Commitment, Vertical Contracts and Dynamic Pricing of Durable Goods
Job Market Paper.
Please find the updated version at www.stanford.edu/~daljord.

Øystein Daljord∗

August 28, 2014

Abstract

Resale Price Maintenance is a vertical contract in which a manufacturer sets the retail price. Traditional motivations for RPM are that it can avoid double marginalisation and provide incentives for complementary service provision. This paper explores a new and complementary role for Resale Price Maintenance (RPM) as part of a price skimming strategy. RPM fixes the price path in dynamic markets. I argue that fixing the price path can improve upon price skimming strategies by providing commitment to future prices. Using a deregulation of RPM in a book market as a natural experiment, the effect of RPM as a commitment device is analysed empirically in two steps. First, using detailed and comprehensive retail sales data from the Norwegian book market, I show that in the absence of RPM, price skimming falters. Prices fall faster over the lifecycle and demand shifts from consumers buying early at high prices to later at lower prices. I then turn to quantify the dynamic effects of RPM. A distinction is made between two effects of RPM: it both eliminates price competition between retailers and coordinates prices over time. To separate out the price discrimination effect, I estimate a dynamic demand model and evaluate the returns to counterfactual vertical contracts in a dynamic oligopoly equilibrium model at the estimated parameters.

1 Introduction

This paper shows how Resale Price Maintenance can improve on price skimming strategies. Price skimming can be a powerful price strategy in markets for durable goods. By gradually lowering the prices over time, it lets consumers with high valuations purchase early, while consumers with lower valuations purchase later at lower prices. An obstacle to price skimming arises when forward looking consumers come to expect future discounts. A consumer that expect a future discount will find a current purchase less attractive. Demand shifts towards future lower prices and price skimming falters. The seller is effectively competing with itself. The phenomenon is known as the Coase conjecture (1972) and has received wide attention in the theory literature (Waldman (2003)). Coasian dynamics has recently been studied empirically in a variety of markets, such as college textbooks (Chevalier & Goolsbee (2009)), consumer electronics (Conlon (2012)), video games (Nair (2007)), fashion goods (Soysal & Krishnamurthi (2012)) and sports event tickets (Sweeting (2012)).

∗Thanks to my advisor Harikesh Nair, my committee of Lanier Benkard, Tim Bresnahan and Wesley Hartmann. Thanks also to Chris Conlon, Jeremy Fox, Pedro Gardete, Jakub Kastl, James Lattin, Jon Levin, Erik Madsen, Josh Moller, Sridhar Narayanan, Christian Riis, Stephan Seiler, Ali Yurukoglu and Pavel Zruynsov for comments. Thanks to Bokdatabasen for access to data, to Henrik Lande and Gjermund Nese with the Norwegian Competition Authority for facilitating access to the sales data. Thanks to Retriever for access to their media archives. Thanks to Mike Saunders and Marcus Edwall for indispensable help with software. Thanks to Shuo Xie for excellent research assistance.
Stokey (1981) shows that a seller faced with forward looking consumers would prefer to commit to a price path. Fixing the future prices can persuade consumers to purchase early by making future purchases less attractive. I first argue that RPM can provide commitment to future prices by being a legally binding contract between a manufacturer and a retailer that regulates the retail price. I then empirically quantify the returns to RPM as a commitment device in the Norwegian book market where a government deregulation imposed changes to the vertical contracts.

The deregulation of the Norwegian book market is a compelling setting for an empirical study of vertical contracts and dynamic pricing. The Norwegian book industry has an oligopolistic structure of both publishers and retailers. Until the deregulation, the industry had employed a time limited RPM for decades. The publisher fixed the retail price for a limited period, followed by a heavy discount sale, after which retailers were to price books at their discretion. Following an alignment of the Norwegian competition law with its European counterpart in 2004, the industry’s use of RPM was deemed illegal. The change in legislation provides an exogenous change in the vertical contracts that help identify its dynamic effects.

Comprehensive retail sales data show two pronounced changes to the sales patterns following the deregulation: prices fell earlier and demand shifted from early at high prices to later at lower prices. The changes in sales patterns both show evidence of forward looking consumers and that in the absence of RPM, price skimming falters as more demand is served at lower prices. The deregulation imply that the vertical contracts helped the industry maintain price strategies that retailers were not able to sustain independently. The changes in sales patterns provide the point of departure for the analysis of vertical contracts as part of a price skimming strategy.

I then turn to quantify the commitment effect of RPM. The main empirical challenge lies in that beyond weakening commitment to future prices, RPM also allows retailers to compete on price. To separate the increased retailer competition from the weakening of the commitment, I turn to series of counterfactual exercises. I first estimate a dynamic demand model and then evaluate various vertical contracts that coordinate pricing across retailers and over time.

The data are at the title-retailer level and the demand estimation allows for demand substitution between retailers and over time. The discount factor that in part determines the inter temporal substitution is usually assumed in dynamic demand models since the identifying variation is hard to find in non-experimental data (Magnac & Thesmar (2002)). By fixing the price paths, RPM arguably also fixed consumers price expectations. Using the deregulation as an instrument shifting the retailers price strategies, and consequently the consumers price expectations, the discount factor can be estimated off the sales data. A machine learning algorithm is used in the estimation routine to extract information from editorial text reviews and serves to control for the endogeneity induced by otherwise omitted demand side variables.

The estimates of the dynamic demand model enter into a dynamic supply side model of oligopolistic retail competition that allows for various types of vertical contracts. The focal contract is the time limited RPM that was used in the Norwegian book market. Three reference benchmark counterfactual contracts are then constructed.

- A baseline contract where pricing is neither coordinated over time nor between retailers. The contract is solved as a dynamic oligopoly game where retailers use time consistent price strategies that satisfy Markov Perfect Nash equilibrium conditions.
- Vertical integration without commitment. The vertical unit uses RPM to coordinate prices between retailers, but has no commitment power. The retailers price strategies are jointly time consistent.
• Vertical integration with commitment. The vertical unit uses RPM to both fully coordinate pricing both across retailers and over time.

The first benchmark is intended as a reference against which to measure various levels of price coordination. The third is a contract that exhausts the scope for coordination of prices both across retailers and over time. The difference between the second and the third contract measures the scope commitment beyond retailer coordination. Finally, the time limited contracts employed in the Norwegian market is compared to the counterfactual benchmark. Preliminary results show that the RPM strategy employed in the Norwegian book market comes quite close to the vertically integrating contract with commitment.

This is primarily a paper about vertical restraints and dynamic pricing, yet the findings have implications for an ongoing policy discussion on RPM. RPM has been illegal in many countries and product markets. The last decades there has however been a shift towards increasing acceptance and endorsement of RPM. Though RPM is per se illegal in the EU, book markets are often made exemptions. Having been per se illegal in the US for about hundred years, RPM is now subject to a rule-of-reason. A rule-of-reason implies that the legal status of RPM depends on its competitive effects. From a policy perspective, it is therefore important to understand the competitive effects of RPM in retail markets, yet little is known empirically, in part because RPM has been an illegal contract form. This paper adds to the scarce empirical evidence on RPM. Results in this paper have particular policy relevance by showing that improved price discrimination may be an effect of RPM, an effect that so far has not been documented in the literature. It is for instance not a priori clear that more efficient price discrimination is welfare reducing.

The Norwegian market RPM studied in this paper was implemented through a Trade Agreement that included the majority of publishers and retailers. The Trade Agreement is not unique in its kind. Fixed price agreements, either regulated through trade organisations or by government law, are found in other European book markets as well. Other countries that practice fixed price agreements are France, Germany, Austria, The Netherlands, Italy, Spain, Portugal, Greece, Hungary, Israel, Slovenia, Argentina, South Korea, Japan and Mexico, according to the International Publishers Association. The widespread use of RPM make the book markets interesting in their own right.

The motivation for fixed price agreements is typically the belief that they stimulate bibliodiversity, that is, a culturally diverse literature and a dense distribution network of book stores. Though neither the supply of variety nor the entry and exit of book stores are the focus of this paper, it points to an effect of RPM that can improve industry profitability so as to sustain bibliodiversity. There is however no clear consensus on the use of RPM as a policy tool in book markets. For instance, while RPM is illegal in the UK book market, RPM is government enforced in the French book market. The variation in policies across countries reflects a divided

---

1For an agreement that is per se illegal, the existence of the agreement is illegal, regardless of its effects. Under a rule-of-reason, either the plaintiff must show that RPM is anti-competitive, or the defendant must show that RPM is pro-competitive.

2See *Leegin Creative Leather Products, Inc.* vs. *PSKS, Inc.*, 551 US 877. Following *Leegin*, a minimum price restraint was no longer per se illegal. A maximum price was absolved from per se illegality status 10 years earlier with *State Oil Co.* v. *Kahn*, 522 US 3. Along with *Continental T.V.*, Inc. vs. *GTE Sylvania Inc.*, 433 US 36 where vertical restraints on territories on exclusivity clauses were made subject to a rule-of-reason from being per se illegal in 1977, these cases collectively represent a significant and consistent change towards an increasingly friendly policy on vertical restraints in the US.

3See Telser (1960) for a contemporary analysis of the Fair Trade institution itself. Following *Leegin*, some evidence has emerged. In the seemingly most comprehensive empirical study, McKay & Smith (2013) use variation in federal precedents across states and a diff-in-diff approach on grocery scanner data before and after *Leegin*. It finds evidence of price increases, but the study does not consider the dynamic pricing of durable goods explicitly.

4Regarding distribution, the UK has experienced sharp decline in the number of bookstores, whereas France has not. See the press release of the UK Booksellers Association of Oct. 3rd 2011, see
opinion on the merits of RPM as a policy instrument. Despite the controversy surrounding the use of RPM, the recent liberalisation of the US legislation and the growing European acceptance of RPM renews its relevance as a managerial pricing tool.

The main managerial content of this paper lies in demonstrating how RPM can coordinate downstream price incentives and mitigate Coasian dynamics at plausibly low costs. Commitment can be achieved through other means than RPM, for instance through reputation. Apple’s fairly consistent policy across their product lines of offering very limited discount over time is an example that can persuade consumers to purchase early. Acquiring commitment through reputation can however be both costly and hard. JC Penney, a clothing retail chain, famously adopted an Every Day Low Price policy to avoid losing demand to future discounts. The retail chain however failed to convince consumers, and was forced to revert to frequent sales.\(^5\) To the extent managers can improve on price skimming through writing vertical contracts, it may prove to be a cost efficient strategy.

The vertical restraints analysed in this paper have become popular in digital markets where it is known as the agency pricing model. Under the agency model, the manufacturer sets the retail price and the retailer gets a share of the revenue. The agency pricing model is used in the App Store and on eBay, and it received wide attention following the recent e-books antitrust case against Apple and a set of US publishers.\(^6\) Coupled with a Most Favoured Nation clause, the publishers along with Apple forced through an industry wide migration from the classic wholesale pricing model to the agency model. The result was an industry pricing model that closely resembles the RPM in the Norwegian market studied in this paper. De los Santos & Wildenbeest (2014) gives an exposition of the antitrust case and show evidence of substantial increases in prices following the adoption of the agency model. According to case documents, the move to the agency model was motivated in part by publisher’s concern that Amazon’s heavy discount e-book price policy would lead consumers to expect low prices on all types of books. The role of consumers expectations, and how they relate to vertical restraints, is the topic of this paper.

