WORKING PAPERS IN ECONOMICS

No. 07/18

TOMMY STAAHL GABRIELSEN, BJØRN OLAV JOHANSEN AND GREG SHAFFER

WHEN IS DOUBLE MARGINALIZATION A PROBLEM?

Department of Economics
UNIVERSITY OF BERGEN

When is Double Marginalization a Problem?*

Tommy Staahl Gabrielsen† Bjørn Olav Johansen‡

University of Bergen *University of Bergen*

Greg Shaffer§

University of Rochester

August 2018

Abstract

Double marginalization refers to the distortion caused by the successive markups of independent firms in a distribution channel. The implication that this both reduces firm profits and harms consumers is known as the double-marginalization problem. Many solutions have been proposed to help sellers mitigate this pricing problem, and it is arguably one of the main reasons why quantity discounts in the distribution channel are as prevalent as they are. Surprisingly, however, the implication that end-user prices will be distorted upward has only been shown under a very restrictive set of circumstances (successive monopoly). Whether and under what conditions double marginalization is a problem in other, more realistic settings is generally unknown. In this paper, we show that double marginalization need not be a problem when an upstream firm sells its product through competing intermediaries and shelf space is costly. When this is the case, we find that there will often be a role for slotting fees, minimum resale price maintenance (min RPM), and minimum advertised pricing (MAP) policies.

Keywords: slotting fees, resale price maintenance, distribution channels.

*We thank the "Det alminnelige prisreguleringsfond" through the Norwegian Competition Authority for financial support.

†University of Bergen, Department of Economics; tommy.gabrielsen@uib.no.

‡University of Bergen, Department of Economics; bjorn.johansen@uib.no.

§Simon School RC 270100, University of Rochester, Rochester, NY 14627; shaffer@simon.rochester.edu.

1 Introduction

Many firms use intermediaries to sell their products.[1] Although the many potential benefits of doing this are widely recognized — e.g., lower distribution costs, better access to consumers — a downside is that firms risk losing control over their end-user prices. The reason is that each member of the distribution channel typically adds a markup to the markups of all channel members above it, and the accumulation of these markups raise a host of challenging problems. One well-known problem is that the successive markups of the different members are thought to "distort end-user prices upward to a higher level than would otherwise be charged, resulting in lower demand than would otherwise be enjoyed, all else equal."[2]

Among the first to voice this concern was Spengler (1950), who showed in a setting of successive monopoly (where an upstream monopolist sells its product to a single downstream firm, which then resells the product to final consumers) that when the downstream firm adds its own markup to the markup of the upstream firm, the resulting final price will be higher than what a monopolist selling directly to consumers would charge. The distortion caused by the successive markups (marginalizations) of the independent firms has since come to be known as "*double marginalization*," and the implication that this both reduces firm profits and has adverse effects on consumers is known as the "*double-marginalization problem*."[3]

Spengler's suggestion was that firms should avoid intermediaries where possible by selling directly to final consumers, thus solving the problem through vertical integration.[4] Jeuland and Shugan (1983), Moorthy (1987), Kolay et al (2004), and others have shown that manufacturers can, in theory, also solve the double-marginalization problem by offering multi-part tariffs in which the per-unit prices the retailers see at the margin are kept to a minimum (an example of this is a two-part tariff in which the per-unit price is set equal to the manufacturer's marginal cost of production and the retailer pays the manufacturer a fixed fee). Yet another solution involves contractually forcing the downstream firms to give up their markups by capping the retail prices that can be charged (a practice known as max RPM).[5]

---

[1]Stern et al (1996; 3) define intermediaries to be "distribution-oriented institutions" that "stand between production on the one hand and consumption on the other." The dollar amounts involved in this activity are staggering. Coughlan and Jap (2016) report that U.S. retail sales accounted for \$5.13 trillion in 2015, or about 29% of GDP, while wholesale distribution revenues were even higher, accounting for \$5.35 trillion.

[2]See Coughlan and Jap (2016; 169), especially the chapter on "How Do I Price Through the Channel?"

[3]This nomenclature does not appear in Spengler's seminal work. Its usage, however, has been in the vernacular at least since Scherer (1980). Scherer sometimes also referred to the problem as the "vertical chain monopoly problem." Overstreet (1983) called it the "successive-monopoly problem." Referring to the problem as the double-marginalization problem has, since Tirole's (1988) textbook, become the norm.

[4]Of course, this is not always practical. There are many cost-based and demand-based reasons to sell through intermediaries. Once again, see Coughlan and Jap (2016), especially chapter 13, for a discussion.

[5]Overstreet (1983) was one of the first to recognize this. See also the discussion in Tirole (1988).

These proposed solutions are well known.[6] A puzzle arises, however, in that we observe many more instances of price floors (min RPM) and minimum-advertised pricing (MAP) policies than we do of max RPM,[7] and while multi-part tariffs that emphasize infra-marginal transfers from the retailer to the manufacturer are common,[8] so too are multi-part tariffs in which the retailer is the one receiving the transfers (e.g., through the collection of slotting fees).[9] The latter, which involve the manufacturer setting a relatively high per-unit price at the margin, feature quantity premia (the average price is increasing in the quantity purchased).[10] They, along with min RPM and MAP policies, are the opposite of what one would expect if the manufacturer were attempting to solve a double-marginalization problem.[11]

Not surprisingly, other explanations have been offered for min RPM, MAP policies, and slotting fees (see the literature reviews below). And while we think these other explanations have merit in their particular contexts (our aim is not to criticize them), we also think the literature's understanding of double marginalization is incomplete. The incompleteness arises because the implication that double marginalization is a problem has only been shown under a very restrictive set of circumstances (successive monopoly). Whether and under what conditions it is a problem in other, more realistic, settings is generally unknown. As we will show, double marginalization need not be a problem in these other settings, and when it is not, there is indeed a role for slotting fees — and for min RPM and MAP policies.

In this paper, we extend Spengler's successive-monopoly setting in two directions. First, we introduce a second intermediary to compete with the first in selling the manufacturer's product.[12] This extension cuts to the core of our concern. Once we allow for competition among intermediaries, downstream markups will no longer be as large, and this will have consequences both for the size of the manufacturer's initial markup (which is endogenously determined), as well as for whether double marginalization will be a problem. Second,

[6]Coughlan and Jap (2016; 170) write "the tendency of prices to rise ... is solved ... with pricing strategies such as multi-part tariffs. These strategies blunt the problem of loss of control over end-user prices."

[7]Of the three forms of resale price maintenance (max RPM, min RPM, and fixed-price RPM), max RPM has historically been the least common. See the surveys by Gammelgaard (1958) and Overstreet (1983).

[8]Trade terms that feature quantity discounts (both incremental and all-units) have this form.

[9]See the Federal Trade Commission reports on slotting allowances (FTC 2001) and (FTC 2003).

[10]The per-unit price must be high enough to offset the initial loss in profit from the payment of the slotting fee. The relatively high per-unit prices are also well documented in surveys (see Bloom et al, 2000).

[11]To solve a double marginalization problem, the manufacturer must earn its profit through relatively high infra-marginal markups while endeavoring to minimize the per-unit markup the retailer sees at the margin — i.e., the form of the pricing schedule should be more like a quantity discount than a quantity premium.

[12]To be clear, others before us have looked at distribution channels that feature downstream competition and found there to be a double-marginalization problem (e.g., there are a series of papers by Ingene and Parry (1995a, 1995b, 1998) to name a few). But, by restricting attention to linear demands, and by not incorporating a cost of scarce shelf space, they have unknowingly biased their results in Spengler's direction.

we endow both intermediaries with an opportunity cost of shelf space, so that it is not a given that they will sell the manufacturer's product.[13] This extension adds realism to the institutional settings we wish to focus on (e.g., grocery-retailing), and will prove useful in understanding, for example, why we might observe slotting fees in some instances but not others. Thus, we will assume throughout that there are competing intermediaries, that the intermediaries control a scarce asset (which can be thought of as shelf space), and that this asset must be obtained by the manufacturer before its product can be resold to end users.

We obtain several results. First, we show that adding shelf-space costs to Spengler's successive-monopoly setting does not eliminate the double-marginalization problem. It only mitigates it. This is true regardless of the functional form of demand. Second, we show that the form of demand (along with the size of the shelf-space costs) *does* affect whether there is a double-marginalization problem when we allow for a second, competing intermediary. Specifically, we find that double-marginalization is a problem only if shelf-space costs are sufficiently low and certain restrictions on demand hold.[14] Otherwise, double marginalization need not be a problem. Third, we find that slotting fees can be optimal when the double marginalization distortion is weak (i.e., whenever the retail prices are too low from the perspective of maximizing joint profits). Lastly, we note that min RPM and/or the adoption of MAP policies can also be optimal in this case, and indeed, we show that there is an equivalence between slotting fees and these policies whenever offering slotting fees is optimal.

These findings have broad implications. From a theoretical perspective, they suggest that Spengler's insights do not readily extend beyond the case of successive monopoly. Even when there are no shelf-space costs, we find that double marginalization is a problem in the case of downstream competition only if certain restrictions are placed on end-user demands, restrictions that go beyond what are normally required for the existence and uniqueness of equilibrium. And, when there are shelf-space costs, we find that Spengler's insights do not extend if the costs are large enough — end-user prices may be too low rather than too high.

