Academic Journal Pricing and the Demand of Libraries

By AVIV NEVO, DANIEL L. RUBINFELD, AND MARK MCCABE*

The prices of for-profit academic journals have increased rapidly over the past decade (Barbara Albee and Brenda Dingley, 2001). There remains substantial debate as to the explanation for these increases. Among those put forward are the increased concentration of the journal industry (see e.g., McCabe, 2002) and the relatively recent effort by major publishers to bundle print and electronic journals (Aaron S. Edlin and Rubinfeld, 2004). While both explanations are undoubtedly important, what is missing is the significant role of the primary customers of journal publishers—the academic libraries. As agents of college and university faculties, libraries serve the interests of their principals while having only limited information about faculty journal demands. Facing little or no hard budget constraint, faculty are unlikely or unwilling to make difficult allocative choices. As a result, libraries have been making hard choices for years (between journals and books, and among journals), in a world of increasing budgetary pressure.

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 can be disseminated. In this paper, we formulate a model of library journal demand and suggest how it can be used to analyze the optimal pricing of journals by publishers. This represents part of a larger project whose long-range goal is to explain the pattern of journal pricing over time, and to evaluate two broad policy questions: (i) To what extent have mergers and/or bundling practices been responsible for the increase in journal prices? (ii) To what extent are journal prices likely to change in response to changing library acquisition strategies (reallocation of budgets between serials and books, journal sharing among libraries, setting hard budgets, etc.)?

I. The Model

Libraries purchase journals on behalf of the users of the journals: the readers. They do so by ranking the journals according to both characteristics and price, and purchasing the top-ranked journals until their budget is exhausted. We model this behavior following a framework similar to the discrete choice model (Daniel McFadden, 1981). Unlike more 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 $i \in \{1, \dots, I\}$ gets from journal $j \in \{1, \dots, J\}$ be

$$u_{ij} = \mathbf{x}_j \boldsymbol{\beta}_i - \alpha_i p_j + \varepsilon_{ij}$$

where p_j is the price of the journal, $\mathbf{x}_j$ are observed journal characteristics (e.g., subfield and citations), and ε_{ij} is a random term. The vector $\boldsymbol{\beta}_i$ captures library-specific taste parameters for observed characteristics. The parameter α_i accounts for price and can be considered the shadow price of the library budget constraint.

The library-specific taste parameters, $\boldsymbol{\beta}_i$ and α_i are modeled as

$$\begin{pmatrix} \boldsymbol{\beta}_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta} \\ \alpha \end{pmatrix} + \mathbf{A}\mathbf{D}_i + \mathbf{v}_i$$

where D_i are observed library attributes (for example, characteristics of the institution and the faculty), $\mathbf{A}$ is a parameter matrix, and $\mathbf{v}_i$ is a vector of unobserved library specific taste shocks.

The library has an exogenous budget B_i for

* Nevo: Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2600; Rubinfeld: School of Law and Department of Economics, University of California, Berkeley, CA 94720; McCabe: Department of Economics, Georgia Institute of Technology, 781 Marietta St., Atlanta, GA 30318. We thank Kostis Hatzitaskos and Ben Stearns for their valuable research assistance, and the Andrew Mellon Foundation for financial support.

buying journals. 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

$$u_i^* = \operatorname{argmax} \sum_{k=1}^{J} p_k \times 1(u_{ik} \geq u_i^*)$$

subject to

$$\sum_{k=1}^{J} p_k \times 1(u_{ik} \geq u_i^*) \leq B_i$$

$$u_i^* \in \{u_{i1}, \dots, u_{iJ}\}$$

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

While the model is straightforward, the price effects are unique. 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. If JET is ranked sufficiently high, the demand for JET will be unaffected, since JET will still be above the cutoff. Second, assuming that 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.

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 and bundling on pricing we pose a standard static pricing model of journal publishing. Suppose there are F for-profit publishers, each publishing some subset, $\mathcal{F}_f$, of the $j = 1, \dots, J$ different journals. The profits of publisher f are

$$\Pi_f = \sum_{j \in \mathcal{F}_f} (p_j - \mathrm{mc}_j) q_j(p) - C_f$$

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

The first-order conditions with respect to price are:[2]

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

where

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

These J equations imply markups and marginal costs for each journal:

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

We use these equations in several ways. First, we use estimates of the demand system to compute the implied marginal costs. Second, in order 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

$$p^* = \widehat{\mathrm{mc}} + \Omega^{\mathrm{post}}(p^*)^{-1} q(p^*)$$

where $\widehat{\mathrm{mc}}$ are the marginal costs implied by the demand estimates and the pre-acquisition ownership structure. Third, we will simulate

[1] This basic model ignores ties and issues of nondivisibility. An extension of this model would include the purchase of other materials by the library (e.g., monographs). This would imply that even if the total library budget is exogenous, the amount used to purchase journals will be endogenous.