The impact of RPM as a price skimming strategy is quantified by bringing together concepts and frameworks from the so far mostly distinct literatures on RPM and dynamic pricing. 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. The literature on channel coordination in dynamic markets is however relatively slim. Desai et. al. (2004) is closest to the idea in this paper. It considers a dynamic channel coordinating two-part tariff in a two-period durable goods market with forward looking consumers. In contrast, this paper considers a wider set of contract types, and in particular RPM, it allows for an oligopolistic retail market, it specifies the sources of commitment, and it provides empirical evidence on the effects.

There is little empirical evidence to bear on the effects of vertical contracts. Some empirical papers on vertical contracts have emerged the last few years, e.g. Besanko et al (2005) on retail pass-through, Villas-Boas (2007) on identification of unobserved vertical contracts, Asker & Ljungquist (2013) on the impact of vertical integration in investment banking. By adding vertical restraints to the pricing problem, this paper also contributes to the scarce empirical evidence on vertical contracts and supply chain coordination in general. The two empirical papers typically closest to the current paper are Ho et al. (2012) and Mortimer et al. (2008). The first examines full line forcing in the video rental industry, the second studies revenue sharing in the same industry. Neither however consider explicitly the

---

5See https://www.nytimes.com/2013/04/14/business/for-penney-a-tough-lesson-in-shopper-psychology.html

http://www.booksellers.org.uk/campaigns/keepbooksonthehighstreet
role of vertical contracts for dynamic pricing. Finally, this paper contributes to a nascent empirical literature on dynamic oligopoly pricing of durable goods (Conlon (2012), Goettler & Gordon (2011), Gowrisankaran & Rysman (2009)).

The next section describes the price discrimination problem in a dynamic market with an upstream publisher and oligopolistic downstream retailers. The discussion sketches how vertical contracts improve price skimming in markets with forward looking consumers. Sections (3) and (4) describe the vertical contracts employed in the Norwegian book industry, the deregulation and the data. Section (5) shows reduced form evidence of the impact of the deregulation of RPM on prices and sales. Section (6) describes the counterfactuals in terms of a demand and a supply side model. The estimation routine is described in sections (7) and (8). The empirical results are reported along with the counterfactual simulations in sections (9) and (10).

2 Price skimming and vertical contracts in oligopolistic markets

This section provides an informal discussion of some of the issues raised in implementation of price skimming strategies in oligopolistic retail markets. It explains how the vertical unit, a publisher and a set of retailers, can use vertical contracts to address these issues. The discussion leads to a description of the RPM contracts used in the Norwegian book industry. An important distinction is made between the effect of RPM on coordinating prices between retailers (horizontal coordination), and the effect on coordinating prices over time (inter temporal coordination). Though conceptually different, both levels of coordination, or lack thereof, affect the vertical unit’s ability to price skim.

Some examples are instructive. Suppose first a publisher has a set of books to sell in a market with a fixed number of heterogenous consumers. Each consumer has unit demand. The publisher adopts a price skimming strategy that gradually lowers the price over time. The publisher’s dynamic trade off in any period is between lowering the current price and increase the current profits, at the expense of reducing future profits. Lowering the current price may increase the current profits by increasing the current demand at the expense of tapping into the future demand. The reduced future demand will furthermore be served at lower prices as high valuation consumers are already cleared out of the market. Making the dynamic trade-off, the publisher gradually lowers the price and appropriates the surplus the books generate.\(^7\)

Adding retailers to the supply chain requires some modifications to the simple price skimming strategy. Suppose that the publisher employed a set of competing retailers to sell the books. The dynamic trade-off for any one retailer considering lowering the current price is again between increasing the current profits at the expense of reducing future demand. The future demand is however of less value to the retailer for two reasons. Firstly, the future demand is shared with its rival retailers, which debases its value. Secondly, by lowering the current price, the retailer gets some of the current demand of his rival retailers. In pursuit of market shares, the retailers collectively fail to efficiently price skim the market. RPM can eliminate wasteful competition between the retailers by directly controlling the retail prices, where wasteful is understood from the perspective of the vertical unit.\(^8\)

Forward-looking consumers represent a further challenge. A forward looking con-

---

\(^7\)The returns to optimising price skimming strategies can be substantial in dynamic markets. Lazarev (2013) finds that current airline pricing strategies extract on the order of 90% of consumers surplus.

\(^8\)Other restraints can in principle achieve the same outcome. In a companion paper, I develop a vertical contract that sustains the efficient price path through a path of wholesale prices and transfers, see Daljord (2014b). Nair (2007) report that decreasing wholesale price paths are seemingly common in the video game industry.
sumer that correctly predicts future discounts will find a future purchase more attractive. Then high valuation demand shifts towards future lower prices and price skimming falters. This is the classic durable pricing problem of Coase (1972).

To counteract the inter temporal substitution, the publisher may announce that the future will hold no discounts. The announcement may however not be credible. To see that, suppose the consumers believe the announcement. Consumers with a valuation in excess of the price now have no reason to delay the purchase. As soon as the high valuation consumers are cleared out of the market, the seller has an incentive to reduce the price to capitalise on the low valuation demand left in the market. But then the announced future prices are inconsistent with the actual prices. It is hard to imagine that announcing future prices that consistently fail to realise can form part of a long run, viable price skimming strategy. The seller loses profits to competition from its own future pricing and is left looking for other means to coordinate prices over time.

Suppose now that instead the publisher exposed himself to costly consequences was he not to price along the announced path. The potential consequences can provide the publisher a commitment to the announced price path by counteracting the incentive to discount as soon as the high valuation consumers have left the market (Stokey (1981)).

The discussion emphasises three key components to a price skimming strategy in a market with forward looking consumers. Firstly, the seller needs to announce a price path. Secondly, the price path must persuade high valuation consumers to purchase early rather than wait for future discounts. Thirdly, a counteracting incentive is required to ensure the seller prices along the announced price path. In the next section, the RPM agreement in the Norwegian book market is shown to both announce a price path and have externally enforced sanctions in place to provide the vertical unit commitment to the announced price paths.

3 Intervention

The book industry in Norway employed a trade agreement ("Agreement") dating back to the 1960s up until the deregulation became effective in May 2005. The Agreement specified a time limited retail price restraint.

The old Agreement was a legally binding contract voluntarily entered between the Association of Booksellers and the Association of Book Retailers that specified the terms of sales in the industry. The restraints of the Agreement had two key components. The first component was a time-limited price restraint period where the publisher fixed the retail prices. The second component was an industry coor-dinated clearance sale following the expiration of the price restraint period. The time limited fixed price and the coordinated clearance sale together spell out a price path.

The Agreement had kept the pricing strategies stable in the industry for decades. The fixed price was often hard printed onto the cover of the book, which served as an announcement of the retail price. There was little price promotion in the industry except for the clearance sale, which was trade marked and heavily advertised.

The Agreement also spelled out rules of arbitration in case of a breach of contract. From Clause 5:

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.

Two features of the sanctions stand out. Firstly, the Agreement did not coordinate on a price level, at least not explicitly, but on a price path. Each publisher was at liberty to set any retail price, but once set, the retailers had to respect the fixed price over the restraint period.\textsuperscript{10}

Secondly, though the agreement regulates a bilateral agreement between a publisher and a retailer on the shape of the price path, the Agreement exposes the vertical unit to threats of legal action by rival firms if the unit was to deviate from the specified price path. The arbitration clause therefore provides means of external enforcement of a bilateral agreement, consistent with commitment.\textsuperscript{11} The exact nature of the consequences were not explicitly stated. We will however see in the empirical section that by and large, the Agreement was respected.

Following an alignment of the Norwegian competition law with its European Union counterpart in 2004, the Norwegian Competition Authority deemed the RPM Agreement unlawful and called for abolishment. The industry voiced strong and united opposition against the deregulation, suggesting the Agreement helped solve an industry coordination problem. The Association of Booksellers, the Association of Publishers and the Association of Authors rallied together against the deregulation and called for exemption from the competition law.\textsuperscript{12} A public debate ensued and a political compromise was reached.\textsuperscript{13}

The deregulation provides an exogenous change to the vertical restraints in the industry. The main changes under the new Trade Agreement effective the spring of 2005 were twofold.

- 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 RPM under the old regime was a floor and a ceiling, retailers were given discretion to discount the fixed price by up to 12.5\% under the new regime.

The changes to the price restraints following the deregulation are illustrated in Figure (1).

The new regulation implied a shorter period of commitment to a fixed price path and a weakening of the restraints themselves. We will see in the coming sections that the deregulation led to substantial changes in both retailer price strategies and the demand patterns. The next section documents changes in sales patterns following the change in the vertical restraints.

\textsuperscript{10}The Trade Agreement has been suspected of facilitating horizontal collusion among publishers. The idea is that with RPM, it is easier for publishers to detect deviations on observable retail prices than say on unobservable and flexible wholesale prices, see Jullien & Rey (2007) for one treatment of the argument.

\textsuperscript{11}Beyond allowing for legal actions of rival firms, the Agreement also allowed the Associations to meter out further punishments within the confines of each associations bylaws. A translated excerpt of the Agreement is given in the Appendix.

\textsuperscript{12}Exemptions from the 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 & van der Ploeg (2005).

\textsuperscript{13}To give some context of the media attention devoted to the deregulation, 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 'Salt Lake City Olympics' gives over a comparable period at the time of the contemporaneous winter olympics. The numbers can give some perspective of the media interest the deregulation spurred in a nation which is above average preoccupied with winter sports.
Figure 1: The price restraints followed calendar time. Some title 1 released early in the year and some other title 2 released later in the year would both have their price restraints lifted at the end of the following year under the old regulation. 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 typical title lifecycle and most titles sell little. After the deregulation, the same titles 1 and 2 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.

4 Data

The sales data is scanner data collected from the four largest book retail chains in four month periods over the years 2004 to 2007, bookending the deregulation effective in May 2005. The data make up about 50% of total national sales over the period. The data is aggregated over four months, tertiles, and across stores within each chain. Observations are on title level identified by an Electronic Article Number (EAN) and contains data on a little more than 27000 titles. The EAN identifier allows the sales data to be merged 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. Price policies were mostly uniform within the chains according to industry representatives. The summary statistics of the scanner data are given in Table (1).

### Table 1: Summary statistics matched scanner data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>retail price</td>
<td>181.51</td>
<td>128.14</td>
</tr>
<tr>
<td>quantity</td>
<td>27.58</td>
<td>171.48</td>
</tr>
<tr>
<td>year</td>
<td>2005.68</td>
<td>1.08</td>
</tr>
<tr>
<td>fixed price</td>
<td>218.19</td>
<td>129.47</td>
</tr>
<tr>
<td>N</td>
<td>1127867</td>
<td></td>
</tr>
</tbody>
</table>

5 Impact of deregulation

Figure (2) illustrates the window of data and the deregulation. 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 and all price restraints were lifted both before and after the deregulation. There are two cohorts of full life cycles in the data. These are the titles published in 2004 and the titles published in 2005. I would ideally compare the price paths and demand for a typical title over the lifecycle before and after the deregulation. The data however cover only one year of sales before the deregulation.