Not surprisingly, the prescription for managers is radically different depending on which scenario holds. If conditions are such that retail prices would otherwise be too high, then Spengler's solution, or multi-part tariffs that emphasize infra-marginal transfers from the retailer, would be appropriate all else equal. If, on the other hand, conditions are such that retail prices would otherwise be too low, then keeping the wholesale price relatively high and compensating the retailers by offering slotting fees is likely to be better. This can explain

[13]Beginning in the 1980's, the rate of new-product introductions in this sector soared while the average size of stores struggled to keep up, a trend that has continued to the present day. It has been estimated that while the average supermarket has room for 25,000 to 45,000 products, there are typically more than 100,000 products available at any time. See the FTC reports on slotting allowances (2001) and (2003).

[14]We can show that these restrictions are always satisfied with linear demands.

why slotting fees are more likely to be observed in practice when shelf-space costs are high.

Our findings also have implications for policy. They suggest that dominant manufacturers that offer slotting fees need not be trying to buy up scarce shelf space with the intent to drive out competitors, as some have alleged.[15] Rather, the fees may be being offered simply as a way of compensating retailers for their high costs of shelf-space. Moreover, the equivalence between min RPM and slotting fees when slotting fees are optimal implies that a policy that favors one over the other, whether intentional or not, may have little or no real consequences.[16] Lastly, our findings suggest that whereas policymakers have come to interpret Spengler's insights as justifying a pro-competitive stance for max RPM, they have failed to appreciate that the flip side of the coin is that the distortions caused by double marginalization can also be interpreted as justifying an anti-competitive stance for min RPM.

This research contributes to several strands of literature. The first of these strands is the contracting literature on vertical control.[17] Much of this literature follows Mathewson and Winter (1984) in identifying settings that require multiple instruments to control multiple targets (e.g., pricing instruments and two-part tariffs may be needed to correct for advertising spillover externalities). Here we abstract from issues that arise from non-price competition in order to maintain the focus on double marginalization as a potential driver for some of the practices that we observe even when other elements in the demand mix may be absent.

There is also a long literature that attempts to explain why manufacturers might want to adopt min RPM. Telser's (1960) free-riding story is probably the most famous of these attempts. Others posit that min RPM can be used to foreclose potential upstream rivals (e.g., Asker and Bar-Issac, 2014), and/or support a manufacturer cartel (e.g., Jullien and Rey, 2007), while still others posit that min RPM can be used to certify retailer quality (e.g., Marvel and McCafferty, 1984) and/or induce higher levels of services and promotions (e.g., Mathewson and Winter, 1984; and Winter, 1993). However, none of the prerequisites for these explanations need be present in our case. The explanation we offer here is simply that min RPM may be needed to support higher profit margins for the retailers, so that when selling the manufacturer's product, they can be compensated for the cost of their shelf space.

Lastly, we come to the literature on slotting fees. The pro-competitive stories focus mostly on the use of slotting fees as a signaling device when manufacturers have private information about their demands (e.g., Kelly, 1991; Lariviere and Padmanabhan, 1997; and Desai, 2000). There is also work on the use of slotting fees to subsidize demand-enhancing

[15]This is one of the prominent explanations offered for why we observe slotting fees (see FTC, 2001).

[16]Policy that is permissive toward slotting fees, but that frowns upon min RPM, for example, may simply cause firms to use slotting fees as their more preferred means of compensating retailers, and vice-versa.

[17]See the respective chapter in Tirole (1988), and Katz (1989), and the references that are cited therein.

investments by the retailers (e.g., Raju and Zhang, 2004; and Kolay and Shaffer, 2013) and/or to increase the incentives for demand-enhancing investments by the manufacturers (e.g., Farrell, 2001). The anti-competitive stories focus mostly on the use of slotting fees to foreclose competitors, either upstream (e.g., Shaffer, 2005) or downstream (e.g., Marx and Shaffer, 2007), or to facilitate tacit collusion (e.g., Shaffer, 1991; Piccolo and Miklos-Thal, 2012).[18] None of these explanations, however, play a role in the present case. There are no demand-enhancing investments, and no firm has private information. There are no upstream competitors to foreclose, nor any competition from would-be upstream rivals to dampen.

The closest paper to ours in methodology is Kuksov and Pazgal (2007). Like us, they consider the use of slotting fees to compensate retailers for their fixed costs (actual costs in their model, opportunity costs in ours). Unlike us, however, they do not solve for general conditions on when double marginalization is or is not a problem when there are competing intermediaries (consumers and firms are located on a Hotelling line in their model), nor do they show the equivalence of slotting fees with alternative potential solutions such as min RPM, or solve the model when the manufacturer is restricted to offering a linear contract.

The rest of the paper proceeds as follows. In the next section, we introduce costs of shelf space and solve the model in a setting of successive monopoly. In Section 3, we extend the model to consider the case of downstream competition. In Section 4, we discuss possible solutions for the manufacturer and obtain our equivalence results. Section 5 concludes.

2 Successive monopoly

We begin by modifying the case of successive monopoly (a la Spengler, 1950) to allow for shelf-space costs. Specifically, we assume that selling to consumers requires that the retailer put the manufacturer's product on display, and that this display costs $SS \geq 0$ for the retailer to provide (alternatively, one can think of SS as the retailer's opportunity cost of shelf space).

The game occurs in two stages. In the first stage, the manufacturer chooses its wholesale price w to maximize its profit subject to the retailer earning a payoff of at least SS. The retailer then chooses its retail price in stage two given its contract terms from stage one.

We solve the game using backwards induction, and thus we begin with stage two. Let $D(p)$ denote the demand for the manufacturer's product as a function of the retail price p. Then, the retailer's payoff from selling the manufacturer's product can be written as:[19]

$$\pi_r(p) \;=\; (p - w)D(p).$$

[18] See Sudhir and Rao (2016) for an empirical study that attempts to distinguish among these stories.

[19] We assume that the retailer's only marginal cost is the per-unit price it pays to the manufacturer.

We make the usual simplifying assumptions. Specifically, we assume that (i) there exists a choke price $\bar{p}$ such that $D(p)$ is zero for all $p \geq \bar{p}$ and positive and downward sloping for all $p < \bar{p}$, (ii) D(p) is continuous and differentiable in p, and (iii) $\pi_r(p)$ is strictly concave.

These assumptions ensure that for all $w < \bar{p}$, the retailer's payoff is well behaved and obtains its maximum at the unique price p that solves $\pi_r'(p) = 0$. Let $p^*(w)$ denote this price. Then, the retailer's maximized payoff as a function of w can be written as

$$\pi_r^*(w) \;=\; (p^*(w) - w)D(p^*(w)).$$

It follows that the retailer's best response in stage two depends on the relationship between π_r^* and SS. If $\pi_r^* \geq SS$, then the retailer's best response is to set $p = p^*(w)$ and sell the manufacturer's product. If instead $\pi_r^* < SS$, then the retailer's best response is to decline to sell the manufacturer's product (which we assume it can do by setting a price of $p \geq \bar{p}$).

The comparative statics of this are straightforward and yield no surprises. Both the retailer and consumers will be worse off the higher is the manufacturer's wholesale price. Consumers will be worse off because the retailer's profit-maximizing price (assuming it sells the manufacturer's product) is increasing in w ($dp^*/dw > 0$), and the retailer will be worse off because it will not be able to pass all of the increase in its costs onto consumers ($d\pi_r^*/dw < 0$).

Turning to the first stage of the game, let c denote the manufacturer's marginal cost of production, and assume that it is profitable for the manufacturer to induce the retailer to sell its product. Then, the manufacturer's problem in stage one can be written as

$$\max_{w}\; (w - c)D(p^*(w)) \tag{1}$$

such that

$$(p^*(w) - w)D(p^*(w)) \;\geq\; SS. \tag{2}$$

The objective in (1) is the manufacturer's first-stage profit as a function of w, taking as given the retailer's best response in stage two, and the constraint in (2) ensures that the retailer will earn its opportunity cost of shelf space. Solving for the optimal wholesale price yields

$$w^* \equiv \min\{w^u, w^c(SS)\},$$

where w^u is the (unconstrained) wholesale price that maximizes (1), and $w^c(SS) > c$ is the (constrained) wholesale price that satisfies (2) with equality. Note that w^u is independent of SS, while our assumptions imply that $w^c(SS)$ is strictly decreasing in SS. When $w^u > w^c(SS)$, the manufacturer will be *forced* to lower its wholesale price to $w^c(SS)$ if it is to induce the retailer to sell its product (because the constraint is binding for all $w > w^c(SS)$).

When $w^u < w^c(SS)$, the manufacturer will *want* to set its wholesale price at w^u (because this maximizes the objective in (1) given that the constraint is not binding for all $w < w^c(SS)$).

