DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

PRODUCT SELECTION AS PRICE DISCRIMINATION IN THE MARKET FOR BOOKS

Sofronis K. Clerides

Discussion Paper 99-08

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-2-892101, Fax: ++357-2-750310
Web site: http://www.econ.ucy.ac.cy

Product selection as price discrimination in the market for books

Sofronis K. Clerides*
University of Cyprus

February 23, 1999

Abstract

Book publishers price discriminate among consumers by publishing books in hardcover and paperback formats and introducing them into the market at different points in time. This behavior is hard to justify using conventional models with a unimodal distribution of consumer preferences. A generalization to bimodal preferences eliminates the problem and affords a natural interpretation of the two peaks as representing libraries and individuals. These two types have different valuations for time, quality, and book content. Estimation of a structural econometric model that incorporates the firm's product quality choices confirms the bimodality in consumer tastes and yields realistic and intuitive implications.

Keywords: price discrimination, product selection, quality choice, book publishing industry.

JEL Classification: L1, L82.

*Address: Department of Economics, University of Cyprus, P.O.Box 20537, CY-1678 Nicosia, Cyprus. Email: sofronis@ucy.ac.cy.

I thank Steve Berry, Ariel Pakes, John Rust, Peter Davis, Jessica Holmes, and Nadia Soboleva for helpful comments and suggestions.

1 Introduction

Market segmentation techniques are frequently employed by firms seeking to maximize their profits. The simplest of these techniques is first degree price discrimination, the practice of charging different prices to different consumers for the same product. Practical use of this is limited, however, due to imperfect information on individual consumers' willingness to pay and because of legal constraints. Of greater empirical relevance is the use of second and third degree price discrimination. The latter involves discrimination on the basis of observable consumer characteristics (such as student or senior citizen discounts). The former, which is also known as nonlinear pricing, discriminates based on *unobserved* consumer heterogeneity. The firm offers a menu of products and prices, each of which is targeted towards a specific type of consumer. The products (or bundles of products) are designed in a way that makes it optimal for consumers of each type to choose the product that was designed for them. Thus consumers *self-select* into the appropriate niche of the market. Market segmentation is therefore achieved not only by the choice of prices, but also by the choice of product quality.

Books are a familiar example of price discrimination.[1] Publishers use publication time and binding quality as ways of discriminating among consumers with hetereogeneous valuations of those attributes. They typically publish a new title in hardcover format first, then follow it up with a paperback version several months (or even years) later. In some cases they publish the hardcover and paperback versions simultaneously, while in other cases they choose not to produce a paperback at all. The model I propose in this paper rationalizes this behavior in the context of profit maximization through the choice of product quality.

The delayed introduction of an inferior version of an established product is a frequently encountered phenomenon. New movies are usually released first in luxurious and expensive first-run theaters; several months later they hit the second-run theater circuit and are also made available on videotape. Patient consumers can therefore enjoy them at a significant discount. Other examples are bakers, who often sell their day-old bread at half-price, and fashion designers, who channel unsold items to stores that specialize in marked-down designerwear. In none of these cases can the price drop be attributed to changing costs. The same is not true for, say, computer prices. Even though the price of a new computer falls significantly within months of its introduction, one can argue that it may be falling costs that drive this price decrease. This argument can not be made for books, movies, bread, or designer clothing.[2]

Existing empirical studies of price discrimination fall into two main categories. The first one is a series of papers that test for the presence of price discrimination by looking at differences in price-cost differentials across different products. Borenstein (1991) compares the markups for leaded and unleaded gasoline; Shepard (1991) looks at gas stations with both full-service and self-service pumps versus those that offer only one of the two options; Borenstein and Rose (1994) look at variation in fares paid by different passengers on the same flight; and Giulietti and Waterson (1997) compare markups for several products across Italian supermarkets. All of these papers find evidence of price discrimination in their respective markets. In Borenstein (1991) and Giulietti and Waterson (1997) discrimination is possible because of the cost (mostly transportation cost) of switching to another retailer, hence we can think of these as examples of third degree price discrimination based on geographical location. In Shepard (1991) the

[1] Many introductory economics textbooks give books as an example of price discrimination. Books are also frequently mentioned as an example of intertemporal pricing in the durable good pricing literature; for example, in Stokey (1979) and Bulow (1982).

[2] There is, however, a difference between books and movies on one hand and bread and designer clothing on the other. In the latter case the products that are sold in subsequent periods are leftovers while in the former case additional copies are manufactured.

discrimination is of the second degree, since gas stations offer two service options and consumers self-select into their preferred one. In Borenstein and Rose (1994) both second and third degree discrimination are observed; airlines use requirements like a Saturday night stay to separate business from leisure travelers, but also make special offers to specific groups, such as students.[3]

The second line of empirical work in price discrimination focuses on nonlinear pricing and follows a more structural approach.[4] Ivaldi and Martimort (1994) study the interaction between energy-buying firms in the French dairy industry and energy providers. The authors assume that the observed price-quantity pairs (that is, the ones chosen by buyers) are points on the optimal price-quantity schedule and use this information to estimate demand parameters. Bousquet and Ivaldi (1997) analyze the optimal pricing of telephone usage. They estimate demand for phone usage using data from French residential callers and then solve for the optimal nonlinear price schedule using a simplified (restricted to a single dimension of heterogeneity) version of the estimated demand function. Leslie (1997) follows a similar approach in his study of price discrimination in Broadway theater. He estimates theater demand using data on ticket purchases when seating of varying quality and price is available. He then uses his estimates to compare the revenue generated by price discrimination to that under alternative pricing schemes, and to measure the welfare effects of the firm's ability to discriminate.

The durability of books also relates the publisher's problem to the large theoretical literature on durable good pricing, which dates back to the seminal work of Coase (1972). Empirical work on intertemporal pricing is limited and largely confined to some studies of markdowns and clearance sales. Pashigian (1988) and Pashigian and Bowen (1991) find that higher demand uncertainty for several 'fashion' items is associated with higher initial prices and bigger markdowns in subsequent periods.[5] In a different context, Warner and Barsky (1995) present some empirical evidence of sales observed at times of high demand (such as weekends or big holidays), which they explain in terms of lower search costs for consumers at those times.

Finally, the publisher's problem can be placed in the context of new good introduction. Publishers produce several new titles (products) every year, but most of these titles are completely unrelated to one another, making it extremely difficult to forecast demand based on past experience. This is quite different from the more commonly studied problem of the introduction of a new product that is an improved version of an already existing one, as is the case for technological products. Rather, it resembles the experience of movie production studios, record companies, and fashion designers.[6] On this aspect also, empirical work is limited. Some interesting studies are those by Lu and Comanor (1998), who study the pricing of new pharmaceuticals, Burton (1994), who explores the introduction of new products by multiproduct firms in the insecticide industry, and Hilke and Nelson (1987), who look at the introduction of Folger's ground coffee in the eastern United States. None of these industries share the essential features of book publishing; new product introductions in pharmaceuticals and insecticides are R&D related, while coffee is a fairly homogeneous good.

Work on the book industry in economic literature is scant at best. In a study on competition in the U.S. retail college textbook market, Siegfried and Latta (1998) find that textbook prices are very rigid across bookstores and are not related to the number of competitors, entry barriers, or retail service cost differences. Bittlingmayer's (1992) study of demand for books in Germany is the only paper that I am aware of that utilizes data at the individual book level. Resale price

[3] For a dissenting view on the identification of price discrimination see Lott and Roberts (1991). The authors argue that many of the most commonly cited examples of price discrimination can be explained by alternative, cost-based arguments; they provide such explanations for several markets, including the gas and airline industries.

[4] See Wilson (1993) for a thorough overview of nonlinear pricing and its applications.

[5] A theoretical model with this implication has been developed by Lazear (1986).

[6] See De Vany and Walls (1996, 1997) for some evidence of demand uncertainty in the movie industry.

maintenance is legal in Germany for books, so price data are more reliable than U.S. prices. Unlike their American counterparts, German publishers frequently increase but rarely decrease the price of a book from its initial level; in Bittlingmayer's data (which include price and sales data for three consecutive years), about half of all prices change from one year to the next and almost all changes are increases. He estimates price elasticities ranging from 1.5 to 3, while the Lerner index implies elasticities between 1.7 and 2.5. Unfortunately, the author does not provide much information on temporal sales patterns in his data.

Unlike most of the existing empirical literature on this topic, the aim of this paper is not to establish whether price discrimination is exercised in the book market. Rather, I take this as a fact and focus on uncovering the nature of consumer heterogeneity that makes price discrimination possible. My study differs from those of Bousquet and Ivaldi (1997) and Leslie (1997) in that it incorporates firm quality choices in the model and uses them in the estimation procedure. This turns out to be crucial in helping identify the nature of consumer heterogeneity. In that sense, the work here is closest to Ivaldi and Martimort (1994), which incorporates information on the firms' nonlinear pricing schedule.

Following the recent empirical literature on product differentiation,[7] I employ a characteristic based discrete choice model to describe the consumer's problem. This framework essentially uses market share information to estimate consumer preferences over product characteristics. This is a lot to ask of the data, especially when one is after a rich characterization of preferences. Imprecise estimates of the desired parameters is often the result. The problem is ameliorated by augmenting the model with a behavioral assumption on firm conduct, typically a pricing assumption (such as Nash pricing).

In some instances – this being one of them – the researcher may be reluctant to impose a pricing regime if no such assumption seems appropriate. In such cases the firm's choice of products and of product quality can serve as an alternative behavioral restriction that can not only enhance the precision of demand estimates, but also improve the accuracy of the model and thus our understanding of firm behavior. The usual assumption that the set of available products is exogenous is clearly a simplification that can create difficulties in a dynamic setting, especially when the model is used to make predictions about the future. For example, the BLP authors report elsewhere[8] that their model does very well in predicting demand in the first few years immediately after the 1974 oil shock, but its performance deteriorates rapidly thereafter. The authors interpret that as a failure of the model to anticipate the change in the product set that followed the sharp increase in fuel prices, namely the introduction of more fuel efficient models.

Incorporating the firm's product selection problem in an empirical framework can easily render the model computationally intractable. Nonetheless, the model I present here is both estimable and realistic for the industry I am studying. I avoid the computational burden described above by reducing the firm's problem to a simple choice of when to introduce its products. The model is rich enough to capture all the essential features of the industry and predicts firm behavior that is consistent with empirical observation.

Specifically, I consider a monopolist that produces a single basic good which can come in a high quality and a low quality version (hardcover and paperback respectively). The monopolist can choose whether to introduce the two versions of the product simultaneously or to market the hardcover version first and then, possibly, the paperback. In the theoretical section I argue that if the distribution of consumers is unimodal then observed firm behavior is consistently suboptimal. Therefore, I allow demand to come from two distinct types of consumers, which

[7] See the seminal work by Bresnahan (1987) and Berry, Levinsohn, and Pakes (1995) (henceforth BLP).

[8] Pakes, Berry, and Levinsohn (1993).

have a natural interpretation as being libraries and individuals. I use data obtained directly from a publishing firm. They provide information on sales, prices and various product characteristics for a large number of books. Estimation makes use of GMM techniques based on a conditional moment restriction on an unobserved product characteristic, and on additional moments derived from the firm's problem.

Estimation of the demand side by itself yields imprecise estimates. This is partly due to the nature of the model (as described above) and partly due to insufficient variation in the product set over different markets. When firm choices are included, on the other hand, the coefficients of interest are estimated with greater precision. The estimates confirm the existence of two types of consumers with substantially different price elasticities and different preferences for quality and time. The estimate of the proportion of each type in the population is also in accordance with prior knowledge.

This paper contributes to the existing literature in a number of ways. It is one of only a handful of economic studies of the book publishing industry, and the only one that utilizes extensive, title-specific price and sales information. It contributes to the growing empirical literature on price discrimination and helps improve our understanding of intertemporal pricing of new products. It is the first empirical model that I am aware of that incorporates quality choice in the estimation procedure. Finally, the empirical results I obtain demonstrate the importance of using a flexible demand system (such as one that allows for multi-peaked preferences) and the crucial role that firm choice information can play in the identification of demand systems in differentiated products.

The rest of this paper is organized as follows. Section 2 provides an overview of the book publishing industry and section 3 describes the available dataset. Section 4 discusses the relevant theory and then section 5 presents a model of firm behavior that captures the main features of the publisher's problem and serves as a theoretical introduction to the empirical model that follows. Section 6 lays out the details of that model, outlines the estimation procedure, then presents and discusses the results. Section 7 summarizes and proposes some directions for related future research. Finally, the details of the estimation procedure are included in an appendix.

2 The book publishing industry

Book publishing is not a particularly large industry; the value of total shipments by producers was around $20 billion in 1995. The number of producers, however, is large. The 1992 Census of Manufactures reports 2,644 book publishing firms in the United States, but most industry experts put that figure closer to 20,000. The large discrepancy is attributed to the thousands of tiny publishing houses (basically one-person, do-it-yourself ventures) that do not qualify for inclusion in the Census reports. A few large firms dominate the industry and are responsible for a significant proportion of sales. Twenty firms were responsible for 83.9% of industry revenues in 1993,[9] while the top 7 publishers held 87% of the spots on the weekly Top 30 bestseller lists compiled by Publisher's Weekly in 1995.