To establish the sales patterns before and after the deregulation, I assume that the price and sales patterns are comparable across cohorts within a given year. The assumption is that though individual titles change across years, the mean prices and aggregate sales are drawn from a stable distribution. I then construct a lifecycle price path before the deregulation by taking the mean price path in
2004 of new titles, splice it with the price path of one year old titles in 2004 and lastly with the two year old titles in 2004. The resulting price path serves as a measure of the representative price path before the deregulation. Holding the release schedule fixed, we can then make meaningful comparisons of the differences in sales before and after the deregulation. There are three periods per year: spring, summer and fall.

5.1 Impact of the deregulation on price paths

The mean representative prices before and after the deregulation are graphed in Figure (3). The prices are normalised to the fixed price publishers set at the time of publication. A price of 1 implies that a title on average retailed at the fixed price, whereas a price of say 0.5 means a title retailed at 50% discount of the fixed price. The normalisation allows comparison of price paths across different price points. The prices are plotted against time at the tertile periodisation of the the data and contains a total of nine points. Confidence intervals are linearly interpolated between the data points to display the variance of the means. Since standard errors are relatively tight, the confidence intervals are reported at non-conventional levels to display visually discernible variation over time. A table of the price paths is provided in the Appendix.

The price restraints before and after are denoted below the graphs. After the deregulation, the retailers were allowed to discount the fixed price by up to 12.5% at discretion. The restraint defines a price band and is illustrated by the shaded area. The retailers are seen to have largely respected the fixed price policy before the deregulation. 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 RPM was effectively a minimum price restraint. The restraint period was followed by the industry coordinated clearance sale which saw average discounts on the order of 45%.

After the deregulation, retailers were allowed more flexibility in determining the price. The restraint period was shortened by eight months. The expiration of the shortened restraint period is illustrated by the vertical orange line. Three changes stand out. Firstly, in the absence of RPM, the prices start declining earlier and at the end of shortened restraint period. The Agreement clearly precluded retailers from discounting towards the end of the lifecycle, before the clearance sale. Secondly, prices do not fall in the early phases, despite the retailers discount discretion. The average price lies well within the price band in the shortened restraint period. As a sign of dynamics, the incentive to cut prices evidently changes substantially over the lifecycle. Thirdly, prices fall to about the same level at the clearance sale. Towards the end of the lifecycle, prices are about the same before and after.

---

14 Note that as titles are released over the year of publication, the set of titles in the sample is growing. There is however no evidence of changes in the timing of the release dates following the deregulation.

15 Perhaps surprising since changes in the duration of the restraint period might change the optimal timing of releases. Einav (2007) finds evidence of strategic timing of releases in the U.S. motion picture industry.
Figure (3) shows that the price paths in the early phases were similar before and after the deregulation, but in prices normalised to the introductory prices. It could be expected that the introductory prices would change along with the price paths, there were however no material changes in the price levels. In Table (2), the results from a linear regression of the introductory prices of all title on a deregulation dummy shows that the mean increase was on the order of 2.5% from 2004 to the post deregulation period 2005-2007.\textsuperscript{16} The results are similar broken down over individual years in the post-deregulation period. Controlling for genre fixed effects, the increase is about half a percentage point higher, either way on par with general inflation.

Table 2: Changes to introductory prices.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>post</td>
<td>4.50</td>
<td>2.25</td>
<td>5.80</td>
<td>1.88</td>
</tr>
<tr>
<td>constant</td>
<td>211.79</td>
<td>1.95</td>
<td>228.62</td>
<td>4.60</td>
</tr>
<tr>
<td>genre fx</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>16605</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure (4) plots the corresponding shares of sales over the lifecycle. The key feature is that at comparable early phase prices, demand shifts towards lower

\textsuperscript{16}The central bank inflation target was 2.5% in the same period.
prices mid-cycle when the RPM no longer was in place to sustain higher prices. The evidence is consistent with forward looking consumers. Expecting lower future prices following the deregulation, consumers are more willing to wait for a future discount at otherwise comparable prices after the deregulation. The change in price strategies and the demand response after the deregulation also show that the price skimming strategies falter in the absence of RPM as demand is served at on average lower prices. The changes in both price and demand patterns are qualitatively stable across years and across genres, see the Appendix for additional evidence.

6 Model setup

The goal of the empirical analysis is quantify the commitment effect of RPM, while controlling for its horizontal coordination effect. Ideally, I would compare vertical contracts that provide commitment versus vertical contracts that do not provide commitment. Since the observed RPM contracts both eliminate horizontal price competition and fix the price path, realistic counterfactuals are required to single out the commitment effect of RPM. The modelling approach follows two steps. The first step is to estimate substitution patterns along the horizontal and inter temporal dimensions from the sales data. The second step is to evaluate counterfactual vertical restraints in a dynamic supply side model at the estimated parameters. The counterfactual vertical restraints are constructed to separate the horizontal effect of RPM from the inter temporal effect of RPM.

I start by developing a model of a market with a forward-looking demand side and a forward looking supply side. The supply side allows for vertical contracts between a publisher and a set of retailers. The retail prices and demand are determined in equilibrium depending on the state of the market. Consumers make choices considering the current prices and their beliefs about retailers price strategies. Retailers set prices taking into account the impact of their prices on both current and future demand. In equilibrium, both consumers and retailers have beliefs about the evolution of the market that are rational in the sense of being consistent with the distribution of the actual prices and demand.

6.1 Demand specification

This section describes the demand model that enters the counterfactuals. The demand model is taken to data in Section (7). A publisher is assumed to be a monopolist of a given title and the relevant dimension of competition is taken to be between retailers for the same title. This is a strong assumption on substitution, but it allows me to parsimoniously focus on substitution between retailers and over time. Allowing for substitution between titles would either require an unfeasibly high dimensional state space, there are thousands of titles released every year, or alternatively require strong assumptions on the state transition process, as for instance in Melnikov (2013). The interest in the current application lies specifically in the exogenous changes the deregulation caused to the state transition process.

Assuming away substitution between titles allows a more flexible specification of the state transition process. The demand is modelled along the lines of Arcidiacono & Ellickson (2011). Demand is represented by a finite horizon discrete choice model with discrete heterogeneity. A finite horizon is used as the number of periods is small, it allows for non-stationary policy functions, the typical life cycle is finite, and that it allows for a convenient, yet flexible estimation of the state transition process.

There are $G$ types of consumers. Each consumer $i$ of type $g$ decides at any time $t = 1, \ldots, T$ whether to buy a given title from retailer $j = 1, \ldots, J$ or, postpone the purchase decision to the next period. Consumers form expectations over the future state of the market. Consumers have unit demand. Once a consumer buys
the given title, he leaves the market never to return.

A consumers flow utility is state contingent. The states \( s_t \) are split in endogenous states \( p \) and exogenous states \( x_t = (\xi_t, z_t) \). The exogenous demand shock \( \xi_t \) is assumed observed by both consumers and retailers, but is not in the data, and is conditionally independent and serially uncorrelated. In each period, every consumer draws a vector of conditionally independent, iid \( EV1 \) shocks \( \epsilon_{it} = (\epsilon_{it1}, \ldots, \epsilon_{itJ_t}) \). Suppressing the group index and the superscript on \( s \), the flow utility for consumer \( i \) of buying the title from retailer \( j \) in period \( t \) is given by

\[
u_{ijt}(s_t, \epsilon_{ijt}) = v_j(s_t) + \epsilon_{ijt}
= \gamma_p p_{jt} + \sigma \xi_{jt} + z_t \gamma_z + \epsilon_{ijt}
\]

The value of the outside option is normalised to \( v_{00t} = 0 \), constant across time and consumers. The deterministic states \( z_t = \{z_{i1t}, \ldots, z_{iJ_t}\} \) include both time-invariant characteristics such as genre and retailer fixed effects, and explicit functions of time such as seasons and a taste for novelty. A full description of the variables that enter the empirical model is given in Section (9).

Consumers are forward looking expected utility maximisers that discount future utility by factor \( \beta \in (0, 1) \). The consumer’s value function is

\[
V_t(s_t, \epsilon_t) = \max \left\{ \max_{j=1, \ldots, J} v_j(s_t) + \epsilon_{jt}, \beta \int \int V_{t+1}(s_{t+1}, \epsilon_{t+1})dF_{s}(s_{t+1} | s_t)dF_{\epsilon}(\epsilon_{t+1}) + \epsilon_{0t} \right\}
(1)
\]

where \( F \) represents the beliefs over the state transitions. By conditional independence of \( \epsilon \), the transition process can be factored \( dF(s_{t+1}, \epsilon_{t+1}|s_t, \epsilon_t) = dF_{s}(s_{t+1}|s_t)dF_{\epsilon}(\epsilon) \). The terminal value equates to the static utility maximisation problem

\[
V_T(s_T, \epsilon_T) = \max \left\{ \max_{j=1, \ldots, J} v_j(s_T) + \epsilon_{jT}, \epsilon_{0T} \right\}
\]

Following Rust (1987), define \( \tilde{V}_t(s_t) = \int V_t(s_t, \epsilon_t)dF_{\epsilon}(\epsilon_t) \), the expected value of making the optimal choice conditional on states \( s_t \) before \( \epsilon \) is revealed. From the conditional independence and additive separability of \( \epsilon \), the familiar log sum representation follows

\[
\tilde{V}_t(s_t) = \ln \left( \sum_{j=0}^{J} \exp(v_{jt}(s_t)) \right) + \Gamma
\]

where \( \Gamma \) is Euler’s constant. Then define the conditional value functions as

\[
v_{jt}(s_t) = \begin{cases} 
\beta \int \tilde{V}_{t+1}(s_{t+1})dF(s_{t+1} | s_t) & \text{for } j = 0 \\
v_j(s_t) & \text{for } j = 1, \ldots, J 
\end{cases}
(2)
\]

Starting from the terminal period, the problem is solved by backward induction. Integrating out the private shocks \( \epsilon_T \)

\[
\tilde{V}_T(s_T) = \ln \left( 1 + \sum_j \exp(v_{jt}(s_T)) \right)
\]

as \( v_{0T} = 0 \) by normalisation. In period \( T - 1 \), the option value of waiting is

\[
v_{0T-1}(s_{T-1}) = \int \tilde{V}_T(s_T)dF(s_T | s_{T-1})
\]

By backwards recursion, the value functions in (2) are defined up to \( t = 1 \).

Consumers choice functions can now be defined in terms of the conditional value
functions. Conditional on still being in the market at time \( t \), the probability of consumer \( i \) choosing to buy from retailer \( j \) is given as

\[
d_{ijt}(s_t) = \frac{\exp(v_{jt}(s_t))}{\exp(v_{0t}(s_t)) + \sum_{k=1}^J \exp(v_{kt}(s_t))}
\]

which completes the model of consumer choice.

