B Appendix B: Online appendix: Not for publication

B.1 Strategic consumers and advertising

In the text, we do not allow consumers to be strategic but simply suppose that high-search-cost consumers choose to purchase from the retailer with the lowest advertised price unless prices are the same, in which case, they randomize. In this appendix, we argue that such behavior is an equilibrium of a game that allows for strategic consumers who engage in reasoning consistent with a perfect Bayesian equilibrium.

Specifically, a retailer's advertised retail price is denoted p_j^a. We assume that this price is flexible, and therefore, p_j^a is set contemporaneously with the retailer's transaction price. We make the following assumption, which ensures that the advertised price is not simply cheap talk.

A1. Advertising cannot be fraudulent in the following sense: the advertised price can be no lower than the retail transaction price p_j, i.e., $p_j^a \geq p_j$.

In all other respects, we keep the game form the same as in Section 2.1. Given that some consumers can visit only a single retailer, beliefs as to p_j given p_j^a need to be considered, which results in the relevant equilibrium concept being (a refinement of) perfect Bayesian equilibrium.

In principle, there is considerable flexibility in assigning off-equilibrium beliefs and thereby inducing perverse equilibrium behavior. However, Assumption A1 specifies that a retailer cannot advertise a price below the price that it charges imposes real constraints. For example, suppose that a retailer charging a price of \$4 or charging a price of \$6 were expected to advertise the same price. The advertised price would have to be a price at or above \$6 to conform with this restriction. Suppose that this advertised price is \$20 and that these are the only two kinds of retailers expected to advertise at \$20 in equilibrium. Then, consumers expect some price above \$4 on average (if it were equally likely that retailers charge \$4 and \$6, consumers would anticipate \$5 on average).[66] Suppose that the retailer charging \$4 instead of advertising a price of \$20 advertises a price of \$4. Given the restriction that retailers cannot advertise at prices below what they charge, it must be that consumers anticipate paying a price at or below \$4, *even if this is an off-equilibrium advertised*

[66] Recall that retailers are indistinguishable to consumers, so that if $p_j^a = p_k^a$ then $E(p_j|p_j^a) = E(p_k|p_k^a)$.

price; subsequently, this advertised price attracts more consumers, as their expected surplus must be strictly greater and would, therefore, be more profitable.[67]

When there are no advertising restrictions, the simple logic in the example above has considerable bite. The underlying point is that for every advertised price arising in equilibrium, there is a unique retail transaction price. This situation creates advertised prices that are unambiguous and are *equivalent* to requiring that retailers will advertise the price that they actually charge.[68] The following result makes this argument precise for pure strategy equilibria.[69]

Proposition 9 *Suppose that there are no advertising restrictions. Then, in any (pure strategy) perfect Bayesian equilibrium, each p^a maps to a unique p on the equilibrium path.*

Proof. Suppose, toward a contradiction, that there is an equilibrium in which this is not the case. Then, the equilibrium must involve (at least) two retailers, who charge different transaction prices, choosing the same advertised price a. The other possibility is that one retailer sets the same advertised price for two different transaction prices, but this is ruled out by virtue of each retailer being able to set only one price.

First, it must be that all retailers advertising at a attract some customers. If not, then they can always set their advertised price equal to a level that does attract customers and still make a profit since the manufacturer cannot discriminate at the wholesale level.

Now, consider all retailers who advertise at a. Of this set, let $p_{\min}(a)$ denote the price of the retailer with the lowest actual price. Note that since there is more than one retailer that advertises at a, it must be the case that consumers anticipate a non-zero probability of an actual price strictly greater than $p_{\min}(a)$ and therefore expect a surplus smaller than that generated by receiving $p_{\min}(a)$ with certainty.[70]

[67]If, in equilibrium, competing retailers advertise and charge a price of \$6 but the retailer of interest has an actual price of \$4, any advertised price in the interval $[4, 6)$ is considered equivalent to advertising a price of \$4.

[68]Two mappings from advertised prices to transaction prices are equivalent if changing the set of prices from which advertised prices can be drawn makes no difference to the realized transaction prices or consummated transactions. That is, the language does not matter as long as the message is the same.

[69]The result can easily be extended to mixed strategy equilibria, albeit with a much stronger notation.

[70]Recall that from consumers' point of view, retailers are identical (aside from their advertised prices). Hence, consumers cannot condition their expectations of p_j on j.

Further, for all retailers that advertise at a, high-search-cost consumers randomize over these retailers.

The restriction that a retailer cannot charge a price higher than its advertised price implies that if the retailer charging $p_{\min}(a)$ advertised its actual price $p_{\min}(a)$, consumers could not put any probability on the retailer charging a higher price than $p_{\min}(a)$ (and might even put some probability on a lower price). That is, consumers anticipate an actual price that is equal to or lower than $p_{\min}(a)$. Expected surplus similarly strictly increases.

Taking together the observations in each of the three paragraphs above, we observe that setting the advertised price at $p_{\min}(a)$ would attract strictly greater demand when compared to choosing the advertised price a. This situation generates the required contradiction. ■

In particular, trivially, with no advertising restrictions, there is no loss in supposing that equilibria will have the intuitive property that consumer beliefs are monotone in the following sense.

Definition 1 *Consumers' beliefs are* ***monotone*** *if,* $\forall p_j^a > \widehat{p}_j^a$, $E(p_j|p_j^a) > E(p_j|\widehat{p}_j^a)$.