In the equilibrium of the game, the manufacturer will set $w = w^*$ in stage one, and the retailer will sell the manufacturer's product and set $p = p^*(w^*)$ in stage two. In contrast, an integrated firm, maximizing the joint profit of the manufacturer and retailer, would set

$$p^m \equiv \arg\max_p \, (p - c)D(p).$$

Comparing the equilibrium price $p^*(w^*)$ to the monopoly price p^m that an integrated firm would set, we can show that the former is always greater. This establishes our first result:

Proposition 1 *In the case of successive-monopoly with shelf-space costs, the equilibrium final price always exceeds the price that an integrated firm would set (i.e., $p^*(w^*) > p^m$) whenever it is strictly profitable for the manufacturer to induce the retailer to sell its product.*

Proof: The proof of Proposition 1 proceeds in four parts. First, notice that it is profitable for the manufacturer to induce the retailer to sell its product only if it is profitable for a fully-integrated firm to sell the manufacturer's product, i.e., only if $(p^m - c)D(p^m) > SS$. This follows because if it is not profitable for a fully-integrated firm to sell the manufacturer's product, then at least one of the independent firms must be worse off when the product is sold. Second, notice that $p^m = p^*(c)$, and thus that $(p^m - c)D(p^m) > SS$ implies that (2) is strictly satisfied at $w = c$. It follows that $w^c(SS) > c$ when $(p^m - c)D(p^m) > SS$. Third, notice that our assumption that the retailer's profit is concave implies that $\pi_m^* = (w - c)D(p^*(w))$ is also concave, and thus that $w = w^u$ solves the manufacturer's first-order condition

$$(w - c)D'(p^*(w))\frac{dp^*}{dw} + D(p^*(w)) = 0. \tag{3}$$

It follows that $w^u > c$ (because the first-order condition in (3) is positive when evaluated at $w = c$). Last, notice that the fact that $w^u > c$ and $w^c(SS) > c$ when $(p^m - c)D(p^m) > SS$ implies that $w^* > c$ when $(p^m - c)D(p^m) > SS$, and thus that $p^*(w^*) > p^m$ when $(p^m - c)D(p^m) > SS$ (because $p^m = p^*(c)$ and $dp^*/dw > 0$). It follows that $p^*(w^*) > p^m$ whenever it is profitable for the manufacturer to induce the retailer to sell its product. **Q.E.D.**

Proposition 1 extends the case of successive monopoly to allow for downstream shelf-space costs. In the proof, we establish that it is always optimal for the manufacturer to add a markup on its product when selling to the downstream firm, and that the retailer will then add its own markup, leading to a final price that is always higher than what an integrated firm would charge. Intuitively, the integrated firm's price is lower because it bases its decision on the product's actual marginal cost, not on the inflated marginal cost that the downstream firm sees. Or, as Tirole (1988) puts it, "The retail price is higher in the decentralized structure than in the integrated one, because of two successive mark-ups (marginalizations)." This outcome has come to be known as the problem of double-marginalization. Not only are the firms' joint profits lower than what they would be in the absence of double marginalization, but consumers lose as well because the retailer's price is higher than it would otherwise be.

Another way of thinking of the intuition for why the equilibrium final price always exceeds the fully-integrated monopoly price is to note that when the downstream firm sets the final price, it does not take into account the negative effect that its decision has on the profits of the upstream firm, an externality that a fully-integrated firm would internalize. To see this, note that when the downstream firm chooses its final price p, it will choose p to satisfy

$$\frac{\partial \pi_r}{\partial p} = \frac{\partial \Pi}{\partial p} - \frac{\partial \pi_m}{\partial p} = 0, \tag{4}$$

where $\Pi = (p - c)D(p)$ is the manufacturer and retailer's joint profit, and $\pi_m = (w - c)D(p)$ is the manufacturer's profit. Since $\partial \pi_m / \partial p = (w - c)D'$ is strictly negative when $w > c$, it follows that $\partial \Pi / \partial p < 0$ at the optimum. This implies that as long as it is optimal for the manufacturer to set $w > c$ when selling to the downstream firm, the downstream firm's equilibrium final price p will exceed the price p^m that a fully-integrated firm would charge.

Our finding that double marginalization extends in this case does not imply that the introduction of shelf-space costs has no effect on the outcome. Once shelf-space costs are large enough to affect the manufacturer's optimal choice of w, it is straightforward to see that further increases in SS will lead to a decrease in w^* (because $\partial w^c(SS)/\partial SS < 0$) and thus in the price $p^*(w^*)$ that consumers pay. It follows that although shelf-space costs do not eliminate double marginalization in this setting, they may be able to mitigate the problem.

3 Downstream competition

We now extend the game to allow the manufacturer to sell to competing downstream firms. As before, we assume that resale requires that the manufacturer's product be put on display. The cost of this display is $SS \geq 0$ for each retailer. The retailers' demands are symmetric.

We want to know whether double-marginalization is still a problem, and what role shelf-space costs play in the outcome. To this end, we assume the game occurs in two stages. In the first stage, the manufacturer chooses wholesale price w_1 for retailer 1 and wholesale price

w_2 for retailer 2 to maximize its profit subject to each retailer earning at least SS in profit.[20] The retailers then simultaneously and independently choose their retail prices in stage two.

As before, we begin with the second stage of the game. Let p_1, p_2 denote the retail prices, and let the demand for the manufacturer's product at retailer i be given by $D_i(p_1, p_2)$. Then, the profit that retailer i gets from selling the manufacturer's product can be written as

$$\pi_i(p_1, p_2) = (p_i - w_i)D_i(p_1, p_2).$$

We make the usual simplifying assumptions. Specifically, we assume that:

(i) $D_i(p_1, p_2)$ is continuous and differentiable in both p_1 and p_2;

(ii) $D_i(p_1, p_2)$ is decreasing in retailer i's own retail price p_i, increasing in retailer j's retail price p_j, and has the property that own-price effects dominate cross-price effects:

$$\frac{\partial D_i}{\partial p_i} < 0; \quad \frac{\partial D_i}{\partial p_j} > 0; \quad \text{and} \quad \frac{\partial D_i}{\partial p_i}\frac{\partial D_j}{\partial p_j} - \frac{\partial D_i}{\partial p_j}\frac{\partial D_j}{\partial p_i} > 0. \tag{5}$$

(iii) the second derivatives of $\pi_i(p_1, p_2)$ are such that for all p_i, p_j such that $D_i, D_j > 0$,

$$\frac{\partial^2 \pi_i}{\partial p_i^2} < 0; \quad \frac{\partial^2 \pi_i}{\partial p_i \partial p_j} > 0; \quad \text{and} \quad \Delta \equiv \frac{\partial^2 \pi_i}{\partial p_i^2}\frac{\partial^2 \pi_j}{\partial p_j^2} - \frac{\partial^2 \pi_i}{\partial p_i \partial p_j}\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} > 0. \tag{6}$$

The inequalities in (5) ensure that demands are downward sloping, that the products sold by the two retailers are substitutes, and that each retailer's demand is more sensitive to an increase in the retailer's own price than it is to an increase in the rival retailer's price.

The inequalities in (6) are also standard. The first inequality ensures that retailer i's profit is well behaved and obtains its maximum at the unique p_i that solves

$$\frac{\partial \pi_i}{\partial p_i} = (p_i - w_i)\frac{\partial D_i}{\partial p_i} + D_i = 0. \tag{7}$$

The second inequality ensures that the p_i that solves (7) is increasing in p_j (i.e., that retailer i's best-response function is upward sloping). The third inequality implies that the Jacobian of the system of profit-maximizing first-order conditions is negative definite and ensures that a Nash equilibrium in the pricing game between the two retailers exists and is unique.[21]

Let $p_1^{**}(w_1, w_2)$ and $p_2^{**}(w_1, w_2)$ denote the equilibrium prices. Then, retailer i's equilibrium profit in stage two as a function of the wholesale prices w_1, w_2 can be written as

[20] We will assume throughout this section that the additional demand generated by the second retailer is enough to justify paying the added fixed costs SS. Otherwise, we are back to the case of successive monopoly.

[21] These conditions are sufficient but not necessary. See Friedman (1983).

$$\pi_i^{**}(w_1, w_2) \;=\; (p_i^{**}(w_1, w_2) - w_i) D_i(p_1^{**}(w_1, w_2), p_2^{**}(w_1, w_2)). \tag{8}$$

It follows that if retailer i's rival is selling the manufacturer's product and setting a price of $p_j^{**}(w_1, w_2)$, then retailer i's best response in stage two depends on the relationship between π_i^{**} and SS. If $\pi_i^{**} \geq SS$, then retailer i's best response is to set $p_i = p_i^{**}(w_1, w_2)$ and also sell the manufacturer's product. If instead $\pi_i^{**} < SS$, then retailer i's best response is not to sell the manufacturer's product (which we assume it can do by setting p_i sufficiently high).

The inequalities in (5) and (6) ensure that consumers will be worse off when there is an equal increase in the manufacturer's wholesale prices.[22] Intuitively, one would expect the retailers to pass along some of their cost increases to consumers in the form of higher prices, and this plus our assumption that best-response functions are upward sloping are enough to establish the result. Surprisingly, however, the inequalities in (5) and (6) are not enough to establish that the retailers will also be worse off. For this, Δ must be sufficiently large. That is, it is not enough for the determinant of the retailer's second-order conditions to be positive, it must also be greater than $\frac{\partial D_i}{\partial p_j}\left(\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} - \frac{\partial^2 \pi_j}{\partial p_j^2}\right)$, which itself is greater than zero.

Lemma 1 *Suppose that for all w_i, w_j such that $D_i(p_1^{**}, p_2^{**})$ and $D_j(p_1^{**}, p_2^{**}) > 0$,*

$$\Delta > \frac{\partial D_i}{\partial p_j}\left(\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} - \frac{\partial^2 \pi_j}{\partial p_j^2}\right) > 0. \tag{9}$$

Then, it can be shown that $\frac{\pi_i^{**}(w_1,w_2)}{\partial w_i} < 0$, $\frac{\partial \pi_j^{**}(w_1,w_2)}{\partial w_i} > 0$, and $\frac{\partial \pi_i^{**}(w,w)}{\partial w_i} + \frac{\partial \pi_j^{**}(w,w)}{\partial w_i} < 0$.

