Apt Demand

An appendix to *Pass-through as an Economic Tool*

Michal Fabinger and E. Glen Weyl

October 2009

I. Demand forms

The *Adjustable-pass-through* (Apt) demand class has two distinct forms, one useful for analyzing price games (vertical monopolies) and monopoly pricing along the cost-price dimension, the other useful for understanding quantity games (Cournot and Stackelberg competition) and monopoly pricing along the production dimension. We first develop the *Apt demand class* and then the *Apt inverse demand class*. Note that Apt inverse demand is not in any sense an inverse of Apt demand, but rather a form for inverse demand (mark-ups as a function of quantities) that is the dual of Apt demand's form (quantities as a function of prices).

A. Apt demand

Over the relevant range of prices, Apt demand takes, or is the limit of a demand taking, the form

$$Q(p) = \lambda \left(|\overline{\rho} - 1| \sqrt{|p - \tilde{p}|} - 2\overline{\rho}\alpha \right)^{\frac{2\overline{\rho}}{1-\overline{\rho}}} \tag{27}$$

However Apt demand, like the (Bulow and Pfleiderer, 1983) class, has slightly different precise formulations, which matter particularly outside of this "relevant range", depending on parameter values. The form of these are shown in Table 3 the main text and are pictured in Figure 5.

A few things are worthy of note about the case when $\overline{\rho} = 1$. Apt demand with *limiting constant mark-up* or *constant mark-up-in-the-limit* can take three forms. The first two have decreasing pass-through, while the final has constant pass-through. There is no constant mark-up-in-the-limit Apt demand with increasing pass-through. Again these are all shown in Table 3 above.

B. Apt inverse demand

Under *limiting quantity absorption* price as a function of quantity (inverse demand) is

$$P(q) = \begin{cases} M\left([1-\overline{\rho}_Q]\sqrt{\overline{q}-q} - 2\overline{\rho}_Q\alpha_Q\right)^{\frac{2\overline{\rho}_Q}{1-\overline{\rho}_Q}} + c & \textit{if} \quad q < \overline{q} - \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} \\ c & \textit{if} \quad \overline{q} - \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} \leq q \end{cases} \tag{28}$$

where c is set equal to the (assumed) linear cost[1] of the monopolist (or more generally the assumed symmetric linear costs of the quantity oligopolists). Matters are otherwise exactly as in the Apt demand class, except that a quantity analog replaces each price-related parameter: $\overline{\rho}_Q$ is the limiting *quantity* pass-through, $\overline{q}$ is the notional choke *quantity*, α_Q is the *quantity* shape and M is a mark-up level shift parameter. This means that the firm's mark-up, $m(q) \equiv P(q) - c$, take the same form, as a function of quantity, that Apt demand quantity takes as a function of price. This is why we refer to it as Apt inverse demand. If $\alpha_Q = 0$ this demand reduces to a quantity analog of the Bulow-Pfleiderer class

$$P(q) = \begin{cases} M\left([1-\overline{\rho}_Q]^2[\overline{q}-q]\right)^{\frac{\overline{\rho}_Q}{1-\overline{\rho}_Q}} + c & \textit{if} \quad q < \overline{q} \\ c & \textit{if} \quad \overline{q} \leq q \end{cases}$$

In the case that $\overline{\rho}_Q = \frac{1}{2}$ this yields linear demand. Other than this and a couple of other special cases, we are not aware of this (inverse) demand appearing in the literature. This is unfortunate because it gives rise to the same tractability (linear solutions) in quantity games that the Bulow-Pfleiderer class does in price (double marginalization or Rochet and Tirole (2003) two-sided markets, for example) games, while allowing for a wide range of (quantity) pass-through rates that are not possible under linear demand.

Under *limiting quantity amplification* Apt inverse demand is

$$P(q) = \begin{cases} \infty & \textit{if} \quad q \leq \underline{q} + \alpha_Q^2\overline{\rho}_Q\left(\frac{4\overline{\rho}_Q}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} - 1_{\alpha_Q<0}\right) \\ M\left(\overline{\rho}_Q[\overline{\rho}_Q+1]\right)^{-\frac{\overline{\rho}_Q+1}{\overline{\rho}_Q-1}}\left(\alpha_Q^2\right)^{-\frac{1}{\overline{\rho}_Q-1}}\left(q-\underline{q}+\alpha_Q^2\overline{\rho}_Q\right)^{-1} + c & \textit{if} \quad \underline{q} - \overline{\rho}_Q\alpha_Q^2 1_{\alpha_Q<0} < q \leq \underline{q} + \overline{\rho}_Q^2\alpha_Q^2 1_{\alpha_Q<0} \\ M\left([\overline{\rho}_Q-1]\sqrt{q-\underline{q}} - 2\overline{\rho}_Q\alpha_Q\right)^{-\frac{2\overline{\rho}_Q}{\overline{\rho}_Q-1}} + c & \textit{if} \quad \underline{q} + (\alpha_Q\overline{\rho}_Q)^2\left(1_{\alpha_Q<0} + \frac{4}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0}\right) < q \end{cases} \tag{29}$$

If $\alpha = \underline{q} = 0$ this demand becomes a natural analog of constant elasticity demand for the quantity dimension of the monopolist's problem: constant elasticity of mark-up with respect to quantity. We do not develop this explicitly here for the sake of brevity, but note that this inverse demand is simple for analyzing quantity problems in exactly the way constant

[1] If, in the case when costs are non-linear, c is replaced with the monopolist's average cost as a function of quantity, the same tractability is maintained. Whether and how this can be extended to the case of oligopoly is less clear and an interesting topic for future research.

elasticity demand is simple for analyzing a monopolist's optimal choice of price. We do not explicitly develop Apt inverse demand (or its surplus formulae) in the case of constant limiting quantity here. Explicit formulations are available on request and are embedded into the software accompanying this paper.

One unattractive feature of Apt inverse demand is that it requires a calibration to specific costs and is therefore less "clean". We view the dependence of the demand on cost as largely a technical convenience, rather than as a fundamental problem with Apt inverse demand. However it causes an important practical difficulty: it makes the attractive tractability of fragile to changes in cost. If costs only change locally, this problem does not arise. This problem also makes Apt inverse demand less natural than Apt demand in the case of non-linear or even asymmetric costs across oligopolists. This is a basic challenge in quantity competition models, and not one special to Apt inverse demand: for simple solutions, we need a simple form for mark-ups, rather than prices, and therefore the inverse demand must be based on costs to yield closed-form solutions except in very special cases such as linear demand.

II. Apt Properties

Apt demand and Apt inverse demand have a number of properties that make them useful for analyzing monopoly and sequential double marginalization and oligopoly problems, listed Subsection V.B of the text. In this section, we formally state and prove these properties. First we discuss basic issues of positivity, monotonicity and differentiability over the relevant range of prices. We then explore properties relating to monopoly optimization, particularly second-order conditions and the simple closed form solutions Apt demand and inverse demand yield for monopoly pricing and production. Next we demonstrate that Apt demand and inverse demand obey SPAs. We then develop the sense in which Apt demand is flexible in allowing for a wide range of independent variation in level of demand, market power, pass-through and slope of pass-through at any given equilibrium price. This is illustrated through a simple technique for estimating the Apt class by method of moments, given estimates of these local properties of demand, techniques implemented in accompanying software. The one constraint placed on these moments is that the slope of pass-through cannot be so large as to violate SPAs. We then discuss the simple closed-form of Apt demand consumer surplus.

A. Basic properties

The following proposition shows that Apt demand and inverse demand obey the most basic properties economists typically assume about demand

Proposition 3. *Apt demand is*

1. *Weakly positive*

2. *Strictly decreasing wherever it is strictly positive and finite and therefore globally non-increasing*

3. *Continuous and infinitely differentiable wherever it is strictly positive and finite, except (for some parameter values) at a single point it may be only twice continuously differentiable*

Apt inverse demand is

1. *Weakly above c and therefore weakly positive as long as $c \geq 0$*

2. *Strictly decreasing wherever it is strictly above c and finite and therefore globally non-increasing*

3. *Continuous and infinitely differentiable wherever it is strictly above cost and finite, except (for some parameter values) at a single point it may be only twice continuously differentiable*

Proof. 1. *Positivity*: Clearly $0, \infty \geq 0$ so we need not consider the constant portions of any form of Apt demand. Furthermore $\lambda, \overline{\rho}, \overline{\rho}+1, \alpha^2 \geq 0$, so we only need consider the expressions involving p, not the coefficients on transformations of these. First note that in all these expressions, the terms that are powers of $\sqrt{|p-\tilde{p}|}$ are always positive and are multiplied $|\overline{\rho}-1|$ or its square. Therefore in the case of $\alpha \leq 0$ the result is immediate as any power of a positive number is positive, as is any exponent to a real power. In the case when $\alpha > 0$ monotonicity, which we establish presently, implies weak positivity, so long as under cost-amplification $\lim_{p\to\infty} Q(p) \geq 0$. But clearly the relevant expression (that is raised to a power) increases without bound in p and therefore it must be always strictly positive for large enough p and therefore have a non-negative limit.