Definition 2 *A* ***monotone perfect Bayesian equilibrium*** *is a perfect Bayesian equilibrium in which consumers' beliefs are monotone.*

In the case of advertising restrictions, a retailer may not be able to set the advertised price equal to its actual price. Moreover, when the MAP price is high, the logic in the example above that underlies Proposition 9 can have little bite. Returning to the example above, suppose that there was a MAP price of \$10 and consumers expected that retailers advertising a price of \$20 were equally likely to actually charge \$4 or \$12, supported by the (off-equilibrium) beliefs that any retailer advertising anything between \$10 and \$20 actually charges \$9. Here, since the MAP restriction prevents the retailer charging \$4 from advertising at a sufficiently low price, this retailer does not gain from advertising at a price below \$20. A focus on monotone perfect Bayesian equilibrium rules out such perverse outcomes. Restricting equilibrium to requiring monotone beliefs leads retailers to advertise at as low a price as possible. The following result is immediate.

Proposition 10 *In a MAP regime, in all monotone perfect Bayesian equilibria, each retailer advertises its actual price unless the MAP restriction binds, and in this case, it advertises the MAP price; i.e.,* $p_j^a = \max\left\{p^{MAP}, p_j\right\}$.

In the paper, we assume that firms advertise a price $p_j^a = p_j$, or, if a MAP restriction exists, $p_j^a = \max\left\{p^{MAP}, p_j\right\}$. In this appendix, in the absence of fraudulent advertising (as under Assumption A1), we have shown that imposing MAP restrictionscan capture all the economic richness that a richer strategy space would deliver with only minimal (reasonable) restrictions on the set from which consumer beliefs are drawn. Further, these restrictions are required only when MAP policies bind.

B.2 Additional material for Section 2

The model in Section 2, while providing a clear intuition, is of course rather stylized in several respects. Here, we allow for greater generality in a few respects in order to explore the robustness of this intuition. We characterize the optimal MAP-based discriminatory price scheme by allowing for marginal cost differences between retailers and a richer correlation structure between consumer type (h or l) and search cost. We then show, via numerical simulation, that MAP-based discrimination is the optimal pricing policy over a range of parameter values. Further, the incentive constraints analogous to Equations (1) and (2) need not both bind in this richer setting.

In particular, while Section 2 supposes that only low-value consumers might be "shoppers" who visit both retailers at no cost, here, we also allow high-value consumers to be shoppers. Further, while Section 2 assumes that retailers are homogeneous, it may be more realistic to allow for some heterogeneity. There are, of course, many potential sources of heterogeneity; for example, some retailers with both a physical and an online presence may have different marginal costs than other retailers who are only online. Further, "non-shopper" consumers may find some retailers more salient than others,[71] in contrast to our assumption that non-shoppers are equally divided between the two retailers. Of course, this salience of a well-known retailer is likely to be reinforced by a MAP restriction that prevents rivals from advertising a lower price.

Adding heterogeneity along all these dimensions requires further notation. We therefore focus on the case in which retailers differ only in their marginal costs, in line with our analysis in Section 3. Specifically, we suppose that one of the retailers is a low-cost retailer (whom incurs no cost beyond payments to the manufacturer; and the other retailer incurs an additional cost of c_H per unit ($c_H > 0$). It is convenient to refer to the retailers as R_L and R_H, respectively. Recall that we suppose that a

[71] For example, De Los Santos et al. (2012) observe that in their data, "Amazon was visited in 74 percent of book transactions and that in only 17 percent did Amazon buyers browse any otherbookstore." (p.2961)

fraction λ of consumers have low valuation of the good. We suppose that a fraction of σ_L of low-value consumers are searchers and σ_H of high-value consumers are searchers. Otherwise, we assume the same timing and strategies as in Section 2.1.

The optimal MAP-based discriminatory price scheme is characterized as follows:

Proposition 11 *In the optimal MAP-based discriminatory price scheme (when feasible), the manufacturer sets*

$$T^*(w) = \min \left\{ \begin{array}{c} (h-w-c_H)(1-\sigma_H)\frac{1-\lambda}{2}, \\ (l-w)\left(\sigma_L\lambda + \sigma_H(1-\lambda) + (1-\sigma_L)\frac{\lambda}{2} + (1-\sigma_H)\frac{1-\lambda}{2}\right) \end{array} \right\}.$$

and

$$w^* = \arg\max_w \quad w\left(1-(1-\sigma_L)\frac{\lambda}{2}\right) + 2T^*(w)$$

s.t.

$$\begin{aligned} \left(\sigma_L\lambda + \sigma_H(1-\lambda) + (1-\sigma_L)\frac{\lambda}{2} + (1-\sigma_H)\frac{1-\lambda}{2}\right)(l-w) \;&\geq\; (1-\lambda)\frac{1+\sigma_H}{2}(h-w) \\ \left(\sigma_L\lambda + \sigma_H(1-\lambda) + (1-\sigma_L)\frac{\lambda}{2} + (1-\sigma_H)\frac{1-\lambda}{2}\right)(l-w-c_H) \;&\leq\; (h-w-c_H)(1-\sigma_H)\frac{1-\lambda}{2} \\ w \;&\leq\; \min\{l, h-c_H\}. \end{aligned}$$

R_L sets $p = l$ and R_H sets $p = h$.

Proof. The proof extends the intuition developed in Proposition 1. It is immediate that the discriminatory MAP scheme should involve the retail prices set at l and h. Since R_H has higher marginal costs of production than R_L, more industry profit is generated when the R_L sets a price of l (and sells a greater quantity) and R_H sets a price of h. As argued below, the manufacturer can ensure that this is the equilibrium outcome.