Lemma 1 says that if (9) holds, then a marginal increase in retailer i's wholesale price will decrease retailer i's profit and increase retailer j's profit (i.e., each retailer will be better off when its rival's wholesale price increases and worse off when its own wholesale price increases). Importantly, it also says that the sum of the retailers' profits will decrease (i.e., own effects will dominate cross effects) when they are evaluated at the same wholesale price.[23]

The restriction on Δ is needed because a marginal increase in w_i has both a direct and an indirect effect on retailer i's equilibrium profit. The direct effect stems from the fact that an increase in w_i makes it costlier for retailer i to purchase its inputs from the manufacturer. This harms retailer i for any given quantity purchased. The indirect effect stems from the fact

[22] This can be shown by totally differentiating the retailers' first-order conditions in (7).

[23] The logic of this claim is as follows. A marginal increase in w_i decreases retailer i's profit by some amount, say A, and increases retailer j's profit by some amount, say B, where $A > B$. Similarly, a marginal increase in w_j increases retailer i's profit by some amount, say C, and decreases retailer j's profit by some amount, say D, where $D > C$. At equal starting wholesale prices, symmetry implies that $A = D$ and $B = C$, and thus that the sum of the changes, $C - A$ for retailer i and $B - D$ for retailer j, must be negative.

that an increase in w_i also leads retailer i's rival to charge a higher final price in equilibrium. This benefits retailer i because the retailers' products are substitutes. The two effects on retailer i's profit thus go in opposite directions, and in general, the net effect is ambiguous.

By assuming that condition (9) holds in what follows, we ensure two things. First, we ensure that the direct effect on i's profit outweighs the indirect effect. Second, we ensure that the net effect is sufficiently negative that it also outweighs the gain to retailer j.[24] It should be emphasized, however, that the main reason for taking condition (9) to be the normal case is that it always holds with linear demands (see the Appendix). For other demand systems, however, condition (9) need not hold. In that case, retailer i need not be worse off when w_i increases, and even if it would be worse off, the sum of the retailers' profits need not be less.

Turning to the first stage, the manufacturer's problem is to set w_1, w_2 to solve

$$\max_{w_1,w_2} \sum_{i=1,2}(w_i - c)D_i(p_1^{**}(w_1, w_2), p_1^{**}(w_1, w_2)) \tag{10}$$

such that

$$(p_i^{**}(w_1, w_2) - w_i)D_i(p_1^{**}(w_1, w_2), p_2^{**}(w_1, w_2)) \ \geq \ SS. \tag{11}$$

We assume that the objective in (10) has a unique solution, and note that symmetry implies that $w_1 = w_2$ at the optimum. Solving for the optimal wholesale price then yields

$$w_1^{**} = w_2^{**} = w^{**} \equiv \min\{\hat{w}^u, \hat{w}^c(SS)\},$$

where $w_1 = w_2 = \hat{w}^u$ are the (unconstrained) wholesale prices that maximize (10), and $w_1 = w_2 = \hat{w}^c(SS)$ satisfy the constraints in (11) with equality. That $w^{**} = \min\{\hat{w}^u, \hat{w}^c(SS)\}$ follows from the definition of $\hat{w}^u$ and our assumption that (9) holds in the normal case (because condition (9) along with symmetry implies that $\pi_i^{**}(w^{**}, w^{**})$ is decreasing in w^{**}).

We have thus far implicitly assumed that shelf-space costs are such that the manufacturer will want to serve both retailers. This is equivalent to assuming that shelf-space costs are below some critical level. With a slight abuse of notation, this assumption can be made explicit as follows. Let $\pi_m^*(w) \equiv (w - c)D(p^*(w))$ denote the manufacturer's profit in the first stage when it sells to only retailer i (where $D(p)$ now corresponds to the demand that retailer i would face as a function of its own price p if retailer j did not sell the manufacturer's product). And, let $\pi_m^{**}(w_1, w_2)$ denote the manufacturer's profit in the first stage when it sells to both retailers (i.e., the profit in (10)). Then, there exists a critical threshold level of costs,

[24] To be clear, it is possible that the net effect on retailer i's profit could be negative but not sufficiently negative to outweigh the gain to retailer j. This would be the case if (9) fails but $\Delta > \frac{\partial D_i}{\partial p_j}\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} > 0$.

$\overline{SS} \equiv \pi_m^{**}(w^{**}, w^{**}) - \pi_m^*(w^*)$, such that for all $SS \leq \overline{SS}$, the manufacturer will set $w_1 = w_2 = w^{**}$ in stage one, both retailers will sell the manufacturer's product in stage two, and the resulting equilibrium final prices will be $p_1^{**}(w^{**}, w^{**}) = p_2^{**}(w^{**}, w^{**}) = p^{**}(w^{**}, w^{**})$.[25]

The next step is to compare these prices to the price p^I that a fully-integrated firm would set. Here, p^I maximizes the sum of the profits of the manufacturer and two retailers:

$$p^I \equiv \arg\max_p \sum_{i=1,2} (p - c) D_i(p, p).$$

Unfortunately, neither the comparison nor the intuition is clear cut. In particular, none of the inferences we made previously hold here, which suggests that new intuition is needed.

To see this, note that we cannot infer (as we did before) that the equilibrium final price will be higher than what an integrated firm would charge simply because of the negative externality that each downstream firm imposes on the upstream firm when it raises its price. The reason is that when retailer i chooses its profit-maximizing price, it will set p_i to satisfy

$$\frac{\partial \pi_i}{\partial p_i} \;=\; \frac{\partial \Pi}{\partial p_i} - \frac{\partial \pi_m}{\partial p_i} - \frac{\partial \pi_j}{\partial p_i} \;=\; 0, \tag{12}$$

where $\Pi = \sum_{i=1,2} (p_i - c) D_i(p_1, p_2)$ is the joint profit of the manufacturer and both retailers, $\pi_m = \sum_{i=1,2} (w_i - c) D_i(p_1, p_2)$ is the manufacturer's profit, and $\pi_j = (p_j - w_j) D_j(p_1, p_2)$ is retailer j's profit. Although it is true that an increase in retailer i's price does indeed impose a *negative* externality on the manufacturer's profit (i.e., the term $\frac{\partial \pi_m}{\partial p_i}$ in condition (12) will be negative), it will now also be true that an increase in retailer i's price imposes a *positive* externality on its rival (i.e., the term $\frac{\partial \pi_j}{\partial p_i}$ in condition (12) will be positive). It turns out that either effect can dominate. It depends on what the manufacturer's wholesale prices are.

It is also not possible to infer (as we did before) that the equilibrium final price will be higher than what an integrated firm would charge simply because of the markup that the upstream firm will be adding to its product when selling to the downstream firms (which makes the cost of the goods sold by the downstream firms higher than the cost of production). The reason is that the markups of the downstream firms under competition will not be as high as the markup that an integrated firm would charge when $w_1 = w_2 = c$. This can be understood more formally by noting that an integrated firm's prices, $p_1 = p_2 = p^I$, satisfy

$$\frac{\partial \Pi}{\partial p_i} \;=\; (p_i - c)\,\frac{\partial D_i(p_i, p_i)}{\partial p_i} + D_i + (p_j - c)\,\frac{\partial D_j(p_i, p_i)}{\partial p_i} \;=\; 0, \tag{13}$$

[25] This is established by showing that $\pi_m^{**}(w^{**}, w^{**}) - \pi_m^*(w^*)$ is weakly decreasing in SS, such that at the level of SS for which $\pi_m^*(w^*) = \pi_m^{**}(w^{**}, w^{**})$, further increases in SS imply that $\pi_m^*(w^*) > \pi_m^{**}(w^{**}, w^{**})$.

whereas the prices $p_i = p^{**}(w,w)$ that arise in equilibrium with competing firms satisfy

$$\frac{\partial \pi_i}{\partial p_i} \;=\; (p_i - w_i)\,\frac{\partial D_i(p_i,p_i)}{\partial p_i} + D_i \;=\; 0. \tag{14}$$

Here we can see that at $w_i = c$, the two conditions differ only in the term $(p_j - c)\,\partial D_j/\partial p_i$, which appears in (13) but does not appear in (14).[26] It follows that simply knowing that the manufacturer has set $w_1 = w_2 = w > c$ in the first stage, and that this implies that $p_i = p^{**}(w,w) > p^{**}(c,c)$ in stage two, will not be enough to establish that $p^{**}(w,w) > p^I$ (because the additional term in (13) implies that p^I will also be greater than $p^{**}(c,c)$).

These observations suggest that, unlike in the case of successive monopoly, a finding that there is double marginalization in the case of downstream competition does not imply that there is a double-marginalization problem. Whether or not there is a double-marginalization problem depends on how high the manufacturer's markup is, which in turn depends on the retailers' cost of shelf space.