2. *Differentiability*: First note that away from functional break points, continuity and infinite differentiability are immediate, because the functions are analytic, so long as p never moves to the other side of $\tilde{p}$ and the argument that is raised to a power never changes signs. But we know neither of these occur from the proof of positivity. So we need only consider break points of the piecewise definitions. Furthermore the proposition only claims properties about demand at points where it is weakly positive and finite, so we need not consider the break points where demand is 0 above these points or ∞ below these, as demand is constructed right, rather than left, continuous at these points in the first case and left continuous at these points in the second case. This leaves only the break point between the two (finite) expressions when $\alpha < 0$ for cost-amplifying demand. These can be verified by simple calculations: at the pasting point, the level and first two derivatives of the two forms match, but the third derivative does not in general. We omit these calculations for brevity, but they are available on request. Finally, it is useful to note, though it is not stated in the proposition, that for cost-amplifying demand $\lim_{p \searrow \underline{p} + \overline{\rho}\alpha^2 1_{\alpha<0} + \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2} 1_{\alpha>0}} Q(p) = \infty$. For $\alpha < 0$ clearly $\lim_{p \searrow \underline{p} + \overline{\rho}\alpha^2} \frac{(\overline{\rho}[\overline{\rho}+1])^{-\frac{\overline{\rho}+1}{\overline{\rho}-1}} (\alpha^2)^{-\frac{1}{\overline{\rho}-1}}}{p - \underline{p} + \alpha^2\overline{\rho}} = \infty \cdot \text{sign}\left((\overline{\rho}[\overline{\rho}+1])^{-\frac{\overline{\rho}+1}{\overline{\rho}-1}} (\alpha^2)^{-\frac{1}{\overline{\rho}-1}} \right)$ and we already showed that $(\overline{\rho}[\overline{\rho}+1])^{-\frac{\overline{\rho}+1}{\overline{\rho}-1}} (\alpha^2)^{-\frac{1}{\overline{\rho}-1}} > 0$; note that this establishes the result for the case of cost-amplifying constant limiting mark-up. For $\alpha = 0$ clearly $\lim_{p \searrow \underline{p}} \lambda \left([\overline{\rho}-1]^2 [p - \underline{p}] \right)^{-\frac{\overline{\rho}}{\overline{\rho}-1}} = \infty$ as $\overline{\rho} > 1$. Finally if $\alpha > 0$ then

$$\lim_{p \searrow \underline{p} + \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}} \lambda \left([\overline{\rho}-1]\sqrt{p - \underline{p}} - 2\overline{\rho}\alpha \right)^{-\frac{2\overline{\rho}}{\overline{\rho}-1}} = \lim_{x \searrow \frac{2\alpha\overline{\rho}}{\overline{\rho}-1}} \lambda \left([\overline{\rho}-1]x - 2\overline{\rho}\alpha \right)^{-\frac{2\overline{\rho}}{\overline{\rho}-1}} = \infty$$

3. *Monotonicity*: We begin by considering cost-absorbing demand and then turn to cost-amplifying demand. Under constant mark-up this property is well-known from Bulow-Pfleiderer, so we do not explicitly establish it here.

   I first consider limiting cost absorption and then cost-absorbing constant limiting mark-up. For $\overline{\rho} < 1$ note that $(1-\overline{\rho})\sqrt{\overline{p} - p} - 2\overline{\rho}\alpha$ and $(1-\overline{\rho})^2[\overline{p} - p]$ are clearly decreasing in p and are raised to a positive exponent $\frac{2\overline{\rho}}{1-\overline{\rho}}$. Therefore for $\overline{\rho} < 1$ the main expression is monotonically decreasing. It remains to check that this main expression is always greater than the value (that is 0) of demand above the range where the main expression is valid. We do this separately for the case of $\alpha > 0$ and $\alpha \leq 0$ For $\alpha > 0$ by the same argument as in *Differentiability*, $\lim_{p \nearrow \overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}} Q(p) = 0$; therefore in this case demand is non-increasing and strictly decreasing wherever it is strictly positive. On the other hand we know from *Positivity* that if $p < \overline{p}$ and $\alpha \leq 0$, then $Q(p) > 0$ and that over this range of prices it is monotonically decreasing. Therefore demand must be lower at any $p' \geq \overline{p}$ than at any $p < \overline{p}$ ensuring that demand is non-increasing and monotonically decreasing wherever it is strictly positive. For $\overline{\rho} = 1$, $\alpha < 0$ so clearly $-\frac{\sqrt{\overline{p}-p}}{\alpha}$ decreases in p and therefore so does the exponential of this. An identical argument to the case of $\overline{\rho} < 1$ establishes from this the whole result for cost-absorbing constant limiting mark-up.

Now we turn to cost amplification, first considering limiting cost amplification and then cost-amplifying constant limiting mark-up. For $\overline{\rho} > 1$, note that $(\overline{\rho} - 1)\sqrt{p - \underline{p}} - 2\overline{\rho}\alpha$ is increasing in p while $-\frac{2\overline{\rho}}{\overline{\rho}-1}$ is a negative exponent. Therefore the main expression is decreasing in p, establishing monotonicity. For cost-amplifying $\overline{\rho} = 1$ the same argument applies noting that $\frac{\sqrt{p-\underline{p}}}{\alpha}$ decreases in p and therefore so does the exponential of it, as $\alpha < 0$.

B. Monopoly optimization

Proposition 4. *Apt demand satisfies strict MUC (i.e. $\frac{1}{Q}$ is strictly convex) wherever it is strictly positive and finite except if $\alpha > 0, \overline{\rho} > 1$ and $p \leq \underline{p} + \overline{\rho}^2\alpha^2$, in which case it satisfies weak MUC (i.e. $\frac{1}{Q}$ is weakly convex). Therefore any solution to equation (1) in the text is an optimal price for the monopolist and any such solution is unique if it exists so long as one of the three following conditions are met*

1. $\alpha \leq 0$
2. $\overline{\rho} \leq 1$
3. $p^\star > \underline{p} + \overline{\rho}^2\alpha^2$ *where* $p^\star$ *is the proposed solution*

Proof. First consider the main form of Apt demand

$$Q(p) = \lambda \left(|\overline{\rho} - 1|\sqrt{|p - \tilde{p}|} - 2\overline{\rho}\alpha \right)^{\frac{2\overline{\rho}}{1-\overline{\rho}}}$$

we have

$$\mu(p) = \frac{\overline{\rho} - 1}{\overline{\rho}}(p - \tilde{p}) - 2\alpha\sqrt{|p - \tilde{p}|}$$

It is easy to show that this formula for μ holds under constant limiting mark-up as well (in the Apt range), except in the case of constant mark-up in which μ is constant and obviously its slope is globally below 1. Under cost absorption we have

$$\mu'(p) = \frac{\overline{\rho} - 1}{\overline{\rho}} + \frac{\alpha}{\sqrt{\overline{p} - p}}$$

If $\alpha < 0$ this is clearly always less that 1. If $\alpha > 0$ this is monotonically increasing and therefore is bounded above by its value at the choke point :

$$\mu'\left(\overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}\right) = \frac{\overline{\rho} - 1}{2\overline{\rho}} < 0 < 1$$

Under cost amplification

$$\mu'(p) = \frac{\overline{\rho} - 1}{\overline{\rho}} - \frac{\alpha}{\sqrt{p - \underline{p}}}$$

For $\alpha \geq 0$, μ' (weakly) monotonically increases to $\frac{\overline{\rho}-1}{\overline{\rho}} < 1$ and therefore is globally below 1. For $\alpha < 0$ it monotonically decreases and therefore is globally bounded below its value at the switch point:

$$\mu'(\alpha^2\overline{\rho}^2) = 1$$

Therefore it is always strictly below 1 above this switch point. At the switch point it is equal 1 and continues to be through the intermediate range of demand (the second part of the piecewise definition) by the fact that this middle section is just a horizontally-shifted constant elasticity demand with an elasticity of 1 which it is well-known to have $\mu' = 1$, weak MUC. Note that this implies our last claim so long as the monopolist will never charge a price outside of the relevant range if there is a solution to her first-order conditions in the relevant range. But this is clearly true for cost-absorbing demand: she would give up positive profits for 0 profits. For cost amplifying demand it follows directly from the smoothness of the demand in its limit approaching the explosive price from above, proved earlier.