Many publishing companies have been around for decades, and consolidation in the industry has made them large and powerful. In recent years, however, a new force has emerged as an important and powerful player in the book business. National book chains like Barnes & Noble, Waldenbooks, and B. Dalton have experienced rapid growth and the conventional wisdom in the industry is that they are squeezing publishers' profits margins. In addition, internet retailers like Amazon.com have established themselves almost instantaneously as important

[9]Source: Greco (1997), p. 58.

players in the industry. Large publishers held the upper hand as long as they had to deal with small independent retailers, but the meteoric rise of book chains has reduced their power. An ever-increasing proportion of sales occurs in national chains, and publishers have become very dependent on them. Small publishers, on the other hand, are said to be flourishing under the new order, as the national chains provide a large outlet for their books, and consequently a better shot at commercial success.

Four main types of agents are active participants in the book industry: authors, publishers, retailers, and consumers. I outline here the main features of the interactions among these groups. The publisher and author agree on a contract which specifies the royalty the author is to receive on [usually net] sales. The royalty is typically increasing in units sold, and lower for paperbacks than for hardcovers. Established authors often get an advance on sales as part of the deal. The publisher decides whether to publish the book in hardcover, paperback, or both, and chooses prices and time of publication.[10] He then sends his sales force around the country to try to convince bookstore managers to put the book on their shelves. Retailers place their orders with the publisher, and they receive the books at a substantial discount over list price. This discount is usually increasing in the number of books ordered and varies between 42-50% for trade books. For textbooks the discount is smaller and fairly standard at 20%. Additional discounts may be granted to bookstores that carry a large number of different titles from the same publisher. This, of course, benefits large chain bookstores that carry many titles.

Most publishers make their books available in both hardcover and paperback versions. In trade (general interest) publishing paperbacks usually make their appearance about a year after the hardcover. In scholarly publishing the wait is longer, usually closer to 18 months or 2 years. Some, mostly academic presses occasionally publish hardcovers and paperbacks at the same time. The frequency of this practice varies from press to press, and it will be a focal point for this paper. I will argue that the main reason for the simultaneous (as opposed to sequential) introduction of hardcover and paperback is the firm's belief that delay of the paperback would kill off a large part of the market. This could occur, for example, if a book is about a current event, in which case interest in it may fizzle away in a few months.

3 The data

The dataset comes from a large academic press and includes practically every book published by the press in the last 25 years. Each title may come in different formats, the two most important of which are hardcover (or cloth) and paperback (others are deluxe editions, large-type editions, audio versions). The dataset identifies which formats each title has appeared in; specifically, I know which titles have appeared both in hardcover and in paperback, and which have appeared in only one format. For each format the following information is available: list price, total units sold, time of publication, number of pages, weight, subject matter (history, philosophy, etc.), number of authors, and whether this is a new edition of a previously published title. This information was obtained directly from the publisher. In addition, I have constructed a measure of the number of previous publications by each author. Information for this (admittedly noisy) variable was collected by searching through a number of databases, including Books In Print, Books Out of Print, various library catalogs, and the online catalog of Amazon.com.[11] Note that I only observe cumulative sales data for each format

[10] Contracts may include stipulations on these issues also. For instance, contracts for trade (general interest) books sometimes specify that the paperback must be published within a year of the hardcover.

[11] Construction of this variable was difficult because there seems to be no comprehensive database that *uniquely* identifies authors. Among the sources used the Library of Congress comes the closest, but it has the disadvantage that it includes publications other than books (notably some journal articles).

(hardcover/paperback); no period by period breakup is available.

The database becomes progressively less dependable as it goes back in time, so to avoid selection problems I drop books published prior to 1980. In addition, I drop titles published after 1990 in order to minimize truncation problems with books that continue to sell for many years after they are first published. To the extent that I am able to identify them, I also drop textbooks from the sample because they pose a different choice problem, both for the publisher and for the consumer.

The resulting dataset includes 1108 titles. Figure 1 summarizes the firm's timing choices. In 136 of the cases (12.3%) the hardcover and paperback were published simultaneously. More than half of the remaining titles never made it to paperback. Of the 434 titles that did make it to paperback, 21 were published by other publishers. The practice of selling paperback rights is followed for titles that succeed in gaining a broad appeal that extends beyond academic and scholarly circles. Most academic presses lack the distribution network and expertise to properly market such books, and prefer to sell paperback rights to other publishers when the opportunity arises. Since they are not published by my publisher, I have no sales data for those paperbacks.

Table 1 reports price and sales data broken down according to the firm's choice of publication timing.[12] On average, paperbacks are priced at less than half the hardcover price, and they sell more copies. The latter is particularly true when the two formats are published simultaneously, which is the case for about one in every eight titles. Publishers seem to follow this strategy for two main kinds of publications: books that deal with current events, and books that are reprints or new editions of previously available titles. Hardcovers that never make it to paperback have a slightly higher price tag than the ones published in both formats.

The timing of publication has a large effect on hardcover sales. When hardcover and paperback are introduced simultaneously, hardcover sales are only about 12% of total sales, as opposed to 38% of total when the paperback is delayed. This is strong evidence that some consumers place a high value on having a book immediately upon publication, and are prepared to pay a significant surcharge for that privilege. The mean waiting time for the paperback when introduction is sequential is almost two and a half years; the median time is less than two years. The bottom panel of Table 1 reports the correlation coefficients between prices, sales, and the lag between hardcover and paperback. A longer wait for the paperback is associated with lower overall sales and a higher ratio of hardcover to paperback sales. The correlation patterns show this to be mostly due to lower paperback sales, which suggests that this publisher is willing to delay the paperback and settle for lower overall sales in order to get more consumers to switch to the much more profitable hardcover.

Hardcover-only titles sell markedly fewer total copies than their dual-format counterparts, while books whose paperback rights are sold do much better. A scatter plot of hardcover and paperback sales (Figure 2) confirms this finding. Even though hardcover-only titles ("sitting" on the horizontal axis, since they have no paperback sales) make up half the sample, very few have sales greater than 5,000 copies. Most successful hardcovers are published in paperback also. Note that some titles with "normal" hardcover sales also end up being sold to other publishers, so this is not limited to blockbuster books. The very distinct sales patterns of simultaneous and sequential titles are also clear in these plots. The figure also shows that the distribution of sales is highly skewed. This underscores the unpredictability of the market for books. Much like firms engaging in R&D, publishers undertake many projects (publish many books). Many of these will turn out to be unprofitable, but a handful of big hits are enough to ensure overall profitability.

[12] The price I observe is the list price, which is determined by the publisher and indicated on the actual book. Unlike trade books (like New York Times bestsellers), scholarly titles are typically not discounted.

Figure 3 plots prices against sales for hardcovers and paperbacks respectively. Panels A and B include all available data points, while panels C and D 'zoom in' by plotting only hardcovers and paperbacks that sold fewer than five and ten thousand copies respectively (95.7% and 92.3% of all titles). Either way, there is very little evidence of a negative relationship and it is hard to argue for a negatively sloped demand curve.

There is little evidence that the timing of publication has any effect on prices. One might have expected the spread between hardcover and paperback price to be smaller when they are simultaneously published, reflecting the smaller degree of differentiation between the two products. The data provide little support for this hypothesis. On the contrary, the spread is larger in the simultaneous case, though the difference is not statistically significant.[13] This suggests that product selection plays an important role. If the decision between simultaneous and sequential was arbitrary, and given that prices are roughly equal in the two cases, then simultaneously published titles should have higher overall sales (hardcover plus paperback) than sequentially published ones. This is because any consumer who makes a purchase in the sequential case would also make one in the simultaneous case, while the converse is not true.[14] Again, however, the data point in the opposite direction. Simultaneously published books have slightly lower overall sales, suggesting that books published that way are somehow different from sequentially published ones.

Table 2 presents the results of some descriptive price regressions. They confirm the intuition that bigger (in terms of number of pages) and higher quality (in terms of weight and weight per page) books should be more expensive. Note that the number of pages is a good determinant of paperback price, while the weight variables do better with hardcover price. This is consistent with a policy that prices paperbacks fairly strictly according to the number of pages, while allowing more freedom to price hardcovers according to overall quality and other factors.[15] This can also explain why the regressions explain more of the variation in paperback price (77%) than in hardcover price (68%).

The results also confirm earlier evidence that titles appearing in both formats are cheaper than ones that do not, and that the timing of paperback publication does not affect hardcover price ('simultaneous' variable insignificant). There is some evidence that the paperback's time of introduction has an effect on its price, though the magnitude of that effect is fairly small. This is explored further in the third column, where I include only sequentially published titles. The length of the wait has no effect on price here. I conclude that if there's any effect of timing on pricing, that is small and limited to the simultaneous/sequential dichotomy.

In this last regression I have also included hardcover sales, since they are observed before sequential paperbacks are published. The estimated coefficient is strongly significant and, surprisingly, negative. This is a difficult result to explain; intuition suggests that a paperback should command a higher price when the hardcover that preceded it is successful. One possible explanation (at least for an academic press) is that once some positive profit is assured (through high hardcover sales), the publisher is willing to sacrifice some profits in order to increase unit sales. The insignificance of the 'previous publications' and new editions variables suggests that previous success does not translate to higher prices. This may be an indication of lack of market power, or perhaps of the absence of 'brand name' authors in the scholarly market.

Two additional variables, the price of paper (pulp) and the average price charged by all university presses in that year have little effect on price. The price of paper is an important

[13] However, Table 1 indicates that the *length* of the wait is positively correlated with hardcover price.

[14] This argument works as long as prices are fixed across the two regimes. One cannot rule out the possibility that prices are set optimally in each regime but the systematic differences between different titles are such that *ex post* prices end up looking very similar. This, however, is an unlikely scenario.

[15] Conversations with several publishers lend further support to this interpretation of pricing policy.

determinant of costs, but it is hard to capture any effect because of the lag between ordering, production, and marketing. The price charged by university presses is an estimate constructed by BISG, which cautions about year-to-year comparisons because of changes in the methods used. Hence, the absence of any impact of this variable on price should be interpreted with some scepticism.

The results from descriptive sales regressions are reported in Table 3. As expected, hardcover sales are higher when the paperback is delayed. A previous publishing record has a positive effect on sales. New editions do not seem to be any more successful than titles appearing for the first time. The length (measured in number of pages) of a book has an overall negative effect on sales, most of the impact being on the paperback. Generally, these regressions show that it is hard to predict a book's success from observable characteristics.

4 The theory

Consider a firm that is faced with two types of customers, one with a high valuation for a good (high types) and one with a low valuation (low types). The firm produces two versions of the product, a high quality one designed for the high types, and a low quality one for the low types. The firm chooses the quality of the 'bad' product to be as low as it needs to be to ensure that high types buy the 'good' product. One can then interpret the publisher's decision to publish a paperback (instead of just discounting the hardcover) as an attempt to make the low-quality product undesirable and thus induce as many purchases of the high-quality product as possible. A publisher faces demand from two main sources. Libraries have high valuations for books, and also for quality (represented by hard covers). Individual consumers generally have much lower valuations. This dichotomy may also help explain why discrimination is not continuous (i.e. why we do not observe a declining sequence of prices). If publishers treat demand as coming from two groups of consumers that are sufficiently distinct from each other but fairly homogeneous within them, then two versions of the product are enough to achieve the separation they desire.

The theoretical literature has produced some interesting results that are pertinent to this problem; work on the pricing problem of a durable good monopolist and on quality choice and product selection are particularly relevant. In this section I discuss each of these briefly in the context of the problem I am studying.

Durable good pricing. Any discussion of durable good pricing must inevitably begin with the seminal work of Coase (1972). In this seven-page classic Coase argues that the seller of an infinitely durable good can not credibly price his product at anything above marginal cost. Suppose that such a producer prices his product above marginal cost. Consumers know that once sales are realized in the first period the producer has an incentive to lower the price in subsequent periods. Hence they will not make a purchase unless the cost of deferring consumption by one period is greater than the expected gain from a lower second period price. In what has become known as the *Coase conjecture*[16], Coase argued that if the length of the period goes to zero (or, equivalently, if the consumers' discount factor goes to 1), then the producer will be unable to sell anything at a price above marginal cost and will lose all monopoly power. Coase also points out that a firm can get around this problem by effectively shortening the period over which the services of the good are acquired by the buyer.[17]

[16] The conjecture has since been formally proven by several authors, including Bulow (1982), Stokey (1981), and Gul, Sonnenschein, and Wilson (1986).

[17] This can be accomplished in a number of ways, such as contractual arrangements, by leasing rather than renting the good (Bulow 1982), or by choosing to make the good less durable (Bulow 1986).

Here I am not be concerned with how the publisher gets around the Coase conjecture. I assume away the problem by thinking of the publication of a book as the stage game of an underlying infinitely repeated game where a firm produces a new product every period, and consumers purchase the firm's product as long as the firm does not deviate from the "no-price-decrease" strategy. Given this, the relevant question becomes whether (and when) intertemporal price discrimination is optimal. Using a continuous time model, Stokey (1979) showed that intertemporal price discrimination is no more profitable than a single price policy if consumers differ only in their valuations of the good.[18] If, however, consumers also differ in their rates of time preference *and* those rates are positively correlated with the valuations then discrimination, in the form of a declining price schedule, can be profitable.

Quality choice and market segmentation. The quality choice problem of a multi-product monopolist was first analyzed by Mussa and Rosen (1978). They consider a monopolist who can produce any number of products at different points along the uni-dimensional quality continuum. Consumers with heterogeneous preferences for quality choose among the (product, price) pairs the one that maximizes utility. The authors show that under certain conditions the monopolist may find it optimal to 'bunch' consumers together, that is to offer the same quality to consumers with different valuations. This bunching occurs if the hazard rate $f(\theta)/[1-F(\theta)]$ (where $f(\theta)$ is the density of consumers and $F(\theta)$ is the associated distribution function) is decreasing. This condition appears frequently in mechanism design problems, and it is usually assumed that the hazard rate is *non*decreasing. This is because most commonly used distributions (such as the uniform, exponential, normal, Pareto, beta and gamma) have nondecreasing hazard rates. The standard example of a distribution that can have a decreasing hazard rate is a bimodal distribution. Thus a bimodal distribution of consumer preferences in the Mussa-Rosen model will lead to a 'discretization' of the product space.

