Collection Sales: Good or Bad for Journals?

Mark Armstrong*
Department of Economics
University College London, UK

April 2008

Abstract

This note discusses the impact of collection sales (i.e., the bundling of several journals for sale by publishers to libraries) on journals. The advent of electronic journal distribution implies that bundling is an efficient sales strategy, and can act to extend the reach of a journal. Current arrangements are discussed and shown to lead to tensions between commercial publishers and non-profit journals. The note argues that non-profit journals should not abandon their participation in collection sales programmes. Rather, non-profit journals may benefit from withdrawing from commercial publishers which distribute their own for-profit journals, and from joining together to be distributed by less commercial publishers who set relatively low prices for their collections.

Keywords: Journal pricing; bundling; price discrimination.

JEL classification: D82, L31, L42, L82.

1 Introduction

Publishers perform several valuable tasks which an academic journal cannot usually do well on its own, such as marketing, negotiations with libraries, maintaining a website, and so on. When it acts for several journals, one function a publisher can perform is *bundling*. In this note I discuss the pros and cons of bundling journals for sale to libraries (or "collection sales"). The perspective is very much that of an individual journal, rather than overall welfare. For instance, bundling may be used as a vehicle for exclusion of new

*I am grateful to Ted Bergstrom, Carli Coetzee, Richard Gedye, Martin Green, Jim Hosek and Chris Wallace for helpful comments. All remaining errors and views are my own responsibility.

journals or rival publishers, but this is not discussed here.[1] In particular, the concern about possible high bundle prices from publishers in this note is only that a journal's reach may not be as large as it desires, rather than that libraries may be exploited.

Library expenditure on scientific journals has significantly increased in recent years.[2] This has been the result of two main factors: significant merger activity among publishers, and the increased use of bundling as a marketing device as facilitated by the electronic distribution of journals. Regarding the former, Dewatripont, Ginsburgh, Legros, and Walckiers (2007) show that journal prices tend to be higher in those scientific fields where publishers are more concentrated. However, this note focuses on the second reason for inflation in library expenditure. The introduction of bundling as a business model is a powerful force for price rises, especially when combined with a concentrated market structure in which publishers each distribute many journals. When a publisher distributes many journals, packaging these journals as a bundle acts significantly to reduce the dispersion of libraries' willingness-to-pay for the bundle compared with their willingness-to-pay for individual journals. This reduces the asymmetry of information between publisher and library, which has two effects: first, it means that more libraries will gain access to a given journal (a positive efficiency effect), and second, the publisher will better be able to extract a library's willingness-to-pay for the bundle (a monopoly exploitation effect). It is quite possible that bundling acts to improve overall welfare, but transfers surplus from libraries to publishers (and then possibly onto the journals themselves).[3]

Journals differ in their form of ownership and their objectives. In this note, I say that a journal is "for-profit" if it aims only to maximize its profit. Usually, such journals are owned by commercial publishers. The main alternative discussed here is that of "non-

[1] See Edlin and Rubinfeld (2004) and Jeon and Menicucci (2006) for discussion of these wider issues in the context of journal publishing, and Armstrong (2008) for a general overview of the possible anti-competitive impacts of bundling.

[2] Edlin and Rubinfeld (2004) report that during the period 1984–2001 the price of library subscriptions to periodicals in law, medicine and physics rose respectively by 205%, 479% and 615%. (Overall inflation in the United States was 70% over this period.) However, the advent of widespread bundling has meant that library expenditure *per journal* has fallen from a peak in 2000. See page 13 in the Association of Research Libraries' *ARL Statistics 2005-06* (published 2008, Washington, DC.).

[3] The impact of bundling many products is parallel to so-called first-degree price discrimination. The textbook monopolist of widgets sets price above marginal cost because it does not have accurate information about its buyers' willingness-to-pay for its product. The result is that demand is inefficiently restricted but buyers are left with some surplus. If the monopolist had accurate information about each buyer's willingness-to-pay, it would charge each buyer her reservation price, with the result that demand is expanded to the efficient level but buyers are left with none of the gains from trade.

profit" journals, which are assumed instead to wish to maximize their "reach" subject to a break-even constraint.[4] (Reach can be measured in various ways, but in this note we take it to be the number of libraries which have access to the journal.) Such journals are usually owned independently of commercial publishers, or owned by less commercial publishers (such as some university presses). In practice, the line between for-profit and non-profit journals is not always so easy to draw. For instance, a journal which does not have profit as its ultimate objective may benefit when it receives more revenue, as a better-financed journal may end up publishing better articles (for instance, by being more efficient at processing submissions or by paying authors when they publish a paper in the journal), and this may maximize its reach or impact in the long term.[5] Nevertheless, in the formal model in this paper I ignore these subtleties, and suppose that journals are exogenously partitioned into for-profit and non-profit journals.

Subscription prices can vary hugely between these two kinds of journal. Bergstrom (2001) reports that in 2000 the six most-cited economics journals were all non-profit journals, and their (stand-alone) library subscription prices were then an average of $180. Five of the top 20 most-cited journals were owned by commercial publishers, and their average subscription price was then $1660 per year.[6]

[4] A third kind of journal, related to the non-profit journals, is the category of society journals. Perhaps the main difference between a society journal and other non-profit journals is that the former may wish to ensure there is still a market for personal journal subscriptions, since this historically has been a major reason why people become members of the society. (Other methods are requiring society membership to submit articles or to attend an annual conference.) A wholehearted bundling policy – where essentially all potential members can electronically access the journal for free via their library – would be likely to decimate personal subscriptions and so reduce society membership. Thus, all else equal, a society journal is more likely to be suspicious of bundling policies. Of course, the same point applies also to site licenses for a single journal (where all members of an institution can electronically access the journal when the library has a subscription), but less strongly. With a site license, a journal will lose most of its personal subscribers at subscribing institutions, but with bundling as well it will lose most of its personal subscribers at those institutions which take the collection (as well as those at institutions with full library subscriptions).