It should be clear that all of our claims above hold for the case where $\tilde{q} \neq 0$ as long as the requisite properties are required of the *residual demand* $Q(p) - \tilde{q}$ and it is this residual demand that takes the Apt form. By duality to MUC a natural second-order condition is to require that for all $\tilde{q}$ the monopolist's profits as a function of cost are globally concave. This requires by duality that $\frac{1}{m}$ be convex or that $\kappa' < 1$; we refer to this condition as *production contraction* (PC). By direct analogy, PC holds under Apt inverse demand. Furthermore, by duality, an analog of Theorem 1 from the text is trivial to prove for PC: price controls become production restrictions.

Corollary 3. *Apt inverse demand with a c parameter set by the monopolist's linear costs satisfies strict PC (i.e. $\frac{1}{m}$ is strictly convex) wherever m is strictly positive except if $\alpha >$*

$0, \overline{\rho} > 1$ and $q \leq \underline{q} + \overline{\rho}_Q^2 \alpha_Q^2$, in which case it satisfies weak PC (i.e. $\frac{1}{m}$ is weakly convex). Therefore any solution to equation (4) from the text is an optimal price for the monopolist and any such solution is unique if it exists so long as one of the three following conditions are met

1. $\alpha_Q \leq 0$
2. $\overline{\rho}_Q \leq 1$
3. $q^\star > \underline{q} + \overline{\rho}_Q^2 \alpha_Q^2$

where $q^\star$ is the proposed solution.

By duality of the monopolist's problem, concavity of profits in quantity for a particular cost implies concavity of profits in price given that cost. Therefore Apt inverse demand calibrated to cost c will give rise to a first-order solution to the monopoly pricing problem, just as Apt demand will give rise to a first-order solution to the monopoly production problem with $\tilde{q} = 0$. However Apt inverse demand does not satisfy MUC[2] nor does Apt demand satisfy PC as this concavity fails away from calibrated cost or exogenous quantity levels.

Proposition 4 justifies a first-order approach to solving for monopoly optimal prices (production) under Apt (inverse) demand. Because the two problems are duals, we only here show how the pricing problem is solved. Simple calculations show that under Apt demand (over the relevant range of prices)

$$\mu(p) = \frac{\overline{\rho} - 1}{\overline{\rho}}(p - \tilde{p}) - 2\alpha\sqrt{|p - \tilde{p}|}$$

Solving for the monopolist's optimal price is easy, assuming that an interior solution exist; as we will soon see, the conditions sufficient for this are economically intuitive. Here we explicitly take the reader through the necessary calculations to show how simple they are:

$$p - c = \frac{\overline{\rho} - 1}{\overline{\rho}}(p - \tilde{p}) - 2\alpha\sqrt{|p - \tilde{p}|} \tag{30}$$

so in the limiting cost-absorbing or cost-absorbing limiting constant mark-up cases this becomes

$$-(\overline{p} - p) + 2\alpha\overline{\rho}\sqrt{\overline{p} - p} + \overline{\rho}(\overline{p} - c) = 0$$

which is clearly a quadratic equation in $x \equiv \sqrt{\overline{p} - p}$

$$-x^2 + 2\alpha\overline{\rho}x + \overline{\rho}(\overline{p} - c) = 0$$

[2] A counter-example is available on request.

This equation has a unique, real, positive solution at prices yielding positive demand for x if $\overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2} 1_{\alpha>0} > c$. This is an intuitive condition that we assume. Were it to fail, the monopolist would have an incentive to shut down as any price she could charge yielding weakly positive profits would lead to 0 demand. The quadratic equation yields two solutions, the correct one being

$$\sqrt{\overline{p} - p} = \alpha\overline{\rho} + \sqrt{\alpha^2\overline{\rho}^2 + \overline{\rho}(\overline{p} - c)}$$

or

$$p_M = \overline{p} - \left(\alpha\overline{\rho} + \sqrt{\alpha^2\overline{\rho}^2 + \overline{\rho}(\overline{p} - c)}\right)^2 \tag{31}$$

where p_M is the monopolist's optimal price.

In the case of limiting cost amplification or cost-amplifying limiting constant mark-up we instead have

$$p - \underline{p} + 2\alpha\overline{\rho}\sqrt{p - \underline{p}} - \overline{\rho}(c - \underline{p}) = 0$$

which again is quadratic in $\sqrt{p - \underline{p}}$. For a unique solution interior to the region where demand takes the Apt form (i.e. equation (27)) we require $c > \underline{p} - \overline{\rho}\alpha^2 1_{\alpha<0} + \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2} 1_{\alpha>0}$. This is an equally intuitive condition to that required for limiting cost-absorbing demand: the firm could not earn infinite profits by charging a positive mark-up and still having infinite demand. Under this assumption the solution is given by

$$p_M = \underline{p} + \left(\sqrt{\alpha^2\overline{\rho}^2 + \overline{\rho}(c - \underline{p})} - \alpha\overline{\rho}\right)^2 \tag{32}$$

Thus in both cases we obtain a simple, closed form solution. Precisely the same reasoning shows that the same formulae hold *mutatis mutandis* for Apt inverse demand.

C. Signed pass-through assumptions

Apt demand satisfies the signed pass-through assumptions (SPAs) of Subsection I.E.

Proposition 5. *If $\overline{\rho} < (>)1$ Apt demand is globally strictly log-concave (convex). If $\overline{\rho} = 1$ and $\alpha < 0$, one formulation of Apt demand is globally strictly log-concave and another is globally strictly log-convex. If $\overline{\rho} = 1$ and $\alpha = 0$ then Apt demand is log-linear. Finally $\alpha < (> / =)0$ Apt demand has globally strictly decreasing (strictly increasing/constant) pass-through in the Apt range. These properties hold weakly wherever demand is finite and strictly posiive. In other words, Apt demand satisfies the SPAs.*

Proof. This is obvious for $\alpha \leq 0$. For $\alpha = 0$ we have globally constant mark-up and under cost-absorption pass-through must decrease from its (never-reached) limit of $\overline{\rho} \leq 1$. In the case of cost amplification it must decrease to its (never reached) limit of constant mark-up. Therefore in each case demand is either globally log-concave, globally log-convex or globally log-linear.

Now consider the case of $\alpha > 0$. Pass-through monotonically increases, so in the case of $\overline{\rho} < 1$ demand must be log-concave at any price p below $\hat{p}$ if demand is log-concave at $\hat{p}$. But note that demand is log-concave at the choke point, as shown in the proof of Proposition 4. Similarly in the case of limiting cost-amplifying demand must be log-convex at any price p above $\hat{p}$ if demand is log-convex at $\hat{p}$. But consider the explosive price $\underline{p} + \frac{4\alpha^2\overline{\rho}^2}{(\overline{\rho}-1)^2}$. At this price pass-through is

$$\frac{1}{1 - \frac{\overline{\rho}-1}{\overline{\rho}} + \frac{\alpha}{\sqrt{p-\underline{p}}}} = \frac{\overline{\rho}}{\overline{\rho} - \overline{\rho} + 1 + \frac{\alpha\overline{\rho}}{\frac{2\alpha\overline{\rho}}{\overline{\rho}-1}}} = \frac{2\overline{\rho}}{1+\overline{\rho}} > 1$$

given that $\overline{\rho} > 1$. Therefore limiting cost-amplifying demand is globally log-convex.

From the proof of Proposition 4 we have that, in the Apt range,

$$\mu''(p) = \frac{\alpha}{2\sqrt{|p - \tilde{p}|}}$$

and

$$\rho'(p) = \frac{\mu''(p)}{(1 - \mu'[p])^2}$$

So clearly $\rho'(p)$ has globally the same sign as μ'' in the Apt range. This covers the entire domain where demand is strictly positive and finite except the intermediate range for $\alpha > 0$ and $\overline{\rho} = 0$. In that range $\mu' = 1$ so $\mu'' = 0$ and pass-through is constant (at infinity). Therefore monotonicity of pass-through holds weakly over the full range of finite, strictly positive demand.

It is easy to see that this extends immediately to Apt inverse demand by duality.

D. Estimation

In this subsubappendix we discuss the estimation of Apt demand and inverse demand. We assume this estimation is based on data on second-order quantitively-measurable exogenous cost shocks, though not necessarily on the measurement of the level of cost, or on third-order price shocks. Demand levels are directly observable, elasticities can be measured using the effects of cost shocks, through prices, on demand, and pass-through rates and their slope can be measured through the first and second-order terms of the first-stage regression of prices on cost shocks or by the second- and third-order effects of price shocks on demand. Of course even if exogenous cost or price shocks are available, our estimation procedure requires a (very strong) assumption that the observed market is actually a monopoly, or at least that the calibrated residual demand curve is multiplicatively separable (does not strategically interact) in the prices of other firms. On the other hand if data on the responses of quantities or mark-ups at many levels of a GCS model are available, the data demands of identification can be greatly reduced as discussed in Appendix B, but for simplicity we do not consider this approach formally her. Another alternative approach is structural estimation by maximum likelihood. We eschew this as we believe that it obscures, rather than clarifies, the source of identification for such a simple demand function. However, we are aware that we thereby likely sacrifice some precision in estimation; we hope that in the future other researchers will improve on the approach here, which is implemented in statistical software accompanying this paper.