The fixed fee is set to extract as much surplus as possible, subject to both the low-cost retailer and high-cost retailer being willing to take up the contract; that is, it is equal to the minimum of their (gross of fixed fee) profits.

The manufacturer chooses the input price w^* to maximize its profits.

The penultimate two inequalities at the end of the statement of the proposition correspond to the incentive constraints that guarantee that the two retailers do not wish to deviate from their equilibrium pricing strategies. The first ensures that the low-cost retailer prefers charging a price of l to charging a price of h, and the second ensures that the high-cost retailer prefers charging a price of h to undercutting the low-cost retailer and attracting all searchers. The final inequality ensures that both retailers prefer to make positive sales. ■

Note that Example 1, in the main text suffices to demonstrate that there are parameter values in which the optimal MAP-based discriminatory scheme is feasible. We show that this finding is not a knife-edge result by considering a range of parameters through extending the example.

Figure B.1 below sets $h = 2$, $l = 1$ and $\lambda = 0.5$, as in Example 1. It sets $c_H = 0.2$ and varies σ_L and σ_H. In this example, the optimal non-MAP scheme is realized by setting $w = h$, resulting in a profit to the manufacturer of 1. The black region indicates where the optimal MAP-based discriminatory price scheme dominates the optimal non-MAP scheme for the manufacturer. In this region, it is never the case that both incentive constraints bind in the optimal MAP-based discriminatory price scheme.

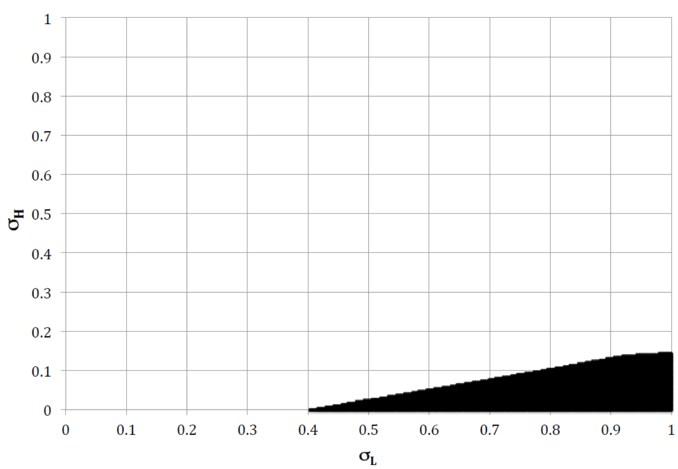

Figure B.1: Region with MAP-based discrimination being optimal, by (σ_H, σ_L)

B.3 Additional material for Section 3

Section 3.2: The manufacturer's problem with RPM

As described in the text, RPM can only play a role in inducing service provision only if it binds for both low and high cost realizations; that is, if $P > w + c_H$.

In this case, where both high- and low-cost firms set a price equal to $P > w + c_H$, demand for a retailer irrespective of its costs is given by $\frac{q(P)}{2}$. Service for low cost and high cost realization can be easily implicitly characterized through the first-order conditions described in the main text in (12) and (11). Finally, we can turn to the problem of the manufacturer who sets T equal to a retailer's expected profits so

$$\begin{aligned} T &= \alpha\left[(P-w-c)\frac{q(P)}{2}(1-(1-s_H)(1-\alpha s_H-(1-\alpha)s_L))-I(s_H)\right] \\ &\quad +(1-\alpha)\left[(P-w)\frac{q(P)}{2}(1-(1-s_L)(1-\alpha s_H-(1-\alpha)s_L)-I(s_L)\right], \end{aligned} \tag{37}$$

and chooses w and $P > w + c_H$ to maximize

$$2T+wq(P)\left[\alpha^2(1-(1-s_H)^2)+2\alpha(1-\alpha)(1-(1-s_L)(1-s_H))+(1-\alpha)^2(1-(1-s_L)^2)\right]. \tag{38}$$

Section 3.3: The manufacturer's problem with MAP

The manufacturer sets T as follows

$$\begin{aligned} T &= \alpha\left[(P-w-c)\frac{q(P)}{2}(1-(1-s_H)(1-\alpha s_H-(1-\alpha)s_L))-I(s_H)\right] \\ &\quad +(1-\alpha)\left[(p^m(w)-w)\frac{q(p^m(w))}{2}(1-(1-s_L)(1-\alpha s_H-(1-\alpha)s_L)-I(s_L)\right], \end{aligned} \tag{39}$$

and chooses w and $P > w + c_H$ to maximize expected profits: the sum of the fixed fee T and expected revenue from the per-unit fee:

$$2T+w\left[\begin{array}{c} \alpha^2(1-(1-s_H)^2)q(P)+(1-\alpha)^2(1-(1-s_L)^2)q(p^m(w)) \\ +\alpha(1-\alpha)(1-(1-s_L)(1-s_H))(q(P)+q(p^m(w))) \end{array}\right]. \tag{40}$$

B.4 Additional material for Section 4

While the main text considers a natural collusive scheme: the cartel mimics the same behavior that a monopolist supply chain selling a single good would set. However, colluding manufacturers could potentially achieve higher profits. This is because, with two retailers through which the manufacturers could sell, the industry is less likely to lose a sale by raising the wholesale price above v than it would in case there is only one retailer. With two retailers a consumer's maximal valuation is

$v + \max\{\xi_1, \xi_2\}$ rather than $v + \xi_1$ with only a single retailer. The optimal collusive scheme would take account of this fact.