A discrete version of the Mussa-Rosen setup that is perhaps more better suited for empirical work is explored by Itoh (1983). Discreteness in the product space enables the author to investigate the effects of changes in the degree of product differentiation (such as the introduction of new products or the withdrawal of existing ones) on prices and welfare. A number of interesting results come out of this analysis. The base result is that the introduction of a new good leaves the prices of all *lower* quality goods unchanged, while causing the prices of all *higher* quality goods to change (positively or negatively) by the exact same amount. The direction of the change turns out to depend on the hazard rate. One particularly interesting corollary is that if a new product is introduced that is of higher quality than all existing products then the prices of those products will remain unchanged. Consequently, consumer welfare will always increase when a high-end product is introduced.

When a low quality product is introduced, on the other hand, the effect on prices (and welfare) is ambiguous. Itoh shows that the introduction of a new good will raise (leave unchanged, lower) the prices of higher quality goods if the hazard rate function is concave (linear, convex) over the relevant interval. This has markedly different implications for different distributions. The uniform, exponential, and Pareto distributions all have linear hazard rates. Introduction of a new good in that case will have no effect on the prices of the existing goods. The hazard rate of the normal distribution is convex, so prices of higher quality goods will drop, while that of the lognormal is concave, hence prices will rise.

A ditribution that is of special interest is the extreme value distribution, since this is fre-

[18] This result also depends critically on the assumption that all agents – consumers and the firm – have the same discount rate. Landsberger and Meilijson (1985) showed that in Stokey's framework the firm stands to benefit from intertemporal price discrimination if and only if its discount rate is lower than that of the consumers (that is, if the firm is more patient).

quently used in empirical discrete choice models, including the one in this paper. The hazard rate of the extreme value distribution is initially convex and then concave, so no definite conclusion can be drawn about the direction of price changes. Generally, however, if the qualities are such that only consumers near the high end of the distribution (where the hazard rate is concave) are affected by the new product, then the price of the high quality product should be expected to rise.

This case is of particular interest since the introduction of a lower quality version of a successful product is a practice frequently employed by firms trying to capture the lower end of the market. Paperbacks are an example of that, as are student (or reduced capability) versions of computer software. Interestingly, these lower quality products are not necessarily cheaper to produce. It has been widely reported, for example, that Intel used to produce the 486SX chip by taking the 486DX and destroying the math co-processor.[19] Publishers themselves often turn unsold hardcovers into paperbacks by – literally – stripping off the hard covers and putting on paper ones. Clearly, in both those cases it would be cheaper to just discount the original product, but then the firms would be subject to the pitfalls of the Coase conjecture.

5 A two-type, two-period model

The main implication of the theoretical work discussed above is that the shape of the distribution of consumer preferences is critical in determining firm choices and market outcomes. The distinction between unimodal and multimodal distributions is particularly crucial, and it plays an important role in the empirical work that is the main focus of this paper. In this section I explore this issue further in a simple two-period, two-type model that brings out the essential features of the publisher's problem.

Consider a monopolist producing a single product that it can sell in two periods. The firm can market the product in two versions, a 'basic' one, and a 'deluxe' one (corresponding to our paperback and hardcover respectively). Consumers have the option of buying at most one of the versions available, or nothing at all (the outside option). Consumer utility takes the characteristic-based form proposed by Lancaster (1966). Let ζ_i denote consumer i's valuation for the basic version if it is consumed today, and let ψ_i be the value assigned to the extra quality of the deluxe version. Finally, let ρ_i denote the value lost when the product is received next period instead of the current one. Formally, the utility of consumer i from purchasing product j is

$$\begin{aligned} u_{ij} &= \zeta_i + \psi_i H_j - \rho_i T_j - \alpha_i P_j + \varepsilon_{ij} \\ &= \nu_{ij} + \varepsilon_{ij} \qquad\qquad j = 1, 2; \end{aligned}$$

where H_j is a dummy for the high quality (deluxe) version, and T_j is 0 if product j is consumed today and 1 if it is consumed tomorrow.

I define the two types as follows: "high types" (indexed by superscript h) have valuations described by the quadruple $(\zeta^h, \psi^h, \rho^h, \alpha^h)$, while the valuations of "low types" (indexed by superscript l) are given by $(\zeta^l, \psi^l, \rho^l, \alpha^l)$. High types make up a proportion λ of the population. In other words,

$$\nu_{ij} = \begin{cases} \nu_j^h & \text{with probability} \quad \lambda \\ \nu_j^l & \text{with probability} \quad 1-\lambda \end{cases} \tag{1}$$

We can think of $\nu_j^\tau = \zeta^\tau + \psi^\tau H_j - \rho^\tau - \alpha_i P_j T_j$, $\tau = h, l$ as the mean quality (value) of product j for consumers of type τ and ε_{ij} as each individual's deviation from that mean. Note that consumers are homogeneous *within* their type. I further assume that production costs are zero.

[19] See Deneckere and McAfee (1996) for more on this and for other interesting examples of this practice.

Following what the data seem to suggest, I assume that demand for books is such that prices depend on binding quality but not on time of publication. Thus prices are set at p_1 and p_2 for a hardcover and paperback respectively, regardless of when they are published.[20] If ε_{ij} has the extreme value distribution then the probability that a consumer of type τ will purchase good j takes the well-known logit form

$$s_j^\tau = \frac{e^{\nu_j^\tau}}{1 + e^{\nu_1^\tau} + e^{\nu_2^\tau}}, \quad j = 1, 2; \quad \tau = h, l. \tag{2}$$

The overall market shares are simply

$$s_j = \lambda s_j^h + (1 - \lambda) s_j^l, \quad j = 1, 2. \tag{3}$$

Since prices are set at p_1 and p_2 regardless of when each product is published, the firm's only choice variable is the timing of introduction of each product. I normalize the time of hardcover introduction to be 0, so the firm just chooses T_2, the time of paperback introduction, with $T_j \in \{0, 1\}$. Total firm profit as a function if its choice variable can be written as

$$\Pi(T_2) = p_1\, s_1(T_2) + [1 - T_2(1 - \delta)]\, p_2\, s_2(T_2)$$

where δ is the firm's discount factor. I will hereafter denote $\Pi^{sim} = \Pi(0)$, $\Pi^{seq} = \Pi(1)$, and similarly for the shares.

The single-type case. Let us first suppose that there is only one type of consumer, that is $\nu_j^h = \nu_j^l \equiv \nu_j$. Consider the choice between simultaneous and sequential introduction of the two goods. Note that if $\rho = 0$ then market shares are the same under each scenario. Hence, simultaneous profit can be no less than sequential profit, and will be strictly larger if the firm's discount factor, δ, is less than 1. If publication is simultaneous, market shares do not depend on ρ, and so neither does profit. On the other hand, profit does depend on ρ when publication is sequential; the derivative of sequential profit with respect to ρ is

$$\frac{\partial \Pi^{seq}}{\partial \rho} = p_1 s_1^{seq} s_2^{seq} - \delta p_2 (1 - s_2^{seq}) s_2^{seq}. \tag{4}$$

Since simultaneous profit does not depend on ρ, sequential publication can only be profitable if sequential profit is (at least somewhere) increasing in ρ. Rearranging equation (4) we see that $\partial \Pi^{seq} \partial \rho > 0$ when

$$\frac{s_1^{seq}}{s_0^{seq}} > \frac{\delta p_2}{p_1 - \delta p_2}. \tag{5}$$

This says that sequential profit is increasing in ρ only when the market share of the hardcover is large (which would occur if consumers are very impatient, that is if ρ is large). For a typical book the right hand side of the condition above is around 0.7, which means that the share of the hardcover would have to be at least 70% of the share of the outside good for the condition to be satisfied. This is unlikely to occur. In fact, evaluating (5) (using $\delta = .95$) reveals that the condition holds for only 1 out of 413 sequentially published titles. Thus I conclude that sequential profit in this model can only be increasing in ρ for implausibly large values of the

[20] This is not as restrictive as it may sound. It could, for instance, be derived as an implication from a model where books compete against each other in a differentiated product market.

latter, and therefore sequential publication is almost never optimal. This implication casts serious doubt on the validity of the model.[21]

The firm's choice may also be driven by variation in ζ, the valuation for the basic product. It is possible to show that in this model simultaneous publication will be preferred for very low values of ζ. The intuition is simple. Given that the loss from waiting is fixed at ρ, a low ζ means that many consumers will drop out of the market rather than purchase in period 2. As ζ increases, the second period purchase becomes more and mroe attractive for the consumer, and hence so does sequential introduction to the firm.

This condition, however, is only relevant if it becomes binding at "reasonable" levels of demand. For example, the condition is empirically uninteresting if it implies that the sequential menu becomes more profitable only when almost every consumer makes a purchase. It turns out that this is indeed the case. One can derive a condition similar to (5) which fails to hold for *any* of the 136 simultaneously published titles in the dataset. Thus, we can conclude that variation in ζ in this model (just like variation in ρ) can not explain the choice of the sequential menu. The single-type model fails to deliver the kind of firm behavior that we observe empirically.

The two-type case. The introduction of a second consumer type can get us around this problem. Market shares for each consumer type and overall market shares are given by (2) and (3) respectively. To simplify matters, let $\rho^h = \alpha_0 + \alpha_1 \rho^l$, where $\alpha_0 > 0$ and $\alpha_1 > 1$. Then the derivative in (4) for the two-type case becomes

$$\frac{\partial \Pi^{seq}}{\partial \rho^l} = \lambda \alpha_1 \left[p_1 s_1^h s_2^h - \delta p_2 (1 - s_2^h) s_2^l \right] + (1 - \lambda) \left[p_1 s_1^l s_2^l - \delta p_2 (1 - s_2^l) s_2^l \right]. \tag{6}$$

If a large enough proportion of high types purchases the deluxe good then the first term in this expression will be positive (because of condition (5)), and can outweigh the negative second term. As ρ increases further, the first term evens out because the high type market is exhausted. The second term, however, continues to increase until it eventually outweighs the first term. Thus with two types sequential profit can be initially increasing and then decreasing in ρ, making sequential publication optimal for moderate discount factors but not for very large ones.

A graphical representation makes this point more clearly. Figure 7 plots the firm's revenue premium from choosing sequential (SEQ) over simultaneous (SIM) publication as a function of ρ^h and ρ^l, given ζ and plausible values for the remaining parameters. Sequential introduction is profitable when high types are impatient and low types are patient. Impatient high types will tend to buy in period 1, boosting the market share of the deluxe good and making the derivative in (6) positive, as required for SEQ to be optimal.

Figure 7 plots the same revenue premium for different values of ζ and ρ^l, keeping ρ^h fixed at 3. The prediction of the single-type model that SIM is less profitable than SEQ for high values of ζ carries through to the two-type version. For moderately low values of ζ (the ones that may be empirically relevant in the sense that they result in reasonable market shares), one can see again that SEQ is profitable if low types are patient (ρ^l close to 0), but SIM becomes more profitable as low-type consumers become impatient.

To understand where this implication comes from, note that by choosing to publish sequentially instead of simultaneously, the firm gets some consumers to switch from hardcover to paperback. But it also causes some paperback buyers to drop out of the market. For the former effect to outweigh the latter there must be a fairly large mass of consumers at the high

[21] A similar condition determines the choice between producing the deluxe good only or both goods. The former is only optimal if $s_1/(1 - s_1) < \delta p_2/(p_1 - \delta p_2)$, where s_1 is the share of the deluxe good when only that is produced. This condition is also very unlikely to hold.

end of the distribution. This is unlikely to be the case if the distribution of consumers is unimodal. Introducing two types of consumers solves this problem, by the presence of a large mass of consumers (the high types) at the high end of the distribution. When discount factors are very small the firm achieves no segmentation by delaying the paperback, so simultaneous introduction is preferred. When discount factors are moderately large, separation is achieved with sequential introduction. For very large discount factors the delay would completely kill off the paperback market, so simultaneous publication is again optimal.

6 An empirical model of product selection

6.1 Demand specification

I model the consumer as facing a discrete choice problem. For every title published, he chooses between purchasing one of the versions of the book made available by the publisher (hardcover and possibly paperback), or purchasing nothing. The discrete choice framework is appropriate since consumers are unlikely to want to buy more than one product.[22]

Consider the market for a single book. Let the instantaneous utility that individual i derives from consuming format j of this title be given by

$$U_{ij}^0 = \begin{cases} e^{x'\beta_{x,i}+\beta_{h,i}H_j-\beta_{p,i}P_j+\xi_j+\eta_i+\varepsilon_{ij}} & j=1,2; \\ e^{\varepsilon_{i0}} & j=0; \end{cases}$$

where products 1 and 2 are the high and low quality goods (hardcover and paperback) respectively, and product 0 is the outside alternative. The utility from consuming the same item at time T_j from now is discounted according to the rate of time preference ρ_i:

$$U_{ij} = e^{-\rho_i T_j} \cdot U_{ij}^0$$

Taking logs of both sides delivers the linear form

$$u_{ij} = \ln(U_{ij}) = x'\beta_{x,i} + \beta_{h,i}H_j - \beta_{p,i}P_j - \rho_i T_j + \xi_j + \eta_i + \varepsilon_{ij}.$$

Conveniently, the time of publication enters linearly as just another product characteristic. The vector x includes characteristics that are common to both formats, such as content and author information. The variable H_j is a dummy that takes the value of 1 if product j is a hardcover and 0 otherwise, P_j is the price of product j, and T_j is the time of publication (measured as time since hardcover publication, so that $T_1 = 0$ always).