Proposition 2 *In the case of downstream competition with shelf-space costs, there exists a critical level of costs, $\widehat{SS} > 0$, such that $p^{**}(w^{**},w^{**}) > p^I$ if (9) holds and shelf-space costs are below this level. For higher shelf-space costs, or if the inequality in (9) is reversed, it will be optimal for the manufacturer to induce final prices that are below the integrated prices.*

Proof: Shelf-space costs under downstream competition are bounded below by zero and above by $\overline{SS}$. Consider the case in which they are zero. In this case, we know that w_1^{**}, w_2^{**} are optimally chosen to maximize the objective in (10). Letting $\Pi^{**}(w_1,w_2) =$

$$\sum_{i=1,2} (p_i^*(w_1,w_2) - c) D_i(p_1^*(w_1,w_2), p_2^*(w_1,w_2))$$

denote the joint profit of the manufacturer and retailers, we can rewrite this objective as

$$\Pi^{**}(w_1,w_2) - \pi_1^{**}(w_1,w_2) - \pi_2^{**}(w_1,w_2).$$

It then follows that the manufacturer's wholesale prices w_i, w_j, will be chosen to satisfy

$$\frac{\partial \pi_m^{**}}{\partial w_i} \;=\; \frac{\partial \Pi^{**}}{\partial w_i} - \frac{\partial \pi_1^{**}}{\partial w_i} - \frac{\partial \pi_2^{**}}{\partial w_i} \;=\; 0. \tag{15}$$

Since (9) implies that $\frac{\partial \pi_1^{**}}{\partial w_i} + \frac{\partial \pi_2^{**}}{\partial w_i} < 0$, we have that $\frac{\partial \Pi^{**}}{\partial w_i}$ must be strictly negative at the optimum. This means that the manufacturer's optimal wholesale prices will exceed the

[26] This term is strictly positive for all p_1, p_2 such that $p_j > c$ and $D_j(p_i,p_i) > 0$.

wholesale prices that would maximize overall joint profits, which in turn implies that the induced final prices, $p^{**}(w_1^{**}, w_2^{**})$, will exceed the prices that an integrated firm would charge.

Now consider the case in which $SS = \overline{SS}$. In this case, we know that w_1^{**}, w_2^{**} satisfy (11) with equality, and are given by $w_1^{**} = w_2^{**} = \hat{w}^c(\overline{SS})$. There are two subcases. Either the induced final prices at these wholesale prices are greater than or equal to p^I or they are less than p^I. If they are greater than or equal to p^I, then it follows that for all shelf-space costs between zero and $\overline{SS}$ inclusive, the induced final prices will equal or exceed p^I. But if they are less than p^I, then there will necessarily be some critical threshold in the interior between 0 and $\overline{SS}$ such that for shelf-space costs that are equal to or below this threshold, the induced final prices will equal or exceed p^I, and for all shelf-space costs that are above this threshold but still less than or equal to $\overline{SS}$, the induced final prices will be less than p^I.

If the inequality in (9) is reversed, then the induced final prices will be less than p^I regardless of shelf-space costs because then $\frac{\partial \pi_1^{**}}{\partial w_i} + \frac{\partial \pi_2^{**}}{\partial w_i} > 0$ implies that $\frac{\partial \Pi^{**}}{\partial w_i}$ will be positive. **Q.E.D.**

Proposition 2 establishes that the double-marginalization problem that was first identified in Spengler (1950) extends to the case of downstream competition if the retailers' shelf-space costs are sufficiently low and condition (9) holds. Thus, for example, there will be a double-marginalization problem in the absence of shelf-space costs when demands are linear. As we have shown, the proof of this follows by establishing that when condition (9) holds, an unconstrained manufacturer will impose a negative externality on its retailers when it raises its wholesale prices, and thus will choose wholesale prices that exceed the wholesale prices it would choose if it were trying to maximize overall profits. Proposition 2 also establishes, however, that for sufficiently high shelf-space costs, the manufacturer will eventually be constrained to charge lower wholesale prices, and that at some point, these lower wholesale prices may become sufficiently low that double marginalization will no longer be a problem.

To those who see the proverbial glass as half full, these results can be seen as establishing that there exist plausible settings under which the results in Spengler (1950) (and the implications that follow from them) extend to cases other than successive monopoly. To those who see the glass as half empty, however, Proposition 2 can alternatively be seen as establishing that Spengler's insights do not readily extend — because condition (9) implies restrictions on demand that go beyond what is necessary to ensure the existence and uniqueness of the downstream Nash equilibrium. In the absence of (9) holding generally, the externality on the retailers could go the other way when evaluated at the equilibrium quantities. Higher wholesale prices would then benefit the retailers, and there would then no longer be a double-marginalization problem, even in the absence of shelf-space costs. The other caveat is that whether or not (9) holds, shelf-space costs must be sufficiently low for there to be a problem. It is easy to come up with numerical examples in which shelf-space costs are such that the induced equilibrium final prices are strictly below the integrated level. In these cases, as we will now show, the implications for managers have an unexpected twist.

4 Contractual solutions and an equivalence result

Spengler (1950) advocated vertical integration as a way of solving the double-marginalization problem. He suggested that vertical integration should be allowed by competition authorities because it would benefit consumers in the form of lower prices. What we have shown, however, is that while Spengler's insights extend beyond the case of successive monopoly under some circumstances, they do not extend under other circumstances. Allowing vertical integration when shelf-space costs are sufficiently high, for example, can harm consumers.

Others have noted that more sophisticated contracts and/or price restraints on resale can sometimes be used to proxy for vertical integration. In the case of successive monopoly, for example, it is well known that the manufacturer can use either two-part tariffs or max RPM to induce the integrated outcome by setting the terms to eliminate one of the markups.[27]

Things are more complicated in the case of downstream competition with shelf-space costs, however, because the strategy of eliminating either the wholesale or retail markup then no longer works. Nevertheless, as we will now show, it is still possible for the manufacturer to induce the integrated outcome. Consider the manufacturer's problem with two-part tariffs:

$$\max_{w_1,w_2,F_1,F_2} \sum_{i=1,2} \left((w_i - c)D_i(p_1^{**}(w_1,w_2), p_1^{**}(w_1,w_2)) + F_i\right)$$

such that

$$(p_i^{**}(w_1,w_2) - w_i)D_i(p_1^{**}(w_1,w_2), p_2^{**}(w_1,w_2)) - F_i \;\geq\; SS.$$

Assuming that it is optimal to serve both retailers, the solution to this problem is to set $w_1 = w_2 = w^I$ such that $p_1^{**}(w^I,w^I) = p_2^{**}(w^I,w^I) = p^I$. The fixed fees will then be set at $F_1 = F_2 = F^I \equiv (p^I - w^I)D_i(p^I,p^I) - SS$, yielding a maximized profit for the manufacturer of $\Pi^I - 2SS$, where $\Pi^I \equiv (p^I - c)D_i(p^I,p^I)$ is the profit that a fully-integrated firm earns.[28]

Now consider the manufacturer's problem when it can impose price restraints (RPM):

[27] With a two-part tariff, for example, the manufacturer obtains the vertically-integrated outcome by setting $w = c$, thereby inducing the retailer to set $p = p^m$, and then using its fixed fee to extract the retailer's surplus. With a max RPM policy, the manufacturer obtains the vertically-integrated outcome by first setting $w = p^m$ and then prohibiting the retailer from setting a final price that is above this level.

[28] The non-integrated manufacturer earns less because it must give each retailer at least SS in profit.

$$\max_{w_1,w_2,p_1,p_2} \sum_{i=1,2}(w_i - c)D_i(p_1, p_2)$$

such that

$$(p_i - w_i)D_i(p_1, p_2) \;\geq\; SS.$$

The solution to this problem is to set $p_1 = p_2 = p^I$ and $w_1 = w_2 = \overline{w}$ such that $(p^I - \overline{w})D_i(p^I, p^I) = SS$. This also yields a maximized profit for the manufacturer of $\Pi^I - 2SS$.

The manufacturer obtains its maximum profit in both cases by first inducing the independent retailers to price at the vertically-integrated level and then extracting as much of the surplus from them as possible, subject to each retailer earning at least SS in profit.

As we now show, however, this equivalence between RPM and two-part tariffs is more precisely understood as an equivalence between certain forms of RPM (max or min) and certain forms of two-part tariffs (positive or negative fixed fees). The particular form of each that is needed to induce the integrated outcome depends on whether the manufacturer wants to increase or decrease the retail prices relative to the status quo under linear contracts.

Proposition 3 *In the case of downstream-competition with shelf-space costs, if* $p^{**}(w^{**}, w^{**}) > p^I$*, then the manufacturer maximizes its profit by adopting an RPM clause with max RPM or, equivalently, specifying a two-part tariff in which the fixed fee is positive. If instead* $p^{**}(w^{**}, w^{**}) < p^I$*, then the manufacturer maximizes its profit by adopting an RPM clause with min RPM or, equivalently, specifying a two-part tariff in which the fixed fee is negative.*

Proof: We have already shown that the manufacturer can obtain the maximum profit with an optimally chosen RPM contract, and we have also shown that the manufacturer can obtain the maximum profit with an optimally chosen two-part tariff contract. We now establish that max RPM is the optimal form of RPM if and only if it is optimal to offer a two-part tariff that has a positive fixed fee. By definition, $F^I > 0 \longleftrightarrow SS < (p^I - w^I)D_i(p^I, p^I)$. But we know that $SS = (p^I - \overline{w})D_i(p^I, p^I)$. Rearranging gives $F^I > 0 \longleftrightarrow w^I < \overline{w}$. Since $\partial p^{**}(w, w)/\partial w > 0$ and $p^{**}(w^I, w^I) = p^I$, it follows that $F^I > 0 \longleftrightarrow p^{**}(\overline{w}, \overline{w}) > p^I$. Thus, at the wholesale prices $w_1 = w_2 = \overline{w}$ under RPM, both retailers will want to price above p^I. It follows that under RPM, the manufacturer will have to impose a price ceiling at p^I to prevent this. A two-part tariff with a positive fixed fee thus corresponds to max RPM. By analogous reasoning, a two-part tariff with a negative fixed fee corresponds to min RPM.

