THE PRICING OF ACADEMIC JOURNALS

Mark McCabe, Aviv Nevo, and Daniel L. Rubinfeld*

November 17, 2006

Abstract

In this paper we investigate the claim that academic journals are too expensive. We estimate library demand for academic journals and ask if short run profit maximization by publishers can explain observed prices. Libraries purchase a portfolio of journals so to estimate demand we extend the standard discrete choice model, and estimation methods, to allow for a choice consisting of a subset of a larger set of journals. Unlike the discrete choice model, the model allows for both positive and negative cross-price effects. We estimate the model using library holdings data and find that on average prices in the industry are lower than what static pricing models predict. Furthermore, we simulate the effects of mergers and find that the likely unilateral effect of a merger is to lower prices.

---

* Assistant Professor of Economics, Georgia Institute of Technology, Professor of Economics, Northwestern University, and Robert L. Bridges Professor of Law and Professor of Economics, the University of California, Berkeley, respectively. We wish to thank Kostis Hatzitaskos and Ben Stearns for their valuable research assistance, and the Andrew W. Mellon Foundation for financial support. We received valuable comments from seminars at University of Wisconsin, NYU's Stern School of Business, the Norwegian School of Economics and Business, the 2006 annual meeting of the European Economics Association, and Duke University.

1. INTRODUCTION

The prices of for-profit academic journals have increased rapidly over the past two decades (Albee and Dingley, 2001). For example, the price of library subscriptions for business and economics titles rose by 393 percent between 1984 and 2001, whereas the CPI grew at only 70 percent. This has lead to concerns from librarians and researchers regarding the implications for the dissemination of knowledge and the tension between for profits firms that receive free labor from academics (Bergstrom, 2001). Given that electronic transmission of knowledge is becoming increasingly important, an understanding of the reasons for the increases in journal prices is a vital element in the ongoing discussion of best mechanisms by which scholarly communications should be disseminated.

Previous work has offered several explanations for the increase in prices. One view credits a series of mergers that substantially increased concentration in the journal industry as the primary explanation for the rapid price increases (McCabe, 2002). Edlin and Rubinfeld (2004) emphasize the relatively recent effort by major publishers to bundle print and electronic journals. Both explanations are undoubtedly important, but they cannot tell the whole story. For one thing, bundling is a relatively recent innovation in the industry, and consequently cannot account for much of the historical price increases in the industry.

In our view, what is missing is an understanding of the strategic nature of the pricing of for-profit journals in a world whose primary customers are academic libraries. As agents of college and university faculties, libraries serve the interests of their principals and have only limited information about faculty journal demands. Unlike their libraries, individual faculty do not face budget constraints, and are therefore unlikely or unwilling to make difficult choices how to allocate funds between journals. As a result, libraries' demands for academic journals are likely to be inelastic. As the result of increasing journal prices and increasing budgetary pressures, libraries have been substituting journals for monographs and books, often without explicit faculty guidance or recognition.

In this paper we ask if the inelasticity of library demand can explain the seemingly high prices. In particular, we use library holding data to estimate the library demand model sketched out by Nevo, Rubinfeld, and McCabe (2005). Using the demand estimates we compute the profit maximizing markups under a variety of assumptions concerning journal pricing. We compute implied marginal costs by subtracting the computed markup from observed prices. We then compare the implied costs to rough estimates of costs, which allows us to check if firms set prices according to the profit-maximizing models. Finally, we use the estimated demand elasticities and computed marginal costs to simulate the likely effects of mergers.

Our exercise follows a long tradition in the industrial organization literature. However, we offer a different demand system and a unique way to estimate it. In our model, libraries choose to hold a subset of available journals. This choice cannot be modeled as a discrete choice, because libraries hold more than one journal, or as a choice among a continuum of options, because of journal heterogeneity (i.e., there is information in which journals are chosen and not simply the number of chosen journals). We therefore offer an extension of the standard discrete choice model to handle model library choice. Libraries rank all journals, as in the standard discrete choice model, but here they choose a group of journals representing their top choices. The number of chosen journals can be set a-priori, or, as we assume below, until a budget is exhausted.

We estimate the parameters of the model using information on library holdings and by choosing the parameters that maximize the likelihood of observed portfolios choices. Computation of the likelihood is difficult even in cases where the standard discrete choice model yields a closed form expression for the choice probabilities (for example, the well-known logit model). To compute the likelihood of the observed choices we sum over all the (unobserved) rankings that yield that choice. The probability of any ranking is straightforward to compute, but because there are many rankings we use a simulator estimator to evaluate this probability.

The results support two main findings. First, as we suspected we find that "high" prices are due to inelastic demand. Indeed, prices are lower than implied by most pricing models, given the demand estimates. This suggests that the real question is not why prices of journals have been so high, but why they were not higher. We do not believe that a change in demand elasticity can explain the rise in prices over the last two decades. Rather it is likely that firms were gradually moving their prices upward to reach the profit maximizing levels.

Second, unlike the standard discrete choice model, our model allows (indeed expects) choices to be complements as well as substitutes. The intuition is as follows. As a price of a journal decreases its ranking by all libraries increases. As a result, it is more likely to be chosen, which in turn makes other journals less likely to be purchased. This is the standard effect in discrete choice models, and it implies that options are substitutes. In our model, there is an additional effect. As the price of a journal decreases, libraries that already purchased the journal might now have sufficient funds to also purchase one of the other journals. This generates a positive cross-price effect. As a theoretical matter, we cannot say which effect dominates. To the extent that the second effect dominates and choices are complements, a merger can be expected to lower prices, not raise them as in the typical case. Indeed, when we simulate the likely effect of a number of mergers in the industry we find that the unilateral effect was to lower prices, not increase them as in the standard model.

We are not the first that examine extensions of the discrete choice model. Hendel (1999) studies firm's choices of computers. He observes firms simultaneously buying several different brands (and multiple units of each brand). He models each brand as satisfying the need of an independent task. Thus, each task chooses a unique brand and the firm is just an aggregation of tasks. In contrast we allow for interaction between the demand for each journal through a budget constraint. In his study of media outlets, Gentzkow (2006) models choices of various options: online and offline version of a newspaper. He models the choice as buy offline, online, both or neither. Thus, the econometric model is still a discrete choice model, but now one of the options is a bundle of both online and offline. Gentzkow's model is more general because he allows for interaction between the utility from the different alternatives. However, the possibility of modeling bundles as part of the choice set is feasible in his case because he only has a small number of options. This is not a realistic option for us. For the same reason the model of Allenby, Kim and Rossi (2002) is not applicable to our data.