[2] The pricing behavior of nonprofit journals can be treated in a similar way by assuming they set prices equal to average costs.

the optimal pricing assuming journals are bundled.

II. Data and Estimation

We estimate the demand model using a panel data set of price data for over 200 economics and management journals for the libraries of all two- and four-year Georgia colleges and universities for the period 1988–2001.[3] We merged the pricing data with time-series information on library holdings, journal quality (as measured by the number of articles and the number of citations), and a budgetary measure (the total library budget of each institution).

We estimate the parameters, $\boldsymbol{\beta}$, α, the matrix $\mathbf{A}$, and the parameters of the distribution of $\mathbf{v}$ and $\boldsymbol{\varepsilon}$, denoted $F(\mathbf{v}, \boldsymbol{\varepsilon})$, by maximizing the probability of a library choosing its observed portfolio. Assuming independence of the errors across libraries, the log likelihood is given by

$$\mathrm{LL}(\theta) = \sum_{i=1}^{I} \Pr(\mathcal{P}_i | \mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta})$$

$$= \sum_{i=1}^{I} \Pr(u_{ij} \geq u_i^* \ \forall j \in \mathcal{P}_i \text{ and}$$

$$u_{ik} < u_i^* \ \forall k \in \mathcal{J} \backslash \mathcal{P}_i | \mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta})$$

where $\mathcal{P}_i$ is the observed portfolio held by library i, $\mathbf{X}$ is a matrix capturing the characteristics of all journals, $\boldsymbol{\theta}$ denotes the parameters of the model, and u_i^* is defined above.

Computation of this likelihood is difficult. We have explored two approaches. First, we explored a simulation approach. For each library we made several thousand draws from $F(\mathbf{v}, \boldsymbol{\varepsilon})$, 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).

[3] We chose not to include data on K–12 institutions. We expect to use a similar data set for Minnesota to test more complex versions of this model.

Unfortunately, this approach proves cumbersome when library portfolios contain a large number of journals. 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 $\mathcal{P}$ denote a particular portfolio (subset) of the journals and let $\#(\mathcal{P})$ denote the number of journals in this subset. Denote the journals not included in $\mathcal{P}$ by $\mathcal{J} \backslash \mathcal{P}$. 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, $\#(\mathcal{P})$. Then

$$\Pr(\text{ranking})$$

$$= \Pr(1 \succeq 2 \cdots \succeq \#(\mathcal{P}) \succeq \mathcal{J} \backslash \mathcal{P})$$

$$= \Pi_{\mathcal{J}}(1) \Pi_{\mathcal{J} \backslash 1}(2) \cdots \Pi_{\mathcal{J} \backslash \{1,2,\ldots,\#(\mathcal{P})-1\}}(\#(\mathcal{P}))$$

where $\Pi_{\mathcal{P}}(k)$ denotes the probability that option k is ranked above all other options in the choice set $\mathcal{P}$. The probabilities $\Pi_{\mathcal{P}}(k)$ are given by the standard discrete choice formulas and can be computed either analytically or by simulation.

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

$$\Pr(\mathcal{P} | \mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta})$$

$$= \sum_{k=1}^{(\#\mathcal{P})!} \Pr(\text{ranking}_k | \mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta}).$$

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

$$\sum_{k=1}^{(\#\mathcal{P})!} \Pr(\text{ranking}_k|\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta}) = (\#\mathcal{P})!$$

$$\times \frac{1}{(\#\mathcal{P})!} \sum_{k=1}^{(\#\mathcal{P})!} \Pr(\text{ranking}_k|\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta}) \equiv (\#\mathcal{P})!$$

$$\times \bar{\Pi}(\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta}).$$

We compute $\bar{\Pi}(\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta})$ by simulation. We draw L rankings from the set of possible rankings, and we use these to compute an estimate of the average probability, that is,

$$\hat{\Pi}(\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta})$$

$$= \frac{1}{L} \sum_{k=1}^{L} \Pr(\text{ranking}_k|\mathbf{X}, \mathbf{A}_i, B_i, \boldsymbol{\theta}).$$

In Monte Carlo experiments even with $L = 1$ the estimator seemed to perform well.

We estimated a simple version of the model using one year of the Georgia data (1998), and a set of libraries that held at least five titles. We assumed no heterogeneity in the taste parameters (i.e., $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ and $\alpha_i = \alpha$) and that ε_{ijt} are independently and identically distributed, extreme value. The observed characteristics included number of papers published in the journal per year and the number of citations to articles published in the journal.[4] The estimated effect of price is negative and statistically significantly different from zero. The effects of citations and number of papers are both positive and significantly different than zero as well. All these coefficients are of the expected signs. In the empirical work that follows, we will use these parameters to compute demand price elasticities and to simulate the effects of acquisition and bundling strategies, as we describe below.