The aggregate demand is simply the sum of individual demand. The initial shares of each consumer type is \( w^0 \). Starting from period 1, the residual demand \( R^0 \) of type \( g \) is the share of consumers that is still in the market. It is defined recursively

\[
R^0_{t+1}(s_t) = R^0_t d^0_{ijt}(s_t)
\]

with \( R^0_1 = 1 \) for all \( g \). The market shares of type \( g \) is then

\[
D^0_{jt}(s_t) = R^0_t d^0_{ijt}(s_t)
\]

Finally, aggregate demand sums the demand over all groups

\[
D_{jt}(s_t) = \sum_g w^0 D^0_{jt}(s_t)
\]  

(3)

Note that the consumer expectations \( F \) have so far been left unspecified. The expectations in the counterfactuals and in the estimation enter differently. In the counterfactuals, we will see that the consumer expectations are assumed structurally consistent with the price strategies, and price strategies with consumer expectations. Both sides expectations are determined in equilibrium, more detail is given in Section (6). No equilibrium restrictions are however imposed in the demand estimation. Instead, a flexible reduced form specification for expectations that is consistent with the observed price process is employed. The empirical specification of the transitions are given in more detail in Section (8).

6.2 Assessing coordination by contracts

I make one key simplifying assumption for the counterfactual models

**Assumption 1.** The joint profits of the vertical unit can be split arbitrarily between its members.

Vertical contracts that maximise the joint profits are called *sufficient* following Matthewson & Weiner (1984). With arbitrary profit splits, there will always exist a sufficient vertical contract that weakly Pareto dominates any insufficient vertical contract. Some sufficient contract will therefore be preferred of all parties regardless of the bargaining protocol. The assumption allows me to abstract away from the bargaining process between publisher and retailers. I can first focus on finding the optimal price paths for the vertical unit, and then find a vertical contract that implements those price paths.

There may be more than one sufficient contract. A classic example is the case of double marginalisation where both a two-part tariff and RPM are sufficient. The contracts maximise the joint profits by different means: RPM controls the retail prices directly, whereas a two-part tariff controls the retail prices indirectly through the wholesale price. By controlling the retail price level directly, RPM can in principle implement any price path, including the sufficient restraints. The implementation is therefore by RPM where I abstract away from the profit splits under Assumption (1). The profit splits are anyhow unobserved.

The assumption of unrestricted profit splits is reasonable in the sense that profit transfers are common in the contracts used in the industry. Interviews with industry representatives point out that the vertical contracts employed often contain
various kinds of fees that do not directly depend on the quantity sold. These fees serve as vehicles to transfer profits within the vertical unit, without affecting the retailers price incentives. The fees are subject to bargaining. There may still be reasons why the actual contracts in the industry are insufficient. Without data on the vertical contracts themselves, I can not directly test the assumption.

The time limited RPM contract used in the Norwegian book industry is compared to a set of counterfactual vertical contracts.

- No coordination NC. The publisher supplies the retailers books at the marginal cost of production. The retailers engage in dynamic, unilateral price competition.
- Horizontal coordination HC. The vertical unit coordinates the prices across retailers, but has no commitment power.
- Full commitment FC. The vertical unit is endowed with commitment power and coordinates prices both across retailers and over time.

The retailer’s price setting will be defined in terms of a set of value functions for each counterfactual contract scenario. Note that the retailers control variables $p_t$ are among the consumers state variables and retailers state variables will be the composition of demand $D_t$, derived from the consumers control variables. Both $p$ and $D$ are determined in equilibrium, along with both sides expectations.

The objects of interest from the counterfactual exercises are the equilibrium price strategies and associated profits in the above scenarios. The vertical unit maximises profits subject to various vertical contract restrictions. The periodisation is discrete following the format of the data. The pay off relevant market conditions for the retailers are fully summarized by the current state $s_t \in S$. The retailers per-period profit function depends on the state of the market and the rival retailers prices

$$\pi_j(p_t, s_t) = (p_{jt} - c)D_{jt}(p_t, s_t)$$

where $p_t$ is the vector of prices at times $t$ and where $c$ is the constant marginal cost of production. The retailers state variables $s_t = (D_t, x_t)$ are the composition of demand $D_t$ and exogenous state variables $z_t$. For a given set of price paths $p_t = \{p_{1t}, \ldots, p_{Jt}\}_{t=1}^T$, the retailers state contingent expected profits are recursively defined as

$$\Pi_{jt}(p, s_t^r) = \pi_j(p_t, s_t^r) + \rho \int \Pi_{jt+1}(p_{t+1}, s_{t+1})dF(s_{t+1}|s_t, p_t)$$

with the terminal period profit given by $\Pi_{jT}(s_T) = \pi_j(p_T, s_T)$. The games are complete information in the sense that consumers are assumed to know the firms policy functions and firms know the consumers demand functions, this is common knowledge, and all state variables are commonly observed by all agents.

6.3 **Rational expectations equilibrium**

The statistics of interest in the counterfactuals are the price strategy profiles and the associated profits of the vertical unit. An equilibrium consists of

- A collection of non-stationary, state contingent price policy functions $p_{jt} : S \rightarrow \mathbb{R}_+$ for all $j, t$.
- A set of demand functions $D^g_{jt} : S \rightarrow [0, M]$, for all consumer groups $g = 1, \ldots, G$.
- Expectations over the state transition $F^r, F^d$ for retailers and consumers, respectively. The actual transition process is denoted $F$. In rational expectations equilibrium $F^d = F^r = F$. 

14
The price policy functions satisfy optimality conditions derived from the value functions below and demand will satisfy the optimality conditions implied by (1). The beliefs $F$ are restricted to be rational in the sense that they are consistent with the distribution of prices and demand induced by the equilibrium behaviour of both consumers and retailers.

The retailers optimal price strategies will depend on the vertical contracts employed. Consumers expectations of future prices will in turn depend on the price strategies through $F$. The distributions $F$ capture that consumers dynamic purchase decisions may be affected by the price strategies in the industry through their impact on consumers expectations. For instance, consumers that know there is a time-limited

As an example, the conditional value functions from (2) in the demand specification are in the counterfactuals given as

$$v_{jt}(s_t) = \begin{cases} 
\beta \int \tilde{V}_{t+1}(p_{t+1}(s_{t+1}), s_{t+1})dF(s_{t+1}|s_t) & \text{for } j = 0 \\
u_j(p_t(s^f_t), s^f_t) & \text{for } j = 1, \ldots, J
\end{cases} \quad (4)$$

consistent with the policy functions $p(s)$. Attention is restricted to pure strategy Markov Perfect Equilibria (MPE) where the current state $s_t$ summarises all the pay-off relevant information for all agents (Maskin & Tirole (2001)). A Markov strategy depends only on the current state, regardless of the history that led to it. MPE allows for a parsimonious parametrisation and a relatively small state space, both of which are convenient features for both computation and estimation purposes.

I do not have existence theorems, but I find pure strategy MPE at the empirically relevant parameters. Multiplicity of equilibria is almost certain given the non-linearity of the model. Yet, preliminary sensitivity tests of the calculated equilibria to perturbations of the parameter space local to convergence do not reveal issues of discontinuous jumps in policy functions. I have found reaction curves to intersect locally at convergence. Further discussion on solution methods are relayed to Section (10) on results.

### 6.4 No coordination (NC)

In the baseline counterfactual, publishers send retailers books at the constant marginal cost of production $c$. The retailers then engage in dynamic price competition to extract surplus in the market. The profits and price paths of this counterfactual serves as a baseline case in that there is neither coordination of prices between retailers nor coordination of prices over time within the vertical unit. It is not entirely clear how to define a baseline in terms of the per-unit wholesale price. In an alternative benchmark, the publisher could charge a monopoly wholesale price. That would however raise the issue of double marginalisation. Though double marginalisation is an interesting problem in itself, it is also an issue about profit splits, which I deliberately abstract away from. The baseline is not meant to correspond to an observed policy, simply provide a benchmark to measure the returns to coordination against.

The retailers are assumed to maximise the discounted profit stream, taking into account the prices of its rivals, their own future price setting and the impact of their price setting on consumers expectations. The marginal cost of production is $c$, equal across retailers within the vertical unit.\(^{17}\) It is common knowledge that retailers set prices simultaneously at the start of every period, conditional on the state $s_t$.

\(^{17}\)The assumption is consistent with the Robinson-Patman Act that prohibits manufacturers to price discriminate retailers through wholesale prices.
The retailers discount profits by factor $\rho$, set to the inflation target of 2.5% in Norway at the time. The retailers are assumed to price according to state contingent strategy profiles $p_t : S \to \mathbb{R}_+^J$. A No-Coordination strategy profile $p^{\text{NC}} = \{p^{\text{NC}}_{jt} \}$ is an MPE if, for each $t = 1, \ldots, T$, each retailer $j$ prefers $p^{\text{NC}}_{jt}$ to any alternative Markov strategy $p_j$, conditional on all rival retailers $-j$ following $p^{\text{NC}}_t$. Then $p^{\text{NC}}$ is an MPE if
\[
\Pi_{jt}(s_t | p^{\text{NC}}) \geq \Pi_{jt}(s_t | p_{jt}^{\text{NC}})
\]
\[
= \pi_j(p_{jt}^{\text{NC}}, s_t) + \rho \int \Pi_{jt+1}(s_{t+1} | p^{\text{NC}}) dF^{r}(s_{t+1} | s_t, p_j, p^{\text{NC}}_{jt})
\]
s.t. (3) & $F^{r} = F$
for all $j = 1, \ldots, J$ for all $t = 1, \ldots, T$, for all $s \in S$.

### 6.5 Horizontal coordination (HC)

The vertically integrated unit can coordinate prices between retailers, but not over time. It improves on price discrimination over the baseline by eliminating wasteful competition between retailers. The integrated unit can not commit and is assumed to re-optimise prices in each period. The resulting price strategy is time-consistent. By time-consistency, the vertical unit has an incentive to self-enforce the agreement in any period. RPM is a sufficient restraint to implement HC.

The problem is equivalent to finding the time-consistent price path of a multi-product monopolist. Define the integrated per-period profits as $\pi(p_t, s_t) = \sum_{j=1}^{J} \pi_j(p_t, s_t)$. A strategy $p^{\text{HC}} = \{p^{\text{HC}}_1, \ldots, p^{\text{HC}}_J\}$ is an MPE if it solves
\[
\Pi_t(s_t) = \max_{p_t} \pi(p_t, s_t) + \rho \int \Pi_{t+1}(s_{t+1}) dF^{r}(s_{t+1} | s_t, p_t)
\]
s.t. (3) & $F^{r} = F$
for all $t = 1, \ldots, T$, for all $s \in S$. The resulting policy functions are denoted $p^{\text{HC}}_t(s_t)$.

### 6.6 Full commitment (FC)

In the final counterfactual, the vertical unit is endowed with commitment power. The vertical unit can with commitment achieve profits beyond what it achieves employing the strategies in (6). The vertical unit announces a price path in the introductory period. The resulting commitment price paths are time-inconsistent, they depend only on time, and not on the states of the market.\(^{18}\) Because of time-inconsistency, the vertical unit will generally have an incentive to deviate from optimal policy. The vertical unit needs a source of commitment to sustain prices along the optimal path. The source of commitment is assumed and not modelled directly.