In both the (q_i, p_i, c_i) and (q_i, p_i) formulations, our estimate is anchored at a(n estimated) baseline equilibrium price $p^\star$. First when second-order cost shocks are available it results from estimating a two-stage regression. The first is a first-stage regression of prices (near $p^\star$) on exogenous cost variations (near $c^\star$)

$$p_i = p^\star + \rho(p^\star)(c_i - c^\star) + \frac{\rho(p^\star)\rho'(p^\star)}{2}(c_i - c^\star)^2 + \epsilon_i \tag{33}$$

where $c^\star$ is the average of the cost (shock), $\rho(p^\star)$, $\frac{\rho(p^\star)\rho'(p^\star)}{2}$ are the estimated coefficients, $p^\star$ is an estimated intercept, ϵ_i is an error term[3] with $E[\epsilon_i] = E\left[\epsilon_i(c_i - c^\star)\right] = E\left[\epsilon_i(c_i - c^\star)^2\right] = 0$, $(c_i - c^\star, p_i - p^\star)$ are observations of cost shocks and price responses. In the second stage, a regression of quantities on prices is estimated (presumably by the method of moments), using cost shocks as an instrumental variable. In particular the estimated equation is

$$q_i = q^\star - \frac{q^\star}{\mu(p^\star)}(p_i - p^\star) + \eta_i \tag{34}$$

where $\frac{q^\star}{\mu(p^\star)}$ is the estimated coefficient, $q^\star$ is an estimated intercept, η_i is an error term[4] with $E[\eta_i] = E[\eta_i(c_i - c^\star)] = E[\eta_i(c_i - c^\star)^2] = 0$, $(p_i - p^\star, q_i)$ are observations of price shocks and quantity responses. Because our focus here is not on the mechanics of estimation[5], we do not

[3] Note that ϵ_i is most easily formalized, by equations (31) and (32), as a linear approximation to the effect of a shock to $\tilde{p}$ (except in the case when $\alpha = 0$ and $\overline{\rho} = 1$) or as an unobserved cost shock. However it can also be formulated as a shock to $\overline{\rho}$ and/or α. We leave out these more structural interpretations, and stick with the reduced-form approach here, simply to save space. Notes making more rigorous the structural interpretation of ϵ_i are available on request.

[4] Again, we neglect a structural interpretation here, but a η_i is most easily thought of a linear approximation to the effect of a shock to λ.

[5] As described in the documentation attached to it, the accompanying computational toolkit uses the

discuss further the details of any particular estimation procedure. However it is worth noting that if a second-order term in the second-stage regression (34) can also be estimated, we have an over-identifying restriction, namely that the coefficient on this second-order (demeaned) term should be $\frac{(2\rho(p^\star)-1)q^\star}{2\mu(p^\star)^2\rho(p^\star)}$.

Alternatively if one only has data on exogenous price shocks and quantity responses (as is often the case) the appropriate estimating equation is

$$q_i = q^\star - \frac{q^\star}{\mu(p^\star)}(p_i - p^\star) + \frac{(2\rho(p^\star) - 1)\, q^\star}{2\mu(p^\star)^2\rho(p^\star)}(p_i - p^\star)^2 + \frac{1 - q^\star\rho'(p^\star) + 2\,(\mu[p^\star] - \rho[p^\star] - 2\mu[p^\star]\rho[p^\star])}{6\mu(p^\star)^2\rho(p^\star)}(p_i - p^\star)^3 + \nu_i \quad (35)$$

where again $\frac{q^\star}{\mu(p^\star)}$, $\frac{(2\rho(p^\star)-1)q^\star}{2\mu(p^\star)^2\rho(p^\star)}$ and $\frac{1-q^\star\rho'(p^\star)+2(\mu[p^\star]-\rho[p^\star]-2\mu[p^\star]\rho[p^\star])}{6\mu(p^\star)^2\rho(p^\star)}$ are the estimated coefficients, $q^\star$ is an estimated intercept, $p^\star$ is now the mean value of p, ν_i is an error term now with $E[\nu_i] = E[\nu_i(p_i - p^\star)] = E[\nu_i(p_i - p^\star)^2] = E[\nu_i(p_i - p^\star)^3] = 0$ and $(p_i - p^\star, q_i)$ are observations of exogenous price shocks and quantity responses.

Once we have estimated these regressions, we simply match the parameter values derived from solving the equations generated by the coefficients for $q^\star, p^\star, \mu(p^\star), \rho(p^\star)$ and $\rho'(p^\star)$ to those that would be generated by an appropriate Apt demand at that price. A primary virtue of Apt demand is that its parameters can be set to match, at any given price, an arbitrary (positive) level of demand, arbitrary (positive) market power, arbitrary (positive) pass-through and a wide range of slope of pass-through. There are only two constraints on this flexibility. The first is that demand must be positive, decreasing and satisfy MUC; thus $q^\star, \mu(p^\star), \rho(p^\star)$ must be greater than 0. This is hardly a restriction as we would likely conclude there was a flaw in our identifying restrictions if they gave negative estimates of these. The second restriction is more substantive, namely that Apt demand cannot match slopes of pass-through that are too large in absolute value, because this would lead to violation of the identifying assumptions proved in the previous section. This constraint binds more strongly for increasing than decreasing pass-through. In particular for Apt demand to match an observed slope of pass-through at price $p^\star$ it must be the case that $0 < \rho(p^\star) \leq 3$ and

$$-\frac{(\rho(p^\star) - 1)^2}{\mu(p^\star)} \leq \rho'(p^\star) < \frac{(\rho(p^\star) - 1)^2}{8\mu(p^\star)} \quad (36)$$

or $3 \leq \rho(p^\star)$ and

$$-\frac{(\rho(p^\star) - 1)^2}{\mu(p^\star)} \leq \rho'(p^\star) < \frac{\rho(p^\star) - 2}{2\mu(p^\star)} \quad (37)$$

If the estimated coefficients from equations (33) and (34) or equation (35) fail to satisfy

generalized method of moments and runs a second-order regression in both stages, imposing cross-equation coefficient restrictions corresponding to the structural connection between demand properties and pass-through rates, as discussed presently in the text.

one of these pairs of restrictions when estimated in a naïve independent manner, one would have to use a more sophisticated approach explicitly incorporating these constraints, such as maximum likelihood estimation. Again, in the interests of brevity, our focus is not on estimation here so we will simply assume that the coefficients are such that one pair of conditions does in fact hold and Apt demand can give a perfect (local) fit. In this case the parameters of the appropriate Apt demand take a simple closed form. If $\rho(p^\star) = 1$ then $\alpha = 0, \overline{\rho} = 1, \mu = \mu(p^\star)$ and $\lambda = q^\star e^{\mu p^\star}$. If $\rho(p^\star) < 1$ and $\rho'(p^\star) \neq 0$ then demand is limiting cost-absorbing or cost-absorbing limiting constant mark-up (if $-\frac{(\rho(p^\star)-1)^2}{\mu(p^\star)} = \rho'(p^\star)$) and

$$\overline{p} = p^\star + \frac{\rho(p^\star)}{4\rho'(p^\star)}\left(1 - \rho(p^\star) - \sqrt{(1-\rho(p^\star))^2 - 8\mu(p^\star)\rho'(p^\star)}\right) \quad (38)$$

$$\overline{\rho} = \frac{2\rho(p^\star)}{3 - \rho(p^\star) - \sqrt{(1-\rho(p^\star))^2 - 8\rho'(p^\star)\mu(p^\star)}} \quad (39)$$

$$\alpha = \frac{\text{sign}\left(\rho'[p^\star]\right)}{4\sqrt{\rho(p^\star)|\rho'(p^\star)|}}\left(1 - \rho(p^\star) - \sqrt{(1-\rho(p^\star))^2 - 8\mu(p^\star)\rho'(p^\star)}\right)^{\frac{3}{2}} \quad (40)$$

$$\lambda = q^\star\left([1-\overline{\rho}]\sqrt{\overline{p} - p^\star} - 2\overline{\rho}\alpha\right)^{-\frac{2\overline{\rho}}{1-\overline{\rho}}} \quad (41)$$

If $\rho'(p^\star) = 0$ and $\rho(p^\star) < 1$ then demand is Bulow-Pfleiderer cost-absorbing and

$$\overline{p} = p^\star + \frac{\mu(p^\star)\rho(p^\star)}{1-\rho(p^\star)} \quad (42)$$

$$\overline{\rho} = \rho(p^\star) \quad (43)$$

$$\alpha = 0 \quad (44)$$

$$\lambda = q^\star\left([1-\overline{\rho}]^2[\overline{p} - p^\star]\right)^{-\frac{\overline{\rho}}{1-\overline{\rho}}} \quad (45)$$