The term ξ_j accounts for an unobserved (by the econometrician) product characteristic. A large part of the variation in ξ_j is likely to be due to factors common to the hardcover and paperback, so I model it as the sum of two terms,

$$\xi_j = \zeta + \omega_j.$$

The ζ part is a catch-all term for a book's popularity. It denotes the valuation for the book's content, and it must vary across titles in order to explain their different degrees of success. Variation of ω_j across titles will help explain variations in the relative success of hardcovers and paperbacks. In addition, the parameter ρ_i is also expected to vary across titles, because some books become obsolete faster than others. The generalized utility function over individuals $n = 1, \ldots, M$, products $j = 1, \ldots, J$ and markets $n = 1, \ldots, N$ is

$$u_{ijn} = x_n'\beta_{x,i} + \beta_{h,i}H_{jn} - \beta_{p,i}P_{jn} - \rho_{in}T_{jn} + \xi_{jn} + \eta_{in} + \varepsilon_{ijn}.$$

[22] This is not strictly true. Libraries often purchase multiple copies of the same title.

6.2 Estimation

This demand model can be estimated independently of the supply side. Estimation is based on the set of moment conditions

$$E[\xi_{jn}|z_{jn}] = E[\zeta_n + \omega_{jn}|z_{jn}] = 0,$$

where the n subscripts denote different markets (book titles), and z_{jn} is a vector of appropriately chosen instruments. That is, I assume that the unobserved product characteristic is mean independent of the variables included in the instrument set. I also assume that ζ_n and ω_{jn} are i.i.d. across titles and independent of each other, while the ω_{jn} are also independent across formats. These assumptions are not imposed on the model, but are used in the calculation of the optimal weight matrix and the standard errors. The conditional variance of ξ_{jn} is $E[(\zeta_n + \omega_{jn})^2|z_{jn}] = \sigma_\zeta^2 + \sigma_\omega^2$. Letting $\xi_n = [\xi_{1n} \quad \xi_{2n}]'$,

$$E[\xi_n \xi_n'|z_n] = \begin{bmatrix} \sigma_\zeta^2 + \sigma_\omega^2 & \sigma_\zeta^2 \\ \sigma_\zeta^2 & \sigma_\zeta^2 + \sigma_\omega^2 \end{bmatrix} \equiv \Sigma. \tag{7}$$

Also, if $\xi = [\xi_1 \quad \xi_2 \quad \ldots \quad \xi_N]'$ then

$$E[\xi\xi'|z] = I_N \otimes \Sigma \equiv \Omega.$$

Single-type nested logit. In order to estimate the model I need to impose some structure on the distribution of consumer preferences. The nature of the heterogeneity in this problem can be captured well in the nested logit framework developed by McFadden (1978). This assumes that all taste parameters are fixed across individuals; that is, $\beta_{x,i} = \overline{\beta}_x$, $\beta_{h,i} = \overline{\beta}_h$, $\beta_{p,i} = \overline{\beta}_p$, and $\rho_{in} = \overline{\rho}$ for all i, n. The individual-specific term $\eta_i + \varepsilon_{ij}$ is restricted in a way that allows us to introduce correlation among 'similar' products in a simple way. Specifically we have (suppressing n subscripts for clarity)

$$\begin{aligned} u_{ij} &= x'\overline{\beta}_x + \overline{\beta}_h H_j - \overline{\rho} T_j - \overline{\beta}_p P_j + \xi_j + \eta_i(\psi) + \psi\varepsilon_{ij} \\ &\equiv \delta_j + \eta_i(\psi) + \psi\varepsilon_{ij}. \end{aligned}$$

where $\delta_{jn} \equiv x'\overline{\beta}_x + \overline{\beta}_h H_j - \overline{\rho} T_j - \overline{\beta}_p P_j + \xi_j$ is the mean consumer valuation of the good's quality. This form of the individual-specific term implies the existence of two 'nests' or groups, one including all versions of the product, and the other including just the outside good. The parameter ψ measures the correlation among consumers' valuations for the hardcover and paperback. Cardell (1997) has derived the unique distribution $C(\psi)$, $0 \leq \psi \leq 1$, that is conjugate to the extreme value distribution. In other words, if $\eta_i(\psi)$ is distributed according to $C(\psi)$ and ε_{ijn} has the extreme value distribution, then the sum $\eta_i(\psi) + \psi\varepsilon_{ijn}$ also has the extreme value distribution. The parameter ψ ranges between 0 and 1 and it measures the relative importance of the η_i term in this sum. As ψ tends to 0, heterogeneity across titles dominates, while if ψ goes to 1 the title-specific η_i vanishes and the nested logit reduces to the simple logit. If the disturbance ε_{ij} has the extreme value distribution, then the market share functions take the form

$$s_j = \frac{e^{\delta_j/\psi}}{\left[e^{\delta_1/\psi} + e^{\delta_2/\psi}\right]^{1-\psi} \left[1 + \left(e^{\delta_1/\psi} + e^{\delta_2/\psi}\right)^{\psi}\right]}, \quad j = 1, 2. \tag{8}$$

Berry (1994) shows this implies the following linear equation that can be estimated (I re-introduce n subscripts):

$$\ln(s_{jn}/s_{0n}) = x_n'\overline{\beta}_x + \overline{\beta}_h H_{jn} - \overline{\rho} T_{jn} - \overline{\beta}_p P_{jn} + \psi \ln(s_{j|\sim 0,n}) + \xi_{jn}.$$

This is a straightforward linear equation of the log-odds ratio on product characteristics and the log of the *within* share (the share of format j conditional on a purchase being made, $\ln(s_{j|\sim 0,n})$; the coefficient on this variable is the parameter ψ). Least squares estimation, however, will produce inconsistent estimates because the disturbance term ξ_{jn} is likely to help determine both the book's price and the timing of publication, and is therefore correlated with P_{jn} and T_{jn}; moreover, the within share clearly depends on xi_{jn}, so that also is endogenous. Instrumental variable estimation can be used to circumvent this problem; the choice of instruments is discussed later on in this section.

Two-type logit. The logit demand model with multiple consumer types was first used by Berry, Carnall, and Spiller (1996) (henceforth, BCS) to study the market for airline tickets. The model introduces heterogeneity in consumer tastes in a simple and intuitive way. It is particularly useful in problems – such as this one – where the correlation between an individual's tastes for different attributes is likely to be high. A drawback of the multiple-type model is that it restricts this correlation to be 1 for consumers of the same type. It also restricts the variance of the preferences to be the same across types. It is possible to use a less restrictive procedure that specifies a fully flexible distribution (incorporating both bimodality and arbitrary correlation patterns) for the entire vector of taste parameters. Such a procedure, however, would be very intensive computationally and possibly too demanding of the data, so I prefer the multiple-type version.

Let $\beta_i = [\beta_{x,i} \quad \beta_{h,i} \quad \beta_{p,i} \quad \rho_i]'$ be the vector of all taste parameters that enter the utility function. The two-type model allows the parameter vector β_i to vary across types, but not within types, as in (1). Specifically, letting superscripts denote types,

$$\beta_i = \begin{cases} \beta^h & \text{with probability} \quad \lambda \\ \beta^l & \text{with probability} \quad 1-\lambda \end{cases} \tag{9}$$

This defines two δ_j's for each format: δ_j^h, the mean quality of product j among consumers of type h, and δ_j^l, the mean quality among consumers of type l:

$$\delta_j^\tau = x'\beta_x^\tau + \beta_h^\tau H_j - \rho^\tau T_j - \beta_p^\tau P_j + \xi_j, \quad \tau = h, l. \tag{10}$$

These δ_j^τ's imply market shares s_j^τ for each type of consumer. The overall market share is the weighted average of the shares among consumers of each type:

$$s_j = \lambda s_j^h + (1-\lambda) s_j^l \tag{11}$$

This simple extension makes it impossible to obtain an analytic expression for the unobservable product characteristic ξ_j. I resort to an iterative procedure, proposed in BLP, that solves for the δ_j^τ's by making use of the contraction mapping theorem. Define the overall (unconditional on individual consumers) mean quality δ_j to be the weighted average of the mean qualities of the two types:

$$\begin{aligned} \delta_j &= \lambda \delta_j^h + (1-\lambda)\delta_j^l \\ &= \lambda x_j'\beta^h + (1-\lambda) x_j'\beta^l + \xi_j. \end{aligned}$$

Recall that I am interested in finding the δ_j's that equate observed shares with predicted shares, that is the δ's that solve $s_j = s_j(\delta, x, \theta)$. Let $P(\cdot)$ denote the distribution function of consumer preferences defined in (9). BLP have shown that the operator $T(s_j, \theta, P)$ defined pointwise by

$$T(s_j, \theta, P)\,[\delta_j] = \delta_j + \ln(s_j) - \ln[s_j(x, \delta, P; \theta)] \tag{12}$$

is a contraction mapping with modulus less than one (θ is a vector that includes all the parameters of the model). This implies that I can solve for δ_j recursively by evaluating the right hand side of (12) at some initial guess for δ_j, obtaining a new $\delta_j{}'$, substituting the latter back into the right hand side, and continuing until our convergence criterion is satisfied. The resulting estimate of δ is then used in (10) to yield the ξ_j. These are interacted with the vector of instruments z_j to create the GMM criterion function to be minimized:[23]

$$G(\theta) = \xi' Z \hat{W}^{-1} Z' \xi$$

The matrix $\hat{W}$ is an estimate of the variance $W = Z'\Omega Z$ of the moment vector $Z'\xi$ and is the optimal weighting matrix in the sense that it delivers efficient estimates given the choice of instruments. White (1980) has shown that, if the ξ_j are independent, an estimate of this matrix can be obtained using:

$$\hat{W} = \frac{1}{N_f^2} \sum_{j=1}^{N_f} \xi_j(x_j, \hat{\theta})^2 z_j z_j',$$

where N_f is the number of observations (the number of formats published) and z_i' is the ith row of the matrix of instruments Z. The ξ_j of a particular title are not, however, independent but have the covariance structure described in (7), which is similar to that of a panel dataset. To account for that I use the estimator proposed by Newey and West (1987):

$$\hat{W} = \frac{1}{N_f^2} \sum_{j=1}^{N_f} \xi_j^2 z_j z_j' + \frac{1}{N_f^2} \frac{1}{2} \sum_{j=2}^{N_f} \xi_j \xi_{j-1} (z_j z_{j-1}' + z_j z_{j-1}').$$

The required initial estimate $\hat{\theta}$ is obtained by setting $\hat{W} = I$.

Instruments. The choice of instruments is critical in the estimation of demand systems. I assume that the unobserved product characteristic is mean independent of all observed characteristics except price and the time of publication of the paperback. In estimating oligopolistic models of differentiated products one can typically use the characteristics of other firms as instruments. This option is not available here since I have a monopoly. I do, on the other hand, have a number of separate markets, so I can use characteristics of books in those markets as instruments. Hence, to instrument for price I use the average price of other books of the same format by year and by category. In addition I use the weight of the book, the price of pulp, wages, and real GDP. Finding suitable instruments for the timing variable is more difficult. The firm's timing choice is based on the book's content; but anything that is correlated with that probably also belongs in the demand function, and hence cannot be used as an instrument. I use the mean waiting time of books of the same format by year and by category. Finally, to instrument for the within share I use the ratio of hardcover to paperback price of other books.

[23] The GMM method of estimation is due to Hansen (1982).

Results. The results for the single-type logit model are reported in Table 4. As one might expect, a good previous publishing record has a positive effect on sales, and delaying the paperback lowers its sales significantly. The coefficient on the within share suggests that about 60% of the variation in utility is common to hardcovers and paperbacks.

However, the results are unsatisfactory in a number of ways. The coefficient on price is very small, with implied elasticities much smaller than unity. The coefficients on the hardcover dummy and on the number of pages are both negative, suggesting that consumers *dislike* the higher quality associated with the hardcover and need to be compensated for acquiring additional pages. These results, along with the implication that firm choices are suboptimal, cast serious doubts on the validity of the single-type model.

Estimation of the two-type model did not improve the results. Essentially, the demand side by itself fails to identify the two types; it predicts that all consumers belong to one of the types with probability 1, and the preferences of that type are the same as those in the single-type model. For this reason I do not report the results from the two-type model.

The failure of the demand side to identify the two types is not surprising given the structure of the problem. Precise demand estimates are generally difficult to obtain from discrete choice models that use only product-level data. Even without this problem, identification of more than one types requires substantial variation in the number of available products in each market. To see why, note that a well known limitation of the simple logit model is that it implicitly assumes the independence of irrelevant alternatives (IIA) axiom. This IIA assumption leads to counterintuitive results, such as the "blue bus/red bus paradox" first pointed out by Debreu (1960). Suppose that individuals have a choice between driving their own car or taking the bus, and the probability of each is equal to 1/2. Now suppose two buses can be used, which are identical except in that one is red and the other is blue. If individuals do not care about color, one would expect the two buses to compete for the same consumers, while the car-riding crowd will stay with their choice of driving. The IIA axiom, however, implies that the probabilities of each of the three choices (car, red bus, blue bus) should be the same. By adding more bus choices, one can drive the car probability further down towards zero.[24]

It is this kind of counterintuitive prediction that random coefficient models take advantage of in order to identify consumer heterogeneity. The two-type model of BCS works well for two reasons: the Nash pricing assumption, and variation in their data. The number of products per market in their dataset varies tremendously, from just one product in some markets to several hundreds in others.[25] By contrast, the number of products in my dataset is always either one or two. Consequently, the counterintuitive implications of the IIA axiom do not become apparent in the market share data, and estimation of the demand side alone fails to identify the two types. In other words, consumer choices are perfectly consistent with the single-type model. Firm behavior, on the other hand, is inconsistent with the single type model, as seen in the previous section. The information contained in firm choices can help pin down otherwise imprecise demand estimates. Accordingly, in the next section I extend the model to include the firm's optimization problem and derive implications that can be used to enhance estimation results.