[5] This note does not address why journals have the aims they do. To do this, we would need to model the author side of the market as well. Authors, all else equal, prefer a wide-reaching journal to one with more limited readership. This may be because they like the idea of many people reading their paper, or because wide-reaching journal is typically a prestigious journal. (More readers will, all else equal, result in more citations and a higher impact factor. And of course causation goes the other way too, and readers are more likely to read prestigious journals.) Thus, a journal's ultimate aim may be to publish the "best papers", and the way to attract to best papers is to have a wide readership. See Jeon and Rochet (2007) for more detailed discussion about how journal pricing (both to institutions and to submitting authors) and the quality of articles interact.

[6] See Dewatripont, Ginsburgh, Legros, and Walckiers (2007) for more recent data for a much wider range of scientific disciplines, which also shows that for-profit journals typically are significantly more expensive than non-profit journals.

The rest of the note goes as follows. In the next section, I discuss current arrangements for collections sales, and how this system leads to various tensions for non-profit journals. In section 3, some of the options for the future of journal pricing and remuneration are discussed. In section 4 a formal model is presented in an attempt to analyze the long-run evolution of the market. Here, bundling is assumed to be widespread and publishers are assumed to compete for journals along two dimensions: the *revenue* they offer journals and the *reach* of their collection to libraries. For-profit journals care only about the former dimension, while non-profit journals care only about the latter (as long as they receive enough revenue to cover their fixed costs). Publishers are assumed to make lump-sum payments to attract journals, as opposed to current arrangements which are a messy compromise between stand-alone subscription payments and *ad hoc* revenue allocations from collection sales. In this framework, all journals prefer to participate in collection sales programmes, but the two kinds of journals are distributed by publishers with different (induced) pricing strategies: for-profit journals are attracted to (or are owned by) publishers who offer them high revenue but relatively low reach, while non-profit journals go to publishers who market their collections with relatively low prices and high reach.

2 Tensions in Current Publishing Arrangements

There are a number of methods for pricing institutional subscriptions and remunerating journals. Historically, all journals were disseminated only in hard copy and bundling was not significantly used.[7] The result was an extremely simple and transparent process for setting subscription prices and allocating revenue to journals. Each journal sets the price a library must pay to subscribe to the journal, regardless of which other journals the library buys. (There may be some explicit price discrimination, say to developing countries.) An independent journal typically has the authority to set its subscription price (with the advice of the publisher), and so the journal has full control over its reach and revenue. For non-profit journals we expect to see the quality of the journal being inversely related to its stand-alone subscription price, at least so long as the fixed cost of running a journal does not vary much with quality.[8] An important feature of this

[7] Of course, though, a journal can itself be regarded as a bundle of individual articles.

[8] Suppose a high-quality journal has library demand $Q_H(p)$ with subscription price p, while a lower-quality journal has demand $Q_L(p) < Q_H(p)$. If both have the same fixed cost of operation, then

"non-bundled" market is that a journal does not really care about which other journals its publisher markets, or how it markets them. So long as its publisher performs its marketing tasks diligently and offers a competitive rate for its services, a journal (for-profit or non-profit) is indifferent about its publisher's interactions with other journals since that has no impact on libraries' decisions about whether to subscribe.

However, the move to electronic distribution of journals makes bundling an efficient and profitable strategy. Since it costs virtually nothing for a publisher to supply a library with an extra journal, a publisher will wish to strike a deal with a library to take all the journals to which it does not already have full subscriptions. Current publishing arrangements are a somewhat messy mixture of stand-alone full subscriptions and "top-up fees" for the right to access electronically the remaining journals in the collection. In more detail, a publisher obviously knows which libraries have full institutional subscriptions to its various journals. The publisher may then offer a library access to all the publisher's additional journals (to which the library does not currently subscribe) for an additional top-up fee.[9] Thus, a journal to which the library already subscribes continues to receive its subscription fee, plus a share (to be discussed) of the top-up fee. Similarly, if a library has not recently subscribed to the journal, the publisher may persuade the library to take its whole bundle of journals, in which case the library will then have access to the journal and the journal will receive just its share of the top-up fee.

There are a number of concerns with this arrangement: First, there is a limit to how long this system can work: why should a library's subscription decisions from many years ago determine how much it should pay for a publisher's collection of journals?

Second, and relatedly, a publisher may have a difficult job policing this arrangement. What if a library cancels its stand-alone subscription to a journal (as it obviously has every right to do), but then next year has access to the journal as part of its top-up bundle? Publishers claim they can pursuade libraries not to do this. For instance, if a library cancels one full subscription the publisher can threaten to increase its top-up fee by the corresponding amount the next year. However, there are two problems with this

the high-quality journal can set a lower break-even price than the lower-quality journal. The impact of quality on for-profit journal prices is less clear-cut, since the impact of a demand expansion on a monopolist's price is ambiguous (without further assumptions on demand elasticity and the nature of the demand expansion).

[9] See section B3 in Edlin and Rubinfeld (2004) for more details of Elsevier's various 2002 bundling offers to libraries.

policy:

- If the (commercial) publisher has the ability to increase the top-up fee, we expect that it would do this in any case. Relatedly, it might not be credible for the publisher to promise to keep the overall charge for the collection unchanged if a library ceases a full subscription to a journal in the collection. For instance, suppose the library originally placed a value v on having access to a particular journal, while the full subscription price for this journal is p. Originally, $v > p$ and the library decided it was worthwhile subscribing. But now for some reason the value v falls below p, and the library decides to cancel its subscription. If the publisher is confident the library is willing to bear a price rise of p in the top-up fee, it seems it could have increased the price of the top-up fee before v declined. If not, then the library might cease subscribing to the whole collection when faced with the price rise, which the publisher would not wish to risk.

- Carrying out this threat may benefit the publisher (via its own journals in the collection) rather than the cancelled journal itself. Indeed, for this reason the publisher may have an incentive to allow libraries to cancel full subscriptions to independent journals and to load more of the total charge onto the top-up fee.

