



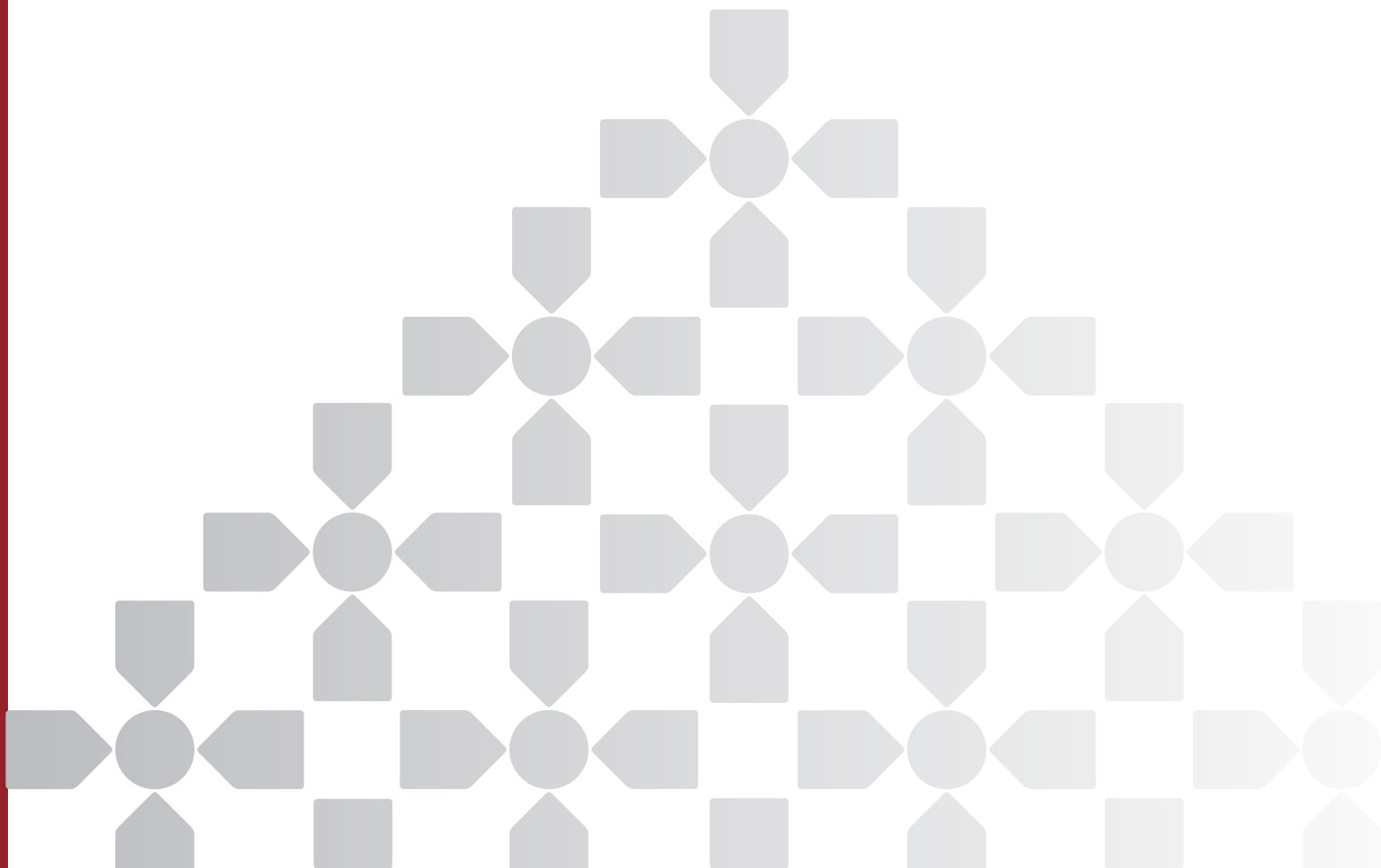
KONKURRANSE
TILSYNET
Norwegian Competition Authority

Rapport 4/2018

Double Marginalization and the Cost of Shelf Space

Tommy Gabrielsen, Bjørn Olav Johansen og Greg Shaffer

*Prosjektet har mottatt forskningsmidler fra
det alminnelige prisreguleringsfondet.*



Double Marginalization and the Cost of Shelf Space*

Tommy Staahl Gabrielsen[†] Bjørn Olav Johansen[‡]
University of Bergen *University of Bergen*

Greg Shaffer[§]
University of Rochester

May 2018

Abstract

We extend the theory of double marginalization to the case of downstream competition and show that whether a double-marginalization problem arises depends on the cost of the retailers' shelf space. When it is low, double-marginalization is a problem in the normal case. Inducing the integrated outcome in these instances leads to lower prices for consumers. In contrast, when it is sufficiently high, there is no double-marginalization problem. In these instances, inducing the integrated outcome leads to higher prices for consumers. We further show that there is an equivalence between (a) max RPM and two-part tariff contracts with positive fixed fees, and (b) min RPM and two-part tariff contracts with negative fixed fees (i.e., slotting allowances). The former are used when there is a double-marginalization problem. The latter are used when there is no double-marginalization problem.

*We thank the Norwegian Competition Authority for its support and encouragement of this research.