For $\rho(p^\star) > 1$ and $\rho'(p^\star) \neq 0$ then demand is limiting cost-amplifying or cost-amplifying constant limiting mark-up (if $-\frac{(\rho(p^\star)-1)^2}{\mu(p^\star)} = \rho'(p^\star)$) and

$$\underline{p} = p^\star - \frac{\rho(p^\star)}{4\rho'(p^\star)}\left(\rho(p^\star) - 1 - \sqrt{(\rho(p^\star)-1)^2 - 8\mu(p^\star)\rho'(p^\star)}\right) \quad (46)$$

$$\overline{\rho} = \frac{2\rho(p^\star)}{3 - \rho(p^\star) + \sqrt{(\rho(p^\star)-1)^2 - 8\rho'(p^\star)\mu(p^\star)}} \quad (47)$$

$$\alpha = \frac{\text{sign}\left(\rho'[p^\star]\right)}{4\sqrt{\rho(p^\star)|\rho'(p^\star)|}}\left(\rho(p^\star) - 1 - \sqrt{(\rho(p^\star)-1)^2 - 8\mu(p^\star)\rho'(p^\star)}\right)^{\frac{3}{2}} \quad (48)$$

$$\lambda = q^\star\left([\overline{\rho}-1]\sqrt{p^\star - \underline{p}} - 2\overline{\rho}\alpha\right)^{\frac{2\overline{\rho}}{\overline{\rho}-1}} \quad (49)$$

but if $\rho'(p^\star) = 0$ and $\rho(p^\star) > 1$ then demand is Bulow-Pfleiderer cost-amplifying and

$$\begin{aligned}
\underline{p} &= p^\star - \frac{\mu(p^\star)\rho(p^\star)}{\rho(p^\star) - 1} && (50)\\
\overline{\rho} &= \rho(p^\star) && (51)\\
\alpha &= 0 && (52)\\
\lambda &= q^\star \left([\overline{\rho} - 1]^2[p^\star - \underline{p}]\right)^{\frac{\overline{\rho}}{\overline{\rho}-1}} && (53)
\end{aligned}$$

Of course, as emphasized above, any other estimate of local pass-through rates and slope of pass-through rates would suffice for this estimation approach. The toolkit attached to this paper allows the user to either direct enter local properties or to estimate them from data.

The formulae above can be verified by tedious algebra that is available on request, but omitted here for brevity. However we do take the space demonstrate that, as long as the matched moments obey the restrictions we imposed by inequalities (36) or (37) the resulting parameters are consistent with the parametric and range restrictions required for demand to be in the Apt range .

For all the cases where $\overline{\rho}'(p^\star) = 0$ this is immediately clear by inspection and requires no proof. For the case of $\rho(p^\star) < 1$ and $\rho'(p^\star) \neq 0$ we want to show that $p^\star < \overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}$, $\overline{\rho} \in (0, 1], \alpha \neq 0$ and $\lambda > 0$. The last two are immediately clear by definition. For the first we need to show that $p^\star < \overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}$, which by equation (38) is equivalent to

$$\frac{\rho(p^\star)}{4\rho'(p^\star)}\left(1 - \rho(p^\star) - \sqrt{(1-\rho(p^\star))^2 - 8\mu(p^\star)\rho'(p^\star)}\right) > \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}$$

If $\rho'(p^\star) < 0$ this holds because $\alpha < 0$ and the term inside the root is strictly greater than $1 - \rho(p^\star)$ if and only if $\rho'(p^\star) < 0$; this also relies on the fact that the term inside the root is always positive, which is guaranteed by assumption (36). On the other hand when $\rho'(p^\star) > 0$ this becomes

$$\frac{\rho}{4\rho'}\left(1 - \rho - \sqrt{(1-\rho)^2 - 8\mu\rho'}\right) > 4\frac{1}{16\rho\rho'}\left(1 - \rho - \sqrt{(1-\rho)^2 - 8\mu\rho'}\right)^3 \frac{4\rho^2}{\left(3[1-\rho] - \sqrt{(1-\rho)^2 - 8\mu\rho'}\right)^2}$$

where we abbreviate $\rho \equiv \rho(p^\star), \rho' \equiv \rho'(p^\star)$ and $\mu \equiv \mu(p^\star)$. This is equivalent, by a bit of algebra

$$\left(3[1-\rho] - \sqrt{(1-\rho)^2 - 8\mu\rho'}\right)^2 > 4\left(1 - \rho - \sqrt{(1-\rho)^2 - 8\mu\rho'}\right)^2 \iff$$

$$3(1-\rho) > 2(1-\rho) - \sqrt{(1-\rho)^2 - 8\mu\rho'} \iff 1-\rho > \sqrt{(1-\rho)^2 - 8\mu\rho'}$$

which clearly holds as $\rho' > 0$. To see that $\overline{\rho} \in (0,1]$ note that by assumption $\rho > 0$ so $\rho \in (0,1]$ so long as

$$3-\rho-\sqrt{(1-\rho)^2-8\mu\rho'} \geq 2\rho \iff 3(1-\rho) > \sqrt{(1-\rho)^2-8\mu\rho'} \iff$$

$$(1-\rho)^2 \geq -\mu\rho' \iff -\frac{(1-\rho)^2}{\mu} \leq \rho'$$

as assumed in inequality (36). Thus the assumed restrictions generate the required parameter restrictions when $\rho' \neq 0$ and $\rho < 1$.

Now we consider the case of $\rho > 1$, again with $\rho' \neq 0$. First, we must show $\overline{\rho} \geq 1$. Then we need to establish that if $\rho' < 0$ then $p^\star - \underline{p} > \overline{\rho}^2\alpha^2$ and if $\rho' > 0$ that $p^\star - \underline{p} > \frac{4\overline{\rho}^2\alpha^2}{(\overline{\rho}-1)^2}$. To see that $\overline{\rho} > 0$ note that $\rho > 0$ by assumption and therefore

$$\overline{\rho} > 0 \iff \sqrt{(\rho-1)^2 - 8\rho'\mu} > \rho - 3$$

This holds automatically if $\rho \leq 3$. If $\rho > 3$ then it is true if and only if

$$(\rho-1)^2 - 8\rho'\mu > (\rho-3)^2 \iff \rho^2 - 2\rho + 1 - 8\rho'\mu > \rho^2 - 6\rho' + 9 \iff \rho' < \frac{\rho-2}{2\mu}$$

which is assumed in the case of $\rho > 3$ by inequality (37). Next $\overline{\rho} \geq 1$ is equivalent to

$$3-\rho+\sqrt{(\rho-1)^2-8\rho'\mu} > 2\rho \iff \sqrt{(\rho-1)^2-8\rho'\mu} > 3(\rho-1) \iff \rho' > -\frac{\rho-1}{\mu}$$

which is assumed by both inequalities (36) and (37). Now we show that if $\rho' < 0$ that $p^\star - \underline{p} > \overline{\rho}^2\alpha^2$. This is equivalent to

$$\frac{\rho}{4\rho'}\left(\rho-1-\sqrt{(\rho-1)^2-8\mu\rho'}\right) > \frac{4\rho^2}{16\rho\rho'}\frac{\left(\rho-1-\sqrt{(\rho-1)^2-8\mu\rho'}\right)^3}{\left(3-\rho+\sqrt{(\rho-1)^2-8\mu\rho'}\right)^2} \iff$$

$$\left(3-\rho+\sqrt{(\rho-1)^2-8\mu\rho'}\right)^2 > \left(\rho-1-\sqrt{(\rho-1)^2-8\mu\rho'}\right)^2$$

because $\rho' < 0$ this is equivalent to

$$3-\rho > 1-\rho$$

which clearly holds.

Finally in the case of $\rho' > 0$ we need to show that $p^\star - \underline{p} > \frac{4\alpha^2\overline{\rho}^2}{(\overline{\rho}-1)^2}$. We omit the proof of this because it is identical to the proof that $\overline{p} - p^\star > \frac{4\alpha^2\overline{\rho}^2}{(\overline{\rho}-1)^2}$ in the case of $\rho' > 0, \rho < 1$.

The procedure for fitting an Apt inverse demand is analogous. It would be exactly dual if data on exogenous quantity ($\tilde{q}$) shocks were available. However we below assume that instead we calibrate an Apt inverse demand, as Apt demand, based on cost shock data as we believe this is more realistic. This forces a translation between observed cost-price pass-through and its slope and the quantity pass-through (and its slope) that would hypothetically be observed. This translation is (perhaps surprisingly) easy, given our results on the connection between cost-price and quantity pass-through established in Subsection I.C.