The rest of the paper is organized as follows. Section 2 offers a brief overview of the journal publishing industry and its evolution. We begin the formal analysis in Section 3 with a model of library journal demand. Libraries are viewed as ranking journals and then buying a sufficient number of journals to exhaust their budgets. In Section 4 we offer a method to estimate model using data on holdings of business and economic journals by college and university libraries in Georgia. In Section 5, we present the demand estimates and ask whether not-for-profit prices are consistent with a short-run profit maximizing strategy by publishers. We then go on to evaluate the extent to which journal prices increases can be attributed to the consolidation in the industry. To do so, we use a merger simulation methodology to analyze the likely unilateral effects of journal mergers. Section 6 offers concluding remarks about the likely effects of the bundling of electronic and print journals and of changes in library purchasing policies.

2. THE INDUSTRY AND THE DATA

2.1 Journal Industry

The academic journal industry includes of a substantial number of both for-profit and not-for-profit firms. While there are numerous journal publishers of each type, consolidation over time has led to the vast majority of for-profit journals being owned by a a relatively small number of publishers: Elsevier (2211 journals), Springer (1574), Blackwell (863), and John Wiley (776) are among the five largest.[1] Many of the not-for-profit journals are sponsored by academic associations. Thus, the American Economic Association currently sponsors three journals and is in the process of significantly expanding its journals list.

While the format and editorial policies and production costs appear to be similar for both for-profit and not-for profit companies, for-profit prices and profit margins are higher, despite the fact that the top-ranked journals in many subject areas are published by not-for-profit publishers. Not-for-profit journals are priced less than for profits, and the divergence between the two has grown over time. Using our data set, we find that the average ratio of 2000 to 1990 prices for non-profits

---

[1] These journal totals were obtained from each publisher's website on November 14, 2006. Taylor & Francis' website reports that they publish "more than 1000 journals." Although it is not easy to determine the number of business and economics titles published by these firms, Thomson Scientific, a subsidiary of the Thomson Corporation, tracks citations to more than ten thousand scientific journals and categorizes these titles by field. In 2006, according to Thomson, Elsevier published 68 business and economics journals, Blackwell 41, Springer 27, Taylor & Francis 14 and John Wiley 7.

is 2.03 (for 75 titles), whereas the for-profit ratio is 3.77 (115 titles). There is heterogeneity across for-profits publishers. Some of the for-profits, such as Elsevier (the largest), have been aggressive in their efforts to increase prices, others, such as Blackwell have been less aggressive.

There have been two particularly significant changes in the journal industry over the past two decades. First, concentration of ownership has increased substantially, in part because the largest publishers have a disproportionate share of new journals, and in part because the publishers have grown through acquisitions. The largest commercial publishers – Elsevier, Blackwell, Springer (created by the merger of Kluwer-Academic and Bertelsmann-Springer), Taylor & Francis, and John Wiley – have all grown through acquisitions and/or through internal growth (starting new journals).[2]

The second change has been the move towards Internet distribution of journals. In the past five years, a substantial number of journals have gone online. Indeed, this shift has been so significant that Elsevier now claims to have the third largest Internet (sales) revenues behind only Amazon and AOL. The move towards electronic publishing has lowered the cost of publication and has consequently threatened to lower the cost-based barriers to entry into the journal publication industry. However, as Edlin and Rubinfeld (2004), have suggested, the use of a variety of bundling practices in the past five years appears to have created a strategic barrier to entry that is equally or more significant than the classic cost-based barriers. Because the library holdings dataset we study ends in 1998, the focus of our paper will be on the first phenomenon (acquisitions); we leave an empirical study of the second (bundling) to future work.

It is noteworthy that Elsevier acquired the JAI Press and Academic Press titles in 1997 and 2001, respectively. We will evaluate the likely predicted price effects of those two acquisitions in

[2] See Munroe (2004) for a detailed description of the history of commercial publisher acquisitions.

Section 4B of the paper.

2.2 The Data

Our analysis focuses on business and economics journals. We know price and characteristics of economics and business journals. We also know holdings and characteristics of libraries in the state of Georgia. The complete journal data set is a panel with price data for over 200 economics and business titles for the period 1988-2000; the holdings (and other data) are available for 86 academic and public libraries for the period 1988-1998. Journal characteristics and institutional prices for the period 1988-1998 were initially collected by the Department of Justice, and further updated by using data collected by Thomson Scientific (for characteristics) and price information reported by individual journals. Library holdings data were collected primarily from two online sources: the Online Computer Library Center's *Worldcat* database and the University System of Georgia's *GALILEO* library catalogs. Library characteristics were obtained from the National Center for Education Statistics.

We estimate the demand model using data for the year 1998. Because the holding date are stable over time, we did not think that adding additional years was justified. We return to the economic implications of this below. We restricted the sample to libraries with 5 or more titles in their collection and to titles with 5 or more subscriptions over all libraries. This yielded 40 libraries, thirty-four 4-year colleges and universities, five 2-year colleges, and one public library, and a total of 120 journals. Just over half of these titles (62) were published by 16 commercial firms. Fifty titles were published by just six firms: Elsevier (20 titles), Blackwell (13), Academic Press (5), Kluwer (5), John Wiley (4) and JAI Press (3). The remaining 58 titles were managed by 43 non-profit publishers, including the University of Chicago (5 titles), the Institute for Operations Research and Management Science (4), MIT (4), and Oxford (3).

We used three journal characteristics: the number of articles published in 1998, the number of citations in 1998 to articles published in each journal during the period 1994-1998, and the journal's first year of publication (to control for age). Library characteristics include expenditures on all serial subscriptions in 1998 and the (calculated) budgetary expenditures on business and economic journals.