[†]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

Antitrust law in the middle of the last century was a mess. With little or no useful theory to guide them – the term “double marginalization” had not yet been coined – courts were beginning to look upon vertical integration as illegal per se, lumping it together with horizontal integration as something that necessarily reduces competition “unreasonably.” Motivated by this perceived mistreatment, Spengler (1950) opined 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.”¹

Spengler illustrated his points in a model of successive monopoly (an upstream monopolist sells its product to a single downstream firm, which then resells the product to final consumers). He showed 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. He theorized that allowing the firms to vertically integrate would eliminate one of the markups and thus lead to a lower retail price. The successive markups (marginalizations) of the independent firms has since come to be known as “double marginalization,” and the implication that this both lowers profits and has adverse effects on consumers is known as the “double-marginalization problem.”²

After a time, it was recognized that there are other solutions to the problem of double-marginalization.³ Many of these solutions have the upstream firm either voluntarily agreeing to give up its own markup, or contractually forcing the downstream firm to give up his. For example, it is known that with a two-part tariff, the upstream firm can obtain the vertically-integrated outcome by setting its wholesale price equal to its marginal cost and extracting the downstream firm’s surplus with a fixed fee. This solution effectively amounts to selling off the vertical structure to the downstream firm, which then becomes the residual claimant of all profit. It is also known that the upstream firm can obtain the vertically-integrated outcome by setting its wholesale price equal to the monopoly retail price less any marginal distribution costs the retailer may have and then prohibiting the

¹Spengler made it clear in his seminal work that he did not mean to imply that the courts’ treatment of horizontal integration was necessarily correct. Rather, in contrasting vertical and horizontal integration, he was simply pointing out that there are important qualitative differences between the two.

²This nomenclature does not appear in Spengler’s seminal work. Its usage, however, has been in the vernacular at least since Scherer (1980). Scherer sometimes referred to the problem as the “vertical chain monopoly problem.” Overstreet (1983) called it the “successive-monopoly problem.” Since Tirole’s (1988) textbook, however, referring to the problem as the double-marginalization problem has become the norm.

³Overstreet (1983) was one of the first to recognize this. See also the discussion in Tirole (1988).

downstream firm from adding his own independent markup (e.g., by imposing a ceiling on the retail price that the downstream firm can charge). The latter solution is a form of resale price maintenance known as maximum resale price maintenance (max RPM).⁴

The theory, simple as it was, implied that vertical integration, and these alternative contractual solutions, were procompetitive. It helped to stem the tide of public opinion against vertical integration and caused an outcry against the per-se illegal treatment of max RPM, an outcry that ultimately led to a reversal of its per se treatment in *State Oil v. Khan* (1997).⁵ Indeed, it is hard to overestimate the impact that the theory has had. Max RPM is now considered to be at worst benign and mostly beneficial,⁶ and even a quick perusal of the latest U.S. Merger Guidelines makes it clear that the economic effects of vertical integration are no longer lumped together with those of horizontal integration.⁷

Nevertheless, the theory is highly incomplete. It has been worked out only under a very restrictive set of circumstances (successive monopoly), and the theory such as it is cannot explain the practice of minimum resale price maintenance (min RPM), whereby an upstream firm imposes a floor on the retail price that can be charged for its product. Both are shortcomings. The former is concerning because one does not see many cases of successive monopoly in practice. There is thus a need to establish the robustness of the insights in more common settings. The latter is concerning because, among the different forms of resale price maintenance, min RPM has historically been the most commonly observed form.⁸ Not being able to account for this form limits the theory's applicability.

Despite these shortcomings, little has been done since Spengler's time to make the theory less restrictive and/or to expand its applicability. Early contributions sought to understand the effects of allowing the downstream sector to be perfectly competitive. It was found that there was no double marginalization in this case, and thus no need for vertical integration.⁹ Surprisingly, however, the case of imperfect downstream competition was not considered. There has also been no attempt to broaden the theory's applicability.

⁴Bowman (1952) defines resale price maintenance as "a system of pricing a trade-marked, branded, or otherwise identified product for resale in which a manufacturer restricts the price at which its product may be resold." It can be a maximum price (max RPM), a minimum price (min RPM), or a fixed-price.

⁵The U.S. Supreme Court ruled in *State Oil v. Khan*, 522 U.S. 3 (1997) that max RPM was not inherently unlawful, thereby overruling its decision in *Albrecht v. Herald*, 390 U.S. 145 (1968).

⁶Echoing this sentiment, Scherer and Ross (1990) wrote "The most plausible reason why a manufacturer would wish to prescribe price ceilings binding its resellers is to avoid repeated marginalization by vertically pyramided monopolists it is hard to see how such behavior could harm competition or consumers."

⁷See, for example, the joint DOJ/FTC 2010 Merger Guidelines, issued August 19, 2010.

⁸See the surveys by Gammelgaard (1958) and Overstreet (1983). See also the references cited therein.

⁹See Bork (1954) and the discussion of his and other Chicago School writings in Hovenkamp (2014).

Attempts to explain min RPM have all shined the spotlight elsewhere. Proponents of the Chicago school, for example, quickly latched onto Telser’s (1960) free-riding story as a possible explanation for min RPM (a story that has nothing to do with double marginalization),¹⁰ and even opponents of the Chicago school have tacitly accepted the view that min and max RPM are altogether different and require very different institutional settings.¹¹

In this paper, we will nevertheless attempt to do both: make the theory less restrictive (by extending it to the case of downstream competition) and expand the theory’s applicability (by accounting for both max and min RPM in the same institutional setting). Throughout, we will assume that the downstream firms are in control of a scarce asset (which can be thought of as shelf space), which must be obtained by the upstream firm before its product can be resold to consumers. We have in mind an institutional setting, such as in the grocery-retailing sector, in which obtaining the necessary space to display one’s product has become a routine cost of doing business for many consumer-goods manufacturers.¹² In this setting, it is understood that retailers must be allowed to earn at least their opportunity cost of shelf space if a manufacturer’s product is to be stocked.