It is difficult to find public data to see whether journals in collection sales have better or worse institutional renewal rates than those which choose to be sold outside a publisher's collection.[10] But in any event a journal can only trust that its publisher is doing its best to maintain its base of historical subscriptions, against each library's obvious incentive to exploit bundling discounts (and the publisher's possible incentive to load more onto the top-up fee).

Third, for as long as this arrangement persists, it is virtually impossible for a journal to secure a new full library subscription. (Such a library will surely ask for the journal as part of the top-up package.) Thus, once a subscription has been lost, it will not come back. In addition, in a new market such as China, where there are historically few full

[10] One data-point is *Econometrica*'s institutional subscription figures—see Table 1 in "Report of the Secretary" (2007), *Econometrica*, 75(1), page 292. Here it is reported that, starting from a base of 2438, the journal lost 124 institutional subscribers in the year leading up to the end of 2001, 93 in 2002, 3 in 2003, 189 in 2004 and 80 in 2005. The journal was sold as part of Blackwell's collection only in the single year 2004, which is when it incurred its largest subscriber loss. However, there is so much noise in this time series that it is hard to conclude much from this with confidence.

subscriptions, a journal will be unable to generate many full subscriptions. Of course, for a really new institution the top-up fee for the bundle of all journals may be much larger than for an older library with many historical institutional subscriptions. In this case, the thorny issue of how the top-fee for the bundle is allocated among journals (see the next point) becomes all the more important.

Fourth, how should the top-up fee be allocated between journals? If a publisher's journals are fairly homogenous, then it does not seem unreasonable to divide the fee equally. (E.g., Elsevier's *Handbooks in Economics* currently all have the same price, and revenues from bundled electronic dissemination are allocated equally to all books.) But usually there is significant journal heterogeneity in terms of (i) subscription fees, (ii) impact and prestige, (iii) whether the journal is independent or owned by the publisher, and (iv) the number of historic full subscriptions to the journal. A common arrangement is that the top-up fee is allocated in proportion to the journal's stand-alone subscription fee. This arrangement is troubling for a number of reasons:

- A publisher has an incentive to start its own low-quality journals, and to set a very high subscription fee for these. Few libraries would choose to have stand-alone subscriptions to such a journal, but the publisher will syphon off a large fraction of the top-up fee merely as a result of having these journals in its collection.

- Independent journals then have an extra incentive to raise their subscription fees, purely in order to obtain a larger slice of the top-up fee, which could lead to inflationary pressure in the market.

- There is little evidence of positive correlation between a journal's subscription price and the quality of a journal in the collection – if anything, the reverse.[11]

[11] A look at Wiley-Blackwell's website *Synergy* (visited on September 17, 2007) yields the following data for economics journals: *Agricultural Economics* has impact factor 0.584 and subscription fee £364; *Contemporary Economic Policy* has impact 0.444 and subscription £151; *Econometrica* has impact 2.402 and subscription £278; *Economic Journal* has impact 1.629 and subscription £336; *Journal of Economics and Management Strategy* has impact 1.0 and subscription £217; *Journal of Economic Surveys* has impact 0.75 and subscription £410; *Journal of Industrial Economics* has impact 1.152 and subscription £146; *Kyklos* has impact 0.627 and subscription £303, and *Review of Economic Studies* has impact 2.0 and subscription £204. (Subscription fees are the "online only" institutional subscription fee in the UK for the 2007 volume.) For further discussion of this point, see Bergstrom (2001) and Dewatripont, Ginsburgh, Legros, and Walckiers (2007).

Note, though, that impact factor is not the best measure of a library's willingness-to-pay for a journal, since it expresses a journal's impact on a per-article basis. A better measure would be impact factor

Note that in established regions, where a journal historically has good reach, this injustice is not as great as it may at first appear (at least in the short run). For so long as the publisher is adequately policing the renewals of its institutional subscribers (the second point above), the journal will be receiving its full institutional subscription fee in any case, *plus* a small top up. For instance, consider a simple world in which there are two kinds of journal: the top tier, which all relevant libraries subscribe to on an individual basis, and a lower tier, to which all libraries consume only as part of a bundled package. In such a situation, the top tier journals do not drive the demand for the top-up bundle, and it is not clear why they should receive *any* of the top-up fee. Any revenue which a top-tier journal receives from the top-up fee might be regarded as being "paid twice" for the subscription. Nevertheless, the number of full library subscriptions is likely to continue falling (despite the publishers' claimed best efforts to stop this). In future, therefore, the top-up fee is sure to be an ever larger fraction of even a top-tier journal's revenue, and this allocation issue will assume ever greater importance. Moreover, in new regions where libraries are starting from scratch, this issue is important right now. An important point is that an allocation rule for a "top-up" fee which is broadly acceptable to an established journal in a mature market may not be remotely acceptable for allocating revenues from an entirely new institution, yet some publishers currently apply the same rule in both situations.

Finally, current arrangements leave independent journals with relatively little control over their reach and revenues. In contrast to the previous regime of stand-alone journal sales, where a journal knew exactly how much it would receive from each subscribing institution, a journal is now left to the mercy of a publisher in its (private) negotiations with libraries. If a publisher claims to have struck a good deal to supply a consortia of, say, Chinese universities, the journals in its collection have to take that at face value (even if they accept the publisher's rule for how it divides collection revenue).

3 Possible Future Arrangements

For all of these various reasons, it is hard to imagine how current arrangements can persist for long, and the industry will move to some other method (or methods) for remunerating

multiplied by the number of articles per year. See the useful website *www.eigenfactor.org* for rankings of journals in terms of (improved) measures of per-article impact, total impact, and cost-effectiveness.

journals. Some journals have reacted to these current tensions simply by reverting to the historical system of stand-alone pricing and distribution, and ceasing to be included in their publisher's collection sales programme. Such a response has some advantages: it eliminates the danger of the journal cross-subsidizing the publisher's own second-rate journals, and it gives the journal full control over the price libraries pay. However, it also eliminates all efficiency gains from the use of bundling. (These efficiency gains, which can be dramatic, are described in more detail in section 4.) For this reason, this response cannot easily be recommended.