[4] We used data for 43 colleges and universities with holdings of five or more titles (121 titles were represented in the subsample). The average institutional price was \$319.57 (for profits, \$518.62; nonprofits, \$110.40), and the maximum was \$1,769 (the *Journal of Econometrics*, an Elsevier journal). The mean number of citations (measured over a five-year window) was 433.28 (for profits, 364.35; nonprofits, 505.71). The average holding was 40.4 titles, with the maximum holdings, 202, represented by the University of Georgia.

III. Analysis of Publisher Pricing

Given the estimated library demand parameters, it is possible to evaluate the nature of journal pricing. Our ultimate goal is to answer several questions: (i) How are for-profit journal publishers setting prices generally? (ii) What effect does the acquisition by one publisher of a set of journals of another publisher have on journal pricing? (iii) How will journal publishers respond if libraries alter their purchasing behavior? (iv) What is the likely effect of bundling on journal pricing?

Using the estimated parameters one can address these questions empirically. Here we discuss informally some of the potential effects on pricing of acquisition and bundling.

A. *The Unilateral Effects of Acquisition*

In principle, one can use the model discussed in Section I to simulate the likely price effects of a hypothetical acquisition in a similar way to traditional merger simulation (e.g., Nevo, 2000). The difference stems from the different demand system we introduce and the implications it has for cross-price effects. In a traditional unilateral effects analysis, the 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 separate 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 analysis. Second, if A raises its price, some libraries might decide to continue purchasing it, and 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, which 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 be apply equally if our previous example refers to JET and the *International Journal of Industrial Organization* rather than JET and JPET.)

B. *Bundling*

We now analyze the likely effects on pricing if a publisher with market power were to follow a pure bundling strategy (and there is no price discrimination).[5] Suppose as before that a publisher owns two journals, A and B, and that initially the publisher has chosen profit-maximizing prices for the two journals without accounting for bundling. Suppose also 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 saw in the previous section, it is the potential loss of revenue on the sale of B that 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. However, several implications seem likely in an oligopolistic world with a competitive fringe (the nonprofits). First, the bundling of some or all print journals is likely to be an equilibrium strategy for all firms with substantial market power. Second, such a bundling strategy may in some cases deter the entry of new for-profit journal publishers, whose journals are more likely to be marginal in a bundling world than in a non-bundling world.

IV. Concluding Remarks

We estimate a basic model of the demand of libraries for economics journals. Using the basic model structure, we have explained why (absent efficiencies), the acquisition of one publisher's journals by another may in theory lead to either higher or lower prices. We have also explained why bundling is likely to unambiguously lead to higher journal prices.

Where do we go from here? We are currently applying and extending the analysis in several ways. First, we are estimating demand functions that allow for more library heterogeneity. Second, we are simulating the likely unilateral price effects of various acquisitions. Third, we plan to simulate the likely equilibrium price effects of various bundling strategies.

Our approach can serve as the basis for an analysis of a number of policy issues. First, publishers have begun to bundle print, electronic, and both print and electronic journals. What likely price effects can be expected as bundling strategies evolve? Second, how does the analysis of pricing and bundling change when one accounts for endogenous library budgets, and if 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, what are the efficiencies that result from acquisitions and/or journal bundling? Accounting for efficiencies, what are the likely effects of each of these strategies? Fourth, are there too many journals? If so, what are the

[5] For an explanation of why the advent of electronic publishing has made bundling a more appealing strategy for publishers, see McCabe (2004).

social welfare implications if individual journals exit the industry?

REFERENCES

Albee, Barbara and Dingley, Brenda. "U.S. Periodical Prices—2001." *American Libraries*, 2001, *2* [available online: ⟨http://www.ala.org/al_onlineTemplate.cfm?Section=selectdarticles&Template=/ContentManagement/ContentDisplay.cfm&ContentID=7282⟩].

Edlin, Aaron S. and Rubinfeld, Daniel L. "Exclusion or Efficient Pricing? 'Big Deal' Bundling of Academic Journals." *Antitrust Law Journal*, 2004, *72*(1), pp. 119–57.

McCabe, Mark J. "Journal Pricing and Mergers: A Portfolio Approach." *American Economic Review*, 2002, *92*(1), pp. 259–69.

McCabe, Mark J. "Information Goods and Endogenous Pricing Strategies: The Case of Academic Journals." *Economics Bulletin*, 2004, *12*(10), pp. 1–11.

McFadden, Daniel. "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden, eds., *Structural analysis of discrete data.* Cambridge, MA: MIT Press, pp. 198–272.

Nevo, Aviv. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." *RAND Journal of Economics*, 2000, *31*(3), pp. 395–421.