Table 1 provides the mean, standard deviations, minimum, and maximum of each of the variables used in the analysis. The range of prices for individual journals is extremely wide. The average institutional price was $322 (for profits $594.30, non-profits $111.70), and the maximum was $1,769 (the Journal of Econometrics). The mean number of citations (in 1998 for articles published during 1994-1998) was 436 (for profits 341, non-profits 514). To the extent that citations are a measure of quality, these numbers suggest that even though for-profits journals are more expensive, on average they are of lower quality.

The average library holding was about 36 titles (of the 120 titles studied here); the University of Georgia had the largest number of titles, 119. An interesting feature of holding is that they are not proper subsets of each other. Roughly 45 percent of the average library's holding are not held by the library just above it in the ranking by the number of journals. If all holdings were subsets of each other than we would expect no cases like this. Overall, in the data there roughly 35 percent of holdings violate the subset conditions, suggesting that the smaller libraries have more violations. This has implications for the modeling approach below.

3. THE MODEL

The libraries in our dataset purchase multiple journals. Furthermore, as we saw in the previous section the holdings are not proper sets of each other: journal and library heterogeneity are important. Consequently, we cannot model the holding decisions as simply a decision on the number of journals to hold. We now propose a model that will capture the key features of the choice data.

Libraries are assumed to purchase journals so as to best satisfy their perception of the demands of academic faculty. Their budgetary decision is two-tiered. First, they allocate their annual budget between (a) journals and (b) books and monographs.[3] Second, within the journal budget category, libraries choose individual journals by ranking them according to both characteristics and price. They then purchase the top-ranked journals until their journals budget is exhausted.[4] We model this behavior following a framework similar to the discrete choice model (e.g., McFadden, 1981). Unlike traditional discrete choice demand models, we allow consumers to choose not just their top option, but their top *n* options, where *n*, the number of journals that fit into the budget, is endogenous.

Formally, let the "utility" a library $i \in \{1,...,I\}$ gets from journal $j \in \{1,...,J\}$ be

$$u_{ij} = x_j \beta_i - \alpha_i p_j + \epsilon_{ij} \tag{1}$$

where p_j is the price of the journal, x_j are observed journal characteristics (for example, sub-field and citations), and ϵ_{ij} is a random term. The vector β_i captures library-specific taste parameters for

[3] In the analysis below we focus on economics and business journals, so there is an additional layer of allocating the journals budgets between fields.

[4] We assume also that no journals are offered as a bundle, which was the case during the estimation period.

observed characteristics. The parameter α_i accounts for price sensitivity in the rank; it can be considered the shadow price of the library budget constraint.

The library-specific taste parameters, β_i and α_i are modeled as

$$\begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} + AD_i + \nu_i$$

where D_i are observed library attributes (for example, characteristics of the institution and the faculty), A is a parameter matrix, and ν_i is a vector of unobserved library specific taste shocks.

The library has a budget B_i for buying journals. The budget is given by the first tier of the demand analysis, and is taken as exogenous to the journal demand model. A library will purchase journals by going down the list of journals (ranked according to u_{ij}) and buying all journals until it cannot afford the next journal on the list. Define the cutoff quality, u_i^*, for each library, as

$$u_i^* = argmax \sum_{k=1}^{J} p_k \cdot 1(u_{ik} \geq u_i^*) \quad s.t. \; \sum_{k=1}^{J} p_k \cdot 1(u_{ik} \geq u_i^*) \leq B_i \text{ and } u_i^* \in \{u_{i1}, ..., u_{iJ}\}$$

where 1(·) is the indicator function. The library then purchases all journals such that $u_{ij} \geq u_i^*$.

We ignore issues of non-divisibility. To see the importance of ignoring non-divisibility, consider the following: the first *k* journals cost more than the library's budget, but the first *k*-1 journals *and* the *k*+1 journal cost less than the budget. Our model suggests that the library buys only the first *k*-1 journals, but one might imagine a different model in which the library would also buy the *k*+1 journal. More generally, the solution to our choice problem does not necessarily correspond to the solution to the choice problem where the libraries choose journals to maximize utility subject to a budget, or other, constraint.

The model directly generalizes the standard discrete choice model. The equivalent of the outside, or no choice, option arises in our model when the prices of all options are higher than the budget. In this case no journals are purchased.

Finally, we assume that there are no interactions in utility between journals. One could imagine that, for example, the utility from the Journal of Econometrics depends on whether Econometric Theory is also purchased. The model could in principle deal with this extension but requires an adjustment of the estimation method discussed below.

While the library demand model follows closely the standard discrete choice model the price effects are richer. As the price of a journal, say the Journal of Economic Theory (JET), increases, there are two possible effects. First, the price increase might lower the ranking of JET, possibly dropping it below the cutoff level and promoting the purchase of another journal. This is the usual cross-price effect found in standard discrete choice models. However, if JET is ranked sufficiently high by a particular library, the demand for JET by that library will be unaffected, even if it drops below another alternative, since JET will still be above its budgetary cutoff. This does not imply that demand will be less price sensitive than the standard discrete choice model. The price sensitivity is determined not only by the shape of the distribution of ϵ_{ij}, as in the discrete choice model, but also by the distribution of budgets.

In our model there is an additional cross-price effect. If the budget is fully expended on journals, an increase in the price of JET will crowd out a marginal journal, say (solely for purposes of discussion) the Journal of Public Economic Theory (JPET). If JET is ranked higher than JPET, a library will continue to purchase JET, despite the price increase, but the library will no longer purchase JPET. Thus, a positive cross price effect is possible in the model

For a given ranking and a given budget this model generates non-smooth price responses: no effect on demand almost everywhere, with occasional discrete jumps in demand as a marginal journal is dropped when there is no slack in the budget. With heterogeneity in either budgets or the ranking, given by a continuous distribution, the demand is smoothed out and is well-behaved.

In order to address the effect of acquisitions on pricing we pose a standard static pricing model of journal publishing. Suppose there are F for-profit publishers, each publishing some subset, $\mathscr{F}_f$, of the $\boldsymbol{j=1,...,J}$ different journals. The profits of publisher $\boldsymbol{f}$ are

$$\Pi_f = \sum_{j \in \mathscr{F}_f} (p_j - mc_j)\, q_j(p) - C_f,$$

where $\boldsymbol{q_j(p)}$ is the quantity sold of journal $\boldsymbol{j}$, which is a function of prices of all brands, $\boldsymbol{mc_j}$ is the constant marginal cost of production, and $\boldsymbol{C_f}$ is the fixed cost of production.