Alternatives which are more in the way of incremental adjustments to current arrangements include:

- Dividing the top-up fee on the basis of, say, impact factor (multiplied by the number of articles per year—see footnote 10) instead of the subscription fee. This would largely remove the danger of cross-subsidy from prestigious non-profit journals to the publisher's own less good journals. However, in mature markets at least, there is real danger of the prestigious journals being "paid twice".

- Dividing the top-up fee on the basis of "usage" (say, the number of downloads of the journal at all institutions).

- Specifying a larger payment (e.g., £50, as opposed to something like £5 or less under typical current arrangements) a journal receives if it is included in a bundle for a library which did not previously subscribe. In such cases, libraries would probably have to have the right to opt out of receiving the journal within the collection. Such a system is far from transparent, however, and involves complex calculations for how important a particular journal is in a publisher's collection.

- Implementing distinct allocation rules in mature markets and emerging markets. Thus, in mature markets, established prestigious journals already receive the revenue from their historical full subscriptions, and for them the details of how the relatively small top-up fee is allocated are not so important. But in new markets, prestigious journals should receive an allocation of the collection fee which reflects their importance in the collection (and not in proportion to their individual library subscription rate, which has little relevance in the new market).

More radical arrangements are also possible to imagine. Many people have suggested than an "open access" model is the best way for the market to evolve, and that journals should not charge for access at all, but should cover their fixed costs by levying fees on published authors. This idea has some attractions: if electronic access to journals is free then the whole vexed issue of bundling goes away; and clearly open access journals have the widest conceivable reach. But the debate about open access is sometimes confused between the normative and the positive. While some people might want the market to move in this direction, this might just be wishful thinking unless there are reasons to think that there are underlying forces pushing in this direction. For one example for why the open access model might not be sustainable, consider the likely size of the author fee needed to allow journals to break even. A journal with operating costs of, say, £300,000 per year which publishes 50 articles per year would need to levy a publication fee of £6,000 per article. It seems plausible that fees of this magnitude could deter some people from publishing their research, which is a market failure arguably worse than that involved in some people not being able to read published research. Moreover, many journals may be tempted to accept more articles than they otherwise would in order to receive this substantial payment, with a result that quality could fall relative to a "reader pays" business model. While the open access model may work for some journals, it seems far-fetched that it will become universal in the market in the medium term (at least without major external pressure from some source).[12]

The remaining alternative considered in this paper is the widespread use of collection sales but without using historical data on library subscriptions. As is clear from the discussion so far, once bundling is involved, it is hard to establish (i) a journal's contribution to the overall demand for the collection and (ii) how to determine the price for the collection as a whole. Instead of using data on past stand-alone subscriptions, publishers could compete to offer an independent journal an annual fixed payment for the right to market the journal as the publisher sees fit. Such an arrangement bypasses the allocation issue altogether, and it gives journals no explicit control over the total charge for the bundle. Implicit in this method is a promise from the publisher to attain a specified reach for its bundle. While for-profit journals do not care about reach, non-profit journals will need to consider both the publisher's remuneration and reach when choosing its

[12] See McCabe and Snyder (2007) and Jeon and Rochet (2007), and the references contained therein, for further discussion of the impact of the open access model and the likelihood of it emerging.

publisher. This arrangement is hardly ever seen now, but it is surely a good candidate for the long run norm in the industry as pure bundling becomes commonplace. The outcome would be different for profit-maximizing journals and for non-profit journals. We expect to see two kinds of publisher in this market: "commercial publishers" aiming to serve for-profit journals and less commercial publishers (perhaps university presses, and so on) aiming to serve the non-profit journals. Commercial publishers would set a profit-maximizing bundle price to libraries, and then pay this back to their journals in proportion to their value in the bundle. Less commercial publishers would charge a lower bundle price which just allows their journals to break even but which expands their reach. I present a stylized model of such a market in the next section.

4 A Model of Collection Sales

In this section I provide a model of collection sales and of how the market might evolve once the current transitional arrangements have ceased.[13] The model presumes that bundling is the norm (as is efficient with electronic dissemination), that historic records of which libraries used to have full subscriptions are irrelevant, that publishers pay lump-sum fees to journals for the right to market their journals, and that publishers charge libraries a lump-sum fee for the right to access their entire collections of journals.

Suppose the various journals are indexed by $i = 1, 2, ..., N$, and that a particular library is willing to pay v_i to obtain access to journal i.[14] Journals are assumed not to substitute for each other, and a library's willingness-to-pay for one journal is unaffected by which other journals it buys. That is to say, if the library has access to the set of $\mathcal{I}$ journals, its total willingness-to-pay for the collection is $\sum_{i \in \mathcal{I}} v_i$. We assume that publishers cannot observe a particular library's list of valuations $(v_1, ..., v_N)$ but they know the distribution of $(v_1, ..., v_N)$ in the population of libraries. Suppose that libraries are not budget constrained, and can afford to buy any journal for a price up

[13] Related analysis is found in Bergstrom and Bergstrom (2004). Their interest is in the situation where a single journal is sold to many readers (via a site license), unlike my focus which is on selling many journals to a single "reader" (the library). However, in economic terms the analysis of the two situations is almost identical. In addition, their focus is more on the impact of bundling arrangements on the welfare of readers, while here only the interests of journals are discussed.

[14] With the advent of electronic distribution, an institution's value for a journal is more accurately known than in the print era (when rarely-answered questionnaires were sent around to faculty and students to gauge their need for the journal), since the number of downloads from a journal can be monitored.

to its willingness-to-pay, v_i.[15] Also, suppose the marginal cost of supplying a journal to an additional library is zero, and the costs in the publishing layer of the market are normalized to zero.

In this model we expect to see two kinds of publisher emerge in order to cater to the different aims of the two kinds of journals: commercial publishers will cater to (or own) for-profit journals, and concentrate their efforts on extracting as much revenue as possible from libraries; "non-profit" publishers will attract the non-profit journals by setting a low subscription price for their collections (but enough to cover the journals' costs) and thereby enhance their reach.