Modifying the theory of double marginalization to examine explicitly the case of downstream competition and the role of shelf-space costs, we obtain three main results. First, we show that when shelf-space costs are added to the standard case of successive monopoly, there continues to be a double-marginalization problem (a la Spengler). The addition of shelf-space costs in this case mitigates but does not eliminate the adverse effects on consumers when the manufacturer is restricted to offering a linear wholesale price. Second, we show that when the theory is extended to allow for downstream competition, establishing that there is double marginalization in equilibrium (i.e., that all firms will have a markup) no longer necessarily implies that there is a double-marginalization problem. When shelf-space costs are sufficiently low (e.g., when they are zero), we find that in the “normal” case (given some additional restrictions on the properties of demand) there is always a double-marginalization problem. Allowing the manufacturer to obtain the integrated profit in these settings, either by vertically integrating with the downstream firm, or by engaging

¹⁰The story holds that min RPM can be used to prevent free riding on the provision of dealer services.

¹¹Summarizing the prevailing view that explanations of min and max RPM require different institutional settings, Scherer and Ross (1990) wrote “One type of manufacturer intervention, for example, the setting of *maximum* resale prices, may advance the producer’s interests best under one constellation of retailing conditions and a quite different one, for example, the setting of *minimum* resale prices, under another.”

¹²Beginning 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).

in other contractual solutions, necessarily leads to lower prices for consumers. In contrast, when shelf-space costs are above this threshold, there is no double-marginalization problem. Allowing the manufacturer to obtain the integrated profit in this case leads to higher prices for consumers. Third, we establish that there is an equivalence between (a) max RPM and two-part tariffs in which the retailers pay the manufacturer a fixed fee, and (b) min RPM and two-part tariffs in which the manufacturer pays the retailers a fixed fee (slotting allowance). The former arise when double marginalization is a problem. The latter arise when there is double marginalization, but no double-marginalization problem.

These findings have broad implications. 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 additional restrictions are placed on demand, restrictions that go beyond what is 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 these costs are large enough. Our findings also suggest that whereas policymakers have come to interpret Spengler’s insights as justifying a pro-competitive stance for max RPM (an interpretation that also holds in our model), they have failed to appreciate that the flip side of the coin is that the same theory can be interpreted as justifying an anti-competitive stance for min RPM — made possible because of our finding that not all double marginalization leads to a double-marginalization problem when shelf-space costs are sufficiently high. Lastly, our findings predict that there is an equivalence between min RPM and the offering of slotting allowances in the sense that both can be used under similar circumstances to compensate retailers for the opportunity cost of their shelf space. It follows that a policy that favors one over the other, whether intentional or not, may have little or no real consequences.¹³

Much of the literature on RPM has sought to explain why manufacturers might want to impose min RPM on their retailers. Some authors posit that min RPM can be used to foreclose potential rivals (e.g., see Asker and Bar-Issac, 2014), dampen competition among manufacturers (e.g., see Shaffer, 1991; and Rey and Verge, 2010), and/or support a manufacturer cartel (e.g., see Jullien and Rey, 2007), while others posit that min RPM can be used to certify retailer quality (e.g., see Marvel and McCafferty, 1984) and/or induce higher levels of services and promotions (e.g., see Telser, 1960; Mathewson and Winter, 1984; and Winter, 1993). None of these explanations apply in the present case. We have abstracted from the former by assuming a static setting in which there is only one active manufacturer, and we have abstracted from the latter by assuming that retail

¹³Policy that is permissive toward slotting allowances, but that frowns upon min RPM, for example, may simply cause firms to choose slotting allowances as their preferred means of compensating retailers.

prices are the only arguments in demand (implying that there is no need for additional services or promotions). The explanation we offer here is simply that min RPM may be needed to support higher profit margins for the retailers, so that in the course of selling the manufacturer’s product, they can be compensated for the cost of their shelf space.

Our explanation also differs from the “outlets hypothesis” explanation for RPM that was espoused in a series of papers by Mathewson and Winter (1983), Bittlingmayer (1983), Gallini and Winter (1983), and Perry and Groff (1985). In these models, retailers incur fixed start-up costs, the downstream market structure is determined by a free-entry condition, and the task of the manufacturer is to determine whether it wants to adopt a min RPM policy (in order to induce more retailers to enter), or a max RPM policy (in order to discourage the number of retailers that enter). They showed that, depending on the way demand is modeled (e.g., whether it has a spatial interpretation, or is viewed as coming from a representative consumer), either effect can dominate. Here, we differ in that the downstream market is not subject to free entry. Importantly, we also differ in that we find that the manufacturer’s decision as to whether to adopt max or min RPM depends only on the magnitude of the retailers’ shelf-space costs, not on the way demand is modeled.¹⁴

The literature on slotting allowance has grown substantially in recent years. As in the RPM literature, there are both pro-competitive stories and anti-competitive stories. The pro-competitive explanations focus mostly on the use of slotting allowances as a signaling device when manufacturers have private information about their demands (e.g., see Kelly, 1991; Lariviere and Padmanabhan, 1997; and Desai, 2000). There is also work on the use of slotting allowances to subsidize demand-enhancing investments by the retailers (e.g., see Raju and Zhang, 2004; and Kolay and Shaffer, 2013) and/or to increase the incentives for demand-enhancing investments by the manufacturers (e.g., see Farrell, 2001). The anti-competitive stories focus mostly on the use of slotting allowances to dampen competition (e.g., see Shaffer, 1991; Foros and Kind, 2008; and Innes and Hamilton, 2006), foreclose upstream competitors (e.g., Shaffer, 2005), foreclose downstream competitors (e.g., Marx and Shaffer, 2007), and facilitate tacit collusion (e.g., Piccolo and Miklos-Thal, 2012). None of these explanations, however, play a role in the present case. There are no demand-enhancing investments in our model, and no firm has private information. There is no foreclosure of competitors, nor is there any competition from upstream rivals to dampen.

The closest article in this literature to our paper is Kuksov and Pazgal (2007). Like us,

¹⁴We also share some common ground with Shaffer (1995). But the focus in his paper is on the use of RPM to reduce the profit that a retailer can earn by selling its most profitable alternative to the manufacturer’s product, when that alternative is a substitute for the manufacturer’s product. Because two-part tariffs are the benchmark in his model, the problem of double marginalization does not arise.

they consider the use of slotting allowances to compensate retailers for their fixed costs (actual costs in their model, opportunity costs in ours). Unlike us, however, they do not allow the manufacturer to impose RPM, nor do they solve the model for the case in which slotting allowances are not allowed. They also restrict attention to Hotelling demands.