Therefore estimation proceeds in essentially the same way. Throughout what follows we will, for simplicity, assume $\tilde{q} = 0$, though the procedure can easily be changed to accommodate the case when $\tilde{q} \neq 0$. Given a set of observations of cost shocks and corresponding price and quantity responses $(c_i - c^\star, p_i, q_i)$ we first estimate the regression

$$p_i = p^\star + \rho_Q(q^\star)(c_i - c^\star) - \frac{\rho_Q(q^\star)\rho_Q'(q^\star)q^\star}{2(p^\star - c)}(c_i - c^\star)^2 + \epsilon_i \tag{54}$$

where $c^\star$ is the average of the cost (shock), $\rho(p^\star)$, $\frac{\rho(p^\star)\rho'(p^\star)}{2(p^\star - c)}$ are the estimated coefficients, $p^\star$ is an estimated intercept, ϵ_i is an error term with $E[\epsilon_i] = E\left[\epsilon_i(c_i - c^\star)\right] = E\left[\epsilon_i(c_i - c^\star)^2\right] = 0$, $(c_i - c^\star, p_i - p^\star)$ are observations of cost shocks and price responses. Note that this is entirely analogous to the Equation (33); the only change is that the coefficients have been relabeled. In the second stage, a regression of quantities on prices is estimated, using cost shocks as an instrumental variable. In particular the estimated equation is

$$q_i = q^\star - \frac{q^\star}{p^\star - c}(p_i - p^\star) + \eta_i \tag{55}$$

where $\frac{q^\star}{p^\star - c}$ is the estimated coefficient, $q^\star$ is an estimated intercept, η_i is an error term with $E[\eta_i] = E[\eta_i(c_i - c^\star)] = 0$, $(p_i - p^\star, q_i - q^\star)$ are observations of price shocks and quantity responses. As with equation (54) *mutatis mutandis*, this is equivalent to equation (34), simply changing the expressions for the coefficients. Again if a second-order term in the second-stage regression (34) can also be estimated, we have the over-identifying restriction that the coefficient on this second-order (demeaned) term should be $\frac{(2\rho_q(q^\star)-1)q^\star}{2(p^\star-c)^2\rho_Q(q^\star)}$. If only data on exogenous production shocks and price responses are available, an analogous procedure to above is available, but we omit it here for brevity. However, it is implemented in the toolkit attached to this paper.

Once obtained these local properties can be used to fit an Apt inverse demand assuming that certain restrictions on the estimated coefficients are satisfied, just as in the case of Apt demand. We must have $p^\star > c, \rho_Q(q^\star) > 0$ and, of course, $q^\star > 0$. The estimated coefficients must also satisfy the quantity analog of conditions (36) and (37) *mutatis mutandis* as in the list below. So long as these hold, Apt inverse demand is fit exactly as Apt demand is, with c given by its estimated value and the rest of the values given by equations (38-53) with the following substitutions being made:

1. $\overline{p} \to \overline{q}$ or $\underline{p} \to \underline{q}$ as appropriate
2. $\overline{\rho} \to \overline{\rho}_Q$
3. $\alpha \to \alpha_Q$ or $\mu \to q^\star$ as appropriate
4. $\lambda \to M$
5. $p^\star \to q^\star$
6. $\rho(p^\star) \to \rho(q^\star)$
7. $\rho'(p^\star) \to \rho_Q(q^\star)$
8. $\mu(p^\star) \to q^\star$
9. $q^\star \to p^\star$

We omit all proofs as these are essentially identical to what was shown earlier in this subsection for the case of Apt demand.

E. Surplus

Consumer surplus takes a simple, closed form when demand is of the Apt or Apt inverse form.

Under limiting cost-absorption, consumer surplus

$$V(p) \equiv \int_p^{\overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}} Q(x)dx =$$

$$\begin{cases} \frac{Q(p)\left([1-\overline{\rho}^2][\overline{p}-p]-4\alpha\overline{\rho}^2\left[\alpha+\sqrt{\overline{p}-p}\right]\right)+\lambda(-2\overline{\rho}\alpha)^{\frac{2}{1-\overline{\rho}}}1_{\alpha<0}}{1+\overline{\rho}} & if \quad p \leq \overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0} \\ 0 & if \quad \overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0} < p \end{cases} \tag{56}$$

In the case of constant pass-through ($\alpha = 0$) this reduces to

$$V(p) = \begin{cases} (1-\overline{\rho})Q(p)(\overline{p}-p) & if \quad p \leq \overline{p} \\ 0 & if \quad \overline{p} < p \end{cases}$$

or more simply $V(p) = \overline{\rho}Q(p)\mu(p)$, paralleling Theorem 3 in Subsection I.D. While the Apt demand-based estimate, which also incorporates information about the slope of pass-through, does not have such a straight-forward justification, it seems a reasonable way to quantitatively refine locally-based estimates of consumer surplus.

Under limiting cost amplification, the formula is given by

$$V(p) = \begin{cases} \infty & \textit{if} \quad p \leq \underline{p}\overline{\rho}\alpha^2\left(\frac{4\alpha^2\overline{\rho}}{(1-\overline{\rho})^2}1_{\alpha>0} - 1_{\alpha<0}\right) \\ Q(p)\mu(p)\left(\frac{\overline{\rho}(\overline{\rho}+5)}{\overline{\rho}+1} + \log\left(\alpha^2\overline{\rho}[\overline{\rho}+1]\right) - \log(p - \underline{p} + \alpha^2\overline{\rho})\right) & \textit{if} \quad \underline{p} + \overline{\rho}\alpha^2 1_{\alpha<0} < p \leq \underline{p} - \overline{\rho}^2\alpha^2 1_{\alpha<0} \\ \frac{Q(p)\left([\overline{\rho}^2-1][p-\underline{p}]+4\alpha\overline{\rho}^2\left[\alpha-\sqrt{p-\underline{p}}\right]\right)}{1+\overline{\rho}} & \textit{if} \quad \underline{p} + \overline{\rho}^2\alpha^2\left(1_{\alpha<0} + \frac{4}{(1-\overline{\rho})^2}1_{\alpha>0}\right) < p \end{cases} \tag{57}$$

Again in the case of $\alpha = 0$ this converges to $V(p) = \overline{\rho}Q(p)\mu(p)$. The formulae for constant limiting mark-up are respectively for the cases of cost absorption, cost amplification and constant mark-up

$$V(p) = 2\lambda\alpha\left(\alpha - e^{-\frac{\sqrt{\overline{p}-p}}{\alpha}}\left[\alpha + \sqrt{\overline{p}-p}\right]\right) \tag{58}$$

$$V(p) = \begin{cases} \infty & \textit{if} \quad p \leq \underline{p} - \alpha^2 \\ \frac{2\lambda\alpha^2}{e}\log\left(\frac{2(e\alpha)^2}{p-\underline{p}+\alpha^2}\right) & \textit{if} \quad \underline{p} - \alpha^2 < p \leq \underline{p} + \alpha^2 \\ 2\alpha Q(p)\left(\alpha - \sqrt{p-\underline{p}}\right) & \textit{if} \quad \underline{p} + \alpha^2 < p \end{cases} \tag{59}$$

$$V(p) = \lambda\mu e^{-\frac{p}{\mu}} \tag{60}$$

We now establish the validity of these formulas. Surplus is simply given by the integral of demand so it is sufficient to first establish that at all points the derivative of the relevant function surplus formula is equal to the negative of demand and then to show that the formula converges to 0 as demand does.

First consider the case of limiting cost-absorption. From the formulae in Section V.B of the text we have that in the relevant range

$$V'(p) = \frac{Q'(p)\left([1-\overline{\rho}^2][\overline{p}-p] - 4\alpha\overline{\rho}^2\left[\alpha + \sqrt{\overline{p}-p}\right]\right) + Q(p)\left(\frac{2\alpha\overline{\rho}^2}{\sqrt{\overline{p}-p}} - 1 + \overline{\rho}^2\right)}{1+\overline{\rho}} =$$

$$Q(p)\left(\frac{\overline{\rho}\left[4\alpha\overline{\rho}^2(\alpha+\sqrt{\overline{p}-p}) - (1-\overline{\rho}^2)(\overline{p}-p)\right] + \sqrt{\overline{p}-p}\left[(1-\overline{\rho})\sqrt{\overline{p}-p} - 2\overline{\rho}\alpha\right]\left[\frac{2\alpha\overline{\rho}^2}{\sqrt{\overline{p}-p}} - 1 + \overline{\rho}^2\right]}{(1+\overline{\rho})\sqrt{\overline{p}-p}\left([1-\overline{\rho}]\sqrt{\overline{p}-p} - 2\overline{\rho}\alpha\right)}\right) =$$

$$Q(p)\left(\frac{\left(4\alpha\overline{\rho}^3 + 2\alpha\overline{\rho}[1-\overline{\rho}^2] + 2\alpha\overline{\rho}^2[1-\overline{\rho}]\right)\sqrt{\overline{p}-p} - (1-\overline{\rho}^2)(\overline{p}-p)}{(1+\overline{\rho})\sqrt{\overline{p}-p}\left([1-\overline{\rho}]\sqrt{\overline{p}-p} - 2\overline{\rho}\alpha\right)}\right) =$$

$$Q(p)\left(\frac{2\alpha\overline{\rho}(1+\overline{\rho})\sqrt{\overline{p}-p} - (1-\overline{\rho})(1+\overline{\rho})(\overline{p}-p)}{(1+\overline{\rho})\sqrt{\overline{p}-p}\left([1-\overline{\rho}]\sqrt{\overline{p}-p} - 2\overline{\rho}\alpha\right)}\right) = -Q(p)$$