I divide the analysis into two parts. First, I focus on the case where all libraries have a similar willingness-to-pay for large bundles of journals (even though they may have very different demands for individual journals). In this situation, it turns out that journals do not face a meaningful trade-off between revenue and reach. Second, I discuss the more realistic situation in which libraries differ systematically in their demand for bundles of journals, in which case a journal must decide whether to sacrifice revenue for greater readership.

4.1 Similar Libraries

For the sake of simplicity consider the case where each v_i is independently and uniformly distributed on the interval $[0, 1]$. We discuss the cases of stand-alone journal pricing and bundling in turn.

Stand-alone pricing: Consider the case of a for-profit journal. Without bundling, the optimal price for such a journal maximizes $p(1-p)$, i.e., $p = \frac{1}{2}$ is the subscription charge for a stand-alone journal. The profit to be shared between the journal and the publisher is $\frac{1}{4}$, and the journal sells to 50% of libraries. If the publisher market is competitive, all that profit goes to the journal and the outcome is just as if the journal sold directly to libraries (e.g., via its own website), so long as the journal was equally good as a publisher at marketing activities. For a non-profit journal, the non-bundled price which just covers its fixed cost will obviously depend on that fixed cost: for example, if the fixed cost is

[15] Clearly this is a significant simplification. If a library faces an exogenous and *ad hoc* budget constraint, journals will have to compete with each other to be purchased, even if they are not substitutes in the library's "utility function". See Nevo, Rubinfeld, and McCabe (2005) and Jeon and Menicucci (2006) for analysis of this point.

$k = \frac{1}{10}$, the lowest break-even price is $p \approx 0.11$, and so this non-profit journal reaches 89% of libraries.

Bundling two journals: Turning next to bundling arrangements, suppose first that a publisher distributes just two journals.[16] Suppose that the publisher charges libraries a price p_i for subscribing to journal i on its own, and the price p_{12} for subscribing to the collection of both journals. Then, under the realistic assumption that there is a discount for consuming the bundle (i.e., $p_{12} \leq p_1 + p_2$), the pattern of library demand for the publisher's journals is as depicted on Figure 1.

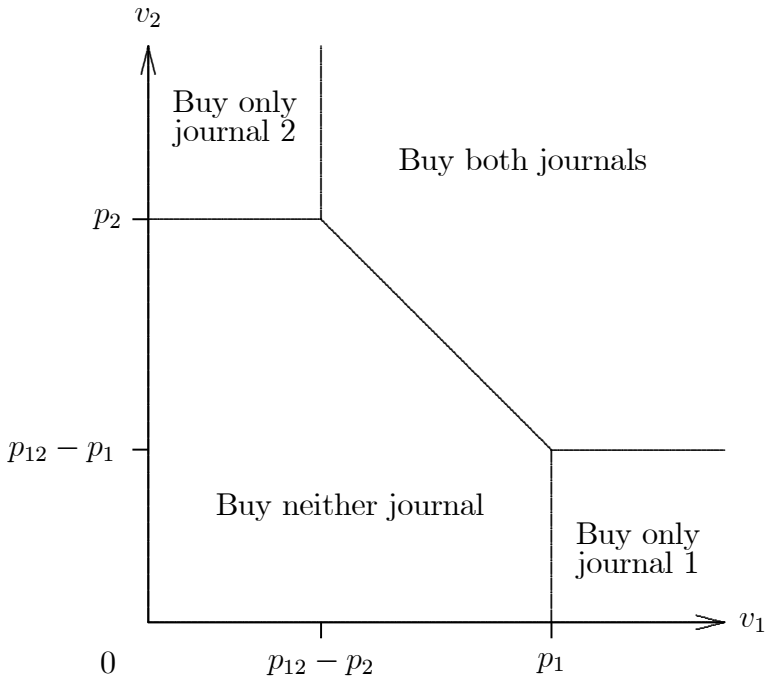

Figure 1: Pattern of Library Subscriptions with Mixed Bundling

When these journals are for-profit, the publisher will aim to maximize the profit from marketing these journals. Brute-force calculations using the areas of the various regions on Figure 1 show that the most profitable way to sell these journals to libraries is "mixed bundling" of the form:

$$p_1 = p_2 = \frac{2}{3} \; ; \; p_{12} = \frac{4}{3} - \frac{\sqrt{2}}{3} \approx 0.86 \; .$$

Thus the price for a single journal rises relative to the situation with stand-alone pricing

[16] This discussion of two-product bundling is taken from sections 12.5 and 12.8.2 in Armstrong (2008).

(where the price was $\frac{1}{2}$), while the total price for the bundle falls. Bundling generates total profit of approximately 0.55, so that the incremental profit for the "second" journal is 0.3, instead of 0.25 with stand-alone pricing. One can show that each journal now sells to more than 60% of libraries (compared to 50% with stand-alone pricing), and so bundling acts to expand the reach of these journals (even though these for-profit journals do not care about their reach).[17]

Suppose instead that the journals are non-profit, so that the publisher chooses its pricing policy to maximize the journals' reach, subject to allowing the journals to cover their fixed cost, say k, for each journal. In this example at least, it turns out that *pure* bundling is optimal, and the journals are not made available individually. Again, the price for the bundle depends on the size of k. For instance, if $k = \frac{1}{10}$, the bundle price for the two journals is approximately 0.2 and the collection's reach is about 98%. Thus, bundling again acts to boost the reach of a journal. (With stand-alone pricing and $k = \frac{1}{10}$, a non-profit journal reached 89% of libraries.) Of course, it is generally true that bundling can expand the reach of a non-profit journal, since it adds to the range of pricing instruments which could be employed: the publisher could sell its journals without bundling, but except in knife-edge cases it can do strictly better with a degree of bundling. (The only case in which bundling cannot do better is when a library's willingness-to-pay for one journal is *perfectly* correlated with that for the other.)