The rest of the paper proceeds as follows. In the next section, we present the model and solve it for the case 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 the paper.

2 Successive monopoly

We begin by considering the case of successive monopoly (a la Spengler, 1950) modified to allow for shelf-space costs. In this case, an upstream manufacturer sells its product to a downstream retailer, which then resells the product to final consumers. We assume that resale requires that the product be on display, and that this display costs $SS \geq 0$ for the retailer to provide (one can think of SS as the retailer's opportunity cost of shelf space).

The game takes place in two stages. In the first stage, the manufacturer chooses its contract terms to maximize its profit subject to the retailer earning at least SS in profit. Initially, we will assume that the manufacturer is constrained to offering a linear contract with a per-unit wholesale price of w . Later, we will allow for two-part tariff contracts (w and a fixed fee F) and/or an RPM clause in the contract that allows the manufacturer to specify a retail price floor or ceiling. The retailer chooses its retail price in stage two.

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 profit from selling the manufacturer's product can be written as:¹⁵

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

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

¹⁵As is standard in the vertical-contracting literature, we assume for simplicity that the retailer incurs no marginal costs other than the per-unit price w that it pays the manufacturer for its product.

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

It follows that the retailer's best response in stage two depends on the relationship between π_r^* and SS . If $\pi_r^* \geq SS$, the retailer's best response is to set $p = p^*(w)$ and sell the manufacturer's product. If instead $\pi_r^* < SS$, 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 in this case 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 profitably be able to pass all of the increase in its costs onto consumers ($d\pi_r^*/dw < 0$).

Turning to the first stage, let $c \geq 0$ 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 maximand $(w - c)D(p^*(w))$ 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 ensures that the retailer will earn non-negative profit. 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 $w^c(SS)$ 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 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 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 (1) given that the constraint is not binding for all $w < w^c(SS)$). These observations follow from the definition of w^u and the fact that $\pi_r^*(w)$ is decreasing in w .

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 yields our first main result:

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

Proof: The proof 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 concave, and thus that $w = w^u$ satisfies the manufacturer's first-order condition

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

It follows that $w^u > c$ (because the manufacturer's first-order condition is positive when evaluated at $w = c$). Last, notice 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 Spengler's results to the case in which the downstream firm has shelf-space costs. The method of proof establishes that it is optimal for the manufacturer to impose a markup on its product when selling to the downstream firm, and that the downstream firm will then add its own markup, leading to a final price that is always higher than what a fully-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 is well known in the literature and has come to be known as the double-marginalization problem. 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 greater than it would otherwise be.

Another way to think 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. More formally, when the downstream firm chooses its final price p , it will choose p to satisfy

$$\frac{\partial \pi_r}{\partial p} = \frac{\partial \Pi}{\partial p} - \frac{\partial \pi_m}{\partial p} = 0, \quad (3)$$

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 share. 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 Spengler’s results extend 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 can at least mitigate the problem.

3 Downstream competition

We now extend the successive-monopoly game to one with multiple competing downstream firms. Specifically, we now focus on the case of an upstream monopolist that sells its product to two competing retailers, which then resell the product to final consumers. As before, we assume that resale requires that the manufacturer’s product be put on display, and that the cost of this display is $SS \geq 0$ for each retailer. The retailers are symmetric.

We want to know whether double-marginalization is still a problem, and what role the retailers’ shelf-space costs play in the outcome. To this end, we assume as before that the game takes place in two stages. In the first stage, the manufacturer chooses its wholesale price for retailer 1, w_1 , and its wholesale price for retailer 2, w_2 , to maximize its profit

subject to each retailer earning at least SS in profit.¹⁶ Once chosen, the contract terms become observable to all, and the retailers then simultaneously and independently choose their retail prices in stage two. We will assume initially that only linear contracts are feasible. Later, we will allow for the possibility of two-part tariffs and RPM provisions.

As before, we solve the game using backwards induction, and thus we begin with stage two. 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, retailer i 's profit from selling the manufacturer's product assuming it incurs no other marginal costs of distribution can be written as

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

We make the following assumptions on the retailers' demands and profits:

- (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:

$$\frac{\partial D_i}{\partial p_i} < 0; \quad \frac{\partial D_i}{\partial p_j} > 0; \quad \text{and} \quad \frac{\partial D_i}{\partial p_i} \frac{\partial D_j}{\partial p_j} - \frac{\partial D_i}{\partial p_j} \frac{\partial D_j}{\partial p_i} > 0. \quad (4)$$

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

$$\frac{\partial^2 \pi_i}{\partial p_i^2} < 0; \quad \frac{\partial^2 \pi_i}{\partial p_i \partial p_j} > 0; \quad \text{and} \quad \Delta \equiv \frac{\partial^2 \pi_i}{\partial p_i^2} \frac{\partial^2 \pi_j}{\partial p_j^2} - \frac{\partial^2 \pi_i}{\partial p_i \partial p_j} \frac{\partial^2 \pi_j}{\partial p_j \partial p_i} > 0. \quad (5)$$

The assumptions in (4) ensure that demands are downward sloping, the products sold by the two retailers are substitutes, and 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 firm's price.