Next, we establish that if (9) holds and $p^{**}(w^{**}, w^{**}) > p^I$, then the manufacturer will optimally want to offer a two-part tariff in which the fixed fee is positive (or a max RPM contract). First, note that $p^{**}(w^{**}, w^{**}) > p^I$ implies that $p^{**}(w^{**}, w^{**}) > p^{**}(w^I, w^I)$, and therefore that $w^{**} > w^I$ (given that (5) and (6) ensure that $p^{**}(w, w)$ is increasing in w).

Since we know that the retailers are earning at least SS in profit when they are faced with the wholesale prices $w_1 = w_2 = w^{**}$, it follows that they will be earning a flow payoff that is strictly in excess of SS when they are faced with the wholesale prices $w_1 = w_2 = w^I$ (because their profits are decreasing in w when (9) holds). Thus, it follows that $F^I > 0$.

Lastly, we establish that if (9) holds and $p^{**}(w^{**}, w^{**}) < p^I$, then the manufacturer will optimally want to offer a two-part tariff in which the fixed fee is negative (or a min RPM contract). In this case, note that $p^{**}(w^{**}, w^{**}) < p^I$ implies that $p^{**}(w^{**}, w^{**}) < p^{**}(w^I, w^I)$, and therefore that $w^{**} < w^I$. Since we have shown that $w^{**} = \hat{w}^c(SS)$ in this case, we know that the retailers are earning exactly SS in profit under the status quo. It follows that they would then be earning a flow payoff that is strictly less than SS when they are faced with the wholesale prices $w_1 = w_2 = w^I$. Thus, it must be that $F^I < 0$ in this case. **Q.E.D.**

Proposition 3 establishes that when (9) holds and $p^{**}(w^{**}, w^{**}) > p^I$, there is an equivalence between max RPM and two-part tariff contracts in which the retailers pay the manufacturer a fixed fee. These contracts benefit consumers and thus are pro-competitive for the same reason that they are pro-competitive in Spengler's world of successive monopoly — they correct for a double-marginalization problem that would otherwise arise. At the same time, however, Proposition 3 also establishes that the manufacturer may sometimes want to *increase* the retail prices *above* the status quo that would arise with linear contracts, and that when this is the case, there is an equivalence between min RPM and two-part tariff contracts in which the retailers are paid a fixed fee (i.e., a slotting fee). These contracts, in contrast to the previous contracts, may harm consumers and thus may be anti-competitive.

This last finding has no counterpart in Spengler's world of successive monopoly. The reason is that, unlike in Spengler's world, there need not be a double-marginalization problem. A double-marginalization problem arises if and only if the sum of the manufacturer and retailer markups under linear contracts exceeds the overall markup that an integrated firm would charge. While this condition is always satisfied for any positive markup by the manufacturer in Spengler's world, it is not always satisfied when there is downstream competition. Rather, the manufacturer's markup in the case of downstream competition must be large enough to offset the fact that the markups of the downstream firms under competition are less than the markup that an integrated firm with the same marginal cost would charge.

Shelf-space costs matter critically in this determination. For low levels of shelf-space costs (e.g., when they are zero), we have shown that double marginalization will be a problem as long as (9) holds. Higher shelf-space costs, however, imply weakly lower markups (strictly lower markups when the retailers' profit constraints bind), and so at some point, when shelf-space costs become high enough, double marginalization will no longer be a problem. It is at this point (and for even higher shelf-space costs) that we would expect to observe the manufacturer offering slotting fees, or alternatively, adopting MAP or a policy of min RPM.

Both practices have received a lot of scrutiny in recent years. Slotting fees in particular have been the focus of several government investigations, and are said to account for billions of dollars annually in the U.S. alone.[29] Existing studies tend to focus on the use of slotting fees to foreclose competitors or facilitate collusion, or posit that manufacturers have better information about their products and use slotting fees as a signaling device. However, none of these common explanations account for, or depend on, shelf-space costs to make them work. In contrast, our results suggest that there is a positive relationship between the retailers who receive slotting fees and their opportunity cost of shelf space. Specifically, slotting fees are more likely to be observed when the retailers' shelf-space costs are relatively high. These findings accord with the views of practitioners as reported in the surveys by Bloom et al. (2000) and Wilkie et al. (2002), and are widely accepted as stylized facts. Here, we find that there is a causal relation. Slotting fees arise because shelf-space costs are relatively high.

Min RPM has also been the focus of much academic study and, along with MAP policies,[30] have seen a revival in recent years thanks to the Supreme Court's ruling in *Leegin* (2007).[31] The ruling removed the per-se ban on RPM that had been in place since 1975 (min RPM continues to be a hard-core restraint in the EU) and has spawned numerous articles contesting whether the ban on min RPM would have been better left in place, or whether the Court should have sided with those who believe the practice is almost always pro-competitive.[32] To this debate, our findings offer a new wrinkle. They suggest that whereas courts and policymakers have come to interpret double marginalization as justifying a pro-competitive stance toward the practice of max RPM, they have failed to appreciate that the flip side of the coin is that the same theory in essentially the same institutional setting might be interpreted as justifying an anti-competitive stance toward min RPM.

While one might argue that an anti-competitive stance toward min RPM is not justified when relatively simple two-part pricing with slotting fees can achieve the same outcome, the same could also be said about the eagerness of the courts and policymakers to adopt a pro-competitive stance toward max RPM when relatively simple two-part tariffs with positive

[29]The Federal Trade Commission has conducted its study of the practice (FTC, 2001; FTC, 2003) as has the Canadian Bureau of Competition (2002) and the Norwegian Competition Authority (2005). Slotting fees are also discussed in the U.K.'s Competition Commission's investigation of grocery markets in 2008, and are the focus of the UK's *Groceries Supply Code of Practice*. See Shaffer (2013) for a survey of the literature.

[30]Papers on MAP policies include Israeli (2015) and Israeli, Anderson, and Coughlan (2016).

[31]See *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 551 U.S. 877 (2007).

[32]For starters, see the recent symposium in the *Review of Industrial Organization* on resale price maintenance and the tenth anniversary of *Leegin*, vol. 50, pp. 129-261, and the various papers cited therein.

fixed fees could be used in their place. Or, one might argue that an anti-competitive stance toward min RPM is not justified when there are pro-competitive explanations for min RPM that cannot easily be replicated with two-part tariff contracts. But by the same reasoning one could also argue that a pro-competitive stance toward max RPM is not justified when there are anti-competitive explanations for max RPM that also cannot easily be replicated.[33]

The Supreme Court in *Leegin* (2007) sidestepped much of the debate by settling on a rule-of-reason approach. Under this middle-ground approach, lower courts and policymakers would be required to determine which explanations best fit the facts in any given instance (e.g., was the product subject to free-riding by discounters, was the manufacturer intending to use RPM to foster non-contractible demand-enhancing investments, did the practice have the effect of facilitating collusion, was the imposition of RPM aimed at foreclosing competitors, etc.). Here, the theory need not take a backseat to these other more recognizable explanations. In the past, when RPM was legal, and before the advent of slotting fees in the 1980's, min RPM was frequently observed on commonly purchased grocery items such as aspirin, pens, pencils, toothpaste, soap, shaving cream, and milk (see Overstreet, 1983, and Bowman, 1955, for a more extensive list). In proposing a theory of min RPM that is based on compensating retailers for the opportunity cost of their shelf space, and that does not hinge on the existence of externalities in non-price competition, we are able to explain why manufacturers might want to adopt min RPM on these and many other kinds of products.

5 Conclusion

A central tenet in the management of distribution channels is that pricing through intermediaries is likely to be inefficient unless steps are taken to correct for double marginalization. The purpose of this paper is to understand when this inefficiency is likely to lead to end-user prices that are too high (referred to as the "double marginalization problem"), and when it is likely to lead to end-user prices that are too low. To this end, we have extended Spengler's canonical setting of successive monopoly (one upstream and one downstream firm) to allow for costly shelf space and competing downstream firms, and found that whereas double marginalization is indeed always a problem in the case of successive monopoly, even when shelf space is costly, it is not always a problem when there is downstream competition. In particular, we found that double marginalization is a problem in the sense of prices being too high only when shelf-space costs are sufficiently low and demand satisfies certain conditions.

These findings have implications for the kinds of policies that managers should adopt.

[33] See O'Brien and Shaffer (1992) and Gabrielsen and Johansen (2017).

They suggest that rather than always striving to keep per-unit markups as low as possible (as is normally recommended), or looking for ways to cap retail prices (max RPM), managers might do well to do the opposite. For example, they suggest that when the cost of shelf space is relatively high, a manufacturer can instead earn higher profits by keeping its wholesale prices relatively high and offering slotting fees to its retailers. Alternatively, it could opt to buttress retail prices by insisting on retail price floors (min RPM), or adopting MAP policies.