6.3 Incorporating firm choices

In order to focus on the product selection aspect of the publisher's problem, and to keep estimation simpler, I treat prices as fixed (conditional on product characteristics). With this

[24] For a concise and in-depth analysis of the logit model and its implications see Anderson, de Palma, and Thisse (1992).

[25] See BCS (page 18) for a relevant discussion of the identification issue in their model.

simplification, the publisher's problem reduces to a simple choice problem among three alternative menus: simultaneous (SIM), sequential (SEQ), and hardcover-only (HON). Equivalently, the publisher chooses T_2, the time of publication of the paperback, where T_2 can take the values 0 (SIM), 1 (SEQ), or ∞ (HON). The publisher knows the parameters that enter the consumer utility function and the value of the unobservable (to the econometrician) ξ_j.

Let p_1 and p_2 be the *net* price (after unit cost, bookseller discount and author royalty) for the hardcover and paperback respectively. For every title published the firm chooses $T_2 \in \{0, 1, \infty\}$ to maximize profits:

$$\max_{T_2 \in \{0,1,\infty\}} \Pi(T_2|x, \xi, \theta).$$

Profits are given by

$$\Pi(T_2|x, \xi, \theta) = M\ [p_1\, s_1(T_2|x, \xi, \theta) + I\,(T_2 < \infty)\ (1 - T_2(1 - \delta_f))\, p_2\, s_2(T_2|x, \xi, \theta)] - FC(T_2),$$

where δ_f is the firm's discount factor, and $FC(T_2)$ are the fixed costs associated with each choice. Thus the choice of T_2 determines profits through its effect on market shares and fixed costs.

The choices of publishing the paperback immediately in some cases while delaying it or not publishing it at all in others suggests an inherent difference in the nature of the demand for particular titles. I model this difference as arising from two sources: the different degrees of popularity across titles, and the different rates of decay of that popularity.

The success of a book is hard to explain using observable physical product characteristics. Nonetheless, some of the variation in popularity is captured in this model by variables that are title-specific but not format-specific, such as number of pages, weight per page, and variables that pertain to the author and to the book's content. All remaining variation in popularity is represented by the ζ term.

The second source of demand variation across titles is in the different rates of decay of that popularity. Some books have shorter lives than others. For example, books that deal with current issues may become irrelevant or outdated within a few months of publication. In that case the publisher would not want to delay a paperback issue. I model this heterogeneity as variation in the rate by which consumers discount future consumption of the book. So the discount rate is heterogeneous *both* across consumers *and* across titles. In terms of the model this is implemented as variation in the mean of the distribution of consumers' discount rates. Some of this variation may be determined by observable characteristics.

The publisher makes his timing decision based on the values of ζ, ρ^h and ρ^l for each title. He calculates expected sales and profits under each menu, and picks the most profitable one. The model must deliver predictions of firm action that match the observed actions as closely as possible. This involves estimating expected profits under each menu. This is fairly straightforward for menus that are actually chosen by the firm, since I observe both prices and market shares in those cases. I do not, however, observe those variables for menus that were *not* chosen by the firm. For example, if the firm chooses SIM I observe prices and market shares when the two formats are published simultaneously, but I do not know what prices and shares would have been had publication been sequential. The model needs to have enough structure to enable us to predict these prices and shares from *within* the model.

In the remainder of this section I outline the procedure used to calculate profits and predict firm behavior. I start by assuming that from the contraction mapping I get estimates of the mean quality of every format published, for each type; that is, I know δ_j^τ, for $j = 1, 2$ and $\tau = h, l$ (the details of the contraction are discussed in the appendix). I refer to these as the *actual*

δ's, because they correspond to the book formats that were actually published. Accordingly, I refer to δ's that correspond to unpublished formats as *counterfactual* δ's.[26] Also, to keep computation simpler, I assume that each title gets one draw of ρ from which the discount rates of both consumer types are determined, so that the relationship between ρ^h and ρ^l is deterministic and linear in form:

$$\left.\begin{array}{l}\rho^l = \rho \\ \rho^h = \alpha_0 + \alpha_1 \rho\end{array}\right\} \Longrightarrow \rho^h = \alpha_0 + \alpha_1 \rho^l.$$

In order to estimate counterfactual shares I need to create "counterfactual products". I do this by perturbing actual δ's by the increment that would result from a change in the relevant product characteristic. For example, the only difference between available products under SIM and SEQ is that the paperback has a value of 0 for the time-of-publication characteristic ($T_2 = 0$) in the SIM case and a value equal to the length of the publication lag in the SEQ case. Therefore, if the value of ρ was known one could add $\rho \cdot T_2$ to the actual δ_2 and thus "transform" the menu from SEQ to SIM. Note that this perturbation is only needed for paperbacks. Hardcovers are part of every actual menu, so an estimate of their quality is always available.

For books published only in hardcover there is no actual paperback to create a counterfactual paperback from, so some approximations are required. I create the paperback from the hardcover by subtracting the difference in valuations due to the hard cover, the price difference, the publication lag, and the difference between the unobservable product characteristics $\xi_1 - \xi_2$. Of course, since the paperback was never published, the values of the variables for the paperback are not known. To obtain paperback price estimates I run a regression of paperback price on observed product characteristics and use those parameters to predict the price of paperbacks had they been published. For the length of the publication lag, I simply use the median of observed waiting times as an estimate. Finally, I have nothing to estimate ξ_2 from, so I integrate out over it. The transformation rules for each case are described in Table 5.

Once I calculate counterfactual δ's I plug them into the market share equation (8) to produce the counterfactual shares. With these market share estimates in hand, I proceed to compute profits under each menu. I have to approximate actual costs and the actual price received by the publisher, since I do not observe them. I estimate unit and fixed costs for each title by making use of the cost data presented in an earlier section. I regress actual costs for the subsample on observed characteristics and use the resulting coefficients to estimate costs for the entire dataset. This produces estimates $\hat{c}_1$ and $\hat{c}_2$ respectively for the marginal cost of the hardcover and paperback, and $\widehat{FC}(T_2)$ for the fixed costs.

Unfortunately, similar data on a subsample are not available for publisher price, so my estimates of it are less informed. Based on information provided by the publisher, I assume that the firm receives 50% of the list price, and that the author's royalty payment is 10% and 5% of that respectively for hardcovers and paperbacks. Thus , predicted prices are $\hat{p}_1 = .90 * .5 * p_1 - \hat{c}_1$, $\hat{p}_2 = .95 * .5 * p_2 - \hat{c}_2$. I also set the firm's discount factor, δ_f, to 0.95.

The remaining details are relegated to the appendix. A summary of the entire estimation procedure is as follows (n superscripts denote iteration number):

1. Carried over from previous iteration: δ^n, θ^n, $f^n(\rho|T)$.
2. Contraction to get δ^{n+1}:
   (a) use δ^n, θ^n to calculate δ^h and δ^l as functions of ρ;
   (b) plug δ^τ's into (8) to get market shares conditional on ρ;

[26] I believe that my use of counterfactuals is not subject to McAfee's (1983) critique of the reconstruction of history.

(c) use $f^n(\rho|T)$ to integrate out the ρ;
(d) plug shares into (12) to obtain updated δ;
(e) repeat cycle until convergence to δ^{n+1}.

3. use δ^{n+1}, $f^n(\rho|T)$ to integrate over ρ and obtain $\xi(\theta^n, \delta^{n+1})$;
4. From δ^{n+1}, create counterfactual δ's using rules in Table 5. Plug these into (8) to get counterfactual shares.
5. Use actual and counterfactual shares to calculate the firm's choice probabilities conditional on ρ and ω_0.
6. Integrate out to get unconditional probabilities and construct moments; obtain updated $f^{n+1}(\rho|T)$ in the process.
7. Evaluate GMM criterion and update θ^{n+1}.

6.4 Results and implications

The results from the model are presented in Table 7. For this set of estimates I used an exponential distribution on the ρ's and fixed the parameter α_1 to be 0, so that $\rho^h = \alpha_0$. Despite the large standard errors for some of the parameters, the results are very much in accordance with intuition and with the predictions of the model discussed earlier. The model is able to identify two types of consumers with markedly different preferences with respect to time and price. The preferences of high type consumers are not estimated very precisely, while those of low types are much better (the estimate of the hardcover dummy being an exception). The estimate of ψ suggests that about three quarters of the variation in the unobserved product characteristic is due to factors common to hardcovers and paperbacks. The proportion of high types in the population is 2.3% (parameter λ). The mean of the distribution of discount rates is also reasonable, even though it is not not precisely estimated.

The implications of these results are very reasonable. These are presented in Table 7. The model estimates the number of high types to be 2,317, a realistic figure. High types make a purchase about two thirds of the time, but the kind of purchase depends very much on the available choices. If hardcover and paperback are published simultaneously only 20.0% of high types buy the hardcover, compared to 67.4% when they are published sequentially. YUP estimates that it sells close to one thousand copies of a typical book to libraries and institutions. This model predicts 465 and 1,562 hardcover copies sold to high types in the SIM and SEQ cases respectively, certainly a reasonable estimate. Low types exhibit similar purchasing patterns in their choice between hardcover and paperback, although in absolute terms they make purchases much less frequently. Note that in either case, about half of all hardcovers are purchased by low types.

The model predicts firm choice correctly about two thirds of the time: 82.9% of the time when publication is simultaneous and 65.7% of the time when it is sequential. The conditional means of ρ confirm that higher ρ^l (low types more impatient) cause the firm to choose simultaneous over sequential publication.

The price elasticities implied by the model for the median title are given in Table 7. These should be read as the percentage change in demand of the row product for a 1% change in the price of the column product. The three numbers given in each cell are, from top to bottom, elasticities for type 1's, for type 2's and total market elasticities. Type 2's clearly have more elastic demands than type 1's. Perhaps surprisingly, the demand for hardcovers is much more elastic than the demand for paperbacks. This is because roughly half of all hardcovers are bought by type 2's, who are very price sensitive, and a change in price will cause them to switch. The overall pattern of substitution is very reasonable and reassuring. Discrete choice

models typically underestimate own price elasticities, but here only 8% of paperbacks (2.7% of all products) have elasticities that are less than unity.

The distribution of consumers implied by these results is shown for a particular title in a contour plot in Figure 7. Note the two peaks representing high and low types. The straight lines divide up the space into three regions, which represent the consumers who make different purchases (assuming the outside good has a value equal to the mean of its distribution). Consumers who find themselves in the bottom left rectangle make no purchase. Consumers in the region to the right are the ones who buy the hardcover, while those up and to the left buy the paperback.

In the case of this particular title the publisher achieved the separation of the two types very well, as the point of congruence of the three regions is very close to the point where the two masses drift apart. This is not the case for every title. This picture also helps gives a good understanding of how the publisher can take advantage of different preferences of the two types. Suppose he could find a way to change a paperback so that it would become less appealing to high types, but would not affect the low types' valuation of it. Then the publisher would be able to shift the mass of high type consumers down further into the hardcover region, without causing any of the low types to drop out of the market, thus increasing his profits.

7 Conclusions and extensions

This paper provides an analysis of the book publishing industry and particularly of the issue of price discrimination as practiced by book publishing firms. Heterogeneity in consumer preferences prompts publishers to discriminate between consumers by publishing a book in two versions, a hardcover and a paperback. More interestingly, these two versions are not always the same, but they may differ in the time of availability of the paperback: sometimes it is introduced at the same time as the hardcover, while other times it is published several months later. In this paper I develop a model that explain this practice as profit-maximizing product choice on behalf of the firms. Different choices can be explained in these models in terms of the varying degrees of popularity of different books, and, more importantly, in terms of the different rates of decay of this popularity.

In the empirical part I estimate this model. I use a discrete choice framework to describe the consumer's problem and I show that using a unimodal distribution of consumer preferences in a standard discrete choice model implies that observed firm behavior is inconsistent with profit maximization. Hence I allow the distribution of consumer preferences to be bimodal. That is, there are two types of consumers; preferences vary across type, but consumers of the same type have identical preferences up to an idiosyncratic error term.

I estimate the demand side of the model using method of moment techniques. The estimates I obtain are imprecise because of insufficient variation in the demand data. In order to improve the precision of the estimation procedure I use a product selection assumption that turns out to be critical in identifying the parameters of the model. This serves as an alternative to the Nash pricing assumption that is usually imposed in this kind of model. The kind of product selection assumption employed here may serve as a useful identifying restriction in other models where the researcher is reluctant to make a pricing assumption.

Estimation of the full model yields results that are intuitive and consistent with empirical observation. The results suggest that there are indeed two types of consumers with substantially different preferences for books and for price and quality. These two types correspond naturally to libraries and individuals, and their existence serves to explain observed firm choices. This finding is also consistent with the diminishing importance of the hardcover in the book industry.

During the 1980s and 1990s library budgets have shrunk significantly throughout the United States; moreover, the proportion of those budgets that is earmarked for books has also shrunk as prices of journals (especially scientific ones) have soared. This has led to a substantial drop in the demand for hardcovers. The trend towards simultaneous introduction of hardcover and paperback is thought to be part of the industry's reaction to this development. Moreover, there is a small but growing trend towards bypassing hardcovers altogether and publishing straight in paperback; this can also be attributed to declining library budgets. In terms of the model presented here, this can be interpreted as a change over time in the preferences of high types (or as a reduction in their numbers). In future work I intend to explore this possibility by allowing preferences to vary over time.