The assumptions in (5) are also standard. The first condition ensures that retailer i 's profit function is well behaved and obtains its maximum at the unique p_i that solves

$$\frac{\partial \pi_i}{\partial p_i} = (p_i - w_i) \frac{\partial D_i}{\partial p_i} + D_i = 0. \quad (6)$$

The second condition ensures that the p_i that solves (6) is increasing in p_j (i.e., that retailer i 's best-response function is upward sloping). The third condition is the condition for the Jacobian of the system of profit-maximizing first-order conditions to be negative

¹⁶We will assume throughout this section that the additional demand from the second retailer is enough to justify paying the added fixed costs SS . Otherwise, we are back to the case of successive monopoly.

definite and ensures that a Nash equilibrium of the pricing game exists and is unique.¹⁷

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

$$\pi_i^{**}(w_1, w_2) = (p_i^*(w_1, w_2) - w_i)D_i(p_1^*(w_1, w_2), p_2^*(w_1, w_2)). \quad (7)$$

It follows that if retailer j 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 $\pi_i^{**} < SS$, 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).

Our assumptions in (4) and (5) ensure that consumers will be worse off when there is an equal increase in the manufacturer's wholesale prices.¹⁸ 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, our assumptions in (4) and (5) are not enough to establish that the retailers will also be worse off when there is an equal increase in the manufacturer's wholesale prices. For this, Δ must be sufficiently large.

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

$$\Delta > \frac{\partial D_i}{\partial p_j} \left(\frac{\partial^2 \pi_j}{\partial p_j \partial p_i} - \frac{\partial^2 \pi_j}{\partial p_j^2} \right) > 0. \quad (8)$$

Then, a marginal increase in the manufacturer's wholesale price w_i will (i) decrease retailer i 's equilibrium profit, (ii) increase retailer j 's equilibrium profit, and (iii) decrease the sum of the retailers' equilibrium profits when they are evaluated at symmetric final prices.

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 more costly for retailer i to purchase its inputs from the manufacturer. This harms retailer i for any given quantity purchased. The indirect effect arises because the increase in w_i leads retailer i 's rival to charge a higher final price in equilibrium. This benefits retailer i because the retailers' offerings are substitutes. The two effects thus go in opposite directions, and in general, the net effect is ambiguous. With

¹⁷These conditions are sufficient but not necessary. See Freidman (1983).

¹⁸This can be shown by totally differentiating the retailers' first-order conditions in (6).

the condition in (8), we ensure two things. First, we ensure that the net effect on retailer i 's profit is negative, i.e., that the direct effect on retailer i 's profit outweighs the indirect effect on retailer i 's profit. Second, we ensure that the net effect is sufficiently negative that it also outweighs the gain to retailer j when retailer i 's wholesale price increases.¹⁹

In what follows, we take condition (8) to be the normal case.²⁰ When it holds, it implies that retailers will be better off when their rival's wholesale price increases and worse off when their own wholesale price increases. Importantly, condition (8) along with symmetry also implies that own effects dominate cross effects in the sense that both retailers will be worse off when the manufacturer increases both wholesale prices equally and $w_1 = w_2$.²¹

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

$$\max_{w_1, w_2} \sum_{i=1,2} (w_i - c) D_i(p_1^*(w_1, w_2), p_2^*(w_1, w_2)) \quad (9)$$

such that

$$(p_i^*(w_1, w_2) - w_i) D_i(p_1^*(w_1, w_2), p_2^*(w_1, w_2)) \geq SS. \quad (10)$$

We assume that second-order conditions are such that there is a unique solution to (9), and note that symmetry implies that $w_1 = w_2$ at the optimum. Solving 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 (9), and $w_1 = w_2 = \hat{w}^c(SS)$ satisfy the constraints in (10) with equality. That $w^{**} = \min\{\hat{w}^u, \hat{w}^c(SS)\}$ follows from the definition of \hat{w}^u and our assumption that (8) holds in the normal case (because (8) together with symmetry implies that $\pi_i^{**}(w^{**}, w^{**})$ is decreasing in w^{**}).²²

We have thus far implicitly assumed that the retailers' shelf-space costs are such that the manufacturer will want to serve both retailers (i.e., not so high that the manufacturer would only want to serve one retailer, and not so high that the manufacturer would not want to serve either retailer). With some abuse of notation, this assumption can be made

¹⁹To 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 (8) fails but $\Delta > \frac{\partial D_i}{\partial p_j} \frac{\partial^2 \pi_j}{\partial p_j \partial p_i} > 0$.

²⁰We show in the Appendix that condition (8) is always satisfied with linear demands.

²¹The 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 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 , is negative.

²²This is the first instance where we have made use of the condition in (8).

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 the abuse of notation is that $D(p)$ 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 that is to be maximized in (9)). Then, it can be shown that there exists a critical threshold level of costs, $\overline{SS} \equiv \pi_m^{**}(w^{**}, w^{**}) - \pi_m^*(w^*)$, 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 equilibrium final prices will be $p_1^*(w^{**}, w^{**}) = p_2^*(w^{**}, w^{**}) = p^*(w^{**}, w^{**})$.²³

As we did before, we are now ready to compare the equilibrium final prices (assuming that $SS \leq \overline{SS}$) to the prices $p_1 = p_2 = p^I$ that a fully-integrated firm would set, where

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

is the price that maximizes the joint profit of the manufacturer and both retailers.

Unfortunately, the comparison here is not as clear cut as it was in the previously considered case of successive monopoly. Neither is the intuition. We cannot simply infer as we did in the paragraphs following Proposition 1 that because of the negative externality that a downstream firm imposes on the upstream firm when it raises its price, the final price will necessarily be higher than what a fully-integrated firm would charge. When a downstream firm, say retailer i , chooses its profit-maximizing price, it will set p_i to satisfy

$$\frac{\partial \pi_i}{\partial p_i} = \frac{\partial \Pi}{\partial p_i} - \frac{\partial \pi_m}{\partial p_i} - \frac{\partial \pi_j}{\partial p_i} = 0, \quad (11)$$

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 taking its wholesale prices w_1, w_2 , as given, and $\pi_j = (p_j - w_j)D_j(p_1, p_2)$ is retailer j 's profit. Although it is easy to see that an increase in retailer i 's price does indeed impose a *negative* externality on the manufacturer's profit π_m when $w_1, w_2 > c$ (the second term in (11) is negative), it will now also be true that an increase in retailer i 's price imposes a *positive* externality on its rival (the third term in (11) is positive). Depending on the equilibrium prices, which in turn depend on the wholesale prices charged by the manufacturer, either effect can dominate.