Now consider

$$\lim_{p\nearrow\overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}} V(p)$$

In the case of $\alpha < 0$ we have this as

$$\lim_{p\nearrow\overline{p}} \frac{Q(p)\left([1-\overline{\rho}^2][\overline{p}-p] - 4\alpha\overline{\rho}^2\left[\alpha+\sqrt{\overline{p}-p}\right]\right) + \lambda\left(-2\overline{\rho}\alpha\right)^{\frac{2}{1-\overline{\rho}}}}{1+\overline{\rho}} = \lambda\frac{-4\alpha^2\overline{\rho}^2(-2\overline{\rho}\alpha)^{\frac{2\overline{\rho}}{1-\overline{\rho}}} + (-2\overline{\rho}\alpha)^{\frac{2}{1-\overline{\rho}}}}{1+\overline{\rho}} = 0$$

In the case of $\alpha = 0$, $V(p) = Q(p)\mu(p)\overline{\rho}$ and both $Q(p)$ and $\mu(p)$ head to 0 as p approaches $\overline{p}$, so clearly $V(p)$ also approaches 0. If $\alpha > 0$ then $\lim_{p\nearrow\overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}} Q(p) = 0$, so $\lim_{p\nearrow\overline{p}-\frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}} V(p) = 0$.

Next consider the case of limiting cost amplification. In the Apt range the envelope condition $V'(p) = -Q(p)$ follows from the same reasoning as in the limiting cost-absorbing case. In the intermediate range (for $\alpha < 0$) the envelope condition again holds:

$$V'(p) = (Q'[p]\mu[p] + \mu'[p]Q[p])\left(\frac{\overline{\rho}(\overline{\rho}+5)}{\overline{\rho}+1} + \log\left(\alpha^2\overline{\rho}[\overline{\rho}+1]\right) - \log(p+\alpha^2\overline{\rho})\right) - \frac{Q(p)\mu(p)}{p-\underline{p}+\alpha^2\overline{\rho}}$$

$$= -Q(p)$$

Here we have two conditions to check: continuity at the break point and approach to 0 as $p\to\infty$. At the break point when $\alpha < 0$ we have

$$\lim_{p\nearrow\underline{p}+\overline{\rho}^2\alpha^2} V(p) = \overline{\rho}^2\alpha^2(\overline{\rho}+5)Q(\underline{p}+\overline{\rho}^2\alpha^2)$$

while

$$\lim_{p\searrow\underline{p}+\overline{\rho}^2\alpha^2} V(p) = Q(\underline{p}+\overline{\rho}^2\alpha^2)\frac{(\overline{\rho}^2-1)\alpha^2\overline{\rho}^2 + 4\alpha\overline{\rho}^2(\alpha+\alpha\overline{\rho})}{1+\overline{\rho}} = \alpha^2\overline{\rho}^2 Q(\underline{p}+\overline{\rho}^2\alpha^2)(\overline{\rho}+5)$$

Furthermore

$$\lim_{p\to\infty} V(p) = \lim_{p\to\infty}\frac{\lambda}{1+\overline{\rho}}\frac{(\overline{\rho}^2-1)(p-\underline{p}) + 4\alpha\overline{\rho}^2\left(\sqrt{p-\underline{p}}-\alpha\right)}{\left([\overline{\rho}-1]\sqrt{p-\underline{p}}-2\overline{\rho}\alpha\right)^{\frac{2\overline{\rho}}{\overline{\rho}-1}}} =$$

by L'Hôpital's rule

$$\lim_{p\to\infty}\frac{\lambda}{2\overline{\rho}(1+\overline{\rho})}\frac{(\overline{\rho}^2-1)+\frac{2\alpha\overline{\rho}^2}{\sqrt{p-\overline{p}}}}{\frac{\left([\overline{\rho}-1]\sqrt{p-\underline{p}}-2\overline{\rho}\alpha\right)^{\frac{\overline{\rho}+1}{\overline{\rho}-1}}}{\sqrt{p-\underline{p}}}} = \lim_{p\to\infty}\frac{\lambda}{2\overline{\rho}(1+\overline{\rho})}\frac{(\overline{\rho}^2-1)\sqrt{p-\underline{p}}+2\alpha\overline{\rho}^2}{\left([\overline{\rho}-1]\sqrt{p-\underline{p}}-2\overline{\rho}\alpha\right)^{\frac{\overline{\rho}+1}{\overline{\rho}-1}}} =$$

applying L'Hôpital again yields

$$\lim_{p\to\infty} \frac{\lambda}{2\overline{\rho}} \frac{\overline{\rho}^2 - 1}{\left([\overline{\rho} - 1]\sqrt{p - \underline{p}} - 2\overline{\rho}\alpha\right)^{\frac{2}{\overline{\rho}-1}}} = 0$$

Thus the formula for limiting cost amplification is valid.

Finally consider the case of constant limiting mark-up. Under cost absorption

$$V'(p) = -\lambda e^{-\frac{\sqrt{\overline{p}-p}}{\alpha}} = -Q(p)$$

and

$$\lim_{p\nearrow\overline{p}} V(p) = 2\lambda\alpha(\alpha - \alpha) = 0$$

Under cost amplification, we first verify the envelope condition in the Apt range:

$$V'(p) = 2\alpha Q'(p)\left(\alpha - \sqrt{p - \underline{p}}\right) - \frac{\alpha Q(p)}{\sqrt{p - \underline{p}}} = Q(p)\left(\left[\frac{\alpha}{\sqrt{p - \underline{p}}} - 1 - \frac{\alpha}{\sqrt{p - \underline{p}}}\right]\right) = -Q(p)$$

In the intermediate range the envelope condition is

$$V'(p) = -\frac{2\lambda\alpha^2}{e(p - \underline{p} + \alpha^2)} = -Q(p)$$

To see that continuity holds note that

$$\lim_{p\nearrow\underline{p}+\alpha^2} V(p) = \frac{4\lambda\alpha^2}{e} = \frac{2\lambda\alpha(\alpha + \alpha)}{e} = \lim_{p\searrow\underline{p}+\alpha^2} V(p)$$

Further the limiting condition also holds=

$$\lim_{p\to\infty} V(p) = \lim_{p\to\infty} \frac{2\alpha}{\lambda} \frac{\alpha - \sqrt{p - \underline{p}}}{e^{-\frac{\sqrt{p-\underline{p}}}{\alpha}}} =$$

by L'Hôpital

$$\lim_{p\to\infty} \frac{2\alpha^2}{\lambda} \frac{1}{e^{-\frac{\sqrt{p-\underline{p}}}{\alpha}}} = 0$$

Finally in the case of constant mark-up demand, this formula is well-known from Bulow and Pfleiderer and extremely easy to derive. We therefore omit the proof in this case.

Now we turn to the formulae for consumer surplus under Apt inverse demand. The natural way to calculate surplus from inverse demand is to integrate price with respect to quantity, the opposite of the technique used earlier. In particular $V(q) = \int_0^q P(r)dr - P(q)q$. Under limiting quantity absorption the expression for consumer surplus can be greatly simplified by using the function

$$\tilde{V}(q) \equiv \frac{(P[q]-c)\left([1-\overline{\rho}_Q^2][\overline{q}-q] - 4\alpha_Q\overline{\rho}_Q^2[\alpha_Q + \sqrt{\overline{q}-q}]\right)}{1+\overline{\rho}_Q}$$

this is given by

$$V(q) = \begin{cases} \tilde{V}(0) - \tilde{V}(q) - (P[q]-c)\,q & \textit{if} \quad 0 < q \le \overline{q} - \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} \\ \tilde{V}(0) - \frac{M\left(-2\overline{\rho}_Q\alpha_Q\right)^{\frac{2}{1-\overline{\rho}_Q}}1_{\alpha_Q<0}}{1+\overline{\rho}_Q} & \textit{if} \quad \overline{q} - \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} < q \end{cases} \tag{61}$$