We begin characterizing the solution to this form of cartel behavior by considering what would maximize cartel profits.

Proposition 12 *Suppose that the two manufacturers merge but continue to sell through the two retailers. The merged manufacturer optimally sets*

$$w^*_{NR} = \frac{3v + \sqrt{v^2+4}}{4}$$

and

$$T^*_{NR} = \frac{4v\left(v^2+3\right) - 4\left(1+v^2\right)\sqrt{v^2+4}}{384} + \frac{1}{6}$$

the merged manufacturer earns $\frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{48}$. *In this case observed retail prices will be in the interval* $\left[w^*_{NR}, w^*_{NR} + \frac{1}{2}\right]$.

Proof. First suppose that the manufacturer sets $w < v$ then following Proposition 5, the retailers set $p_j = w + \frac{\xi_j}{2}$.

Since $v > w$ it is immediate that the manufacturer sells with probability 1 and would benefit from raising w. Thus the manufacturer optimizes by setting $w \geq v$.

When it does so then its profits from the per-unit charge can be calculated by considering the probability of sale as a function of w. It is easier to consider the probability of no sale at either retailer, which requires that $\xi_1 < p_1 - v$ and $\xi_2 < p_2 - v$ or equivalently $\xi_i < \frac{w+v}{2} + \frac{\xi_i}{2} - v$ iff $\xi_i < w - v$ for $i = 1, 2$. It follows that the Probability that there is no sale is that $(w-v)^2$ and the probability that there is a sale is $1 - (w-v)^2$.

Next consider each retailer's profit (excluding any fixed fee) when the wholesale price is w. This is $E\left[\frac{\xi_j}{2} | \xi_j > \max\left\{\xi_{-j}, w - v\right\}\right]$ since retailer j makes a sale only when the consumer prefers this retailer to the other (which requires $\xi_j > \xi_{-j}$) and to the outside option (which requires that $w - v$); further, when retailer j makes a sale its profit is $\frac{\xi_j}{2}$.

$$\begin{aligned} E\left[\frac{\xi_j}{2}|\xi_j > \max\left\{\xi_{-j}, w-v\right\}\right] &= \frac{1}{2}E\left[\xi_j|\xi_j > \max\left\{\xi_{-j}, w-v\right\}\right] \\ &= \frac{1}{2}\int_0^1\int_{\max\{w-v,z\}}^1 ydydz \\ &= \frac{1}{2}\int_{w-v}^1\int_z^1 ydydz + \frac{1}{2}\int_0^{w-v}\int_{w-v}^1 ydydz \\ &= \frac{v^3-w^3+1}{6} - \frac{v^2w+vw^2}{2}. \end{aligned}$$

This can be extracted as a fixed fee and the manufacturer earns a per unit profit of w. Thus in the range $w \geq v$ (which we assume) and $w \leq v+1$ (which we will later verify must be the case) the manufacturer's overall profits are given by:

$$\Pi = w(1-(w-v)^2) + 2\left(\frac{v^3-w^3+1}{6} - \frac{v^2w+vw^2}{2}\right).$$

The SOC is given by $\frac{d^2}{dw^2}\Pi = 2v - 8w < 0$ where the inequality follows on noting that $w > v$.

The FOC is given by $\frac{d}{dw}\Pi = 0$ which has one solution with $w > v$ (and it is immediate that this solution involves $v+1 > w$) namely $w = \frac{3}{4}v + \frac{1}{4}\sqrt{v^2+4}$.

Manufacturer profits and fixed fee can be calculated through simple substitution to complete the result. ■

Proposition 12 establishes the maximal cartel profits and the retailer contracts that implement them. We now turn to consider the conditions under which the cartel can sustain them.

Proposition 13 *A collusive equilibrium in which manufacturers set $w_t = w^*_{NR} = \frac{3v+\sqrt{v^2+4}}{4}$ if $p^a_j \in \left[w^*_{NR}, w^*_{NR} + \frac{1}{2}\right]$. for all j in all past periods, and $w_t = 0$ otherwise, is supportable if*

$$\delta > \underline{\delta}_{NR} = \max\left\{\frac{12(w^*_{NR} - \pi^c_{NR})}{6(w^*_{NR} - \pi^c_{NR}) + 6w^*_{NR} - 1}, \frac{w^*_{NR}}{\pi^c_{NR} - \pi^p + w^*_{NR}}\right\}$$

where π^c_{NR} is the per-period profit earned by each manufacturer, in such an equilibrium and is given by $\pi^c_{NR} = \frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{96}$.

Proof. There are three possible deviations. The first is $w \in [0, w^*_{NR} - 1]$, denoted D1, note that this also implies that $w < v$. The second is $w \in \left(w^*_{NR} - 1, w^*_{NR} - \frac{1}{2}\right)$, denoted D2 and again here $w < v$. The third is $w \in \left[w^*_{NR} - \frac{1}{2}, w^*_{NR}\right]$, denoted D3.