The problem is now equivalent to an integrated monopolist who ex ante announces price paths for all products and can sustain pricing along the announced path. The commitment price path solves
\[
p^{\text{FC}} = \arg \max_{\{p_t\}_{t=1}^{T}} \sum_{j=1}^{J} \sum_{t=1}^{T} p^{t-1} \int \pi_{jt}(p^{\text{FC}}_t, s_t) dF(s_t | s_{t-1}, p_{t-1})
\]
s.t. (3) & $F^{r} = F$
\(^{18}\)A time-inconsistent price strategy need not be state-independent. In a more general setting, a retailer could for instance commit to a more sophisticated price strategy that depends on only a subset of the payoff relevant states. For instance, a retailer could commit to condition its price strategy only on demand shocks and not the demand composition. Motivated by the price strategies that were employed in the Norwegian book industry, I restrict attention to price strategies that depend only on time, and does not respond to the state of demand.
Since the commitment is credible by assumption, consumers will correctly expect prices to evolve along the announced path. The resulting strategy $p^{FC}$ coordinates pricing both across retailers and over time.

The counterfactual contract forms can be ranked in terms of joint profits. A horizontally coordinated unit $HC$ can do at least as well as an uncoordinated unit $NC$. A coordinated unit with commitment power $FC$ can do as least as well as a coordinated unit $HC$ without commitment power, since $FC$ can commit to any price path. By Assumption (1), the contract types are Pareto-dominated in the sense that for any contract of the form $NC$, there exists a contract of the form $HC$ that Pareto-dominates it. Similarly, for any contract of the form $HC$, there exists a Pareto dominating contract of the form $FC$.

6.7 Fixed price regime

In line with the Trade Agreement, the publisher sets a fixed price $\bar{p}$, constrained equal across retailers, for a restraint period of $\tau$ periods. Following the expiration of the restraint period, retailers set prices freely. The price regime combines the full commitment $FC$ strategy and the rivalrous retailer price strategies in $NC$.

The publisher set a fixed price $p^{FP}$ in the restraint period, taking into account the price strategies $p^{NC}$ that retailers will employ at the end of the expiration period.

$$p^{FP} = \arg\max_p \sum_{j=1}^J \sum_{t=1}^{\tau} \rho^{t-1} \int \pi_{jt}(p, s_t)dF(s_t|s_{t-1}, p) + \sum_{j=1}^J \sum_{t=\tau+1}^{T} \rho^{t-1} \Pi_{jt}^{NC}(s_t|p^{NC})$$

s.t. (3)

The retailers marginal cost, the wholesale prices, is set equal to the marginal cost of production.\(^{19}\)

7 Estimation

This section discusses the identification of the horizontal and the inter temporal substitution informally before turning to the estimation routine

7.1 Identification

The discount factor which in part determines the inter temporal substitution is usually not identified in dynamic models, i.e. there is not a unique vector of discount factor $\beta$ and utility parameters $\gamma$ that can rationalise the data. Estimation usually proceeds conditional on a discount factor that is fixed prior to estimation. As the value of commitment crucially depends on the discount factor, fixing it in some sense assumes the result. The price variation generated by the deregulation however allows the discount factor to be estimated.

The discount factor is identified by measuring how current demand responds to changes in future utility, holding the current utility fixed (Magnac & Thesmar (2002)). For a myopic consumer, only the current price matters, whereas a forward looking consumer responds to changes to future prices. The identifying experiment is arguably in the data. From Figure (3), prices in the year of publication are comparable both before and after the deregulation. Yet demand in the year of publication shifts towards future lower prices after the deregulation. The price variation generated by the deregulation is close to the identifying experiment and the variation that I will use to identify the discount factor.

Having been in place since the early 1960s, the retailers price strategies were widely known among the book buyers. Consumers could be reasonably sure there would be

\(^{19}\)The wholesale price may however be a choice variable of the publisher too. The joint choice of a fixed price and subsequent wholesale prices is not considered here.
no discounts on the introductory price until the expiration of the restraint period. These expectations are likely to change after the deregulation as price paths are seen to substantially change. The otherwise unobserved expectations of consumers before the deregulation $F^b$ are therefore approximated by the price path traced out by the Agreement itself. The expectations $F^a$ held after the deregulation are estimated well from the data in Section (8).

One issue is that the deregulation did not come as a complete surprise in early 2005. Over the course of fall 2004, it became increasingly clear that the old Agreement would be changed in 2005. The impact of the deregulation on consumers price expectations may therefore already have been incorporated in demand for new titles 2004. Fall is moreover the peak demand season in the industry, which is a reason of some concern. I therefore discard the 2004 cohort from the discount factor estimation. The residual demand for the 2003 cohort left in 2004 is however not subject to the same issue. Firstly, there was hardly any public debate in 2003 about deregulation. Furthermore, the 2003 cohort was still sold under the old agreement in 2004, i.e. the coming deregulation would have no effect on the 2003 cohort in 2004.

The standard endogeneity problem in demand estimation arises if retailers rather than randomly vary the prices, set them in response to changes in demand. There is then something observable to consumers and to retailers that affects both the demand and the price incentives that is unobserved in the data. The usual solution is to find instrument variables that are correlated with retailers price incentives, but do not directly affect demand itself. Standard sources of instruments include marginal cost shifters (Working (1927)), variation in the density of the product space (BLP (1995)) and geographical price variation (Hausman (1996), Nevo (2001)).20 None of standard sources are readily available in the current 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. Since most regular instruments are likely weak or unavailable, prices are led to instrument for themselves.

To ameliorate the endogeneity issue in the absence of instruments, I use a data-on-unobservables approach. The idea is that if there is something observable to both retailers and consumers that affect both pricing and the demand, then there might exist data that can account for the effects. Editorial book reviews are used. A total of 1823 reviews were collected for the period 2002 to 2007 from the three largest national circulation newspapers (Aftenposten, Dagbladet and VG), in print and on the web.21 About a third of the text reviews were editorially scored by the newspaper with a grade ranging from 1 to 6 in addition to the text review. The rest were editorially unscored. Just the fact that a title is reviewed is positively predictive of sales, but higher scores also predict higher sales, as expected.22 To extract more information from the unscored reviews, a simple supervised machine learning algorithm was applied.

The scored reviews were used as a training set to build an algorithm that could predict the unscored reviews. The review samples are very small for text analysis, so a premium was put on simple algorithms and crude scoring that robustly capture the important feature. Unscored reviews were predicted as either positive or negative. A $k$ Nearest Neighbours-approach with simple majority voting was used, with $k$ set somewhat arbitrarily to 5.23 Based on the frequency of indicative words in a (stemmed) unscored review, its five closest reviews from the training set in terms of number of shared indicative words were found. Then the vote of the five

---

20 In its simplest form, the geographical instruments use that whereas demand shocks may be local, cost shocks may be national. The idea is then to decompose the price variation in a national component and orthogonal local variation. The shared national variation is used as to instrument for prices.

21 The reviews were accessed through Retriever, a comprehensive and proprietary Scandinavian media archive.

22 See the Appendix for results.

nearest neighbours was cast, if three or more of the neighbours were positive, the review was classified as positive. The scored reviews were converted to the same scale. The algorithm was validated on a set of titles with both editorially scored and unscored reviews.

I use the approach in Nair (2007) to infer the market size on title level from cumulative sales by appealing to the Bass (1969) model of diffusion. The model specifies the market size as a flexible function of the cumulative sales of each title. The market shares are then constructed from observed sales and the estimated market size, see the Appendix for details.

### 7.2 Data and specification

To avoid issues with observing market shares of exactly zero, only titles with strictly positive sales at all retailers for consecutive periods since introduction were used for the estimation of \( \gamma \) parameters. Two specifications are estimated, one with fixed effects and unobserved demand shocks and one restricted without. Estimates from the restricted version is presented below, estimation of the fixed effects model is ongoing.

There are two consumer types \( g \in \{ h, l \} \), high and low. The deterministic state variables are \( z_{jkt} = \{ x_k, n_t, ssn_t, rvw_t \} \) where the time invariant title characteristics are \( x_k \), characteristics such as weight, genre, number of pages. Following Einav (2007) and Ho et. al. (2012), the utility has component \( \gamma^g n_t \) representing a taste for novelty where \( n_t \) is a function of time. Fixed effects for seasonal variation are \( ssn_t \) and the reviews are represented by \( rvw_t \). Retailer fixed effects are \( \gamma^g j \). The utility of consumer \( i \) from purchasing a title \( k \) from retailer \( j \) in period \( t \) is given as

\[
\begin{align*}
  u^g_{ijkt} &= \delta_k + x_k \delta_x + \sigma \xi_{jkt} + \gamma^g_j + \gamma^g_p p_{jkt} + ssn_t \gamma^g_ssn + \gamma^g_n n_t + rvw_t \gamma^g_rvw + \epsilon_{ijt} \\
  &= \delta_{jkt} + X_{jkt} \gamma + \epsilon_{ijkt}
\end{align*}
\]

The parameter vector is split in \( \delta, \gamma \), where parameters \( \delta \) are shared between groups while \( \gamma^g \) are group specific.

### 7.3 Estimation routine

The two specifications are both estimated sequentially. First the consumers expectations \( F_t \) are estimated in reduced form described below. The estimated transition process then enters the demand specification. The estimation routine is described in reverse order, this section describes the estimation of the demand function, the next section describes the estimation of the transition process. I first describe the estimation of the utility parameters \( \delta, \gamma \), which is fairly standard, and then outline the estimation approach to the discount factor, which is somewhat non-standard.

The structural parameters \( \beta, \delta, \gamma \) are estimated using GMM in both specifications. In the restricted version, parameters \( \sigma \xi = 0, \delta = 0 \) for all \( k \) are imposed. The observed demand is matched to the predicted demand \( m^t(\gamma) = D - D(\gamma) \). Without title fixed effects, the time invariant title characteristics \( x_k \) control for heterogeneity at the title level.\(^{24}\) Additional moments for the discount factor, these are described below. Currently, the discount factor is fixed at \( \beta = 0.7 \) on a yearly level. Both estimators are implemented in a two-step routine where the optimal weighting matrix \( W \) is estimated in the first step. The restricted version conditional on \( \beta \) is essentially Feasible Generalized Non-Linear Least Squares.

The BLP routine is used to estimate the fixed effects model. Though there are no exclusion restrictions, the BLP mean utility inversion is used as a computationally

\(^{24}\)Software disclosure: Both the estimation and the counterfactuals use SNOPT and automatic differentiation extensively. A special thanks to Mike Saunders for making the software available for free and similar thanks to Anders Goran at TomLab for at request adding differentiation modules to MAD.
The moments for the discount factor will be added to following Dube, Fox & Su (2012), with tight tolerances. It is computationally costly in terms of CPU-time, but convenient in terms of coding.