The first-order conditions with respect to price are

$$q(p) - \Omega^{pre}(p)(p - mc) = 0,$$

where

$$\Omega^{pre}_{jr}(p) = \begin{cases} -\partial q_j(p)/\partial p_r, & \text{if } \exists f\text{: } \{r,j\} \subset \mathscr{F}_f; \\ 0, & \text{otherwise.} \end{cases}$$

These J equations imply markups and marginal costs for each journal

$$p - mc = \Omega^{pre}(p)^{-1} q(p) \quad \Rightarrow \quad mc = p - \Omega^{pre}(p)^{-1} q(p).$$

We will use these equations in several ways. First, in we use estimates of the demand system to compute the implied marginal costs. We will then evaluate the extent to which, the actual prices of for-profit journals are consistent with the markups implied by the demand model. Second, we simulate the likely effects of acquisitions on pricing. To simulate the effects of acquisitions on price, we define Ω^{post} in the same way we defined Ω^{pre} using the post-acquisition structure of the industry. The predicted post-acquisition equilibrium price, p^*, solves

$$p^* = \hat{mc} + \Omega^{post}(p^*)^{-1} q(p^*), \qquad \textbf{(2)}$$

where $\hat{mc}$ are the marginal costs implied by the demand estimates and the pre-acquisition ownership structure.

4. Estimation

Our goal is to estimate the parameters, β, α, the matrix A, and the parameters of the distribution of v and ϵ, denoted $F(\nu,\epsilon)$,using the library holding data discussed in the previous section. We first outline the estimation algorithm and then briefly discuss one alternative method.

We propose to estimate the parameters of the model by maximizing the probability of a library choosing its observed portfolio. Assuming independence of the errors across libraries, the log likelihood is given by

$$LL(\theta) = \sum_{i=1}^{I} Pr(P_i | X, A_i, B_i, \theta) = \sum_{i=1}^{I} Pr(u_{ij} \geq u_i^* \;\; \forall j \in P_i \text{ and } u_{ik} < u_i^* \;\; \forall k \in J \backslash P_i | X, A_i, B_i, \theta)$$

where P_i is the observed portfolio held by library i, the journals not included in $\boldsymbol{P}$ is denoted by $\boldsymbol{J \backslash P}$, X is a matrix capturing the characteristics of all journals, θ denotes the parameters of the model and u_i^* is defined above.

This is the direct analog of maximum likelihood estimation in discrete choice models.

However, generally we do not know how to analytically compute this probability.[5] Alternatively, the likelihood can be computed using simulation. Indeed, our initial effort involved taking several thousand draws from $F(\nu, \epsilon)$, and computed the probability of the observed portfolio (by looking at the fraction of simulation draws for each library that predicted the portfolio as the choice). Unfortunately, this approach proved cumbersome when library portfolios contain a large number of journals , since the number of possible portfolios is very large many observed portfolio choices never happened in the simulation. Simulation methods perform notoriously poor when used to compute very low probability events. Therefore, we substituted a more efficient alternative.

The (marginal) probability of choosing a given portfolio is the sum of the probabilities of choosing the portfolio given a particular ranking, summing over all possible rankings. Because there are many possible rankings, we use simulation to compute the sum. Specifically, we choose a small number of rankings of the chosen journals, and then for each ranking, we compute the probability of choosing the portfolio, average the probability over the chosen rankings, and multiply by the number of possible rankings.

Formally, let ***P*** denote a particular portfolio (subset) of the journals and let #(***P***) denote the number of journals in this subset. Denote the journals not included in P by J*P*. Assume that the ranking of the journals within the chosen portfolio is known. Label the highest ranked journal as 1, second as 2, and so forth up to the number of elements in the portfolio, $\#(P)$. Then

$$Pr(ranking) = Pr\left(1 \geq 2 \ldots \geq \#(P) \geq J \backslash P\right) = \Pi_J(1)\, \Pi_{J\backslash 1}(2) \cdots \Pi_{J\backslash \{1,2,\ldots,\#(P_i)-1\}}(\#(P)) .$$

where $\Pi_P(k)$ denotes the probability that option k is ranked above all other options in the choice set

[5] Below we discuss an alternative that allows us to analytically compute the likelihood if the number of alternatives is small.

P. The probabilities $\Pi_P(k)$ are given by the standard discrete choice formulas, and can be computed analytically or by simulation.

For example, if the chosen portfolio includes 10 journals and v and ϵ are independently distributed with the distribution of ϵ being iid extreme value and the distribution of v denoted by $F(\nu)$, then

$$Pr(ranking) = \int \frac{\exp(x_1\beta_i - \alpha_i p_1)}{\sum_{k=2}^{J}\exp(x_k\beta_i - \alpha_i p_k)} \cdot \frac{\exp(x_2\beta_i - \alpha_i p_2)}{\sum_{k=3}^{J}\exp(x_k\beta_i - \alpha_i p_k)} \cdot \ \ldots \ \frac{\exp(x_{10}\beta_i - \alpha_i p_{10})}{\sum_{k=11}^{J}\exp(x_k\beta_i - \alpha_i p_k)} \, dF(\nu)$$

Since the ranking is not known, we sum the probabilities of all possible rankings that yield a choice of the observed portfolio:

$$Pr(P|X,A_i,B_i,\theta) = \sum_{k=1}^{(\#P)!} Pr(ranking_k|X,A_i,B_i,\theta).$$

The summation is over all the $(\#P)!$ permutations of the elements in the observed portfolio. For large portfolios, computing the above probability might not be computationally feasible. We note that

$$\sum_{k=1}^{(\#P)!} Pr(ranking_k|X,A_i,B_i,\theta) = (\#P)! \cdot \frac{1}{(\#P)!}\sum_{k=1}^{(\#P)!} Pr(ranking_k|X,A_i,B_i,\theta) \equiv (\#P)! \cdot \bar{\Pi}(X,A_i,B_i,\theta).$$

We compute $\bar{\Pi}(X,A_i,B_i,\theta)$ by simulation. We draw L ranking from the set of possible ranking we use these to compute an estimate of the average probability, i.e.,

$$\hat{\Pi}(X,A_i,B_i,\theta) = \frac{1}{L} \cdot \sum_{k=1}^{L} Pr(ranking_k|X,A_i,B_i,\theta) \ .$$

In Monte Carlo experiments we have found that even with a low L the estimator performed well. We compute standard errors by the bootstrap method, or by estimating the information matrix by the outer product of the score of gradients of the likelihood function. In cases where we computed both the differences were small and therefore below we report only the results based on the outer product of the gradients.