Given these three types of deviation, there are three conditions that are necessary for collusion to be able to be sustained. These are

$$\frac{\pi^c_{NR}}{1-\delta} \geq \pi^{D1} + \Pr(p_j < w^*_{NR} | w^{D1}) \frac{\delta}{1-\delta}\pi^p + \left(1 - \Pr(p_j < w^*_{NR} | w^{D1})\right) \frac{\delta}{1-\delta}\pi^c_{NR} \quad (41)$$

$$\frac{\pi^c_{NR}}{1-\delta} \geq \pi^{D2} + \Pr(p_j < w^*_{NR} | w^{D2}) \frac{\delta}{1-\delta}\pi^p + \left(1 - \Pr(p_j < w^*_{NR} | w^{D2})\right) \frac{\delta}{1-\delta}\pi^c_{NR} \quad (42)$$

$$\frac{\pi^c_{NR}}{1-\delta} \geq \pi^{D3} + \Pr(p_j < w^*_{NR} | w^{D3}) \frac{\delta}{1-\delta}\pi^p + \left(1 - \Pr(p_j < w^*_{NR} | w^{D3})\right) \frac{\delta}{1-\delta}\pi^c_{NR} \quad (43)$$

where π^c_{NR} denotes the per-period collusive profit, following Proposition 12 $\pi^c_{NR} = \frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{96}$; if the collusion breaks down (which requires that deviation is observed—that is, the observed retail price is below w^*_{NR}), manufacturers earn the one-shot profit, which is denoted by $\pi^p = \frac{1}{6}$.

Note that it is immediate that deviating to $w > w^*_{NR}$ cannot be optimal.

The proof proceeds by establishing properties of the optimal deviation of each type, first by characterizing the retailer's pricing when $w_j < w_k = w^*_{NR}$ and then examining the manufacturer's problem in setting w_j for each type of deviation. Finally, a δ that is sufficient for none of the deviations to be attractive is derived.

Retailer pricing

First, consider the pricing of a retailer with a wholesale unit price of w_j facing a rival that prices in line with the cartel rule, such that $p_k = w^*_{NR} + \frac{\xi_k}{2}$.[72] The retailer's pricing problem is

$$p^*_j(\xi_j) = \arg\max_p \,(p - w_j) \Pr\left(v + \xi_j - p > \max\left\{v + \xi_k - w^*_{NR} - \frac{\xi_k}{2}, 0\right\}\right) \quad (44)$$

Note that $v + \xi_k - w^*_{NR} - \frac{\xi_k}{2} > 0$ iff $\xi_k > 2(w^*_{NR} - v) = 2(\frac{3v+\sqrt{v^2+4}}{4} - v) =$

[72] Recall that retailers are not part of any cartel agreement and compete in a static game.

$\frac{1}{2}\sqrt{v^2+4}-\frac{1}{2}v$ so we can write this as

$$\begin{aligned}
p_j^*(\xi_j) &= \arg\max_p (p-w_j)\Pr\left(v+\xi_j-p>v+\xi_k-w_{NR}^*-\frac{\xi_k}{2}|\xi_k>\frac{1}{2}\sqrt{v^2+4}-\frac{1}{2}v\right) \\
&\quad +(p-w_j)\Pr\left(v+\xi_j-p>0\right)\left(\frac{\sqrt{v^2+4}}{2}-\frac{v}{2}\right) \\
&= \arg\max_p (p-w_j)\Pr\left(2(w_{NR}^*+\xi_j-p)>\xi_k>\frac{1}{2}\sqrt{v^2+4}-\frac{1}{2}v\right) \\
&\quad +(p-w_j)\Pr\left(v+\xi_j-p>0\right)\left(\frac{\sqrt{v^2+4}}{2}-\frac{v}{2}\right) \\
&= \arg\max_p (p-w_j)\left(\max\{1,2(w_{NR}^*+\xi_j-p)\}-\frac{1}{2}\sqrt{v^2+4}-\frac{1}{2}v\right) \\
&\quad +(p-w_j)\left(\frac{\sqrt{v^2+4}}{2}-\frac{v}{2}\right)1_{v+\xi_j-p>0}
\end{aligned}$$

Note that at $\max\{1,2(w_{NR}^*+\xi_j-p)\}=1$ then the retailer would choose to set $p=w_{NR}^*+\xi_j-\frac{1}{2}$ and so $v+\xi_j-p=v+\frac{1}{2}-w_{NR}^*=v-\frac{3v+\sqrt{v^2+4}}{4}+\frac{1}{2}>0$. Trivially then at lower p where $\max\{1,2(w_{NR}^*+\xi_j-p)\}\neq 1$ then $1_{v+\xi_j-p>0}=1$. It follows that it is without loss of generality to write

$$p_j^*(\xi_j)=\arg\max_p (p-w_j)\Pr\left(v+\xi_j-p>v+\xi_k-w_{NR}^*-\frac{\xi_k}{2}\right) \tag{45}$$

Given that ξ_k is uniformly distributed on $[0,1]$, this amounts to maximizing

$$(p-w)\left(2\xi_j+2w_{NR}^*-2p\right) \tag{46}$$

in the region where $(2\xi_j+2w_{NR}^*-2p)\in[0,1]$ (outside of this region, either there is no chance of winning or the retailer wins for certain and is merely forgoing revenue by dropping the price). Ignoring this constraint results in the pricing rule $p^*(\xi_j):=\frac{w_{NR}^*+w_j}{2}+\frac{\xi_j}{2}$. If $2\xi_j+2w_{NR}^*-2p^*>1$, then the pricing rule is $p=w_{NR}^*-\frac{1}{2}+\xi_j$. Further, if $2\xi_j+2w_{NR}^*-2p^*(\xi_j)>1$, then the same inequality holds for all $\xi>\xi_j$. This expression results in the following retailer pricing rule when $w_j<w_{NR}^*$: $p_j=\max\left\{\frac{w_{NR}^*+w_j}{2}+\frac{\xi_j}{2},w_{NR}^*-\frac{1}{2}+\xi_j\right\}$.