We believe that Spengler (1950) never meant to imply that double marginalization always leads to retail prices that are too high; rather his conclusion was simply an artifact of his assumption of one downstream firm, and his desire to show that vertical integration need not always be harmful.[34] Extending his model to allow a second intermediary to compete with the first shows the limitations of his case. It also sheds new light on the observance of min RPM and slotting fees, and can explain why they may be observed in a variety of otherwise nondescript instances. It is not the case, for example, that min RPM need always be associated with the formation of cartels, the facilitation of tacit collusion, or the prevention of free riding (see Overstreet, 1983), nor is it the case that the offering of slotting fees by a dominant manufacturer need imply that it is always trying to exclude its competitors, or communicate private information to the retailer about its product's demand (FTC, 2001).

One drawback, however, is that we have found that min RPM and slotting fees are equivalent in the sense that either can be used to maximize the manufacturer's profits in the absence of a double-marginalization problem. So, while we can describe when one or the other will be used, we cannot explain (in our simple stylized model) which one will be used. The same holds for max RPM and two-part contracts that feature a positive fixed fee.

In future research, we hope to shed light on the incidence of the different types of contracts when the manufacturer has a choice among equivalent instruments. We suspect that the manufacturer's optimal choice of contracts will depend among other things on (a) the prevailing legal climate (i.e., whether there are sanctions on the use of RPM or slotting fees), and (b) the relative enforcement and monitoring costs of the different instruments. With both min and max RPM, for example, retailers will have an incentive to cheat on the agreement in practice. To prevent this, the manufacturer will obviously need to spend resources to detect such cheating and to enforce the appropriate punishment. Two-part tariffs, however, are also not without costs. The offering of slotting fees, for example, may present a moral hazard dilemma. A retailer can accept the money but then shirk on giving the product

[34] The evidence is that Spengler was motivated by the Court beginning to look upon vertical integration as illegal per se, lumping it together with horizontal integration as something that necessarily reduces competition "unreasonably." He wrote that while "Horizontal integration may, and frequently does, make for higher prices ... Vertical integration, on the contrary, does not, as such, serve to reduce competition and may, if the economy is already ridden by deviations from competition, operate to intensify competition."

adequate shelf-space. In the extreme, it may even stock an alternative product instead.

Among RPM contracts, one might expect RPM price floors to have lower overall monitoring costs. This is because unlike with RPM price ceilings, cheating on price floors hurts rival retailers and therefore is more likely to be reported to the manufacturer. Whereas the burden of monitoring price floors can thus be decentralized, cheating on price ceilings will be up to the manufacturer to detect. This can add to the costs. Although it is true that individual consumers can report price-ceiling violations to the manufacturer, any one consumer's incentive to report such violations is likely to be small relative to the individual's expected cost of doing so (e.g., the cost of its time and hassle). Moreover, setting up a monitoring system in which consumers report directly to the manufacturer may entail significant transactions costs (although perhaps this is less so with the emergence of social media).

It follows that because the payment of slotting fees to retailers inevitably involve moral hazard, and because min RPM may be relatively cheap to enforce, one should not always expect to observe slotting fees when both min RPM and slotting fees are feasible. It also follows that the case for max RPM is a more cautious one given that transaction costs would appear to be lower when the fixed fees are positive (payments to the manufacturer) than when they are negative (payments to the retailers), and relatively higher with price ceilings (max RPM) than with price floors (min RPM). This may help to explain why the incidence of min RPM on consumer goods in the U.S. reached up to 10% when RPM was legalized (up to 44% in the U.K.),[35] while the incidence of max RPM has never been very high.

Monitoring costs aside, the incidences of the different practices will also almost surely be affected by the prevailing legal climate. Policy that is relatively permissive towards slotting fees, but that frowns upon min RPM, for example, may simply cause the majority of firms to choose slotting fees over RPM as the preferred means of compensating retailers for the opportunity cost of their shelf space. Relaxing the policy on RPM, as the Supreme Court recently did in *Leegin*, would be expected over time to reduce the incidence of slotting fees.

[35]See the discussion in Overstreet (1983), and the estimates in Herman (1959) and Pickering (1974).

Appendix

Proof of Lemma 1: We need to show that whenever (9) holds, an increase in w_i will decrease retailer i's equilibrium profit, $\frac{\partial \pi_i^{**}}{\partial w_i} < 0$, increase retailer j's equilibrium profit, $\frac{\partial \pi_j^{**}}{\partial w_i} > 0$, and decrease the sum of the retailers' equilibrium profits, $\frac{\partial \pi_i^{**}(w,w)}{\partial w_i} + \frac{\partial \pi_j^{**}(w,w)}{\partial w_i} < 0$.

To establish the first claim, we differentiate $\pi_i^{**}(w_1, w_2)$ with respect to w_i to obtain

$$\frac{\partial \pi_i^{**}(w_1, w_2)}{\partial w_i} = (p_i^{**} - w_i)\frac{\partial D_i(p_1^{**}, p_2^{**})}{\partial p_j}\frac{\partial p_j^{**}}{\partial w_i} - D_i(p_1^{**}, p_2^{**}). \tag{A.1}$$

Next, we differentiate the system of first-order conditions in (7) to obtain

$$\frac{\partial p_j^{**}}{\partial w_i} = -\frac{1}{\Delta}\frac{\partial D_i(p_1^{**}, p_2^{**})}{\partial p_i}\frac{\partial^2 \pi_j(p_1^{**}, p_2^{**})}{\partial p_i \partial p_j}. \tag{A.2}$$

It is easy to see from our assumptions in (5) and (6) that (A.2) is positive. Substituting (A.2) into (A.1), using retailer i's first-order condition, and simplifying, yields

$$\begin{aligned}\frac{\partial \pi_i^{**}(w_1, w_2)}{\partial w_i} &= \frac{1}{\Delta}\frac{D_i}{\partial D_i/\partial p_i}\left(\frac{\partial D_i}{\partial p_j}\frac{\partial D_i}{\partial p_i}\frac{\partial^2 \pi_j}{\partial p_i \partial p_j}\right) - D_i \\ &= \frac{D_i}{\Delta}\left(\frac{\partial D_i}{\partial p_j}\frac{\partial^2 \pi_j}{\partial p_i \partial p_j} - \Delta\right).\end{aligned} \tag{A.3}$$

Since $\frac{D_i}{\Delta}$ is positive, it follows that (A.3) is negative if and only if $\Delta > \frac{\partial D_i}{\partial p_j}\frac{\partial^2 \pi_j}{\partial p_i \partial p_j} > 0$. Since (9) holding implies that $\Delta > \frac{\partial D_i}{\partial p_j}\frac{\partial^2 \pi_j}{\partial p_i \partial p_j} > 0$, we have that $\frac{\partial \pi_i^{**}(w_1,w_2)}{\partial w_i}$ is indeed negative.

To establish the second claim, we differentiate $\pi_j^{**}(w_1, w_2)$ with respect to w_i to obtain

$$\frac{\partial \pi_j^{**}(w_1, w_2)}{\partial w_i} = (p_j^{**} - w_j)\frac{\partial D_j(p_1^{**}, p_2^{**})}{\partial p_i}\frac{\partial p_i^{**}}{\partial w_i}. \tag{A.4}$$

Next, we differentiate the system of first-order conditions in (7) to obtain

$$\frac{\partial p_i^{**}}{\partial w_i} = \frac{1}{\Delta}\frac{\partial D_i(p_1^{**}, p_2^{**})}{\partial p_i}\frac{\partial^2 \pi_j(p_1^{**}, p_2^{**})}{\partial p_j^2}. \tag{A.5}$$

It is easy to see from our assumptions in (5) and (6) that (A.5) is positive. Since the terms $(p_j^{**} - w_j)$ and $\frac{\partial D_j(p_1^{**}, p_2^{**})}{\partial p_i}$ are positive, it follows that (A.4) must also be positive.

To establish the third claim, we first substitute (A.5) into (A.4), and simplify using retailer j's first-order condition, to obtain the first line below. We then evaluate the expression at $p_1^{**} = p_2^{**}$, and substitute $D_j = D_i$ and $\frac{\partial D_j}{\partial p_i} = \frac{\partial D_i}{\partial p_j}$, to obtain the second line.

$$\begin{aligned}
\frac{\partial \pi_j^{**}(w_1, w_2)}{\partial w_i} &= -\frac{1}{\Delta} \frac{D_j}{\partial D_j / \partial p_j} \left(\frac{\partial D_j}{\partial p_i} \frac{\partial D_i}{\partial p_i} \frac{\partial^2 \pi_j}{\partial p_j^2} \right) \\
&= -\frac{D_i}{\Delta} \left(\frac{\partial D_i}{\partial p_j} \frac{\partial^2 \pi_j}{\partial p_j^2} \right).
\end{aligned} \tag{A.6}$$

Adding (A.3) and (A.6), and grouping common terms, yields

$$\frac{\partial \pi_i^{**}(w_1, w_2)}{\partial w_i} + \frac{\partial \pi_j^{**}(w_1, w_2)}{\partial w_i} = \frac{D_i}{\Delta} \left(\frac{\partial D_i}{\partial p_j} \left(\frac{\partial^2 \pi_j}{\partial p_i \partial p_j} - \frac{\partial^2 \pi_j}{\partial p_j^2} \right) - \Delta \right). \tag{A.7}$$