We also explored an alternative estimation method. A key building block in this alternative is the so-called Block-Marschak polynomial. k in **J\P** is the Block-Marschak polynomial, $K_{k,P}$ is the function

$$K_{k,P} = \sum_{i=0}^{\#(P)} (-1)^{\#(P)-i} \sum_{(Q \subseteq P \text{ and } \#(Q)=i)} \Pi_{J \setminus Q}(k) \qquad \textbf{(3)}$$

where $\Pi_P(k)$ denotes the probability that option k is chosen if **P** is the choice set. Barberá and Pattanaik (1986) provide an interpretation of $K_{k,P}$ as the probability of the event that k is ranked behind the elements of **P** and ahead of the remaining elements in **J\P**. This suggests that

$$Pr(\text{choosing portfolio } \mathbf{P}) = \sum_{\{k \in J \setminus P \text{ and } p_k + \sum_{j \in P_i} p_j \geq B_i\}} K_{k,P}. \qquad \textbf{(4)}$$

Using this formula we can compute the desired likelihood analytically , as long as the probabilities $\Pi_P(k)$ can be computed analytically. The problem is that generally this would require computing a very large number of probabilities (the exact number depends on J and #(**P**)). The expression can be simplified somewhat by defining an artificial option that replaces all the options in **J\P** . This still leaves us with many probabilities to compute. While each of these is rather simple to compute (its just a logit probability) there are large number of these. For example, if #(**P**)=20 then there are over one million probabilities to compute.

We explored reducing the number of probabilities to be computed by noting that what we need to compute is $\sum_{Q \subseteq P \& \#(Q)=i} \Pi_{J \setminus Q}(k)$. By definition this will be equal to the number of probabilities we are summing over times the average probability. The number of elements is easy to compute. And we explored computing the average probability by simulation. We choose with equal probability (and with replacement) sets Q(which are subsets of **P** with $\#(Q)$ = i). For each we analytically compute $\Pi_{J\backslash Q}(k)$ and then average across these draws.

We found that generally this method performed poorly even with a large number of simulation draws because the simulated probabilities, in many cases, were either negative or greater than one. We therefore do not report results using this method.

5. Results

5.1 Demand

As mentioned earlier, the model is estimated using one year of the Georgia data (1998). The observed characteristics include the logarithm of the number of papers published in title *j* in 1998, the number of citations in 1998 to articles published in j during 1994-1998, and the first year of publication. We expect the likelihood of journal choice to be positively related to the number of papers in the journal and to the number of citations. We also expect the age coefficient to be negative, since the younger, less entrenched journals are less likely to be demanded by libraries. In all specifications we assume that ϵ_{ijt} are distributed *i.i.d.* extreme value. In different specification we allowed different degrees of library heterogeneity in the taste parameters. The most restrictive specifications assume no heterogeneity in the taste parameters, i.e., $\beta_i = \beta$ and $\alpha_i = \alpha$. We then allow for observed heterogeneity by interacting the coefficients with the library's budget. Further, we allow for unobserved heterogeneity, or random coefficients, by letting the ν terms to be normally distributed with mean zero and a standard deviation to be estimated. The results are given in Table 2.

Note that in all cases the estimated effect of price is negative, and statistically significantly different from zero. The effects of citations is positive and significant, as expected. The papers coefficient is positive in all but one specification (in which is it is insignificant). The coefficient on first publication year is negative: newer publications are valued less than older, more established journals. When we allowed for heterogeneity through budget interactions, we found, as one would expect, that higher budget libraries are significantly less price sensitive than their lower-budget counterparts. All other budget interaction coefficients are statistically insignificant.

Expanding the model to allow for random coefficients did not substantially change any of the proceeding results. The estimated standard deviations of the distribution of the price and citations coefficients were generally significant, except when we allowed for the interaction of all coefficients. Our preferred specifications eliminates three of the budget interactions and two of the standard deviations, with little effect on the log likelihood.

In order to evaluate the economic magnitude of the parameter estimates, we compute the price elasticities of demand. To do so we use the demand estimates to simulate the purchasing behavior of 6000 libraries. For each simulated library we draw an random term, ϵ_{ijt}, which is distributed *i.i.d.* extreme value for each of the 120 journals. We also draw a random normally distributed random coefficient, ν_{ik}, for each library and each characteristic. The library characteristics, for these 6000 libraries, were generated from the characteristics of the 40 libraries in the sample by replicating each 150 times.

To simulate the aggregate demand for journals, we need to assume a distribution for the library budget so that we can compute which journals the library can afford to buy. Note, that this distribution was not needed for the estimation. Using the observed expenditures of the Georgia libraries on business and economics titles, we estimated three different cumulative distributions:

exponential, with parameter equal to 0.6975, log normal, with parameters 8.2 and 1.1284, and a non-parametric empirical distribution function.

The derivatives of the demand system were computed numerically.

The computed price elasticities of demand are given at the top of Table 3. As we speculated, demand is highly inelastic in all three budget distributions. Furthermore, most of the cross-price elasticities are positive. In Section 3 we discussed the intuition that leads to this result, which can occur in our model but not in the standard discrete choice model. Since the reason the cross-price effects arises is different than the typical demand complementarities story (think cereal and milk) there is no reason that the cross price effects will be positively correlated with distance in some characteristics space.

In the section that follows, we will use the demand parameters to evaluate two economic implications of the estimates: the pricing strategies of the journal publishers and to simulate the effects of acquisition strategies. Before we do so we discuss several caveats associated with our estimates.

In estimating the parameters we assumed that the distribution of the error-terms was independent of the observed characteristics. Berry (1994) discusses, in the context of standard discrete choice model, how characteristics unobserved to the researcher, but observed to the market participants, can impact the estimates of price sensitivity. A similar problem can arise here, although in this specific market the bias might go in a direction opposite to what one might expect.