A further way to extend this line of research is to model consumer heterogeneity more generally. In the current model within-type heterogeneity derives solely from an idiosyncratic error term, thus price elasticities and discount rates differ only across types. A more general specification with random coefficients on price and time would provide a more complete picture of the nature of consumer heterogeneity.

An important aspect of the book publisher's problem is the choice of the size of the initial printing and the possibility of additional printings if demand persists. This choice is important because of the large degree of demand uncertainty in this industry, in which every year each firm introduces into the market a large number of mostly unrelated new products. This kind of problem has potential applications in other durable good markets and industries that face the same kind of problem, such as the movie production and music industries. Empirical work on this issue is hindered by the lack of appropriate data, namely data on the size of printings and on period-by-period sales.

Appendix

The contraction. I now describe the calculation of δ_j^τ using the contraction, as defined in equation (12). Equation (10) implies that

$$\delta_j = \lambda \left[x_j' \beta^h - \rho^h T_j \right] + (1-\lambda) \left[x_j' \beta^l - \rho^l T_j \right] + \xi_j, \tag{13}$$

where the vector β^τ includes all parameters in the utility function except ρ^τ (that is, $\beta_x^\tau, \beta_h^\tau, \beta_p^\tau$) and x_j the corresponding variables. Calculation of market shares requires the shares for each type, and therefore also the δ_j^τ's. Equation (13) implies that

$$\begin{aligned} \delta^h(x,T,\rho,\theta) &= \delta + (1-\lambda)\left[(x'\beta^h - \rho^h T) - (x'\beta^l - \rho^l T) \right] \\ \delta^l(x,T,\rho,\theta) &= \delta - \lambda \left[(x'\beta^h - \rho^h T) - (x'\beta^l - \rho^l T) \right] \end{aligned}$$

Using these $\delta^\tau(\delta, x, T, \rho, \theta)$'s and the market share equations (8) and (11) I obtain market shares $s(\delta, x, T, \rho, \theta)$. Since the value of ρ is unknown I need to integrate over it:

$$s^\tau(\delta, x, T, \theta) = \int_\rho s(\delta, x, T, \theta | \rho) dF(\rho | T) \tag{14}$$

Note that integration is not necessary when publication is SIM or HON, because in those cases $T_1 = T_2 = 0$ and the discount rates drop out of the expression. The obtained predicted market shares are then plugged into (12) to get the next value of δ_j. Note that the distribution

of ρ is conditioned on firm choice T_2 because the latter is likely to be largely determined by both the popularity of a book and its rate of decay. This distribution is obtained from within the model using Bayes' rule:

$$f(\rho|T_2, x, \xi, \theta) = \frac{f(T_2|x, \xi, \theta, \rho)f(\rho)}{f(T_2|x, \xi, \theta)}.$$

Solving for the ξ's. The ξ's used to form the GMM criterion are obtained by using (13) and integrating out over the ρ's:

$$\xi_j(\delta_j, x, \theta) = \int_\rho \left[\delta_j - \lambda\left(x_j'\beta^h - \rho^h T_j\right) - (1-\lambda)\left(x_j'\beta^l - \rho^l T_j\right)\right] dF(\rho|T)$$

Predicting firm choices. In order to predict firm behavior, I assume that true profits are equal to estimated profits plus an additive error term, which has the extreme value distribution with variance σ_π^2, which is a parameter to be estimated. Then the conditional (on ρ and, if SEQ, on ξ_2) probability that the firm chooses a given time of paperback publication is simply

$$\begin{aligned}\Pr(T_2 = t|x, \xi, \theta, \rho) &= \Pr\left[\hat{\Pi}(t|x, \xi, \theta, \rho) + \omega_t = \max_{t' \in \{0,1,\infty\}} \left\{\hat{\Pi}(t'|x, \xi, \theta, \rho) + \omega_{t'}\right\}\right] \\ &= \frac{\exp\left[\hat{\Pi}(T_2|x, \xi, \theta, \rho)\right]}{\sum_{T_2'} \exp\left[\hat{\Pi}(T_2'|x, \xi, \theta, \rho)\right]}\end{aligned}$$

I assume that ξ_2 and ρ are observed by the firm, but not by the proverbial econometrician, so I have to integrate over their distribution. I use Gaussian quadrature because it works very well with two-dimensional integrals that involve density functions. The general form of the functions to be integrated is

$$\Pr(T_2|x, \xi, \theta) = \int_\rho \Pr(T_2|x, \xi, \theta, \rho)\, dF(\rho)$$

for SIM titles, while for SEQ and HON titles there is an additional integral because xi_2 is also unobserved:

$$\Pr(T_2|x, \xi_1, \theta) = \int_{\xi_2}\int_\rho \Pr(T_2|x, \xi_1, \theta, \xi_2, \rho) dF_\rho(\rho)\, dF_\rho(\xi_2|\xi_1).$$

Note that I have assumed that ρ is uncorrelated with the ξ's. This is consistent with the interpretation that the rate of decay of a book's popularity is determined by its content, and not by differences between the hardcover and paperback.

Firm moments. The moments used for estimation are defined as follows. Let the two new variables Y_1 and Y_2 be given by

$$Y_1 = \begin{cases} 1 & \text{if} \quad T_2 = 0 \\ 0 & \text{otherwise} \end{cases} \qquad Y_2 = \begin{cases} 1 & \text{if} \quad T_2 = \infty \\ 0 & \text{otherwise} \end{cases}$$

Letting $\hat{Y}_1$ and $\hat{Y}_2$ denote the observed outcomes of the two variables, we form the two moments as

$$E[\hat{Y}_1 - \Pr(SIM|x, \xi, \theta)|z_2] = 0; \qquad E[\hat{Y}_2 - \Pr(HON|x, \xi, \theta)|z_2] = 0.$$

The vector z_2 includes variables that are orthogonal to the prediction error. For the purpose of estimation I use only a vector of ones.

References

Anderson, S. P., A. de Palma, and J.-F. Thisse (1992): *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2), 242–262.

Berry, S., M. Carnall, and P. T. Spiller (1996): "Airline Hubs: Costs, Markups, and the Implications of Customer Heterogeneity," *NBER Working Paper 5561.*

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

Bittlingmayer, G. (1992): "The Elasticity of Demand for Books, Resale Price Maintenance and the Lerner Index," *Journal of Institutional and Theoretical Economics*, 148(4), 588–606.

Borenstein, S. (1991): "Selling Costs and Switching Costs: Explaining Retail Gasoline Margins," *Rand Journal of Economics*, 22(3).

Borenstein, S., and N. L. Rose (1994): "Competition and Price Dispersion in the U.S. Airline Industry," *Journal of Political Economy*, 102(4), 653–683.

Bousquet, A., and M. Ivaldi (1997): "Optimal Pricing of Telephone Usage: An Econometric Implementation," *Information Economics and Policy*, 9, 219–239.

Bresnahan, T. (1987): "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," *Journal of Industrial Economics*, 35, 457–482.

Bulow, J. (1982): "Durable Goods Monopolists," *Journal of Political Economy*, 90(2), 314–332.

——— (1986): "An Economic Theory of Planned Obsolescence," *Quarterly Journal of Econmomics*, 51, 729–750.

Burton, P. S. (1994): "Product Portfolios and the Introduction of New Products: An Example from the Insecticide Industry," *Rand Journal of Economics*, 25(1), 128–140.

Cardell, N. S. (1997): "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13(2), 185–213.

Coase, R. H. (1972): "Durability and Monopoly," *Journal of Law and Economics*, 15, 143–149.

De Vany, A. S., and D. Walls (1996): "Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry," *Economic Journal*, 106, 1493–1514.

——— (1997): "The Market for Motion Pictures: Rank, Revenue, and Survival," *Economic Inquiry*, XXXV, 783–797.

Debreu, G. (1960): "Review of R.D. Luce, Individual Choice Behavior: A Theoretical Analysis," *American Economic Review*, 50, 186–188.

Deneckere, R., and R. P. McAfee (1996): "Damaged Goods," *Journal of Economics and Management Strategy*, 5(2), 149–174.

Giulietti, M., and M. Waterson (1997): "Multiproduct Firms' Pricing Behaviour in the Italian Grocery Trade," *Review of Industrial Organization*, 12(5/6), 817–832.

Greco, A. N. (1997): *The Book Publishing Industry*. Allyn and Bacon, Needham Heights, MA 02194.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 39, 155–190.

Hansen, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hilke, J. C., and P. B. Nelson (1987): "Caveat Innovator: Strategic and Structural Characteristics of New Product Introductions," *Journal of Economic Behavior and Organization*, 8(2), 213–229.

Itoh, M. (1983): "Monopoly, Product Differentiation and Economic Welfare," *Journal of Economic Theory*, 31, 88–104.

Ivaldi, M., and D. Martimort (1994): "Competition under Nonlinear Pricing," *Annales D' Economie et de Statistique*, 34, 71–114.

Lancaster, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74(2), 132–157.

Landsberger, M., and I. Meilijson (1985): "Intertemporal Price Discrimination and Sales Strategy Under Incomplete Information," *Rand Journal of Economics*, 16, 424–430.

Lazear, E. P. (1986): "Retail Pricing and Clearance Sales," *American Economic Review*, 76(1), 14–32.

Leslie, P. J. (1997): "A Structural Econometric Analysis of Price Discrimination in Broadway Theatre," mimeo.

Lott, J. R., and R. D. Roberts (1991): "A Guide to the Pitfalls of Identifying Price Discrimination," *Economic Inquiry*, XXIX, 14–23.

Lu, Z. J., and W. S. Comanor (1998): "Strategic Pricing of New Pharmaceuticals," *Review of Economics and Statistics*, 80(1), 108–118.

McAfee, R. P. (1983): "American Economic Growth and the Voyage of Columbus," *American Economic Review*, 73(4), 735–740.

McFadden, D. (1978): "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlvist, L. Lundqvist, F. Snickars, and J. Weibull. North-Holland, Amsterdam.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Pakes, A., S. Berry, and J. Levinsohn (1993): "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change"," *American Economic Review, Papers and Proceedings*, 83, 240–246.

Pashigian, B. P. (1988): "Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing," *American Economic Review*, 78(5), 936–953.

Pashigian, B. P., and B. Bowen (1991): "Why Are Products on Sale?: Explanations of Pricing Regularities," *Quarterly Journal of Economics*, CVI, 1015–1038.

Shepard, A. (1991): "Price Discrimination and Retail Configuration," *Journal of Political Economy*, 99(1), 30–53.

Siegfried, J. J., and C. Latta (1998): "Competition in the Retail College Textbook Market," *Economics of Education Review*, 17(1), 105–115.

Stokey, N. L. (1979): "Intertemporal Price Discrimination," *Quarterly Journal of Economics*, 94, 355–371.

——— (1981): "Rational Expectations and Durable Goods Pricing," *Bell Journal of Economics*, 12, 112–128.

Warner, E. J., and R. B. Barsky (1995): "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," *Quarterly Journal of Economics*, pp. 321–352.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–838.

Wilson, R. (1993): *Nonlinear Pricing*. Oxford University Press, New York, Oxford, Published in association with the Electric Power Research Institute.

Figure 1: Frequency of firm choices

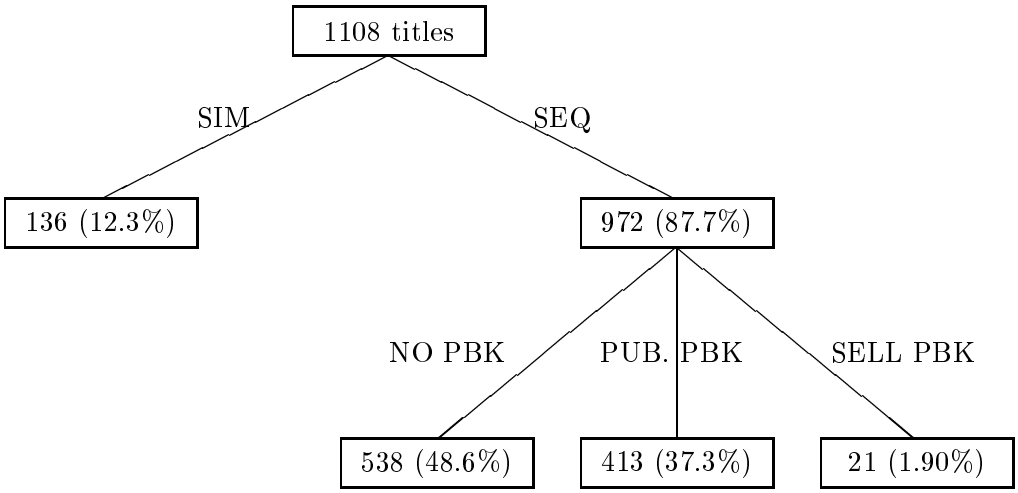

Table 1: Data summary

	N	Price	(s.e.)	Sales	(s.e.)
All books					
Hardcovers	1108	45.96	(17.07)	1,708	(1,863)
Paperbacks	549	19.67	(6.34)	3,980	(5,225)
Hardcovers when:					
Pbk by YUP	549	44.69	(17.16)	2,112	(2,165)
Pbk rights sold	21	42.11	(11.55)	3,809	(5,324)
No pbk	538	47.40	(17.05)	1,215	(855)
Simultaneously published titles					
Hardcovers	136	46.31	(18.79)	668	(667)
Paperbacks	136	20.11	(6.94)	4,958	(6,053)
Sequentially published titles					
Hardcovers	413	44.16	(16.58)	2,587	(2,275)
Paperbacks	413	19.52	(6.13)	3,658	(4,888)

Some simple correlations	
Variable	Correlation with introduction lag
Hardcover sales	-0.079 (0.107)
Paperback sales	-0.244 (0.000)
Hcr-to-pbk sales ratio	0.219 (0.000)
Hardcover price	0.104 (0.035)
Paperback price	0.019 (0.703)
Hcr-to-pbk price ratio	0.174 (0.000)
415 observations; p-values in parentheses.	