Nor can one simply infer the answer by establishing as we did in the proof of Proposition 1 that the upstream firm will always want to add a markup to its product when selling to

²³This 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 $\pi_m^*(w^*) > \pi_m^{**}(w^{**}, w^{**})$.

the downstream firms — because the markups of the downstream firms under competition will be less than the markup that a fully-integrated firm would charge when $w_1 = w_2 = c$. This can be seen by noting that the fully-integrated firm's prices, $p_1 = p_2 = p^I$, satisfy

$$\frac{\partial \Pi}{\partial p_i} = (p_i - c) \frac{\partial D_i(p_i, p_i)}{\partial p_i} + D_i + (p_j - c) \frac{\partial D_j(p_i, p_i)}{\partial p_i} = 0, \quad (12)$$

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

$$\frac{\partial \pi_i}{\partial p_i} = (p_i - w_i) \frac{\partial D_i(p_i, p_i)}{\partial p_i} + D_i = 0. \quad (13)$$

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 (12) but does not appear in (13).²⁴ 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 (12) 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 there is a problem or not depends on how high the manufacturer's markup is, which depends on the cost of shelf space.

Proposition 2 *In the downstream-competition game with shelf-space costs, there exists $\widehat{SS} > 0$ such that $p^*(w^{**}, w^{**}) > p^I$ if shelf-space costs are below this level and (8) holds. For shelf-space costs that are above \widehat{SS} , and for which it continues to be a best-response for the manufacturer to induce both retailers to sell its product, it will be optimal for the manufacturer to induce equilibrium prices that are less than the fully-integrated prices.*

Proof: Shelf-space costs in this game are bounded below by zero and above by \overline{SS} . Consider first the case in which they are zero. In this case, we know that w_1^{**}, w_2^{**} maximize the objective in (9). 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 conveniently 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 when the manufacturer chooses its optimal wholesale prices w_i, w_j , it will choose w_i, w_j , to satisfy

$$\frac{\partial \pi_m^{**}}{\partial w_i} = \frac{\partial \Pi^{**}}{\partial w_i} - \frac{\partial \pi_1^{**}}{\partial w_i} - \frac{\partial \pi_2^{**}}{\partial w_i} = 0. \quad (14)$$

²⁴This term is strictly positive for all p_1, p_2 such that $p_j > c$ and $D_j(p_i, p_i) > 0$.

Since (8) 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 wholesale prices that would maximize overall joint profits, and thus that the induced final prices, $p^*(w_1^{**}, w_2^{**})$, will exceed the prices, p^I , that a fully-integrated firm would charge.

Now consider the case in which $SS = \overline{SS}$. In this case, we know that w_1^{**}, w_2^{**} satisfy (10) 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 less than or equal to \overline{SS} , the induced final prices will be less than p^I .

Q.E.D.

Proposition 2 establishes that the double-marginalization problem extends to the case of downstream competition if the retailers' shelf-space costs are sufficiently low and (8) holds. Thus, for example, there will be a double-marginalization problem in the normal case in the absence of shelf-space costs. The proof of this follows by establishing that when (8) 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 what it would choose if it were trying to maximize overall profits. Proposition 2 also establishes, however, that as shelf-space costs increase, the manufacturer will eventually become constrained to charge lower wholesale prices, and at some point, the 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 circumstances under which Spengler's insights (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 (8) 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 (8) 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 no longer be a double-marginalization problem, even in the absence of shelf-space costs. The other caveat is that

shelf-space costs must be sufficiently low for there to be a problem. It is easy to come up with examples in which shelf-space costs are such that the induced final prices are strictly below the integrated level. In these cases, policy implications have an unexpected twist.

4 Contractual solutions and an equivalence result

Spengler (1950) advocated vertical integration as a way to solve the double-marginalization problem in the case of successive monopoly. He suggested that vertical integration should be allowed because it would benefit consumers in the form of lower prices. What we have shown, however, is that while Spengler’s insights extend under some circumstances, they do not extend under others. Allowing vertical integration in the case of downstream competition when shelf-space costs are sufficiently high, for example, can harm consumers. In these instances, the induced final prices that would arise when the firms are independent would be below the integrated level. Allowing vertical integration would increase them.

Others have noted that two-part tariffs, or alternatively RPM, can sometimes be used to proxy for vertical integration.²⁵ In Spengler’s world of successive monopoly, for example, it is known that the manufacturer can use either two-part tariffs or max RPM to induce the vertically-integrated outcome by setting the terms to eliminate one of the markups.²⁶

Things are more complicated in the case of downstream competition with shelf-space costs, however, because the strategy of eliminating one of the markups 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:

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

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

²⁵Overstreet (1983) was one of the first to suggest that the double-marginalization problem in the successive-monopoly case could be addressed via max RPM or with a two-part tariff. Soon after, Mathewson and Winter (1984) established a methodology for determining when alternative contractual solutions could proxy for vertical integration. They focused on identifying potential externalities affecting retailers’ decisions and coming up with a minimally sufficient set of instruments to neutralize these externalities.

²⁶With 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.

$w_1 = w_2 = w^I$ 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.²⁷

Now consider the manufacturer's problem with RPM:

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

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 $w_1 = w_2 = \bar{w}$ such that $(p^I - \bar{w})D_i(p^I, p^I) = SS$. 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 surplus from them as possible, subject to each retailer earning at least SS in profit. Using the terminology of Mathewson and Winter (1984), the manufacturer needs two instruments to control two targets. Both two-part tariffs and RPM give this flexibility.