The routine proceeds using the inner/outer loop routine of BLP (1995). Conditional on the \( \gamma \) parameters, the shared component \( \delta_{jkt} \) is inverted out from the restriction

\[
\delta_{jkt} : D_{jkt} = D_{jkt}(\delta_{jkt}, \gamma)
\]

where \( D_{jkt} \) is the observed demand and \( D_{jkt}(\delta_{jkt}, \gamma) \) is the predicted demand. The parameters \( \alpha \) of the shared utility component \( \delta \) are concentrated out by a linear regression of \( \delta_{jkt} = \alpha_k + x_k \alpha_x + \xi_{jkt} \), where \( \xi_{jkt} \) is assumed to be serially uncorrelated and mean independent of \( x_k \). The inversion itself is implemented by the BLP contraction mapping

\[
\delta^{it+1} = \delta^{it} + \ln(D_{jkt}) - \ln(D_{jkt}(\delta^{it}, \gamma))
\]

The contraction is run title-by-title until convergence.

In the outer loop, the residuals \( \xi_{jkt} \) are interacted with the remaining covariates \( X \) under the mean independence assumption \( E[\xi|X] = 0 \). The parameters \( \gamma \) are found in an outer loop search as

\[
\hat{\gamma}^{f} = \arg \min_{\gamma} m^{f}(\gamma)Wm^{f}(\gamma)'
\]

where \( m^{f}(\gamma) = \xi(\gamma)'X \), and \( W \) is the optimal weighting matrix.

The empirical strategy for the estimation of the discount factor is less standard. Denote the observed aggregate demand before and after the deregulation \( D^b_t, D^a_t \), respectively. Let the history of prices at time \( t \) be \( h_t = \{p_1, \ldots, p_t\} \). The expectations are still represented by \( F \) and the initial distribution of consumer types \( w \). Changing the notation somewhat, denote the aggregate demand functions \( D_t(\beta, h_t, w, p_t, F) \). The demand composition \( D_t(.) \) at time \( t \) depends on the price history \( h_t \) and the initial distribution \( w \), the current prices \( p_t \), the consumers expectations of future prices \( F \), and the factor \( \beta \) by which they discount future utility.

Following the discussion above, \( \beta \) is estimated by exogenous variation in consumers expectations \( F^b, F^a \) induced by the change of price strategies following the deregulation. The discount factor is estimated off moments that match the current theoretical demand to the observed demand

\[
m_\beta = (D_t(\beta, h^a_t, p_t, F^b, w) - D^b_t, D_t(\beta, h^a_t, p_t, F^a, w) - D^a_t)
\]

are added to (7). The implementation of the estimator faces some challenges in constructing the demand shares before the deregulation. The demand estimation strategy follows titles over a three year lifecycle, yet there is no full lifecycle of data before the deregulation. The first challenge is then to construct both observed and predicted market shares \( D_t(\beta, h^o_t, w, p_t, F^o) \) for the contingency where there is not complete lifecycle data. The Trade Agreement can however be used to reconstruct the historical price paths \( h_t \) since the retail prices were regulated by RPM and there is data on the introductory prices. We also need an estimate of the distribution of expectations.

25 It is computationally costly in terms of CPU-time, but convenient in terms of coding.

26 Following Dube, Fox & Su (2012), with tight tolerances.

27 The moments for the discount factor will be added to \( m^{f}(\gamma) \).
The history of the demand shocks $\xi_{jkt}$ from before 2004 can not be inverted out from a moment restriction as there is no data. The price history is necessary to determine the composition of demand in any period. The demand shocks are therefore set to zero in periods of unobserved demand. The market size for the titles prior to the deregulation is is predicted with diffusion formula, see the Appendix for details. The predicted market shares in unobserved periods are forwardly simulated. Though the estimator is not fully implemented, preliminary tests finds $\beta$ in the area of 0.6 on a yearly level, about ten percentage points lower than Dube et al (2012).

8 State transitions and expectations

The empirical specification of the state transition process is a simple linear-in-parameters Markov process with additive shocks. The specification reflects that consumers may form beliefs about the future prices taking into account the current prices at all retailers and factors such as what phase of the lifecycle the title is in. The specification allows consumers to form expectations from observing variation in price paths across titles, and allow expectations to vary with retailers and different genres of books.

Let the consumer states be partitioned in the stochastic and deterministic processes $s_t = (p_t, z_t)'$, respectively.

$$s_{t+1} = \Theta s_t + \eta_t$$

where $\sigma_{\eta} = 0$, by definition of deterministic $z$. Each retailer is allowed individual coefficients $\Theta_j$. The parameters of the specification can be partitioned

$$\Theta = \begin{bmatrix} \theta_{pp} & \theta_{pz} \\ 0 & \theta_{zz} \end{bmatrix}$$

where the lower left zero matrix follows from the exogeneity of $z$. By forward iteration, we can then write

$$s_{t+r} = \Theta^r s_t + \sum_{\tau=1}^{r} \Theta^{r-\tau} \eta$$

In line with the Markov Perfect Equilibrium concept used in the counterfactuals, the current state $s_t$ is assumed to summarise all the payoff relevant information to the consumer and hence doubles as the information set.

The shocks $\eta$ are assumed mean zero multivariate normal and serially uncorrelated

$$E[\eta_t|s_t] = 0 \text{ for all } t \tag{8}$$

$$E[\eta_t \eta'_t|s_t] = \Sigma \text{ for all } t \tag{9}$$

$$E[\eta_t \eta_{t+r}|s_t] = 0 \text{ for all } t \neq r \tag{10}$$

The assumptions on $\eta$ jointly define a martingale difference sequence adapted to the information set $s_t$. The expectations conditional on $s_t$ are now simple functions of $s_t$ itself.

$$E[s_{t+r}|s_t] = \Theta^r s_t + \sum_{\tau=1}^{r} \Theta^{r-\tau} E[\eta_{t+r}|s_t]$$

$$= \Theta^r s_t$$

\footnote{This is not contradicting the Markov assumptions of the counterfactuals and the demand specification.}
since $E[\eta_{t+r}|s_t] = 0$ for $\tau \geq 1$. Under assumptions (8)-(10), the second moment is

$$V[s_{t+r}|s_t] = V \left[ \sum_{\tau=1}^{r} \Theta^{r-\tau}\eta_{t+r}|s_t \right]$$

$$= \sum_{\tau=1}^{r} \Theta^{r-\tau} \Sigma \Theta^{r-\tau}$$

The specification is estimated by a FGLS for each group. The process is stable at $\hat{\Theta}$ with the characteristic roots of the price parameters all being of modulus less than one. The model fits quite well with $R^2 = 0.96$. See the Appendix for further details.

Following the arguments in Skrainka & Judd (2011), quadrature is used to numerically integrate out the expectations in the demand functions. Normality of $\eta$ motivates the use of multivariate Hermitian quadrature, see Judd (1998) for an exposition. Normality along with rational expectations also implies that the two first moments of the transition process completely describe the expectations $F$.

The option value integrals

$$v_{0t-1}(s_{t-1}) = \int V_t(s_t)dF(s_t|s_{t-1})$$

are solved with tensor product bases with five nodes in each dimension.

9 Estimation results

The discount factor set to 0.7 at yearly level, close to the values found in Dube et al (2012) using experimental variation for consumer electronics. Parameter estimates from the restricted version without demand shocks nor title fixed effects are given in Table (??). The specification is restricted in that it does not accommodate title fixed effects and the discount factor estimator is not yet implemented.

Some patterns emerge

- The high type consumer is less price sensitive than the low type, the price coefficient of the low type is about double the size in absolute value. The price coefficients are not estimated with precision, suggesting substantial heterogeneity.
- High types perceive retailers as differentiated, low types distinguish less between retailers.
- Reviews have a positive effect for high types, low types attach a negative significance to reviews.
- High types prefer to purchase in fall, low types in spring. Since spring is the season of the clearance sale, the seasonal fixed effects may absorb some of the price variation at the expense of the price coefficients.
- Time invariant characteristics weight and pages do not matter much.
- High types have a taste for novelty, low types do not. The taste for novelty of the high type seems too high. In money terms, it implies that the utility to the high type on average decreases by about 3 times the average price of a book. The parameter may pick up the variation of some outlier titles with a particularly high demand.
- There is a slight majority of high types at about $\frac{3}{5}$ths of the market.

The relatively small effects of time invariant characteristics such as pages and weight suggest that the there is substantial variation left for the title fixed effects.

The implied retailer own price elasticity at the mean prices and characteristics are plotted in Figure (5). The own price elasticities are seen to decline over the lifecycle, following the price levels and shifting the mass of demand from the high

---

29 The discount factor used is slightly higher than preliminary estimation results suggest.
types to the low types at lower prices. The net effect is declining price elasticities. The cross price of one retailer elasticity is plotted in Figure (6), the others are similar. The substitution between retailers is increasing over time.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coeff</th>
<th>Std. Err</th>
<th>Coeff</th>
<th>Std. Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>-0.540</td>
<td>0.398</td>
<td>1.041</td>
<td>0.598</td>
</tr>
<tr>
<td>season</td>
<td>0.557</td>
<td>1.246</td>
<td>0.981</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>1.086</td>
<td>1.149</td>
<td>0.083</td>
<td>1.406</td>
</tr>
<tr>
<td></td>
<td>2.432</td>
<td>0.805</td>
<td>-2.869</td>
<td>0.523</td>
</tr>
<tr>
<td>retailers</td>
<td>1.851</td>
<td>0.690</td>
<td>0.046</td>
<td>0.406</td>
</tr>
<tr>
<td></td>
<td>1.469</td>
<td>1.258</td>
<td>0.034</td>
<td>0.931</td>
</tr>
<tr>
<td></td>
<td>1.077</td>
<td>0.232</td>
<td>0.096</td>
<td>0.179</td>
</tr>
<tr>
<td>genre</td>
<td>0.011</td>
<td>0.273</td>
<td>-1.757</td>
<td>0.273</td>
</tr>
<tr>
<td>review</td>
<td>1.301</td>
<td>0.470</td>
<td>-0.526</td>
<td>0.341</td>
</tr>
<tr>
<td>w_{hig}</td>
<td>0.730</td>
<td>0.274</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>novelty</td>
<td>0.970</td>
<td>0.130</td>
<td>-0.008</td>
<td>0.197</td>
</tr>
<tr>
<td>\beta</td>
<td>0.73</td>
<td>-</td>
<td>0.73</td>
<td>-</td>
</tr>
<tr>
<td>weight</td>
<td>0.011</td>
<td>0.109</td>
<td>0.984</td>
<td>0.338</td>
</tr>
<tr>
<td>pages</td>
<td>0.920</td>
<td>0.040</td>
<td>0.070</td>
<td>0.100</td>
</tr>
<tr>
<td>N</td>
<td>12830</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The substitution patterns are somewhat restricted by the structure of the model. For instance, the own-price elasticities have the well known logit property that they are decreasing in prices and increasing in own shares. The variation in own-price elasticities over time follow from changes in prices levels and changes to the demand composition over time. Similarly, the cross price elasticities changes over time as the composition of demand shifts towards higher shares of low types.