Figure 2: Hardcover and paperback sales for all titles

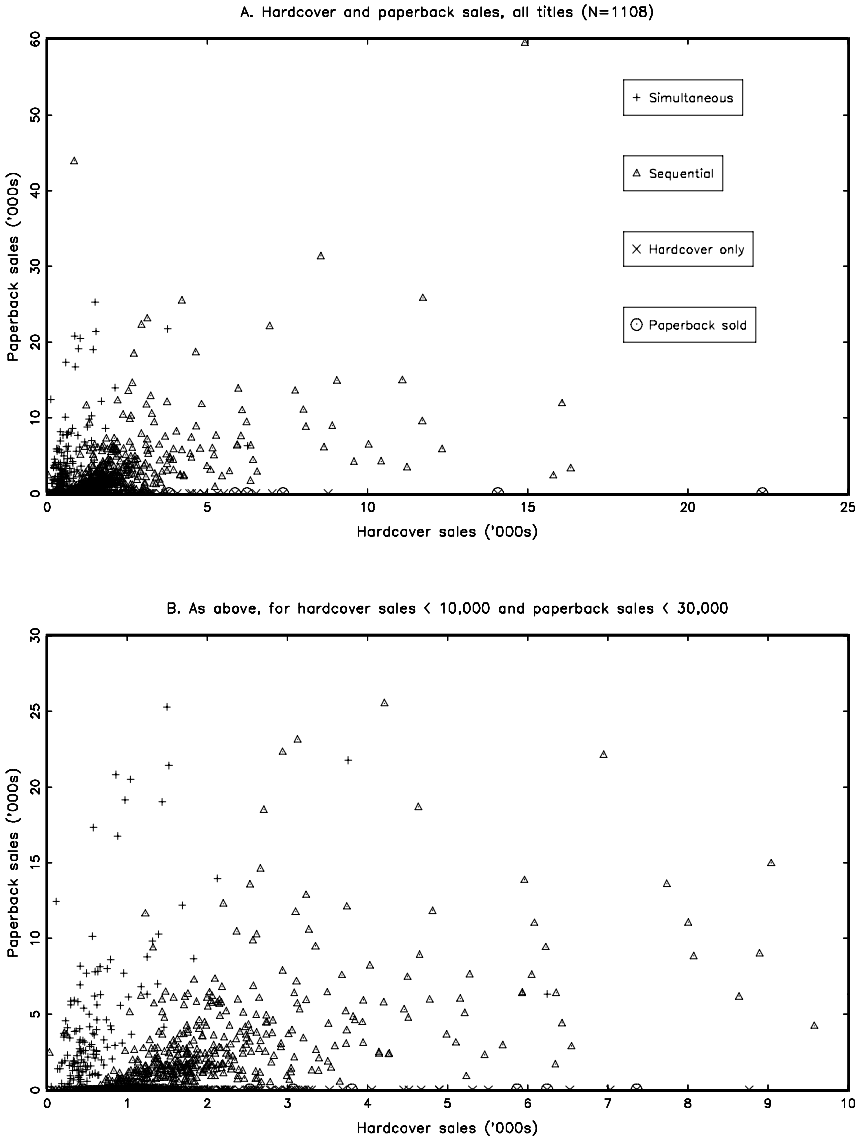

Figure 3: Prices and sales for hardcovers and paperbacks

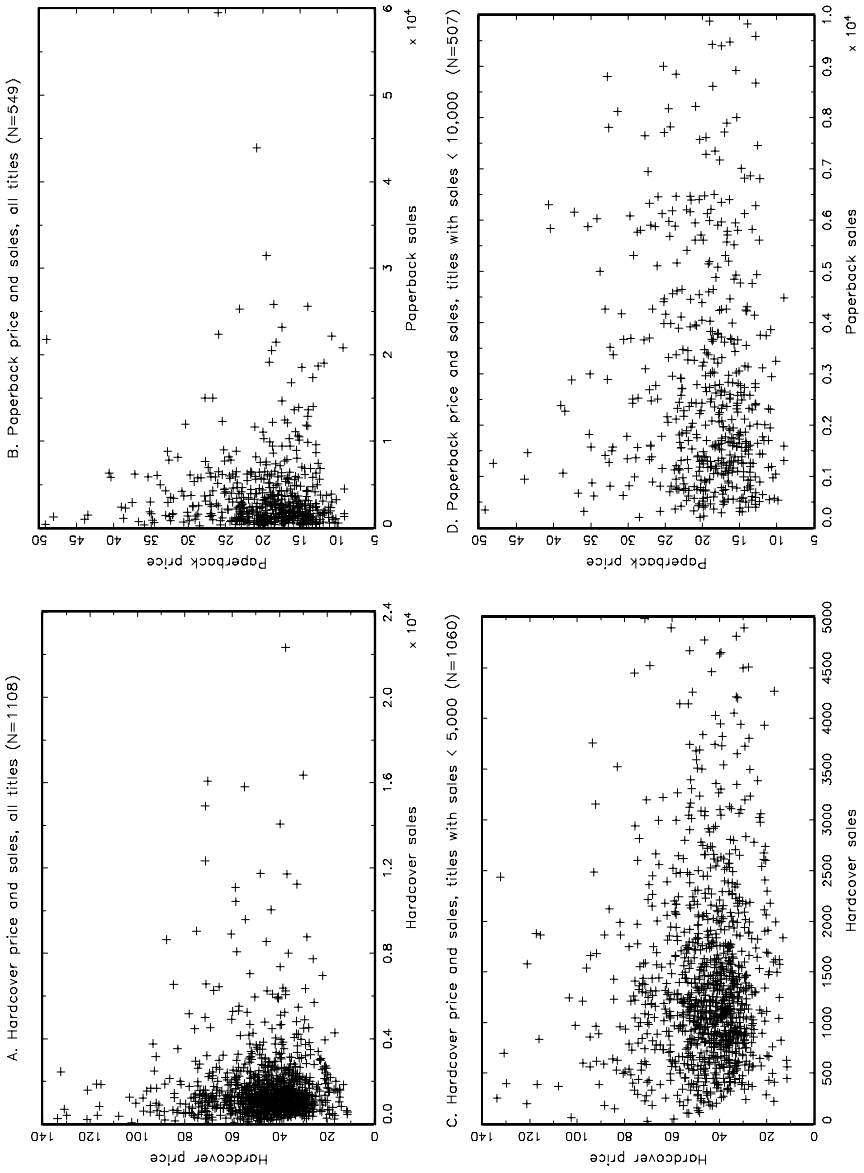

Figure 4: Profit premium of choosing SEQ over SIM for different discount rates of each type

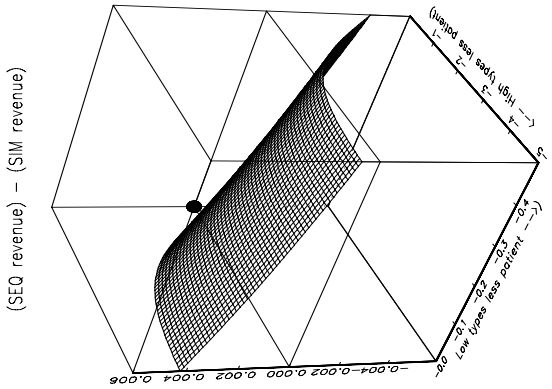

Figure 5: Profit premium of choosing SEQ over SIM for different ζ's and low types' ρ's

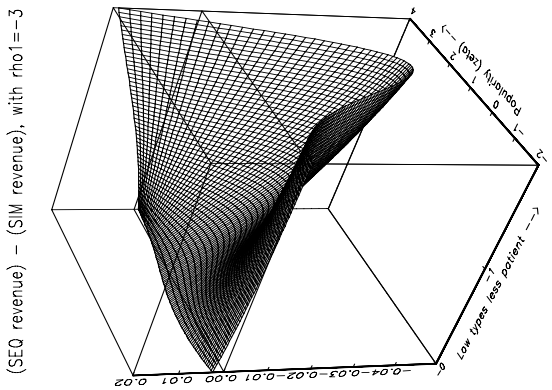

Figure 6: The distribution of consumers for a successful simultaneously published title

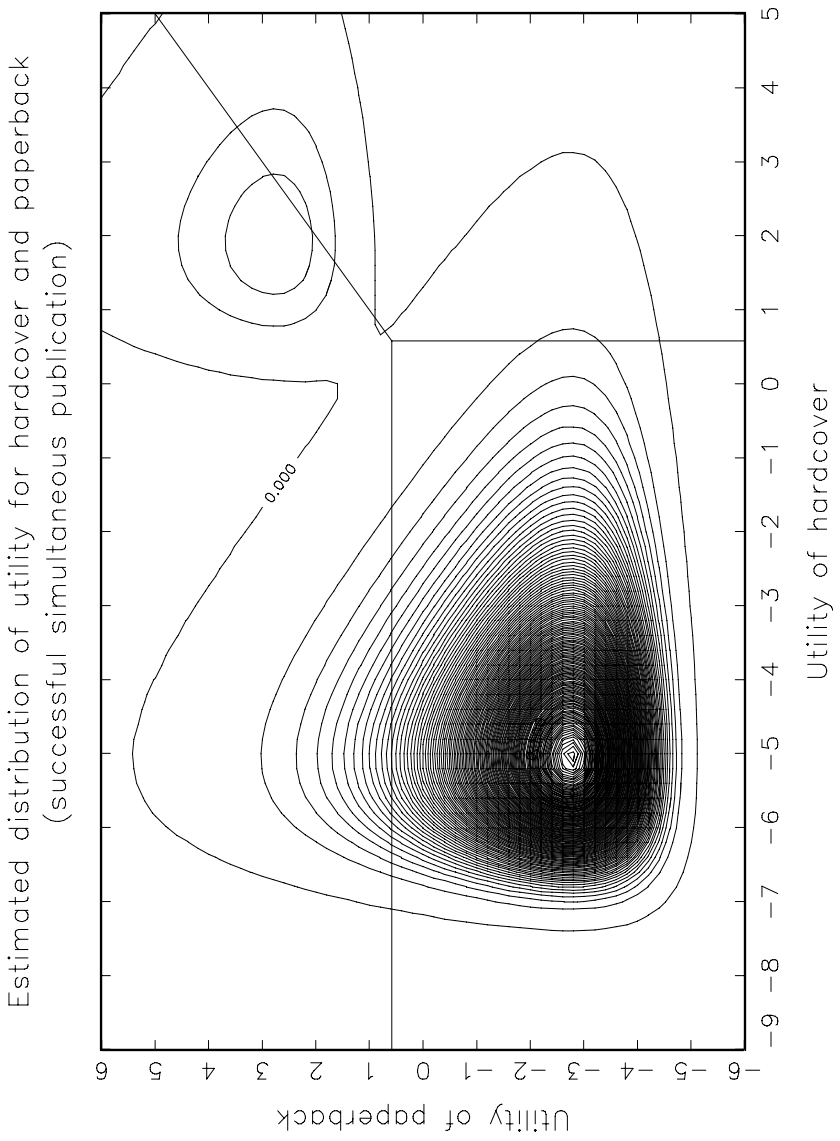

Table 2: Price regressions

	Hardcover Price		Paperback price			
			All titles		Sequential	
Intercept	47.30	(35.20)	20.41	(19.11)	-9.523	(20.59)
Prev. pubs	-.028	(.078)	.031	(.030)	.016	(.032)
Prev. pubs^2	.42e-3	(.46e-3)	.346e-3	(.188e-3)	.283e-3	(.205e-3)
New edition	.348	(1.783)	-.017	(.585)	.224	(1.434)
Authors	2.902*	(.830)	.279	(.325)	.340	(.326)
Simultaneous	1.222	(1.025)	.559	(.436)		
Hardcover only	2.252*	(.625)				
Paperback sold	-4.472*	(1.786)				
Pages	3.755	(2.846)	5.154*	(1.739)	6.501*	(2.491)
Pages^2	-.191	(.170)	-.474*	(.118)	-.497*	(.163)
Weight	19.33*	(5.94)	2.025	(4.796)	-3.197	(6.924)
Weight^2	-1.959*	(.823)	.380	(.917)	1.099	(1.257)
Weight/page	-7.360*	(3.551)	-3.039	(2.527)	-.640	(3.785)
(Weight/page)^2	.653*	(.290)	.871*	(.276)	.848*	(.389)
Pulp price	14.02	(8.39)	-2.533	(4.127)	-.632	(3.935)
UP price	-1.316	(1.970)	-1.265	(4.221)	3.960	(4.419)
Hardcover sales					-.193*	(.065)
Publication lag			.257*	(.105)	.232	(.315)
Publication lag^2					.0017	(.037)
Adj. R^2		.680		.772		.803
Observations		842		435		328
Time and category dummies also included.						
* denotes significance at the 5% level.						