Assume that equation (1) includes a term ξ_j that is unobserved to us but observed to the publishers. In Table 1 we see that not-for-profits have higher citation counts, which may indicate higher quality. Indeed, most of the top ranked journals are published by not-for-profits. Therefore, the not-for profits are likely to have higher unobserved quality. However, they also have lower prices. Consequently, it is not unreasonable to expect a negative correlation between prices and the unobserved term in this market, whereas in other markets we would expect this correlation to be positive.

An additional factor might also suggest that if there is any bias that our estimates are biased towards finding too much price sensitivity. The holdings data are from 1998. However, as mentioned previously, there was little variation in the holding between 1988 and 1998, even though the prices of journals nearly doubled. This suggests that there is a significant amount of state dependence. There are many economic reasons for this. Indeed, when we estimated the model including a lagged holding variable, the lagged variable completely overwhelmed the other variables: it was highly significant and all the other coefficients essentially disappeared. Therefore, the time variation suggests that libraries are even less price sensitive than the cross-sectional variation that was used to estimate the model.

Both of these effects suggest that if anything our estimates over-estimate the price sensitivity and in reality libraries are even less price sensitive. We discuss the economics implications of this below.

5.2 Analysis of Publisher Pricing

Given the estimated library demand parameters, we now evaluate journal pricing. Our goal is to answer two questions: (1) How are for–profit journal publishers setting prices generally? (2) What effect does the acquisition by one publisher of a set of journals of another publisher have on journal pricing?

5.2.1 Are Publishers Profit Maximizing?

As previously described, we use the demand estimates to compute the markups implied by the first-order conditions of various pricing models. Subtracting these markups from the observed prices we calculate marginal costs under various assumptions on the ownership structure. If for-profit publishers were using any of these pricing models we would expect to find reasonable implied marginal costs.

We compute the implied markups using three distributions of budgets. The results are presented in columns of Table 3. Rows present the results from various pricing models, which differ in the ownership structure. The models we examine include single product ownership, multi-product firms, collusion of for-profit journals, and joint profit maximization of all journals. For each model we present average markups (p-mc) and margins ((p-mc)/p), for all journals and with respect to for-profits journals.[6]

Since the estimated own-price elasticities are less than 1, it is not surprising that the computed margins are more than 100 percent, which implies negative marginal costs. This suggests that observed prices are not consistent with profit maximization. This result is consistent across the different distributions. The margins of for-profit publishers are predicted to be lower, despite higher markups. But even for these journals the implied margins are too high to be consistent with observed prices.

The margins continue to be high, and inconsistent with observed prices, in the other pricing models. In all cases, except the average of for-profits margins under monopoly pricing and an exponential distribution of budgets, the margins are over 100 percent. Note that predicted margins decline the more journals are jointly maximizing profits. This is due to the large number of negative

[6] For the purpose of these statistics we did not treat Blackwell as profit maximizing.

cross-price elasticities. In the standard discrete choice model, where the cross-price effects are positive, more products in the joint maximization implies higher predicted markups.

How can one explain the failure of journal publishers to profit maximize, i.e., to increase prices to the point at which demand becomes elastic? It is, of course, possible that our conclusion would change if we were to relax some of the assumptions that were built into the particular specifications utilized here. As we report in the table, we experimented with different budget distributions. We also explored different functional forms for characteristics. Neither of these change the findings that observed prices cannot be rationalized.

Finally, we considered the possibility that the estimates are biased. As we discussed in the previous section, the bias is likely to make our estimates conservative, i.e., libraries are even less price sensitive than we find. In order to rationalize the observed prices using the multi-product pricing model, the price coefficient has to be larger in absolute value by an order of magnitude. It therefore seems some unreasonable that econometric bias can explain our finding. The results appear to support the view that academic journal prices were if anything too low during the period 1988-1998!

In our view, the period of analysis is one in which publishers (a) developed a more sophisticated understanding of the strategic nature of journal pricing, and (b) learned about the demand elasticities of their library customers. The market or markets for journals are complex; most journals do not substitute closely if at all with other journals, and the price sensitivity of demand varies substantially from library to library.

The choice of the profit-maximizing price for a journal is not an easy one. Moreover, the period of analysis was one in which there was substantial growth in the number of journals offered to the public. Publishers were likely cautious in pricing these new journals, to introduce faculty and libraries to the new products and to grow demand over time. Finally, publishers may have found it advantageous to increase prices steadily over time; sharp increases at any point in time bear with them the risk of backlash. Adverse public reaction by libraries could have led libraries to search for cost-reducing strategies (e.g., cutting down the number of copies of journals, reducing the demand for journals directly, or indirectly as libraries find ways to share with other libraries).[7] Adverse reaction by faculty could reduce readership and citation rates, both of which would further reduce demand. Because we have not modeled these possibilities explicitly, we cannot rule out the possibility that pricing during the 1990s was profit-maximizing in the long-run.

5.2.3. The Unilateral Effects of Acquisition

Simulating the likely effect of a merger follows standard merger simulation (for example, Nevo (2000)). The difference stems from the different demand system we introduce and the implications it has for cross-price effects. Merging firms will raise their prices post-merger because they internalize the substitution between their products (or decrease their prices if the products are complements). Since our supply model is the standard model, the same is true here. What is different is the way the cross-price effects are generated. Consider two journals: A and B. If journal A raises its price, there will be two effects. First, some libraries might change their ranking of A relative to other journals. In some cases this change in ranking might be enough for some libraries to no longer purchase journal A. Out of those that no longer purchase A some might decide to purchase B instead. This is the standard substitution effect found in traditional unilateral effects

[7] See for example, "Soaring Prices Spur a Revolt in Scientific Publishing," Dec. 8, 1998, New York Times, and "As Publishers Perish, Libraries feel the Pain," Nov. 3, 2000, New York Times.

analysis. Second, if A raises its price, some libraries might decide to continue purchasing it, but will no longer purchase journal B since they cannot afford it. This creates a negative price effect between the two journals; thus generating complementarities between the products.