Bundling many journals: One can, with difficulty, calculate the precisely optimal mixed bundling tariff if the publisher has three journals to distribute, while four or more journals seems close to intractable. It is technically much easier, and not too unrealistic, to suppose a publisher has "many" journals, and thereby invoke the law of large numbers.[18] When a collection consists of many journals, a library's willingness-to-pay for n journals is tightly distributed around the expected willingness-to-pay, $\frac{1}{2}n$, so that the great majority of libraries will have $\sum_i v_{ij} \approx \frac{1}{2}n$. For simplicity, suppose each library's willingness-to-

[17] In this example, the aggregate surplus of libraries also increases (slightly) when bundling is employed, although those libraries which have little interest in one of the journals are worse off since the stand-alone price rises.

[18] See Armstrong (1999) and Bakos and Brynjolfsson (1999) for details of this analysis. It is important to remember that it is a publisher's entire collection of journals which is usually relevant for an institution's decision to subscribe, not, say, just the number or quality of its economics journals. At the time of writing, for instance, Wiley-Blackwell distributes more than 1400 academic journals and Elsevier distributes more than 2300.

pay is exactly $\frac{1}{2}n$ for a bundle of n journals.

Consider first a publisher which distributes a collection of n for-profit journals. If n is large, the publisher should therefore set the pure bundle price:

$$\text{price for collection of } n \text{ journals } = \tfrac{1}{2}n \ . \tag{1}$$

All libraries will just be willing to pay this, with the result that each journal is available in all libraries (in contrast to the non-bundling situation, where a for-profit journal was taken by just half the libraries). In a competitive publisher market each journal will receive its contribution to the collection revenue, i.e., it will receive the incremental profit $\frac{1}{2}$ (in contrast to the situation with stand-alone pricing, where a for-profit journal obtains just $\frac{1}{4}$). This incremental profit is equal to aggregate willingness-to-pay of all libraries, so that it is just as if the journal (or publisher) could engage in first-degree price discrimination between libraries. In effect, when a library's value for one journal is independent of its value for other journals, bundling with large numbers of journals eradicates a library's private information about its willingness-to-pay. Commercial publishers here extract the entire surplus of libraries. (This contrasts with the previous two-journal example where bundling made the libraries better off.) This example serves to demonstrate the very considerable potential benefits of collection sales in terms of enhancing a for-profit journal's (i) reach (which *doubles* in this example relative to stand-alone pricing) and (ii) revenue (which also doubles).

Consider next the case of non-profit journals. As long as the publisher sets its collection price below that in (1), each library is willing to take the collection. In particular, a non-profit journal benefits from enhanced reach when bundling is employed. (In this example with $k = \frac{1}{10}$, when bundling is used a non-profit journal's reach expands from 89% to 100% coverage, although almost all of this circulation gain is obtained when just two journals are bundled.) Thus, regardless of whether it is for-profit or non-profit, a journal benefits from bundling. An artificial feature of this model is that library demand for a collection of n journals is essentially inelastic up to the choke price (1). As such, non-profit (and for-profit) journals do not face a meaningful trade-off between the revenue they receive and their reach. Therefore, the precise price charged for the collection will depend heavily on whether non-profit journals place any weight on the welfare of libraries (in which case they would set a low price for the collection which just covers the journal's costs), or whether greater revenue for the journal acts to improve the journal's

performance (which might make the journals set a high price for the collection, since it does not impact on their reach). A more sensible model where there is a real trade-off between revenue and reach is presented in section 4.2.

Finally, it is worth pointing out that with large numbers of journals, this analysis works just as neatly when journals differ in their perceived quality. Thus, if a library's willingness-to-pay for journal i comes from some distribution function $F_i(\cdot)$ with mean μ_i, say (where the distribution for v_i is independent of the distribution for v_j) then a library's willingness-to-pay for collection $\mathcal{I}$ is now approximately $\sum_{i\in\mathcal{I}} \mu_i$. In the case of for-profit journals, when the publishing market is competitive, journal i will be paid the entire area under its demand curve ($\mu_i = \int (1 - F_i) dv$) to be included in the collection.

Further remarks: When each v_i is independently and uniformly distributed on $[0, 1]$, this discussion shows that if a for-profit journal is sold on its own it generates profit 0.25, if it sold with one other journal it generates incremental profit 0.3, and if it is sold as a bundle with many journals it generates incremental profit 0.5. Although I have not calculated the incremental profit for other finite collections (with 3, 4, ... journals), it seems plausible that a journal's incremental profit contribution to a collection is increasing in the number of other journals, with limit 0.5. Thus, a publisher with a greater number of other journals will be prepared to pay more to attract an additional journal than a publisher with fewer existing journals. Thus, in this model, the publishing market exhibits a natural tendency towards a concentrated market structure, with large publishers better able to attract further journals than smaller publishers. A similar point applies to competition for non-profit journals: a publisher with a greater number of other journals will be able to offer greater reach to a journal than a publisher with fewer existing journals.

It is possible that a 'niche' journal (either for-profit or non-profit) with low overall demand from libraries would not be able to cover its fixed cost in a world without bundling, and so would not be supplied. But the extra revenue generated with bundling may allow cost-recovery. Thus, it is possible that the use of bundling leads to the introduction of new, smaller, journals. (This effect is quite distinct from the possibility, discussed in section 2, that with bundling publishers may introduce spurious low-value journals in order to boost their share of the top-up fee.) This observation is an instance of the general point that price discrimination can act to open markets which would not

otherwise be served.

4.2 Heterogeneous Libraries

In the analysis of the previous section, a publisher which bundles together many journals in its collection does not face a trade-off between its reach and the price of its collection. Since this seems unrealistic, consider the following extended framework.[19]

Suppose that libraries differ systematically in their willingness-to-pay for collections of journals, captured by the parameter θ. (This parameter might reflect the size of the institution, its degree of research-orientation, or its wealth.) In more detail, a library's willingness-to-pay for journal i is $v_i = \theta w_i$, and if the library subscribes to journal collection $\mathcal{I}$, its total willingness-to-pay for the bundle is $\theta \sum_{i \in \mathcal{I}} w_i$, so that θ shifts a library's willingness-to-pay multiplicatively. Even if w_i and w_j are independently distributed, v_i and v_j are now positively correlated via the common factor θ. Suppose that θ is distributed throughout the population of libraries with the distribution function $G(\theta)$, and that the $N+1$ parameters $\theta, w_1, ..., w_N$ are independently distributed. Since it is not much harder to do the asymmetric journal case, suppose that each library's preference parameter w_i for journal i is drawn from the distribution function $F_i(w)$, with mean μ_i.