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 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 downstream-competition game with shelf-space costs, if (8) holds and $p^*(w^{**}, w^{**}) > p^I$, then the manufacturer can obtain the integrated outcome with either max RPM or, equivalently, a two-part tariff in which the fixed fee is positive. If instead (8) holds and $p^*(w^{**}, w^{**}) < p^I$, then the manufacturer can obtain the integrated outcome with either min RPM or, equivalently, a two-part tariff in which the fixed fee is negative.*

Proof: We have already shown that the manufacturer can obtain the integrated outcome with an optimally chosen RPM contract, and we have also shown that the manufacturer can obtain the integrated outcome with an optimally chosen two-part tariff contract. We now establish that max RPM is the optimal form if and only if it is optimal to offer a two-part tariff that has a positive fixed fee. By definition, $F^I > 0 \iff SS < (p^I - w^I)D_i(p^I, p^I)$. But $SS = (p^I - \bar{w})D_i(p^I, p^I)$. Rearranging gives $F^I > 0 \iff w^I < \bar{w}$. Since $\partial p^*(w, w)/\partial w > 0$ and $p^*(w^I, w^I) = p^I$, it follows that $F^I > 0 \iff p^*(\bar{w}, \bar{w}) > p^I$.

²⁷The non-integrated manufacturer earns less because it must give each retailer at least SS in profit.

Thus, at the RPM wholesale prices $w_1 = w_2 = \bar{w}$, both retailers will want to price above p^I . It follows that under RPM, the manufacturer will have to impose a price cap 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 (8) 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 (4) and (5) 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 (8) holds). Thus, it follows that $F^I > 0$.

Lastly, we establish that if (8) 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 (8) 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. These contracts, in contrast to the previous contracts, may harm consumers and thus may be anti-competitive.

Our finding that min RPM and two-part tariff contracts in which the retailers are paid a fixed fee (also known as a slotting allowance) can arise in equilibrium 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 a fully-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 a fully-integrated firm with the same marginal cost would charge.

Shelf-space costs turn out to matter critically in this determination. For sufficiently low levels of shelf-space costs (e.g., when they are zero), we have shown that as long as (8) holds, the manufacturer's markup will be such that the condition for double marginalization to be a problem will be satisfied. 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, the necessary and sufficient condition for double marginalization to be a problem will no longer be satisfied. It is at this point (and for even higher shelf-space costs) that we would expect to observe the manufacturer offering slotting allowances, or alternatively, adopting a policy of min RPM.

Both of these practices have received a lot of attention in recent years. Slotting allowances in particular have been the focus of several government investigations, and there have been more than eighty academic articles, working papers, and surveys written about them.²⁸ They are said to account for billions of dollars annually in the U.S. alone, and their importance (and alleged abuse) has led to the creation of an ombudsman in the U.K. to adjudicate the practice. Many of the existing studies of slotting allowances tend to focus on the use of slotting allowances to foreclose competitors or facilitate collusion, or posit that manufacturers have better information about their products and use slotting allowances as a signaling device. However, none of these explanations account for, or depend on, shelf-space costs to make them work. In contrast, our results strongly suggest that there is a positive relationship between the retailers who receive slotting allowances and their opportunity cost of shelf space. Specifically, slotting allowances 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 there is a causal relation. Slotting allowances arise because shelf-space costs are relatively high.

²⁸For an in-depth survey, see Shaffer (2013). The Federal Trade Commission conducted its own study of the practice (FTC, 2001; FTC, 2003) as did the Canadian Bureau of Competition (2002) and the Norwegian Competition Authority (2005). Slotting fees are discussed in the Competition Commission's investigation of grocery markets in 2008, and are the focus of the UK's *Groceries Supply Code of Practice*.

Min RPM has also been the focus of much academic study and has seen a revival in recent years thanks to the Supreme Court’s ruling in *Leegin* (2007),²⁹ which 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). The ruling 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.³⁰ To this debate, our findings offer a new wrinkle. They suggest that whereas the courts and policymakers have come to interpret the theory of double marginalization as justifying an unambiguously pro-competitive stance toward 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 tariffs (albeit with slotting allowances) 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 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 cannot easily be replicated with two-part tariff contracts in the absence of service externalities (see O’Brien and Shaffer, 1992; and Gabrielsen and Johansen, 2017).

In *Leegin* (2007), the Court sidestepped much of the debate by settling on a rule-of-reason approach in min RPM cases. 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 foster non-contractible demand-enhancing investments, did the practice have the effect of facilitating collusion, was it aimed at foreclosing competitors). Here, the theory need not take a backseat to these other more recognizable explanations. In the past, when it was legal, and before the advent of slotting allowances 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

²⁹See *Leegin Creative Leather Products, Inc. v. PSKS, Inc.*, 551 U.S. 877 (2007).

³⁰For 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.

compensating retailers for the 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 similar kinds of products.

5 Conclusion

In this paper, we have sought to extend the scope and applicability of the theory of double marginalization *a la* Spengler (1950). We have done so by extending Spengler's model to allow for both competing downstream firms and non-zero costs of shelf space. We found that whereas double marginalization is always a problem (although mitigated by shelf-space costs) in the case of successive monopoly, it is not always a problem when there is downstream competition. In particular, we found that when shelf-space costs are sufficiently high, there is no double-marginalization problem. In these instances, allowing the manufacturer to induce the integrated outcome leads to higher prices for consumers.

We further found that in obtaining the integrated outcome, there is an equivalence between (a) max RPM and two-part tariff contracts with positive fixed fees, and (b) min RPM and two-part tariff contracts with negative fixed fees (i.e., slotting allowances). The former are used by the manufacturer when there is a double-marginalization problem. The latter are used by the manufacturer when there is no double-marginalization problem.