10 Counterfactual results

The counterfactuals are calculated at the preliminary estimates and at the mean of the covariates in the sample. For each of the counterfactual contracts, the price strategies described in Section (6) are calculated along with the corresponding demand and profits. The marginal cost is constant over time and set to 0.5, in line with industry estimates. The HC counterfactual is not yet calculated. This counterfactual serve as the reference level for the estimated value of commitment provided by RPM. I therefore can not currently provide a point estimate of the key concept of the paper. I can however bound the estimate, an argument follows below. Since the parameter estimates are preliminary and based on a restricted specification, the counterfactuals serve mainly as an exposition. Some patterns are however likely to persist in future revisions. One of the results that stand out is that the time limited fixed price regime employed by the industry achieves profits close to the optimal price path. The next pattern is that there are substantial profits to be made from both horizontal and inter temporal substitution.

The counterfactuals are summarised in four plots that display the price paths, the demand paths of the high and the low type and the per-period profits. In the first comparison, the FC contracts are compared to the baseline NC without neither horizontal nor inter temporal coordination. The results are given in Figure (7). The full commitment contracts results in higher overall price levels. It is also clear that price discrimination improves with coordination. The prices travel a greater distance over the lifecycle with under FC, implying improved price discrimination. There is also more rationing in FC. The prices in the terminal period are higher with in FC than in NC, so less demand is served over the lifetime.

The FC price paths improve substantially over the NC contracts. Relative to the baseline, the FC contracts achieve an increase in profits of about 18%.

To separate out the profit increment that is due to the inter temporal coordination,
it must be compared to the returns of the $VI$ contract. That counterfactual has not yet been calculated. Meanwhile, 18% represents an upper bound of the intertemporal coordination effect. It would correspond to assigning the full coordination improvement over the baseline to intertemporal coordination, and none to horizontal.

In Figure (8), the $FP$ strategy that constrain the retailer prices equal in the first six periods is compared to the $FC$ path that allows reoptimization in each period. The profits of the fixed price strategy is within 1 percent of the fully optimal $FC$ price path. As a heuristic rule, it performs surprisingly well. The profits and demand patterns are seen to be very similar over the lifecycle. The $FP$ policy also has the virtue of being simpler to communicate than the $FC$ price path that varies for each period with seasonal variation, which may rationalise its existence as a par ice discrimination device.

10.1 Caveats
An informal test of fit is to see how well the counterfactual model predicts introductory prices under the Fixed Price counterfactual that mimics the actual price policy. The counterfactual fixed price model currently predicts too high introductory prices. The mean introductory price in the (scaled and displaced) data is about 2.10, whereas the predicted price is about 50% higher. Furthermore, the discount at the end of the expiration period in the data is about 45% on average, whereas the the model implies about 20%. These facts together suggest that the
current estimates identify too little heterogeneity between types and too high val-
uations all over.

There are other reasons why the model produce relatively small effects of com-
mitment. The value of commitment in the discrete time models depend crucially
on the periodisation.\footnote{As shown in Stokey (1981), as the length of the period goes to zero, the price of the monopolist goes to cost.} In the model, the periodisation follows the data sampling intervals of four months. The periodisation implies a piecewise commitment. Each retailer by assumption is in the model committed to its price for a period of four months. The periodisation is an artefact of the data sampling and does not cor-
respond to any data on how retailers actually price. In reality, the retailers can
change prices more frequently. In summary, the model therefore likely overstates
the profits of a no-commitment strategy. There is a limit to how much can be done
to ameliorate the problem without data on the actual frequency of price changes.
One simple robustness check is to carve up each four month period in shorter
sub periods and allow the retailers to change prices in every sub-period. This ro-
 bustness check has not been implemented, but the impact is likely to be substantial.

The treatment of uncertainty has important implications for the profitability
of commitment strategies. In the current version, I am abstracting away from id-
iosyncratic demand shocks $\xi$ both in estimation and in the counterfactuals. Ex-
cluding demand shocks inflates the profitability of the commitment strategy over
a time-consistent strategy. It does so since one advantage of the time-consistent
policy is that it can adapt to demand shocks, whereas a fixed price policy can
not. By excluding demand shocks, the potentially profitable flexibility of the time

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{cross elasticity}
\caption{Cross price elasticities.}
\end{figure}
consistent policy is assumed away.

11 Discussion

TBA

12 Summary

This paper studies the role of RPM as part of a price skimming strategy. It starts with the observation that RPM fixes the price path in dynamic markets. By fixing the price path, RPM can act as a commitment device. Comprehensive retail data from a natural experiment in a book market where RPM was deregulated shows a clear impact of RPM on price strategies and demand. Following the deregulation, prices fall faster and demand shifts from early in the lifecycle at high prices to later at lower prices. RPM may however have multiple effects in an oligopolistic retail market. Fixing the price paths also eliminates horizontal competition between retailers. To separate out the impact of RPM on inter temporal price discrimination, I turn to evaluation of counterfactual vertical contracts. Estimates from a dynamic demand model are embedded in a dynamic supply side model that allows for various levels of price coordination, both across retailers and over time. Preliminary results find that the RPM employed in the Norwegian book industry is close to exhausting the scope of inter temporal price discrimination in the market. The counterfactuals calculations are so far incomplete, but an upper bound of 18% is derived at the current estimates. The results in this paper ties into a
policy discussion on the competitive effects of RPM. The results are of managerial relevance by pointing out an effect of vertical restraints that can improve on price skimming strategies.

13 References


Shapiro (1996): "Mergers with Differentiated Products, Antitrust, Spring


A  Calculating market shares

To calculate the market size of title \( k \), run the regression

\[
q_{kt} = a_k + bQ_{kt} + cQ_{kt}^2 + \epsilon_{kt}
\]

with \( q \) aggregated across retailers for each period. The market size is then \( \hat{M}_k = \frac{\hat{a}_k}{x_k} \), where \( x_k \) is the positive root of the quadratic equation \( x_k^2 + bx_k + ac = 0 \).

A.1  Changes in release dates

The release dates are from the Bokbasen database. Bokbasen is a company that supplies the industry with information on titles and is used for logistical purposes across retailers and publishers. The release dates are recorded as the time the publishers register a title in the logistical database. There are known to be some error in the release date data, but data as is show no discernible change in release dates before and after the deregulation. Taking the list of titles and regressing the tertile of release on a dummy for the deregulation, the results are given in Table (??). There are hardly changes in the release dates. The changes in release dates are measured in shares of a tertial and are close to zero. The median title is released in the second tertial, both before and after the deregulation.

Table 3: Change in introductory price levels before and after the deregulation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>post</td>
<td>0.003</td>
<td>0.014</td>
<td>-0.003</td>
<td>0.015</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.063</td>
<td>0.013</td>
<td>2.082</td>
<td>0.036</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>16605</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B  Estimation of the state transitions

The state transition parameters of the price process \( \Theta_p = (\theta_{pp}, \theta_{pz}) \) and \( \Sigma \) are assumed known by the consumers, but unknown to the econometrician and is estimated from the data. The parameters of the deterministic process \( \theta_{zz} \) are known. The transition process is equivalent to SUR specification with laged prices and is estimated by a standard Feasible Generalized Least Squares approach.

Stacking the price equations for each title \( l \), we get

\[
\begin{pmatrix}
  p_1 \\
  \vdots \\
  p_L
\end{pmatrix}
= \begin{pmatrix}
  \theta_{pp}p_1 + \theta_{pz}z_1 \\
  \vdots \\
  \theta_{pp}p_L + \theta_{pz}z_L
\end{pmatrix}
+ \begin{pmatrix}
  \eta_1 \\
  \vdots \\
  \eta_L
\end{pmatrix}
\]

(11)

where \( lp \) are laged prices. We now have standard linear system of equations that by reorganising the parameters \( \Theta \) can be written as

\[
p = S\Theta_{FGLS} + \eta
\]

(12)

The model is estimated with a two-step procedure. In the first stage, OLS is run on (12). The covariance \( \hat{\Omega} \) is estimated as

\[
\hat{\Omega} = I_{LT} \otimes \hat{\Sigma}
\]

where \( \hat{\Sigma} = \frac{1}{T} \hat{\eta}'\hat{\eta} \) is a consistent estimator of covariance from the first stage residuals of OLS. The FGLS estimator is now the familiar

\[
\Theta_{FGLS} = \left( S'\hat{\Omega}^{-1}S \right)^{-1} S'\hat{\Omega}^{-1}p
\]

which completes the estimation of the price transition process. Note that the transition process is estimated independently of the demand system. By standard
arguments, the estimated errors \( \hat{\eta} \) are asymptotically normal with \( \hat{\Sigma} \) a consistent estimator of the covariance.

The estimated moments are given in Table (4) and (5).

<table>
<thead>
<tr>
<th>Table 4: Expectations regressions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( )</td>
</tr>
<tr>
<td>( lp_{11} )</td>
</tr>
<tr>
<td>( lp_{12} )</td>
</tr>
<tr>
<td>( lp_{13} )</td>
</tr>
<tr>
<td>( lp_{21} )</td>
</tr>
<tr>
<td>( lp_{22} )</td>
</tr>
<tr>
<td>( lp_{23} )</td>
</tr>
<tr>
<td>( lp_{31} )</td>
</tr>
<tr>
<td>( lp_{32} )</td>
</tr>
<tr>
<td>( lp_{33} )</td>
</tr>
</tbody>
</table>

| \( \) | \( \) | \( \) |
|-----------------------------------|
| time ffx | yes | yes |
| \( R^2 \) | 0.964 | 0.9565 |
| N | 8730 | 4491 |

The eigenvalues of the matrix of the price coefficients having eigenvalues 0.02, 0.04, 0.09, each well below 1, for the Fiction titles, the Non-Fiction eigenvalues are similar. The estimated coefficients show substantial dynamics and correlations across retailers and over time within titles. The second moments \( \Sigma \) are given in Table (5) in correlation form. After controlling for period fixed effects and lagged prices, there are still substantive positive correlations in pricing across retailers. A multivariate portmanteau test reveal mild signs of serial correlation in \( \eta \).

<table>
<thead>
<tr>
<th>Table 5: ( \eta ) correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
</tr>
<tr>
<td>( p_1 )</td>
</tr>
<tr>
<td>( p_2 )</td>
</tr>
<tr>
<td>( p_3 )</td>
</tr>
</tbody>
</table>

C Review text analysis

A total of 1823 reviews were crawled for the period 2002 to 2007 from three national circulation newspapers, Aftenposten, Dagbladet and VG, in print and on the web. The reviews were accessed through Retriever data base, a comprehensive Scandinavian media archive. About a third of the text reviews were editorially scored, the rest were unscored. Scored reviews are reviews that were given a score from 1 to 6 editorially by the newspaper. Just the fact that a title is reviewed is positively predictive of sales, yet there is additional information in the score itself. A very simple supervised machine learning algorithm was therefore applied to score also the unscored reviews. To get a feel for the impact of reviews on the demand patterns, I run some linear regressions of quantity on price with and without reviews. About 5% of the titles in the 2004 cohort are reviewed. I then run a simple linear regression of quantity on prices with and without reviews. Reviews are seen to be correlated with both prices and quantities. The bias from omitting review is towards zero as expected. The impact of reviews are about two thirds of the mean sales per observation, which seems economically significant. These regressions make no causal claims. Though reviews could certainly spur the sale of books, already well selling books may also be more likely to get reviewed. The
role of the review variable in the analysis is simply to proxy for unobservable perceptions of quality that might affect both demand and retailer pricing.