Case D1

Next, we turn to considering the manufacturer's strategy. The first style of deviation ($D1$) involves choosing the optimal deviation in the interval $w\in[0,w_{NR}^*-1]$.

In this interval, by inspecting the retailer's pricing rule derived above, the probability of detection is equal to 0.5, regardless of w. Hence, the optimal deviation can be derived by maximizing the combined retailer-manufacturer-pair deviation payoff (since the manufacturer can extract the retailer's profit through the fixed fee). It is clear that this action is exactly what the retailer will do when $w_j = 0$. Hence, the profit maximizing $D1$ deviation arises when the retailer sets prices such that $p_j = w^*_{NR} - \frac{1}{2} + \xi_j$. From the manufacturer's point of view, this retailer pricing policy will arise for any $w \in [0, w^*_{NR} - 1]$ and can therefore be implemented in a variety of ways, which all result in the same deviation profit of w^*_{NR} (the lump-sum component of the manufacturer's two-part tariff will be used to extract the remaining expected profits from the retailer). Thus, the D1 deviation yields $\pi^{D1} = w^*_{NR}$ and $\Pr(p_j < v|w^{D1}) = \frac{1}{2}$.

Therefore, we can solve for the minimum δ such that a $D1$ deviation is not attractive. The $D1$ condition requires that

$$\pi^c_{NR}\frac{1}{1-\delta} \geq w^*_{NR} + \frac{1}{2}\frac{\delta}{1-\delta}\frac{1}{6} + \frac{1}{2}\frac{\delta}{1-\delta}\pi^c_{NR} \tag{47}$$

which implies that when $\delta \geq \frac{12(w^*_{NR}-\pi^c_{NR})}{6(w^*_{NR}-\pi^c_{NR})+6w^*_{NR}-1} = \frac{1}{6}\frac{18v-v^2\sqrt{v^2+4}+8\sqrt{v^2+4}+v^3-8}{54v-v^2\sqrt{v^2+4}+20\sqrt{v^2+4}+v^3-16}$, a $D1$ deviation is not attractive.

Case D2

The second style of deviation ($D2$) leaves the probability of detection unchanged at 0.5. To see this, note that for all $w \in \left(w^*_{NR} - 1, w^*_{NR} - \frac{1}{2}\right)$, $p_j = \max\left\{\frac{w^*_{NR}+w_j}{2} + \frac{\xi_j}{2}, w^*_{NR} - \frac{1}{2} + \xi_j\right\} = w^*_{NR}$ when $\xi = \frac{1}{2}$ and that p is monotonic in ξ; therefore, the deviation is detected for $\xi \leq \frac{1}{2}$ and undetected otherwise. Additionally, recall that when $w = 0$, the retailer chooses $p_j = w^*_{NR} - \frac{1}{2} + \xi_j$. Hence, by setting $w \in \left(w^*_{NR} - 1, w^*_{NR} - \frac{1}{2}\right)$, the manufacturer diminishes stage profits, with no compensating return in terms of adjusting the probability around detection (that is, leaving the continuation value unchanged). Hence, a $D2$ deviation must always be dominated by a $D1$ deviation.

Case D3

The third style of deviation ($D3$), involves choosing the optimal deviation in the interval $w \in \left[w^*_{NR} - \frac{1}{2}, w^*_{NR}\right)$. In this interval, a change in w will affect the probability of detection; note that $\frac{w^*_{NR}+w_j}{2} + \frac{\xi_j}{2} > w^*_{NR} - \frac{1}{2} + \xi_j$ iff $w_j - w^*_{NR} + 1 > \xi_j$

and so

$$\begin{aligned}
\Pr(p_j &< w^*_{NC}|w) = \Pr(w^*_{NR} > \frac{w^*_{NR}+w}{2} + \frac{\xi_j}{2} \text{ and } w - w^*_{NR} + 1 > \xi_j) \\
&\quad + \Pr(w^*_{NR} > w^*_{NR} - \frac{1}{2} + \xi_j \text{ and } w - w^*_{NR} + 1 < \xi_j) \\
&= \Pr(w^*_{NR} - w > \xi_j, w - w^*_{NR} + 1 > \xi_j) + \Pr(\frac{1}{2} > \xi_j > 1 + w - w^*_{NR}) \\
&= \Pr(w^*_{NR} - w > \xi_j, w - w^*_{NR} + 1 > \xi_j) \\
&= w^*_{NR} - w
\end{aligned}$$

Consequently, $\frac{d\Pr(p_j<w^*_{NC}|w)}{dw} = -1$.

Given that $w + 1 - w^*_{NR}$ is the value of ξ_j such that $\frac{w^*_{NR}+w}{2} + \frac{\xi_j}{2} = w^*_{NR} - \frac{1}{2} + \xi_j$, the deviation profit can be written as

$$\pi^{D3} = \int_0^{w_j+1-w^*_{NR}} \left(\frac{w^*_{NR}+w_j}{2} + \frac{x}{2}\right)(w^*_{NR} + x - w_j)\,dx + \int_{w_j+1-w^*_{NR}}^{1} \left(w^*_{NR} - \frac{1}{2} + x\right) dx. \tag{48}$$

It will be useful to note that $\frac{\partial \pi^{D3}}{\partial w_j} = -w_j\left(w_j + 1 - w^*_{NR}\right)$ and that Equation (48) also describes π^{D2}.