Now consider a merger of these two journals. The two effects just described impact in opposing directions: the first will generate higher prices post merger, while the second will lead to lower prices. In reality whichever effect dominates will determine if prices go up, down or stay the same post merger. In traditional merger simulation, the size of the cross-price effect, and therefore the impact of the merger, is directly related to how "close" the products are in some characteristics space. That is also the case for the first effect we discuss above: if two products are close in ranking it is more likely that a library that drops one will purchase the other. However, for the second effect closeness is not important. It is more likely to be relevant when one journal is infra-marginal and the other marginal. (Thus, our analysis will apply equally if our previous example refers to JIE and the International Journal of Industrial Organization rather than JET and JPET.)

We performed a number of empirical analyses to evaluate the likely unilateral effects of acquisitions. Our primary goal is to demonstrate the effect of the demand system on merger simulation, and not to make specific predictions about the actual effects of the mergers analyzed. To compute the post-merger price we use the marginal costs predicted in the previous section and solve equation (2). Solving equation (2) by allowing the substitution matrix to vary turned out to be computationally difficult. We therefore left the substitution matrix in its pre-merger price levels. Thus, the only change between Ω^{pre} and Ω^{post} is in the ownership structure. In all cases we left the price of non-for-profits unchanged by the merger, and computed the new post-merger price charged by the for-profits.

The results are presented in Table 4. In different columns we use different budget distributions to generate the cross-price effects. Within each distribution the results differ in the treatment of Blackwell. In the left column we assume that Blackwell does not maximize profits, and thus does not change prices in response to the merger. In the right-hand column we go to the other extreme and assume that Blackwell acts as a for-profit and maximizes profits. As we see in Table 1, the truth is somewhere in between, and therefore the results can be seen as bounding the true effects.

In different rows we simulated two acquisitions: Elsevier's acquisition of JAI and Elsevier's acquisition of Academic Press. We present the predicted price changes for these acquisitions. The effects for the second acquisition are the compounded effects, i.e., the effects of both acquisitions relative to the pre-acquisition observed prices.

We find consistently that the unilateral effect of the mergers is to decrease the price of the merging parties and to slightly increase the prices of non-merging parties. This is consistent with the negative cross-price effects we estimated and shows that the theoretical effect we discussed is empirically relevant.

We note a few caveats to our findings. In the previous section we found it difficult to rationalize observed prices with the price models we examined. Therefore, we would be hard pressed to claim we have realistic predictions on the effects of mergers if we use the same pricing models. Indeed we see this exercise as mainly demonstrating the effect of mergers in the model, and not a realistic prediction of the likely effect of a merger. As we previously noted, it is hard to imagine that rising prices are driven by a change in demand elasticities over time. Instead, we conjectured that the increase in prices was due to for-profit firms moving towards the profit maximizing price. Whatever the constraints are to publishers changing prices too quickly, mergers might ease the process. In other words, mergers might still cause an increase in prices, as claimed by McCabe (2002). However, the mechanism is not the standard unilateral effect; rather it is an effect through the impact on some pricing constraint. Without a model of this constraint and the impact of concentration on it, the effect of mergers cannot be predicted.

6. CONCLUDING REMARKS

We estimate demand of libraries for economics journals. As expected, demand elasticities are low, indeed sufficiently low as to be generally inconsistent with short-run profit maximizing behavior by for-profit journal publishers. We see this as consistent with a period in which publishers were learning about the nature of the demands of individual libraries, and were being conservative in their pricing, so as not to encourage libraries to search aggressively for cost-reducing alternatives.

We also explain why (absent efficiencies) the acquisition of one publisher's journals by another may in theory lead to either higher or lower prices. Our simulations provide direct evidence of the surprising result that journal acquisitions can indeed lead to lower journal prices.

During the period 1988-1998, there was little or no bundling of print and no print and electronic bundling.[8] However, in the past six years, publishers have engaged in aggressive bundling policies. Our results offer some insights into the likely effects of bundling. Suppose that a publisher with market power were to follow a pure bundling strategy (and there is no price discrimination). To keep the example simple, suppose also that the publisher owns two journals,

[8] For example, commercial access to Elsevier's digital journal database, Science Direct, began in February, 1999. Initially, access to Science Direct required maintenance of a set of print subscriptions; over time, as use of electronic journals has become more widespread, this connection between print and electronic has been severed.

A and B and that initially the publisher has chosen profit-maximizing prices for the two journals without accounting for bundling. Suppose further that distribution costs are relatively low and unaffected by the bundling decision. Because A is ranked higher than B, B is the marginal journal for some libraries. As we pointed out previously, it is the potential loss of revenue on the sale of B which effectively constrains the pricing of the publisher. Now assume that the publisher bundles journals A and B so that the purchase of journal A is conditioned on the purchase of journal B. Assuming that the bundle is sufficiently highly valued to be non-marginal for most libraries, journals owned by other publishers will become marginal and a price increase will become unambiguously profitable. The equilibrium implications of such a bundling strategy remain to be developed. We conjecture that bundling can lead to higher prices. Indeed, recent practices of major publishers of bundling print and electronic journals can be seen as a form of price discrimination (Edlin and Rubinfeld (2002), Jeon and Menicucci (2006)).

We believe that the approach put forward in this paper can serve as the basis for an analysis of a number of policy issues. First, it can provide the framework for an analysis of the effects of current and prospective bundling strategies. Second, our analysis could be extended to incorporate endogenous library budgets, and to allow for an evaluation of the price changes that result when libraries alter the means by which they rank journals (e.g., by creating an incentive for faculty to face real prices when making recommendations for journal acquisition and maintenance). Third, the framework could be extended to provide an answer to the question of whether the number of academic journals is optimal. This important question faces many academic societies that are in the midst of deciding whether to encourage and/or support the introduction of new not-for-profit journals.

References

Albee, Barbara and Brenda Dingley (2001), "U.S. Periodical Prices - 2001," *American Libraries* 2, available at http://www.ala.org/al_onlineTemplate.cfm?Section=selectdarticles&Template=/ContentManagement/ContentDisplay.cfm&ContentID=7282.

Allenby, Greg, J. Kim and Peter Rossi (2002), "Modeling Consumer Demand for Variety," , *Marketing Science*, 21, 229-250.