In the columns to the right in (7), period and genre fixed effects are included. The period fixed effects has as expected a substantial impact on the price coefficients as it controls for the declining price path over the lifecycle.

Table 7: Reduced form text analysis with and without fixed effects

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>123.622</td>
<td>6.987</td>
<td>121.211</td>
<td>7.001</td>
<td>198.348</td>
<td>25.354</td>
<td>280.898</td>
</tr>
<tr>
<td>p</td>
<td>-0.774</td>
<td>3.243</td>
<td>-1.489</td>
<td>3.245</td>
<td>-18.523</td>
<td>3.748</td>
<td>-19.278</td>
</tr>
<tr>
<td>review</td>
<td>-</td>
<td>-</td>
<td>81.118</td>
<td>16.229</td>
<td>-</td>
<td>-</td>
<td>80.517</td>
</tr>
</tbody>
</table>

C.1 Validation of the scoring algorithm

The algorithm was validated on a subsample of 82 titles which had at least one scored and one unscored review. Based on the results for the validation set, the algorithm display reasonable performance in adding information about the score. The odds ratio is $\frac{5}{4}$ for negative reviews and $\frac{4}{3}$ for the positive reviews. It is a limit to how well it can do with such a minimum of text that is furthermore typically complex. The score variable however picks up some variation beyond the binary information in being reviewed that has the expected sign and can proxy for unobserved demand shocks.

C.2 Reduced form evidence across genres and over years

Figure (9) shows that the price paths are fairly similar in all years following the deregulation. Prices start declining earlier than with RPM and fall to about the same level, which is indicative of a stable change in price strategies. Similarly,

Figure 9: Lifecycle price paths year by year.
to be stable across years after the deregulation. For all the years following the deregulation, demand shifts towards lower future prices.

![Figure 10: Lifecycle demand year by year.](image)

The price patterns are similar also across genres. Figures (11) and (12) show the price patterns broken down on Fiction and Non-Fiction, respectively. The price pattern is seen to be qualitatively similar.

![Figure 11: Lifecycle price paths Fiction.](image)

### D Demand elasticities and derivatives

The demand derivatives of the inside goods are

\[
\frac{\partial D_{jt}}{\partial p_{kt}} = \begin{cases} 
\gamma_p \left( \sum_g w^g D^g_{jt} \left(1 - D^g_{jt}\right) \right) & \text{if } j = k, j, k = 1, \ldots, J \\
-\gamma_p \left( \sum_g w^g D^g_{jt} D_{kt} \right) & \text{if } j \neq k, j, k = 1, \ldots, J 
\end{cases}
\]

The elasticities are

\[
\eta_{jt} = \frac{\partial D_{jt} \frac{p_{kt}}{D_{jt}}}{\partial p_{kt} D_{jt}}.
\]

The inter temporal elasticities are meant to measure the sensitivity of demand to a change in the expectation of future prices. It is not entirely clear what should be understood by a small change in expectations since the expectations are represented by a distribution \(F\). It is taken to be a change in the first moment.
Figure 12: Lifecycle price paths Non-Fiction.

\[ E[p_{kt+1}|s_t] = \int p_{kt+1} dF(p_{kt+1}|s_t) \] for each group \( g \). The derivatives of the inside goods are then

\[ \frac{\partial D^g_{jt}(s_t)}{\partial E p_{kt+1}} = -\gamma p D^g_{jt} D^g_{jt+1} \left( \ln \left( \sum_{k=0}^{J} \exp(v^g_{kt+1}(s_{t+1})) + \Gamma \right) \right) \frac{\partial f(s_{t+1}|s_t)}{\partial E p_{kt+1}} d_{st+1} \]

for \( j, k = 1, \ldots, J \) and for each group \( g \) individually. The inter temporal elasticities at the market level are the weighted sum of each consumer groups elasticities

\[ E_l E p_{kt+1}, D_{jt}(s_t) = -\gamma p D^g_{jt+1} D^g_{jt} \left( \ln \left( \sum_{k=0}^{J} \exp(v^g_{kt+1}(s_{t+1})) + \Gamma \right) \right) \frac{\partial f(s_{t+1}|s_t)}{\partial E p_{kt+1}} d_{st+1} \]

with weights \( \omega^g_{jt} = \frac{D^g_{jt}}{\sum_g D^g_{jt}} \).

### D.1 Diversion ratios

The diversion ratio \( dr_{jk} \) measures the share of demand of retailer \( j \) that following a price increase of product \( j \) is diverted to retailer \( k \). It is closely related to the cross price elasticity and a common measure of the incentive to coordinate prices between firms in merger analysis (Shapiro (1996), Farrell & Shapiro (2010)). The higher the diversion ratio, the stronger the incentive to coordinate the pricing the products. The diversion ratios for the inside products are

\[ dr_{jkt} = \frac{\partial D_{kt}}{\partial p_{jt}} \left( \frac{\partial D_{jt}}{\partial p_{jt}} \right)^{-1} \]

for \( j = 1, \ldots, J \). The aggregate diversion ratio is \( adr_{kt} = \sum_{j \neq k} dr_{jk} \) and measures the share of demand lost on product \( k \) diverted to all other products in the portfolio.

A similar construct is defined to measure the returns to coordination of prices over time, which is what commitment is about. The inter temporal substitution is measured in terms of substitution from \( D_{kt} \) following a change in \( Ep_{jt+1} \).

\[ dr_{0kt} = \frac{\partial D_{kt}}{\partial E p_{jt+1}} \left( \frac{\partial D_{ot}}{\partial E p_{jt+1}} \right)^{-1} \]
for $k = 1, \ldots, J$ and $t = 1, \ldots, T - 1$

### D.2 Derivatives of the option value

The derivatives of the option value wrt. expectations are used extensively in the counterfactuals. The expectations enter through the option value

$$v_{0t} = \beta \int \left( \ln \left( \sum_{l=0}^{J} \exp(v_{l+1}) \right) + \Gamma \right) dF(s_{t+1}|s_t)$$

with derivatives $\frac{\partial v_{0t}}{\partial \mu_{k,t+1}}$ that are evaluated numerically using the chain rule and quadrature.

A special case of interest is perfect foresight when $\Sigma \rightarrow 0$. The derivative simplifies considerably so

$$\frac{\partial v_{0t}}{\partial \mu_{k,t+1}} = \beta D_{k,t+1} \frac{\partial v_{k,t+1}}{\partial p_{k,t+1}}$$

that is easy to evaluate. An intermediate case where only $\sigma_k \rightarrow 0$ individually is

$$\lim_{\sigma_k \rightarrow 0} \frac{\partial v_{0t}(s_t)}{\partial \mu_{k,t+1}} = \int \frac{\partial g(s_1, \ldots, s_k-1, \mu_k, s_{k+1:t+1}, \ldots)}{\partial s_{k,t+1}} f(s_{-k,t+1}|s_t) ds_{-k,t+1}$$

which simplifies calculation significantly.

A sketch of a proof is provided. The Markovian transition process $F(s_{t+1}|s_t) = MVN(\mu, \Sigma)$. We want to find the limit of the derivative of the continuation value $v_{0t}$ with respect to $\mu_j, t + 1$ as $\sigma_j \rightarrow 0$. Set $j = 1$ for notational convenience. The continuation value is

$$v_{0t}(s_t) = \beta \int \left( \ln \left( \sum_{l=0}^{J} \exp(v_{l+1}(s_{t+1})) \right) + \Gamma \right) dF(s_{t+1}|s_t)$$

$$= \int g(s_{t+1}) f(s_{t+1}|s_t) ds_{t+1}$$

where $F(s_{t+1}|s_t)$ is $MVN(\mu, \Sigma)$. So the desired result is

$$\lim_{\sigma_j \rightarrow 0} \frac{\partial v_{0t}(s_t)}{\partial \mu_{1,t+1}} = \int \frac{\partial g(\mu_1, s_{-1:t+1})}{\partial s_{-1,t+1}} f(s_{-1:t+1}|s_t) ds_{-1,t+1}$$

To see that, suppose for now $F$ is univariate $N(\mu, \sigma)$. Then

$$\frac{\partial}{\partial \mu_{t+1}} \int g(s_{t+1}) f(s_{t+1}|s_t) ds_{t+1} = \frac{\partial g(\mu_{t+1})}{\partial s_{t+1}}$$

Start by noting that from the symmetry of the standardised multivariate normal $f'_\mu = -f'_s$. Take the derivative

$$-\lim_{\sigma_k \rightarrow 0} \frac{\partial v_{0t}(s_t)}{\partial \mu_{t+1}} = \lim_{\sigma_k \rightarrow 0} \int g(s) f'_s(s) ds + \lim_{\sigma_k \rightarrow 0} \int g'_s(s) f(s) ds$$

and use integration by parts in the first line

$$\lim_{\sigma_k \rightarrow 0} \int g(s) f'_s(s) ds + \lim_{\sigma_k \rightarrow 0} \int g'_s(s) f(s) ds = \lim_{\sigma_k \rightarrow 0} g(s) f(s)$$

$$\int \lim_{\sigma_k \rightarrow 0} g(s) f(s) h(s) ds + \int \lim_{\sigma_k \rightarrow 0} g'_s(s) f(s) ds = \lim_{\sigma_k \rightarrow 0} g(s) f(s)$$

$$\lim_{\sigma_k \rightarrow 0} g'_s(s) f(s) ds = -g'_\mu(\mu)$$
The second line uses that since $f$ is normal, it can be written $f'(s) = f(s)h(s)$, where $h(s)$ is linear, and it interchange the order of limits and integration on the left hand side. By symmetry of the normal, $\lim_{s \to \infty} f(s) = \lim_{s \to -\infty} f(s) = 0$. The third line uses that $f$ converges at an exponential rate, faster than $g$. Therefore $\lim_{s \to -\infty} g(s)f(s) = \lim_{s \to \infty} g(s)f(s) = 0$ and the right hand side is zero. Then $\lim_{\sigma \to 0} \int g(s)f(s)h(s)ds = \int \lim_{\sigma \to 0} g(s)f(s)h(s)ds = 0$, since the exponential rate of convergence of $g(s)f(s)$ dominates the linear divergence of $h(s)$. Finally, as $\sigma \to 0$, $f(s)$ becomes a Dirac delta spiking at $\mu$. As $f'_\mu = -f'_s$, the result in the third line now follows.

The extension to multidimensional $s$ is straightforward. In particular, when $\Sigma \to 0$ (perfect foresight over all prices), the partial derivative is just $\frac{\partial v_0}{\partial \mu_{kt+1}} = \frac{\partial g(\mu)}{\partial \mu_{kt+1}}$ with $\mu_{kt+1} = p_{kt+1}$, as expected.