The optimal $D3$ deviation is the solution to

$$\max_{w_j} \quad D3(w_j) \equiv \max_{w_j} \quad \pi^{D3} + \Pr(p_j < w^*_{NR}|w_j)\frac{\delta}{1-\delta}\pi^p + (1 - \Pr(p_j < w^*_{NR}|w_j))\frac{\delta}{1-\delta}\pi^c_{NR} \tag{49}$$

which, taking the derivative with respect to w_j, yields

$$\begin{aligned}
\frac{\partial D3(w_j)}{\partial w_j} &= \frac{\partial \pi^{D3}}{\partial w_j} - \frac{\partial \Pr(p_j < w^*_{NR}|w_j)}{\partial w_j}\frac{\delta}{1-\delta}\left(\pi^c_{NR} - \pi^p\right) && (50) \\
&= -w_j\left(w_j + 1 - w^*_{NR}\right) + \frac{\delta}{1-\delta}\left(\pi^c_{NR} - \pi^p\right), && (51)
\end{aligned}$$

where the last equality follows by substituting for the two derivatives, as calculated above. Note that the second derivative is $\frac{\partial^2 D3(w_j)}{\partial w_j^2} = w^*_{NR} - 1 - 2w_j < 0$ in the range $w \in [w^*_{NR} - \frac{1}{2}, w^*_{NR})$.

Necessary conditions for the existence of an optimal D3 deviation (i.e., an interior solution in $\left[w^*_{NR} - \frac{1}{2}, w^*_{NR}\right))$ are that $\left.\frac{\partial D3(w_j)}{\partial w_j}\right|_{w_j = w^*_{NR} - \frac{1}{2}} \geq 0$ and $\left.\frac{\partial D3(w_j)}{\partial w_j}\right|_{w_j = w^*_{NR}} < 0$.

From Equation (51), $\left.\frac{\partial D3(w_j)}{\partial w_j}\right|_{w_j=w^*_{NR}-\frac{1}{2}} \geq 0$ implies that $\delta \geq \frac{2w^*_{NR}-1}{2w^*_{NR}-1+4(\pi^c_{NR}-\pi^p)}$. Similarly, $\left.\frac{\partial D3(w_j)}{\partial w_j}\right|_{w_j=w^*_{NR}} < 0$ implies $\delta < \frac{w^*_{NR}}{\pi^c_{NR}-\pi^p+w^*_{NR}}$. Hence, for a $D3$ deviation to exist, it must be that $\delta \in \left[\frac{2w^*_{NR}-1}{2w^*_{NR}-1+4(\pi^c_{NR}-\pi^p)}, \frac{w^*_{NR}}{\pi^c_{NR}-\pi^p+w^*_{NR}}\right)$ and so no such deviation can exist if $\delta > \frac{w^*_{NR}}{\pi^c_{NR}-\pi^p+w^*_{NR}}$. This observation, together with the observation that $\delta \geq \frac{12(w^*_{NR}-\pi^c_{NR})}{6(w^*_{NR}-\pi^c_{NR})+6w^*_{NR}-1}$ is required to rule out a deviation of type $D1$, establishes the result. ■

B.4.1 RPM

Next we turn to consider the optimal collusive scheme under RPM (rather than the one that replicates a monopoly supply chain as in the main text).

If the manufacturers set $w = p$ then there is a sale unless both $p > v + \xi_1$ and $p > v + \xi_2$; this suggests that total cartel profits (when $p > v$ and under an interior solution) are given by $p(1 - (p - v)^2)$.

It is easy to verify that this is maximized at $p = \frac{2v+\sqrt{v^2+3}}{3}$ and that $\frac{2v+\sqrt{v^2+3}}{3} \in (v, v + 1)$. Moreover this implies that

$$\pi^c_{RPM} = \frac{1}{2}p(1-(p-v)^2) = \frac{1}{2}\frac{2v+\sqrt{v^2+3}}{3}\frac{3+v-\sqrt{v^2+3}}{3}\frac{3-v+\sqrt{v^2+3}}{3}.$$

This allows us to consider when the optimal RPM collusive scheme is sustainable. Specifically, we can characterize $\underline{\delta}^{RPM}$ through the following:

$$\frac{\pi^c_{RPM}}{1-\delta} \geq \pi^D_{RPM} + \frac{\delta}{1-\delta}\pi^p,$$

where $\pi^D_{RPM} = 2\pi^c_{RPM}$. This yields $\delta \geq \underline{\delta}^{RPM} = \frac{6\pi^c_{RPM}}{12\pi^c_{RPM}-1}$ and, thereby, establishes the following.

Proposition 14 *A collusive equilibrium in which manufacturers set* $w_t = p^{RPM}_t = \frac{2v+\sqrt{v^2+3}}{3}$ *as long as* $p^a_{j,t} = \frac{2v+\sqrt{v^2+3}}{3}$ *for all* j *in all past periods, and* $w_t = 0$ *otherwise, is supportable if* $\delta \geq \underline{\delta}^{RPM} = \frac{6\pi^c_{RPM}}{12\pi^c_{RPM}-1}$ *where* π^c_{RPM} *is the per-period profit earned by each manufacturer in such an equilibrium and is given by,* $\pi^c_{RPM} = \frac{1}{2}\frac{2v+\sqrt{v^2+3}}{3}\frac{3+v-\sqrt{v^2+3}}{3}\frac{3-v+\sqrt{v^2+3}}{3}$.

B.4.2 MAP

Since MAP operates, in effect, as a market division scheme, the optimal collusive MAP contract involves each of the manufacturers operating as a monopolist. Thus, Proposition 8 implements the collusive MAP scheme and determines when it is sustainable.