Since $\frac{D_i}{\Delta}$ is positive, it follows that (A.7) is negative if and only if $\Delta > \frac{\partial D_i}{\partial p_j} \left(\frac{\partial^2 \pi_j}{\partial p_i \partial p_j} - \frac{\partial^2 \pi_j}{\partial p_j^2} \right)$. Since (9) is exactly this condition, we have that $\frac{\partial \pi_i^{**}(w,w)}{\partial w_i} + \frac{\partial \pi_j^{**}(w,w)}{\partial w_i}$ is indeed negative. **Q.E.D.**

Linear demands and condition (9): We have assumed in the text that condition (9) is the normal case. We now show that it is always satisfied with linear demands. With linear demands, $\frac{\partial^2 \pi_i}{\partial p_i^2} = 2\frac{\partial D_i}{\partial p_i}$, $\frac{\partial^2 \pi_i}{\partial p_i \partial p_j} = \frac{\partial D_i}{\partial p_j}$, $\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} = \frac{\partial D_j}{\partial p_i}$, and $\frac{\partial^2 \pi_j}{\partial p_j^2} = 2\frac{\partial D_j}{\partial p_j}$. It follows that

$$\Delta - \frac{\partial D_i}{\partial p_j} \left(\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} - \frac{\partial^2 \pi_j}{\partial p_j^2} \right) = 4\frac{\partial D_i}{\partial p_i} \frac{\partial D_j}{\partial p_j} - 2\frac{\partial D_i}{\partial p_j} \frac{\partial D_j}{\partial p_i} + 2\frac{\partial D_i}{\partial p_j} \frac{\partial D_j}{\partial p_j},$$

which is strictly positive per our assumption in (5) that own effects dominate cross effects. **Q.E.D.**

References

Asker, J. and H. Bar-Isaac, 2014, "Raising Retailers' Profits: On Vertical Practices and the Exclusion of Rivals," *American Economic Review*, 104, 672-86.

Bloom, P., Gundlach, G., and J. Cannon, 2000, "Slotting Allowances and Fees: Schools of Thought and the Views of Practicing Managers," *Journal of Marketing*, 64, 92-108.

Bowman, W., 1955, "The Prerequisites and Effects of RPM," *University of Chicago Law Review*, 22, 825-858.

Canadian Bureau of Competition, 2002, *The Abuse of Dominance Provisions (Sections 78 and 79 of the Competition Act) as Applied to the Canadian Grocery Sector*, Ottawa: Canada.

Coughlan, A. and S. Jap, 2016, *A Field Guide to Channel Strategy: Building Routes to Market*, CreateSpace Independent Publishing Platform.

Desai, P., 2000, "Multiple Messages to Retailers: Signaling New Product Demand," *Marketing Science*, 19, 381-389.

Farrell, J., 2001, "Some Thoughts on Slotting Allowances and Exclusive Dealing," remarks prepared for the American Bar Association Section of Antitrust Law, 49th Annual Spring Meeting, Washington, D.C., March 28, 2001.

Federal Trade Commission, 2001, *Report on the Federal Trade Commission Workshop on Slotting Allowances and Other Marketing Practices in the Grocery Industry*, available at www.ftc.gov/opa/ 2001/02/slottingallowancesreportfinal.pdf.

Federal Trade Commission, 2003, *Slotting Allowances in the Retail Grocery Industry: Selected Case Studies in Five Product Categories*, available at www.ftc.gov/opa/2003/ 11/slottingallowancerpt031114.pdf.

Friedman, J., 1983, *Oligopoly Theory*, Cambridge, Cambridge University Press.

Gabrielsen, T. and B. O. Johansen, 2017, "Resale Price Maintenance with Secret Contracts and Retail Service Externalities," *American Economic Journal: Microeconomics*, 9, 63-87.

Gammelgaard, S., 1958, *Resale Price Maintenance*, Paris, France: The European Productivity Agency of the Organization for European Economic Co-operation.

Herman, E.S., 1959, “A Statistical Note on Fair Trade,” *Antitrust Bulletin*, 4, 583-592.

Ingene, C. and M. Parry, 1995a, Channel Coordination When Retailers Compete, *Marketing Science*, 14, 360-377.

Ingene, C. and M. Parry, 1995b, Coordination and Manufacturer Profit Maximization: The Multiple Retailer Channel, *Journal of Retailing*, 71, 129-151.

Ingene, C. and M. Parry, 1998, Manufacturer-Optimal Wholesale Pricing, *Marketing Letters*, 9, 65-77.

Israeli, A., 2015, “Channel Management and MAP: Evidence from a Natural Experiment,” mimeo.

Israeli, A., Anderson, E., and A. Coughlan, 2016, “Minimum Advertised Pricing: Patterns of Violation in Competitive Retail Markets,” *Marketing Science*, 35, 539-564.

Jeuland, A. and S. Shugan, 1983, Managing Channel Profits, *Marketing Science*, 2, 239-272.

Jullien, B. and P. Rey, 2007, “Resale Price Maintenance and Collusion,” *Rand Journal of Economics*, 38, 983-1001.

Katz, M., 1989, “Vertical Contractual Relations,” in R. Schmalensee and R.D. Willig, eds., Handbook of Industrial Organization, vol 1, Amsterdam: North Holland, 655-721.

Kelly, K., 1991, “The Antitrust Analysis of Grocery Slotting Allowances: The Procompetitive Case,” *Journal of Public Policy and Marketing*, 10, 187-198.

Kolay, S. and G. Shaffer, 2013, “Contract Design with a Dominant Retailer and a Competitive Fringe,” *Management Science*, 59, 2111-2116.

Kolay, S., Shaffer, G., and J. Ordover, 2004, All-Units Discounts in Retail Contracts, *Journal of Economics & Management Strategy*, 13, 429-459.

Kuksov, D. and A. Pazgal, 2007, “The Effects of Costs and Competition on Slotting Allowances,” *Marketing Science*, 26, 259-267.

Lariviere, M. and V. Padmanabhan, 1997, “Slotting Allowances and New Product Introductions,” *Marketing Science*, 16, 112-128.

Marvel, H. and S. McCafferty, 1984, “Resale Price Maintenance and Quality Certification,” *Rand Journal of Economics*, 15, 340-359.

Marx, L. and G. Shaffer, 2007, "Upfront Payments and Exclusion in Downstream Markets," *Rand Journal of Economics*, 38, 823-843.

Mathewson, G.F. and R. Winter, 1984, "An Economic Theory of Vertical Restraints," *Rand Journal of Economics*, 15, 27-38.

Moorthy, K.S., 1987, Managing Channel Profits: Comment, *Marketing Science*, 6, 375-379.

Norwegian Competition Authority, 2005, "Betaling for Hylleplass (Payment for Shelf Space)," Report 2/05, Bergen.

O'Brien, D. and G. Shaffer, 1992, "Vertical Control with Bilateral Contracts," *Rand Journal of Economics*, 23, 299-308.

Overstreet, T., 1983, *Resale Price Maintenance: Economic Theories and Empirical Evidence*, Washington, D.C.: Bureau of Economics Staff Report to the Federal Trade Commission.

Pickering, J., 1974, "The Abolition of Resale Price Maintenance in Great Britain," *Oxford Economic Papers*, 26, 120-146.

Piccolo, S. and J. Miklos-Thal, 2012, "Colluding Through Suppliers," *Rand Journal of Economics*, 43, 492-513.

Raju, J. and Z.J. Zhang, 2005, "Channel Coordination in the Presence of a Dominant Retailer," *Marketing Science*, 24, 254-262.

Scherer, F.M., 1980, *Industrial Market Structure and Economic Performance*, Chicago, Il: Houghton Mifflin, Co.

Shaffer, G., 1991, "Slotting Allowances and Resale Price Maintenance: A Comparison of Facilitating Practices," *Rand Journal of Economics*, 22, 120-135.

Shaffer, G., 2005, "Slotting Allowances and Optimal Product Variety," *The B.E. Journal of Economic Analysis & Policy*, Vol 5: Iss. 1 (Advances), Article 3, available at www.bepress.com/bejeap/ advances/vol5/iss1/art3.

Shaffer, G., 2013, *Impact of Reverse-Fixed-Payments on Competition*, Office of Fair Trading, London: UK, http://webarchive.nationalarchives.gov.uk/20140402142426/http://www.oft.gov.uk/OFTwork/research/economic-research/completed-research.

Spengler, J., 1950, "Vertical Integration and Antitrust Policy," *Journal of Political Economy*, 58, 347-352.

Stern, L., El-Ansary, A., and A. Coughlan, 1996, *Marketing Channels*, 5th Ed., Prentice Hall: Upper Saddle River, NJ.

Sudhir, K. and V. Rao, 2006, "Are Slotting Allowances Efficiency-Enhancing or Anti-Competitive?," *Journal of Marketing Research*, 43, 137-155.

Telser, L., 1960, "Why Should Manufacturers Want Fair Trade?," *Journal of Law and Economics*, 3, 86-105.

Tirole, J., 1988, *The Theory of Industrial Organization*, Boston, MA: MIT Press.

Wilkie, W., Desrochers, D., and G. Gundlach, 2002, "Marketing Research and Public Policy: The Case of Slotting Fees," *Journal of Public Policy & Marketing*, 21, 275-288.

Winter, R., 1993, "Vertical Control and Price Versus Nonprice Competition," *Quarterly Journal of Economics*, 108, 61-76.

Institutt for økonomi
Universitetet i Bergen
Postboks 7800
5020 Bergen
Besøksadresse: Fosswinckels gate 14
Telefon: +47 5558 9200
Fax: +47 5558 9210
www.uib.no/econ/