exp(v_{j,t}(\tilde{R}^e)) \right) + \beta\varphi$$

where φ is Euler's constant. Then write the demand in terms of $v_{0,t}$ conditional on the current state R_t and the expected future state $\tilde{R}^e$

$$D_t(\sigma_t, R_t, v_{0,t}(\tilde{R}^e))$$

In a rational expectations equilibrium, the expectations are consistent with the residual demand they generate. A rational expectations equilibrium is then a pair of functions $R_1^e : [0,1]^2 \to [0,1], R_2^e : [0,1]^2 \to [0,1]$ that are the fixed points to the equilibrium conditions

$$\begin{aligned} D_{0,t}^1(\sigma_t, R_t, v_{0,t}(R_{t+1}^e(R_t))) &= R_{1,t+1}^e(R_t) \\ D_{0,t}^2(\sigma_t, R_t, v_{0,t}(R_{t+1}^e(R_t))) &= R_{2,t+1}^e(R_t) \end{aligned} \tag{11}$$

given the price strategies σ, which are described below. The simultaneous equations in (11) are solved at each grid point and interpolated between. Since the profits relevant to the analysis are those of the integrated vertical unit, wholesale prices are ignored. The marginal cost to the vertical unit is set to 0.2. The counterfactuals are computed at the sample averages of x.

Horizontal Coordination algorithm

The HC algorithm is a multi-product, monopolist pricing problem. Starting in period T, the algorithm solves for equilibria on a grid $\mathcal{R}$ of B points in $[0,1]^2$, indexed by b, with typical element $R^b = [R_1^b, R_2^b]$.

Period *T*: For each point $R^b \in \mathcal{R}$, solve

$$\begin{aligned}\sigma_T(R^b) &= \arg\max_{\sigma \in \mathbb{R}_+^J} \Pi_T(\sigma, R^b) \\ &= \arg\max_{\sigma \in \mathbb{R}_+^J} \sum_{j=1}^{J} (\sigma_j - c) D_{j,T}(\sigma, R^b, v_{0,T}(.))\end{aligned}$$

1. For each guess $\sigma^0 = [\sigma_1^0, \ldots, \sigma_J^0]$
   → Set $v_{0,T}(.) = 0$, by normalization of the outside utility.
   → Calculate $\Pi_T(\sigma^0, R^b)$.

Period *T-1*: For each point $R^b \in \mathcal{R}$, solve for

$$\begin{aligned}\sigma_{T-1}(R^b) &= \arg\max_{\sigma \in \mathbb{R}_+^J} \Pi_{T-1}(\sigma, R^b) \\ &= \arg\max_{\sigma \in \mathbb{R}_+^J} \sum_{j=1}^{J} (\sigma_j - c) D_{j,T-1}(\sigma, R^b, v_{0,T-1}(R^e)) + \rho \Pi_T(\sigma_T(R_T^e), R^e) \\ &\quad \text{such that} \\ &\quad D_{0,T-1}(\sigma, R^b, v_{0,T-1}(R^e)) = R^e\end{aligned}$$

1. For each guess σ^0
   → Solve the system of equations $D_{0,T-1}(\sigma^0, R^b, v_{0,T-1}(R^e)) = R^e$ for the consistent expectations $R_T^e(R^b)$.
   → Calculate $\Pi_{T-1}(\sigma, R^b)$.

Roll back to the first period.

No Coordination algorithm

The NC has the same structure as the HC, but adds an inner loop that in each period iteratively searches for an MPE in prices. The algorithm converges at the relevant range of parameters. See Section (9) for the timing of the game.

Period *T*: For each point R^b in $\mathcal{R}$, search for $\sigma_T = [\sigma_{1,T}^b, \ldots, \sigma_{J,T}^b]$ such that the best response condition

$$\sigma_{j,T}(R^b): \ \pi_j(\sigma_{j,T}(R^b), \sigma_{-j,T}, R^b) \geq \max_{\sigma} \pi_j(\sigma, \sigma_{-j,T}, R^b),$$

holds for all retailers $j = 1, \ldots, J$.

1. Set $v_{0,T}(R^b, .) = 0$, by normalization of the outside utility.
2. For each guess σ^0, solve for the best responses

$$\begin{aligned}\sigma_j^1 &= \arg\max_{\sigma_j \in \mathbb{R}_+} \Pi_{j,T-1}(\sigma_j, \sigma_{-j}^0, R^b) \\ &= \arg\max_{\sigma_j \in \mathbb{R}_+} (\sigma_j - c) D_{j,T}(\sigma_j, \sigma_{-j}^0, R^b, 0),\end{aligned}$$

   sequentially or simultaneously for all $j = 1, \ldots, J$.
3. Repeat until convergence in σ.

Period *T-1*: For each point R^b in $\mathcal{R}$, solve for the best responses

$$\begin{aligned}\sigma_{j,T-1}(R^b) &= \arg\max_{\sigma_j \in \mathbb{R}_+} \Pi_{j,T-1}(\sigma_j, \sigma_{-j}^0, R^b) \\ &= \arg\max_{\sigma_j \in \mathbb{R}_+} (\sigma_{j,T-1} - c) D_{j,T-1}(\sigma_{j,T-1}, \sigma_{-j,T-1}, R^b, v_{0,T-1}(R^e)) + \rho \Pi_{j,T}(\sigma_T(R^e), R^e) \\ &\quad \text{such that} \\ &\quad D_{0,T-1}(\sigma_{T-1}, R^b, v_{0,T-1}(R^e)) = R^e\end{aligned}$$

for all $j = 1, \ldots, J$.

1. For each guess σ^0
   - → Solve the system of equations $D_{0,T-1}(\sigma^0, R^b, v_{0,T-1}(R^{e,0})) = R^{e,0}$ for consistent expectations $R^{e,0}(R^b)$.
   - → Calculate $\Pi_{j,T-1}(\sigma^0, R^b)$ for all $j = 1, \ldots, J$.
   - → Update the best response σ_j^1 sequentially or simultaneously for all $j = 1, \ldots, J$.
2. Repeat until convergence in σ.

Roll back to period 1.

Fixed Price Agreement algorithm

The FPA algorithm solves for the equilibrium to effectively a two-period game. In the first period, the publisher decides on a price restraint p^F which holds for four periods. Following the expiration of the restraint period, the retailers use NC price strategies for the remaining periods. The game is solved by backward induction.

Full Commitment algorithm

The FC algorithm computes the equilibrium as a two stage game between the publisher and the consumers. In the first stage, the publisher announces the prices and, by credible commitment, sets the consumers price expectations. In the second stage, demand is realized along the announced price path.