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 will depend among other things on (a) the prevailing legal climate (i.e., whether there are sanctions on the use of RPM or slotting allowances), 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 need to spend resources to detect such cheating and to enforce the appropriate punishment. Two-part tariffs, however, are also not without costs. Slotting allowances, for example, may present a moral hazard dilemma. A retailer can accept the money but then shirk on the product's shelf-space allocation. In the extreme, she may even stock an alternative product instead.

Among RPM contracts, one might expect price floors to have lower overall monitoring costs. This is because unlike with 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 lump-sum payments to retailers inevitably involve moral hazard, and because min RPM may be relatively cheap to enforce, one might expect to observe price floors on occasion when both min RPM and slotting allowances are feasible. It also follows that the case for max RPM is necessarily a more cautious one given that transaction costs appear to be relatively 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) for the reasons discussed above. Perhaps this may help to explain why the incidence of min RPM on consumer goods sold in the U.S. reached up to 10% in the past when RPM was legalized (up to 44% in the U.K.),³¹ while the incidence of max RPM has never been very high.

Monitoring and enforcement 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 allowances, but that frowns upon min RPM, for example, may simply cause the majority of firms operating in this environment to choose slotting allowances over RPM as the preferred means of compensating retailers for the opportunity cost of their shelf space. Relaxing the policy on RPM, as was recently done by the Supreme Court in *Leegin*, would be expected over time to reduce the incidence of slotting allowances.

³¹See the discussion in Overstreet (1983), and the estimates in Herman (1959) and Pickering (1974).

Appendix

Proof of Lemma 1: We need to show that if (8) holds, then an increase in w_i will (i) decrease retailer i 's equilibrium profit, $\frac{\partial \pi_i^{**}}{\partial w_i} < 0$, (ii) increase retailer j 's equilibrium profit, $\frac{\partial \pi_j^{**}}{\partial w_i} > 0$, and (iii) decrease the sum of the retailers' equilibrium profits, $\frac{\partial \pi_i^{**}}{\partial w_i} + \frac{\partial \pi_j^{**}}{\partial w_i} < 0$, when the individual terms in this sum are evaluated at symmetric final prices $p_1^{**} = p_2^{**}$.

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

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

Next, we differentiate the system of first-order conditions in (6) to obtain

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

It is easy to see from our assumptions in (4) and (5) that (A.2) is positive. Substituting (A.2) into (A.1), using retailer i 's first-order condition, and simplifying, yields

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

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 (8) 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

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

Next, we differentiate the system of first-order conditions in (6) to obtain

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

It is easy to see from our assumptions in (4) and (5) 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.

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

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

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

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 (8) is exactly this condition, we have that $\frac{\partial \pi_i^{**}(w_1, w_2)}{\partial w_i} + \frac{\partial \pi_j^{**}(w_1, w_2)}{\partial w_i}$ is indeed negative.

Q.E.D.

Linear demands and condition (8): We have assumed in the text that condition (8) 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 (4) 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.
- Bittlingmayer, G., 1983, "A Model of Vertical Restriction and Equilibrium in Retailing," *Journal of Business*, 56, 477-496.
- 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., 1952, "Resale Price Maintenance: A Monopoly Problem," *Journal of Business*, 25, 141-155.
- Bowman, W., 1955, "The Prerequisites and Effects of RPM," *University of Chicago Law Review*, 22, 825-858.
- Bork, R., 1954, "Vertical Integration and the Sherman Act: The Legal History of an Economic Misconception," *University of Chicago Law Review*, 22, 157-201.
- 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.
- 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.
- Foros, Ø. and H. Kind, 2008, "Do Slotting Allowances Harm Retail Competition?," *Scandinavian Journal of Economics*, 110, 367-384.

- Friedman, J., 1983, *Oligopoly Theory*, Cambridge, Cambridge University Press.
- Gabrielsen, T. and B. Johansen, 2017, "Resale Price Maintenance with Secret Contracts and Retail Service Externalities," *American Economic Journal: Microeconomics*, 9, 63-87.
- Gallini, N. and R. Winter, 1983, "On Vertical Control in Monopolistic Competition," *International Journal of Industrial Organization*, 1, 275-286.
- 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.
- Hovenkamp, H., 2014, "Robert Bork and Vertical Integration: Leverage, Foreclosure, and Efficiency," *Antitrust Law Journal*, 79, 983-1001.
- Innes, R. and S. Hamilton, 2006, "Naked Slotting Fees for Vertical Control of Multi-Product Retail Markets," *International Journal of Industrial Organization*, 24, 303-318.
- Jullien, B. and P. Rey, 2007, "Resale Price Maintenance and Collusion," *Rand Journal of Economics*, 38, 983-1001.
- 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.
- 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, 1983, "Vertical Integration by Contractual Restraints in Spatial Markets," *Journal of Business*, 56, 497-517.
- Mathewson, G.F. and R. Winter, 1984, "An Economic Theory of Vertical Restraints," *Rand Journal of Economics*, 15, 27-38.
- 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.
- Perry, M.K. and R. Groff, 1985, "Resale Price Maintenance and Forward Integration into a Monopolistically Competitive Industry," *Quarterly Journal of Economics*, 100, 1293-1311.
- 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.
- Rey, P. and T. Verge, 2010, "Resale Price Maintenance and Interlocking Relationships," *Journal of Industrial Economics*, 58, 928-961.
- Scherer, F.M., 1980, *Industrial Market Structure and Economic Performance*, Chicago, Il: Houghton Mifflin, Co.
- Scherer, F.M. and D. Ross, 1990, *Industrial Market Structure and Economic Performance*, Boston, MA: 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., 1995, "Rendering Alternative Offerings Less Profitable with Resale Price Maintenance," *Journal of Economics & Management Strategy*, 3, 639-662.
- 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.offt.gov.uk/OFTwork/research/economic-research/completed-research>.
- Spengler, J., 1950, "Vertical Integration and Antitrust Policy," *Journal of Political Economy*, 58, 347-352.
- 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.