As before, assume that a publisher distributes many journals, so that a library's willingness-to-pay for the collection $\mathcal{I}$ is approximately equal to the average willingness-to-pay:

$$\theta \sum_{i \in \mathcal{I}} w_i \approx \theta \sum_{i \in \mathcal{I}} \mu_i \ .$$

Because the parameter θ remains private information to a library, publishers can no longer accurately predict a library's willingness-to-pay for even a large collection. In particular, commercial publisher cannot fully extract library surplus.

What is the optimal strategy for a commercial publisher who distributes a large number of for-profit journals? If a publisher has the set of $\mathcal{I}$ journals and sets the collection price P, the marginal library type which subscribes has θ such that $\theta \sum_{i \in \mathcal{I}} \mu_i = P$. Therefore, if it chooses to make θ the marginal library, the revenue generated from the collection is

$$(1 - G(\theta)) \times P = (1 - G(\theta)) \times \left(\theta \sum_{i \in \mathcal{I}} \mu_i \right) \ .$$

[19] See section 3 of Armstrong (1999) for more details.

Therefore, regardless of its list of journals, it will chose θ^* to maximize $(1-G(\theta))\theta$. Each of its journals will then (in a competitive market) receive its incremental contribution to the bundle revenue, so that journal i will be paid $(1-G(\theta^*))\theta^*\mu_i$. Notice that a profit-maximizing journal does not care about the quality of the other journals in the same collection (as represented by the other journals' μ_j), since that does not affect its remuneration.

Next, consider competition for non-profit journals. For these journals, a publisher will aim to maximize the reach of its collection (i.e., minimize the marginal library type θ), while covering its journals' fixed costs. It is clear that journals of (approximately) the same quality (as represented by μ), will group together with a publisher, and higher quality journals will have a greater reach.[20] To see this, suppose that a publisher has 100 high-quality journals with $\mu = 3$ and 100 lesser-quality journals with $\mu = 2$, so that the value of this collection to a type-θ library is $500 \times \theta$. If each journal's fixed cost is k, then the publisher must ensure that

$$500(1-G(\theta))\theta \geq 200k \ ,$$

and it will choose the smallest θ which satisfies this break even constraint. But if another publisher "poached" the 100 high-quality journals and offered libraries this high-quality bundle (worth $300 \times \theta$ to the type-θ library), its break even constraint is relaxed to

$$300(1-G(\theta))\theta \geq 100k \ .$$

Thus this rival publisher can offer the high-quality journals an expanded reach, while still covering their costs, relative to the initial case. Thus, just as with the historic case of non-bundled distribution, higher quality journals have lower per-journal prices and greater reach. Nevertheless, bundling leads to greater reach for a non-profit journal of any given quality than does stand-alone pricing.

To illustrate this discussion, suppose each w_i is independently and uniformly distributed on $[0,1]$, so all journals have the same quality, and suppose that θ is also uniformly distributed on $[0,1]$. Suppose a journal's fixed cost is $k = \frac{1}{20}$. Then the commercial publishers, catering to for-profit journals, will choose θ^* to maximize $\theta(1-\theta)$ and so sell to 50% of libraries, and give each of its journals the revenue $\frac{1}{8}$ (well in excess of their

[20] More generally, if cost differs across journals, journals with the same quality/cost ratio μ/k will be distributed together.

costs). By contrast, in the historic stand-alone pricing regime, a for-profit journal in this environment would set the subscription price $p \approx 0.28$, which is attractive to just 36% of libraries. Therefore, each such journal will obtain revenue of around 0.1.[21]

A publisher catering to non-profit journals will choose the marginal subscribing library to have the smallest parameter θ which satisfies the break even constraint

$$\tfrac{1}{2}\theta(1-\theta) = \tfrac{1}{20} \ ,$$

which entails selling the collection to about 89% of libraries. With stand-alone prices, a non-profit journal will set a subscription price of $p \approx 0.06$, which is attractive to around 77% of libraries. To summarise this example from the viewpoint of a non-profit journal:

- Without bundling, the journal can cover its costs by selling to 77% of libraries.
- With bundling, provided the journal uses a "non-commercial" publisher, it can cover its costs and increase its reach to 89% of libraries
- However, if the journal is included in the collection of a commercial publisher which normally deals with for-profit journals, it will be available in just 50% of libraries. In particular, if a non-profit journal feels itself to be locked into dealing with a particular commercial publisher[22], it may be better off withdrawing from the publisher's collection sales programme, and opt instead to be distributed on a stand-alone basis at its own chosen subscription price. In general, however, the comparison of reach between (i) non-bundling sales and (ii) bundling with a commercial publisher is ambiguous and depends on the details of the environment.[23]

In sum, the best outcome for a non-profit journal is to be part of a collection sales programme, but with a non-commercial publisher which typically distributes other non-profit journals of similar quality and which sets a price for the collection which just allows its journals to meet their costs.

[21] A library's willingness-to-pay for a given journal is θw, which is the product of two independent $U[0,1]$ variables. The probability that θv is less than p, where $0 \leq p \leq 1$, is therefore $p(1-\ln p)$. The profit-maximizing stand-alone price therefore maximizes $p[1-p(1-\ln p)]$, which entails $p \approx 0.28$.

[22] For instance, the embarrassment of terminating an existing contract may be a source of inertia. (Those involved in managing non-profit journals are seldom hardened business people.)

[23] For instance, suppose the fixed cost k was much higher, so that only a price close to the profit-maximizing non-bundling price ($p \approx 0.28$) could allow the journal to break-even when the journal was sold on a stand-alone basis. Then the journal's reach is improved (from 36% of libraries to 50%) if it is distributed as a bundled package by a commercial publisher.