B.4.3 Restraints facilitating collusion

First note that in comparing the sustainable of collusion through MAP rather than the RPM, trivially $\underline{\delta}^{RPM} = \frac{6\pi^c_{RPM}}{12\pi^c_{RPM}-1} > \frac{1}{2}$ so the MAP scheme is more sustainable.

As for profitability of the two schemes, it can be shown that the cartel profits under MAP are necessarily higher than the cartel profits under RPM; equivalently $\frac{v}{2} + \frac{1}{3} - (\frac{1}{2}\frac{2v+\sqrt{v^2+3}}{3}\frac{3+v-\sqrt{v^2+3}}{3}\frac{3-v+\sqrt{v^2+3}}{3}) > 0$ in the range $v > 1$. To see this first note that the left hand side of this inequality is positive at $v = 0$ and that its derivative with respect to v is equal to $\frac{(2v^2+3)(\sqrt{v^2+3}-v)-3v}{18\sqrt{v^2+3}} > 0$ where this inequality follows on noting that the numerators is necessarily positive.

It remains to consider the no constraint case. We can compare $\underline{\delta}^{RPM}$ and $\underline{\delta}_{NR}$. Recall that

$$
\begin{aligned}
\underline{\delta}^{RPM} &= \frac{6\pi^c_{RPM}}{12\pi^c_{RPM}-1} = \frac{6(\frac{1}{2}\frac{2v+\sqrt{v^2+3}}{3}\frac{3+v-\sqrt{v^2+3}}{3}\frac{3-v+\sqrt{v^2+3}}{3})}{12(\frac{1}{2}\frac{2v+\sqrt{v^2+3}}{3}\frac{3+v-\sqrt{v^2+3}}{3}\frac{3-v+\sqrt{v^2+3}}{3})-1} \\
\underline{\delta}_{NR} &= \max\left\{\frac{12(w^*_{NR}-\pi^c_{NR})}{6(w^*_{NR}-\pi^c_{NR})+6w^*_{NR}-1}, \frac{w^*_{NR}}{\pi^c_{NR}-\pi^p+w^*_{NR}}\right\} \\
&= \max\left\{\frac{12(\frac{3v+\sqrt{v^2+4}}{4}-\frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{96})}{6(\frac{3v+\sqrt{v^2+4}}{4}-\frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{96})+6\frac{3v+\sqrt{v^2+4}}{4}-1},\right. \\
&\qquad \left.\frac{\frac{3v+\sqrt{v^2+4}}{4}}{\frac{16+36v+2v^2\sqrt{v^2+4}+8\sqrt{v^2+4}-2v^3}{96}-\frac{1}{6}+\frac{3v+\sqrt{v^2+4}}{4}}\right\}
\end{aligned}
$$

These are functions of v and can be simply plotted, we plot the first (RPM) in black and the latter (NR) in red in the graph below.

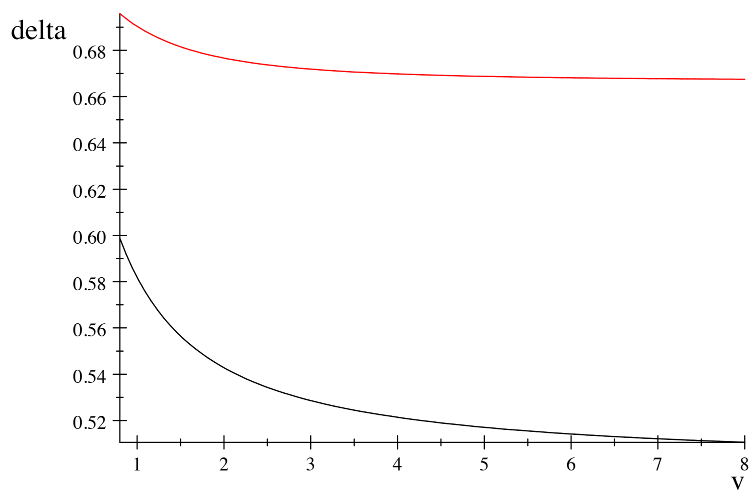

Figure B.2: $\underline{\delta}^{RPM}$ and $\underline{\delta}_{NR}$ against v

In the relevant range ($v > 1$) it follows that $\underline{\delta}_{NR} > \underline{\delta}^{RPM}$ and so it is easier to sustain collusion under RPM than under no constraint. Since as argued above $\underline{\delta}^{RPM} > \frac{1}{2} = \underline{\delta}^{MAP}$, it is also easier to sustain collusion under MAP than under no constraint.

Turning to compare profits, in contrast to the main text it is not immediate that $\pi^c_{MAP} > \pi^c_{NR}$ since under MAP there may be sales to the retailer with the lower-valued shock; however, we can write $\pi^c_{MAP} - \pi^c_{NR} = \frac{v}{2} + \frac{1}{3} - (\frac{4v\left(v^2+3\right)-4\left(1+v^2\right)\sqrt{v^2+4}}{384} + \frac{1}{6}) = \frac{15}{32}v + \frac{1}{96}v^2\sqrt{v^2+4} + \frac{1}{96}\sqrt{v^2+4} - \frac{1}{96}v^3 + \frac{1}{6}$.

Since $v^2\sqrt{v^2+4} > v^3$ for $v > 0$, it follows that this is positive.

Thus MAP is both more profitable and more sustainable than either RPM or no constraints.