Table 3: Sales regressions

	Hardcover		Paperback		Total sales	
Intercept	-8,160	(7,201)	36,547*	(18,133)	17,740	(20,872)
Prev. pubs	111.1*	(15.74)	69.09	(65.61)	162.8*	(46.55)
Prev. pubs^2	-.657*	(.091)	-.908*	(.386)	-1.146*	(.267)
New edition	-137.7	(365.2)	730.5	(1,288)	943.2	(1,066)
Authors	-67.50	(171.2)	404.1	(721.3)	-242.9	(500.0)
Simultaneous	-1170*	(195.7)	-5,183*	(1,382)	-490.1	(601.9)
Publ. lag			-3,365*	(785.5)		
Publ. lag^2			327.8*	(94.33)		
Hardcover only					-4,401*	(371.1)
Price	-19.75	(19.83)	-250.3	(359.4)	67.80	(58.02)
Price^2	-.175	(.170)	1.616	(7.410)	-.945	(.496)
Pages	403.6*	(194.3)	-1,662	(1,144)	-1,027	(567.0)
Pages^2	19.69	(18.80)	308.1*	(117.8)	192.6*	(55.03)
Weight/page	1,064*	(298.9)	834.9	(2,081)	1,522	(872.4)
(Wgt/page)^2	-96.36*	(35.68)	-100.1	(396.3)	-189.3	(104.1)
Pulp price	1776	(1720)	-16,185	(8,271)	2,214	(5,091)
UP price	392.3	(403.7)	-1,354	(3,685)	-2,665	(4,534)
Adj. R^2		.208		.167		.264
Observations		842		435		842

Time and category dummies also included.
* denotes significance at the 5% level.
Price in column 3 is hardcover price.

Table 4: Single-type estimates – nested logit

	OLS				2SLS			
Intercept	-3.027*	(.342)	-2.790*	(.332)	-2.406*	(.304)	-2.699*	(.291)
Prev. pubs	.0296*	(.0070)	.0257*	(.0069)	.0480*	(.0083)	.0479*	(.0081)
Prev. pubs^2	-.18e-3	(.395e-4)	-.155e-3*	(.39e-4)	-.26e-3*	(.448e-4)	-.26e-3*	(.43e-4)
New edition	.102	(.139)	-.0630	(.132)	-.136	(.235)	.00430	(.190)
Authors	-.110	(.0771)	-.0886	(.0761)	-.138	(.0733)	-.120	(.0712)
Pages (100s)	-.153	(.0919)	-.128	(.0905)	-.226*	(.0893)	-.217*	(.0866)
Pages^2	.024*	(.0105)	.0214*	(.0103)	.0230*	(.0081)	.0234*	(.0078)
Weight/page	.458*	(.168)	.503*	(.166)	.164	(.133)	.163	(.129)
(Wgt/page)^2	-.0642*	(.0240)	-.0680*	(.0236)	-.0177	(.0165)	-.0180	(.0161)
Hardcover	-.0518	(.144)	-.358*	(.128)	-1.531*	(.311)	-1.181*	(.200)
Price ($100s)	-.0048	(.0038)	-.00710*	(.0038)	.0129*	(.0063)	.0111	(.0061)
Lag (0/1)	.0949	(.108)			-.692*	(.280)		
Lag (years)			-.142*	(.0278)			-.125*	(.0488)
Within share	.838*	(.0490)	.743*	(.0478)	.591*	(.218)	.605*	(.207)
Adj. R^2		.407		.424		.181		.228
Observations		870		870		1277		1277
Time and category dummies also included.								
* denotes significance at the 5% level.								

Table 5: Counterfactual product creation

Actual choice	Counterfactual choice	Counterfactual $\delta_2^{\tau\prime}$
SIM	SEQ	$\delta_2^\tau - \rho^\tau \hat{T}_2$
SIM	HON	$-\infty$
SEQ	SIM	$\delta_2^\tau + \rho^\tau T_2$
SEQ	HON	$-\infty$
HON	SIM	$\delta_1^\tau - \beta_h^\tau + \beta_p^\tau(\hat{P}_2 - P_1) + \xi_1 - \xi_2$
HON	SEQ	$\delta_1^\tau - \beta_h^\tau + \beta_p^\tau(\hat{P}_2 - P_1) - \rho^\tau \hat{T}_2 + \xi_1 - \xi_2$

Table 6: Two-type estimates with firm's problem.
$\rho \sim exp(\mu_\rho)$

	Type 1	Type 2
Intercept	0.553 (0.572)	-3.880* (0.571)
Hardcover	0.030 (0.563)	-0.422 (0.791)
Price ($100s)	-0.053 (1.538)	-7.859* (3.005)
ρ^h (α_0)	2.814 (3.039)	
Pages (100s)	0.110 (0.089)	
Weight/page	-0.050 (0.081)	
Art	0.982* (0.265)	
Political Science	0.230 (0.153)	
Post-1986	-0.532* (0.103)	
μ_ρ	0.144 (0.158)	
ψ	0.760* (0.269)	
λ	0.023* (0.0062)	
Minimum	408.9	
Deg. of freedom	17	
Observations	1,636	

Table 7: Estimate Implications.

	Hardcover	Paperback
Type 1 shares, SIM:	0.20081	0.45677
Type 1 shares, SEQ:	0.67438	0.04004
Type 1 shares, HON:	0.47101	
Type 2 shares, SIM:	0.00200	0.03713
Type 2 shares, SEQ:	0.01042	0.03746
Type 2 shares, HON:	0.00127	
Type 1 sales, SIM:	465	1,058
Type 1 sales, SEQ:	1,562	93
Type 1 sales, HON:	1,091	
Type 2 sales, SIM:	195	3,627
Type 2 sales, SEQ:	1,018	3,659
Type 2 sales, HON:	124	
Number of type 1's:		2,317

Product choice: % Correct predictions	
When actual choice is SIM:	82.9%
When actual choice is SEQ:	65.7%

Distribution of ρ	
Mean	0.144
Mean given SIM	0.279
Mean given SEQ	0.079

Table 8: Price Elasticities.

	Hardcover	Paperback
Hardcover	-0.0110 -3.7640 -3.6772	0.0006 0.0029 0.0029
Paperback	0.0176 0.0007 0.0012	-0.0112 -1.4580 -1.4245

Fraction of elasticities greater than -1	
Hardcover:	0.000
Paperback:	0.080

SELECTED RECENT PUBLICATIONS

Bertaut, C. and Haliassos, M,. "Precautionary Portfolio Behavior from a Life - Cycle Perspective", *Journal of Economic Dynamics and Control,* 21, 1511-1542, 1997).

Caporale, G. and Pittis, N. "Causality and Forecasting in Incomplete Systems", *Journal of Forecasting*, 1997, 16, 6, 425-437.

Caporale, G. and Pittis, N. "Efficient estimation of cointegrated vectors and testing for causality in vector autoregressions: A survey of the theoretical literature", *Journal of Economic Surveys,* forthcoming.

Caporale, G. and Pittis, N. "Unit root testing using covariates: Some theory and evidence", *Oxford Bulletin of Economics and Statistics*, forthcoming.

Caporale, W., Hassapis, C. and Pittis, N. "Unit Roots and Long Run Causality: Investigating the Relationship between Output, Money and Interest Rates". *Economic Modeling*, 15(1), 91-112, January 1998.

Clerides, K., S. "Is Learning-by-Exporting Important? Micro-Dynamic Evidence from Colombia, Morocco, and Mexico." *Quarterly Journal of Economics* 113(3), pp. 903-947, August 1998, (with Lach and J.R. Tybout).

Cukierman, A., Kalaitzidakis, P., Summers, L. and Webb, S. "Central Bank Independence, Growth, Investment, and Real Rates", Reprinted in Sylvester Eijffinger (ed), *Independent Central Banks and Economic Performance,* Edward Elgar, 1997, 416-461.

Eicher, Th. and Kalaitzidakis, P. "The Human Capital Dimension to Foreign Direct Investment: Training, Adverse Selection and Firm Location". In Bjarne Jensen and Kar-yiu Wong (eds), *Dynamics,Economic Growth, and International Trade,* The University of Michigan Press, 1997, 337-364.

Gatsios, K., Hatzipanayotou, P. and Michael, M. S. "International Migration, the Provision of Public Good and Welfare", *Journal of Development Economics*, 60/2, 561-577, 1999.

Haliassos, M. "On Perfect Foresight Models of a Stochastic World", *Economic Journal*, 104, 477-491, 1994.

Haliassos, M. and Bertaut, C., "Why Do So Few Hold Stocks?", *The Economic Journal,* 105, 1110- 1129, 1995.

Haliassos, M. and Tobin, J. "The Macroeconomics of Government Finance", reprinted in J. Tobin, *Essays in Economics,* Vol. 4, Cambridge: MIT Press, 1996.

Hassapis, C., Pittis, N. and Prodromidis, K. "Unit Roots and Granger Causality in the EMS Interest Rates: The German Dominance Hypothesis Revisited", *Journal of International Money and Finance*, pp. 47-73, 1999.

Hassapis, C., Kalyvitis, S., and Pittis, N. "Cointegration and Joint Efficiency of International Commodity Markets", *The Quarterly Review of Economics and Finance,* Vol 39, pp. 213-231, 1999.

Hassapis, C., Pittis, N., and Prodromides, K. “EMS Interest Rates: The German Dominance Hypothesis or Else?” in *European Union at the Crossroads: A Critical Analysis of Monetary Union and Enlargement*, Aldershot, UK., Chapter 3, pp. 32-54, 1998. Edward Elgar Publishing Limited.

Hatzipanayotou, P. and Michael, M. S. "General Equilibrium Effects of Import Constraints Under Variable Labor Supply, Public Goods and Income Taxes", *Economica*, 66, 389-401, 1999.

Hatzipanayotou, P. and Michael, M. S. “Public Good Production, Nontraded Goods and Trade Restriction”, *Southern Economic Journal,* 63, 4, 1100-1107, 1997.

Hatzipanayotou, P. and Michael, M. S. "Real Exchange Rate Effects of Fiscal Expansion Under Trade Restrictions", *Canadian Journal of Economics,* 30-1, 42-56, 1997.

Kalaitzidakis, P. "On-the-job Training Under Firm-Specific Innovations and Worker Heterogeneity", *Industrial Relations,* 36, 371-390, July 1997.

Kalaitzidakis, P., Mamuneas, Th. and Stengos, Th. “European Economics: An Analysis Based on Publications in Core Journals*.” European Economic Review*,1999.

Lyssiotou, P., Pashardes, P. and Stengos, Th. “Testing the Rank of Engel Curves with Endogenous Expenditure”, *Economics Letters,* 64, 1999, 61-65.

Lyssiotou, P., Pashardes, P. and Stengos, Th. "Preference Heterogeneity and the Rank of Demand Systems", *Journal of Business and Economic Statistics,* Vol 17, No 2, April 1999, 248-252.

Lyssiotou, Panayiota, "Comparison of Alternative Tax and Transfer Treatment of Children using Adult Equivalence Scales", *Review of Income and Wealth*, Series 43, No. 1 March 1997, 105-117.

Mamuneas, T.P. (with Demetriades P.). “Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies”, *Economic Journal*, forthcoming.

Mamuneas T.P. (with Kalaitzidakis P. and Stengos T.). “A Nonlinear Sensitivity Analysis of Cross-Country Growth Regressions”, *Canadian Journal of Economics,* forthcoming.

Mamouneas T.P. (with Bougheas S. and Demetriades P.).”I´nfrastructure, Specialization and Economic Growth”, *Canadian Journal of Economics*, forthcoming.

Mamuneas, Theofanis P. “Spillovers from Publicly – Financed R&D Capital in High-Tech Industries”, *International Journal of Industrial Organization,* 17(2), 215-239, 1999.

Mamuneas, T. P. (with Nadiri, M.I.). “R&D Tax Incentives and Manufacturing-Sector R&D Expenditures”, in *Borderline Case: Interntational Tax Policy, Corporate Research and Development, and Investment*, James Poterba (ed.), National Academy Press, Washington D.C., 1997. Reprinted in *Chemtech*, 28(9), 1998.

Michaelides, A. and Ng, S. “Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of three Simulation Estimators”, *Journal of Econometrics,* forthcoming.

Pashardes, Panos. “Equivalence Scales in a Rank-3 Demand System”, *Journal of Public Economics*, 58, 143-158, 1995.

Pashardes, Panos. “Bias in Estimating Equivalence Scales from Grouped Data”, *Journal of Income Distribution*, Special Issue: Symposium on Equivalence Scales, 4, 253-264, 1995.

Pashardes, Panos. “Abstention and Aggregation in Consumer Demand”, *Oxford Economic Papers*, 46, 502-518, 1994 (with V. Fry).

Pashardes, Panos. “Bias in Estimation of the Almost Ideal Demand System with the Stone Index
Approximation”, *Economic Journal*, 103, 908-916, 1993.

Pashardes, Panos. “What Do We Learn About Consumer Demand Patterns From Micro-Data?”, *American Economic Review*, 83, 570-597, 1993 (with R. Blundell and G. Weber).

Pashardes, Panos. “Non-Linearities and Equivalence Scales”, *The Economic Journal*, 103, 359-368, 1993 (with R. Dickens and V. Fry).

Spanos, Aris " Revisiting Date Mining: ´hunting´ with or without a license”, forthcoming *Journal of Methodology,* July 2000.

Spanos, Aris. "On Modeling Heteroscedasticity: The Student's *t* and Elliptical Linear Regression Models", *Econometric Theory*, 10, 286-315, 1994.

Spanos, Aris. "On Normality and the Linear Regression Model", *Econometric Reviews,* 14, 195-203, 1995.

Spanos, Aris. "On Theory Testing in Econometrics: Modeling with nonexperimental Data", *Journal of Econometrics,* 67, 189-226, 1995.