5 Conclusions

In the past, journals were sold to libraries on a stand-alone basis, and an independent journal usually was free to set its own subscription price. There are two main advantages to this policy (which could still be adopted by a journal now). First, the journal has full control over its revenues and reach. Second, there are no difficult issues to do with how to allocate the publisher's revenues—the journal's contribution to overall revenue is transparent. In such a world, a journal does not have to judge a publisher's position on the "exploitative—non-profit" spectrum: provided it can set its own subscription price, a journal is (all else equal) indifferent to how its publisher otherwise deals with libraries. Similarly, a journal does not care (all else equal) about the quality of other journals distributed by its publisher, since a library makes its purchase decision regardless of which other journals its buys.[24]

The electronic revolution means that it is essentially costless for a publisher to give one more library access to one of its journals. As a result, is it socially efficient for all libraries to have electronic access to all journals. Bundling acts to improve the trade-off between revenue and reach for a journal, so that for a given level of target revenue a journal can expand its reach if it is bundled with other journals. Bundling has the potential to enhance profit (and reach) of for-profit journals and to expand the reach of non-profit journals.

Current arrangements are an awkward mishmash of stand-alone and bundling elements. The price paid by a library for a collection depends on what journals it historically took, and the extra journals are offered to the library in return for a top-up fee in addition to the historical subscription prices. The division of this top-up fee between the participating journals is, almost inevitably, *ad hoc* and perceived as unfair by many journals. Of particular concern is the potential for a commercial publisher to cross-subsidise its own journals from the revenue generated by the independent non-profit (and often more prestigious) journals. In addition, the reliance on historic data about library subscriptions will become ever less tenable. This arrangement is unstable and will likely change to something more sensible. In the meantime, one option for a non-profit journal is to leave its publisher's collection, and to return to the historic business model of stand-alone

[24] Similarly, a journal which participates in a publisher's collection sales programme does not care about the number or quality of the publisher's journals which have chosen to remain outside the publisher's collection.

sales. However, this decision will deny the journal the efficiency and potential circulation benefits involved in collection sales.

This note focussed on one (I hope, plausible) model for the medium-run evolution of the publishing market. This model predicted: the widespread bundling of journals; publishers bid to attract journals using lump-sum fees and promised library circulation; non-profit journals are offered relatively little revenue in return for a wide circulation, while for-profit journals are offered the reverse. The efficiency gains from the use of bundling will flow into higher profits for for-profit journals (often together with higher reach relative to a stand-alone business model), and into higher reach for non-profit journals. The model predicts that non-profit journals of similar quality will group together in the same publisher, and the high-quality collections will have a lower price and greater reach. (If journals of widely disparate quality were published in the same collection, then a rival publisher could poach the higher-quality journals and offer them a preferable revenue/reach combination.)

In the historic regime of stand-alone journal sales there was little tension when a non-profit journal was published by a highly commercial publisher. Now, though, there is a tension, and non-profit journals might benefit from gradually disentangling themselves from the more commercial publishers. Both for-profit and non-profit journals, however, should surely make full use of the powerful instrument of bundling. In particular, relative to any stand-alone sales strategy, a non-profit journal will be better off if it joins the collection sales programme of a noncommercial publisher. The current example of *JSTOR*, which distributes a collection of largely non-profit journals (with a lag), might be a possible guide for how to disseminate the current output of non-profit journals.[25]

References

ARMSTRONG, M. (1999): "Price Discrimination by a Many-Product Firm," *Review of Economic Studies*, 66(1), 151–168.

[25] *JSTOR* (www.jstor.org) distributes a large number of journals from several disciplines, with a preponderance of non-profit and society journals. Articles are distributed with a lag of several years, so as not to unduly cannibalize a journal's library subscriptions. *JSTOR* is available to libraries on a bundled basis (with scope for libraries to choose only particular subject areas). Since *JSTOR* has always distributed collections rather than individual journals, there is no issue about basing its library prices on historical individual subscriptions. *JSTOR* is a non-profit enterprise and sets relatively low prices to libraries, and pays relatively low prices to participating journals for distribution rights.

——— (2008): "Price Discrimination," in *Handbook of Antitrust Economics*, ed. by P. Buccirossi, pp. 433–468. MIT Press, Cambridge, USA.

Bakos, Y., and E. Brynjolfsson (1999): "Bundling Information Goods: Pricing, Profits, and Efficiency," *Management Science*, 45(12), 1613–1630.

Bergstrom, C., and T. Bergstrom (2004): "The Costs and Benefits of Library Site Licenses to Academic Journals," *Proceedings of the National Academy of Sciences*, 101(3), 897–907.

Bergstrom, T. (2001): "Free Labor for Costly Journals," *Journal of Economic Perspectives*, 15(4), 183–198.

Dewatripont, M., V. Ginsburgh, P. Legros, and A. Walckiers (2007): "Pricing of Scientific Journals and Market Power," *Journal of the European Economic Association*, 5(2), 400–410.

Edlin, A., and D. Rubinfeld (2004): "Exclusion or Efficient Pricing? The 'Big Deal' Bundling of Academic Journals," *Antitrust Law Journal*, 72(1), 119–158.

Jeon, D.-S., and D. Menicucci (2006): "Bundling Electronic Journals and Competition Among Publishers," *Journal of the European Economic Association*, 4(5), 1038–1083.

Jeon, D.-S., and J.-C. Rochet (2007): "The Pricing of Academic Journals: A Two-Sided Market Perspective," mimeo.

McCabe, M., and C. Snyder (2007): "Academic Journal Prices in a Digital Age: A Two-Sided Market Model," *The B.E. Journal of Economic Analysis and Policy (Contributions)*, 7(1), Article 2.

Nevo, A., D. Rubinfeld, and M. McCabe (2005): "Academic Journal Pricing and the Demand of Libraries," *American Economic Review*, 95(2), 447–452.