Barberá, Salvador and Prasanta K. Pattanaik (1986), "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," 54 *Econometrica*, No. 3, pp. 707-715.

Bergstrom, Theodore C. (2001), "A Free Labor for Costly Journals?", 15 *Journal of Economic Perspectives,* No. 4, Fall, pp.183-198.

Edlin, Aaron S. and Daniel L. Rubinfeld (2004), "Exclusion or Efficient Pricing? 'Big Deal' Bundling of Academic Journals," 72 *Antitrust Law Journal*., pp. 119-157.

Gentzkow, Matthew (2006), "Valuing New Goods in a Model with Complementarity: Online Newspapers," *American Economic Review,* forthcoming.

Hendel, Igal (1999), "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," 66 *Review of Economic Studies,* April, pp. 423-446.

Jeon, Doh-Shin and Domenico Menicucci (2006), "Bundling Electronic Journals and Competition Among Publishers," *Journal of the European Economic Association,* forthcoming..

McCabe, Mark J. (2002), "Journal Pricing and Mergers: A Portfolio Approach," 92 *American Economic Review,* pp. 259-269.

McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data,* pp. 198-272, Cambridge: MIT Press.

Munroe, Mary (2004), "The Academic Publishing Industry: A Story of Merger and Acquisition," available at http://www.niulib.niu.edu/publishers (last accessed 9/16/06).

Nevo, Aviv (2000), "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," 31 *The RAND Journal of Economics*, No. 3, pp. 395-421.

Nevo, Aviv, Daniel L. Rubinfeld, and Mark McCabe (2005), "Academic Journal Publishing and the Demand for Libraries," *American Economic Review, Papers and Proceedings*, pp. 447-452.

Table 1: Summary Statistics

Variable	mean	min	max
Price ($)	322.0	25	1,769
For-profit	594.3	75	1,769
Blackwell	233.5	75	571
Not-for-profit	111.7	25	455
Citations	436.5	10	3,044
For-profit	341.2	5	1,429
Blackwell	451.8	62	1,870
Not-for-profit	513.5	54	3,044
Papers	47.4	5	261
For-profit	55.1	8	261
Blackwell	39.3	7	86
Not-for-profit	42.8	5	179
% share (of 40 libraries)	29.6 (11.9)	12.5 (5)	97.5 (39)
% held (of 120 titles)	29.6 (35.6)	4.2 (5)	99.2 (119)
Budget ($)	8433	462	37,793
For-profit	8706	0	28,899
Blackwell	861	0	3,035
Not-for-profit	2,494	309	6,481
Journal Age in 1998 (yrs)	38	8	112

Table 2: Estimated Parameters

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(price)	-0.52 (0.02)	-0.38 (0.03)	-0.63 (0.09)	-0.53 (0.10)	-0.84 (0.07)	-0.71 (0.08)	-0.71 (0.08)
log(papers)	0.16 (0.05)	0.02 (0.05)	0.15 (0.07)	0.01 (0.06)	0.27 (0.12)	0.08 (0.12)	-0.01 (0.06)
log(citations)	0.33 (0.03)	0.29 (0.03)	0.34 (0.10)	0.40 (0.05)	0.51 (0.06)	0.45 (0.06)	0.36 (0.06)
log(first year)		-23.85 (2.14)	–	-24.96 (2.24)	–	-28.96 (2.82)	-24.97 (2.12)
budget*log(price)			–	–	0.25 (0.05)	0.22 (0.05)	0.22 (0.05)
budget*log(papers)			–	–	-0.14 (0.15)	-0.08 (0.20)	–
budget*log(citations)			–	–	-0.12 (0.09)	-0.11 (0.09)	–
budget*log(first year)			–	–	–	5.38 (5.16)	–
std of price coeff			0.31 (0.06)	0.41 (0.09)	0.15 (0.07)	0.18 (0.12)	0.23 (0.08)
std of papers coeff			0.05 (0.18)	0.01 (0.17)	0.00 (0.89)	0.01 (0.15)	–
std of citations coeff			0.36 (0.07)	0.22 (0.04)	0.14 (0.08)	0.12 (0.07)	0.25 (0.05)
std of first year coeff			–	0.93 (1.19)	–	0.64 (2.52)	–
log likelihood:	1711.6	1661.8	1679.5	1622.6	1645.3	1592.1	1607.5

Using 40 Georgia libraries, 120 journals, 200 simulated rankings.

Table 3: Elasticities, Margins and Markups

dist of budgets:	exponential		log normal		empirical	
variable:						
mean (own price elas)	-0.43		-0.41		-0.32	
% of cross-price <0	91.8		91.3		91.3	
single product	**markup($)**	**margin (%)**	**markup($)**	**margin (%)**	**markup($)**	**margin (%)**
mean all journals	672.3	252.9	724.6	254.8	970.5	321.3
mean for-profits	1124.5	196.6	1259.6	219.1	1732.4	298.9
multi product						
mean all journals	581.64	238.0	623.4	240.0	728.2	288.0
mean for-profits	917.9	171.3	1025.6	190.2	1163.0	229.3
for-profits collusion						
mean for-profits	666.7	117.5	752.8	130.9	660.4	118.1
monopoly						
mean all journals	315.3	104.8	353.0	109.3	325.8	107.2
mean for-profits	565.0	99.8	656.8	113.8	588.5	105.2

Based on column (7) in Table 2.

Table 4: The Unilateral Effect of Mergers of Prices

dist of budgets:	exponential		log normal		empirical	
variable:	**% change**	**% change**	**% change**	**% change**	**% change**	**% change**
Elsevier's acquisition of JAI						
merging parties	-4.23	-4.25	-4.33	-4.35	-7.79	-7.82
non-merging for-profits	0.50	0.48	0.49	0.46	1.37	1.32
Blackwell	–	0.40	–	0.37	–	0.75
Elsevier's acquisition of Academic Press						
merging parties	-9.72	-9.80	-9.85	-9.92	-18.10	-18.27
non-merging for-profits	2.04	1.95	2.08	1.98	8.15	7.91
Blackwell	–	1.50	–	1.46	–	3.46
Blakwell maximizes profits	no	yes	no	yes	no	yes

Based on column (7) in Table 2.