Under limiting quantity amplification the expression can again be simplified by using a short-hand:

$$\hat{V}(q) \equiv \frac{(P[q]-c)\left([\overline{\rho}_Q^2-1][q-\underline{q}] + 4\alpha_Q\overline{\rho}_Q^2[\alpha_Q - \sqrt{q-\underline{q}}]\right)}{1+\overline{\rho}_Q}$$

$$\ddot{V}(q) = (P[q]-c)\,\kappa(q)\left(\frac{\overline{\rho}_Q(\overline{\rho}_Q+5)}{\overline{\rho}_Q+1} + \log\left[\frac{\alpha_Q^2\overline{\rho}_Q(\overline{\rho}_Q+1)}{\alpha_Q^2\overline{\rho}_Q + q - \underline{q}}\right]\right)$$

I assume that $\underline{q} - \overline{\rho}_Q\alpha_Q^2 1_{\alpha_Q<0} + \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0} \le 0$, as otherwise the monopolist could earn infinite profits by producing $\underline{q} - \overline{\rho}_Q\alpha_Q^2 1_{\alpha_Q>0} + \frac{4\alpha_Q^2\overline{\rho}_Q^2}{(1-\overline{\rho}_Q)^2}1_{\alpha_Q>0}$. Surplus has two different forms. If $0 \ge \underline{q} + \overline{\rho}_Q^2\alpha_Q^2 1_{\alpha_Q<0}$ then (restricting the domain to $q \ge 0$)

$$V(q) = \hat{V}(0) - \hat{V}(q) - (P[q]-c)\,q \tag{62}$$

On the other hand if $0 < \underline{q} - \overline{\rho}_Q^2(|\alpha_Q|\alpha_Q)^-$ then

$$V(q) = \begin{cases} \ddot{V}(0) - \ddot{V}(q) - (P[q]-c)\,q & \textit{if} \quad 0 < q \le \underline{q} + \overline{\rho}_Q^2\alpha_Q^2 1_{\alpha_Q<0} \\ \ddot{V}(0) - \hat{V}(q) - (P[q]-c)\,q & \textit{if} \quad \underline{q} + \overline{\rho}_Q^2\alpha_Q^2 1_{\alpha_Q<0} < q \end{cases} \tag{63}$$

Under constant quantity

$$V(q) = M\kappa\left(e^{-\frac{q}{\kappa}} - 1\right) - (P[q]-c)q \tag{64}$$

With cost-absorbing constant quantity-in-the-limit

$$V(q) = \begin{cases} 2M\alpha_Q\left(e^{-\frac{\sqrt{\overline{q}-q}}{\alpha_Q}}[\alpha_Q + \sqrt{\overline{q}-q}] - e^{-\frac{\sqrt{\overline{q}}}{\alpha_Q}}[\alpha_Q + \sqrt{\overline{q}}]\right) - (P[q]-c)q & \textit{if} \quad 0 < q \le \overline{q} \\ 2M\alpha_Q\left(\alpha_Q - e^{-\frac{\sqrt{\overline{q}}}{\alpha_Q}}[\alpha_Q + \sqrt{\overline{q}}]\right) & \textit{if} \quad \overline{q} < q \end{cases} \tag{65}$$

Finally with cost-amplifying constant quantity-in-the-limit, again assuming that $\underline{q}-\alpha_Q^2 \leq 0$, there are two cases. If $\underline{q}+\alpha_Q^2 \leq 0$ then

$$V(q) = 2M\alpha_Q \left(e^{\frac{\sqrt{-\underline{q}}}{\alpha_Q}} \left[\alpha_Q - \sqrt{-\underline{q}}\right] - e^{\frac{\sqrt{q-\underline{q}}}{\alpha_Q}} \left[\alpha_Q - \sqrt{q-\underline{q}}\right] \right) - (P[q]-c)q \tag{66}$$

On the other hand if $\underline{q}+\alpha_Q^2 > 0$

$$V(q) = \begin{cases} \frac{2M\alpha_Q}{e} \log\left(\frac{\alpha_Q^2+q-\underline{q}}{\alpha_Q^2-\underline{q}}\right) - (P[q]-c)\,q & \textit{if} \quad 0 < q \leq \underline{q}-\alpha_Q^2 \\ 2M\alpha_Q \left(\frac{1}{e}\log\left[\frac{2(e\alpha_Q)^2}{\alpha_Q^2-\underline{q}}\right] - e^{\frac{\sqrt{q-\underline{q}}}{\alpha_Q}} \left[\alpha_Q - \sqrt{q-\underline{q}}\right]\right) - (P[q]-c)\,q & \textit{if} \quad \underline{q}-\alpha_Q^2 < q \end{cases} \tag{67}$$

For brevity we do not formally establish these formulae here, but they can easily be checked by hand and the requisite calculations, which are quite similar to those in the case of Apt demand, are available on request.

III. Generalized Cournot-Stackelberg model

Beyond its tractability in monopoly pricing, Apt (inverse) demand yields closed form solutions for the entire Generalized Cournot Stackelberg (GCS) class of models described in Subsections II.B and II.C. While we focus on this particular class discussed in the text partly because of its own economic interest, our primary goal is not to analyze these situations per se. Rather we hope to demonstrate how Apt demand gives rise to simple closed form solutions to industrial organization models in a range of situations, allowing relatively mechanical analysis. We suspect, and indeed hope, that its most important applications will be in models we are not even aware of.

This section draws heavily on Subsection II.B of the text and appendix *Generalized Cournot-Stackelberg models*. Reading the first and the first page and a half of the second is a prerequisite for following much of this section.

A. Apt solution

The natural method for solving this problem is backward induction. Firms in group k take the mark-ups chosen by all firms in groups weakly above them as given and, conjecturing a particular equilibrium, take the mark-up of all firms strictly below them as a function of the mark-up they choose. Let the conjectured total mark-up of all firms in lower groups be

$\overline{M}_{k-1}\left(c + \sum_{k'=k}^{K} \sum_{i=1}^{N_{k'}} m_{i_{k'}}\right)$. Then the anticipated profits each firm in round k is

$$\pi_{i_k}(m_{i_k}) = m_{i_k} Q\left(c + \sum_{k'=k}^{K} \sum_{i'=1}^{N_{k'}} m_{i'_{k'}} + \overline{M}_{k-1}\left[c + \sum_{k'=k}^{K} \sum_{i'=1}^{N_{k'}} m_{i'_{k'}}\right]\right) \tag{68}$$

If we let $c_k \equiv c + \sum_{k'=k+1}^{K} \sum_{i'=1}^{N_{k'}} m_{i'_{k'}}$ and $Q_k(p) \equiv Q\left(p + \overline{M}_{k-1}[p]\right)$ then we can rewrite equation (68) as

$$\pi_{i_k}(m_{i_k}) = m_{i_k} Q_k\left(c_k + \sum_{i'=1}^{N_k} m_{i'_k}\right)$$

which is simply the profit function for (any) one firm in the simultaneous vertical monopolies problem with cost c_k and demand Q_k. It is well-known that the conditions for a symmetric interior equilibrium in this game are

$$\overline{m}_k^\star = N_k \mu_k(c_k + \overline{m}_k^\star) \tag{69}$$

$$m_{i_k}^\star = \frac{\overline{m}_k^\star}{N_k} \quad \forall i = 1, \ldots, N_k \tag{70}$$

where $\mu_k(p) \equiv -\frac{Q_k(p)}{Q_k'(p)}$. In fact these conditions identify the unique equilibrium of the game (which is stable) for all values of c_k if and only if $\mu_k'(p) < \frac{1}{N_k}$ wherever $Q_k(p)$ is strictly positive (Seade, 1980). Furthermore any equilibrium not given by equations (69) and (70) at a point where $\mu_k'(\overline{m}_k) < \frac{1}{N_k}$ is unstable or on a boundary. In what follows we will not impose restrictions on the domain of Apt demand to guarantee the existence and uniqueness of such an equilibrium, as we believe the appropriate conditions may vary across applications. Instead we assume that such stable, interior equilibrium (optimum) at each level where all firms anticipate firms below them playing a stable, interior equilibrium is played. As we will see, a unique such *stable-anticipating-stable* equilibrium exists under Apt demand.

Note that the solution to equation (69) is simply a monopoly pricing problem with a scaled-up market power. Clearly scaling up market power does not remove demand from the Apt class. This means that so long as Q_k is of the Apt form, the solution for $\overline{m}_k$ takes the form of an Apt monopoly pricing solution. Furthermore, as we show in the proof of the following proposition, whenever the value of this solution, conditional on the total price from all upstream firms, $\overline{m}_{k-1}(p)$ takes the form of an Apt monopoly pricing solution, $Q_k(p)$ is an Apt demand. Thus at a stable-anticipating-stable equilibrium, the attractive tractability of Apt demand is preserved in this far more complicated problem. Anderson and Engers (1992) observed that one very special case of the Bulow-Pfleiderer demand class has this recursive property[6]. The following proposition show this generalizes not only to the entire Bulow-Pfleiderer class, but also to the much broader Apt demand.

Proposition 6. *Assume that $Q(p)$ is Apt and the firms play a stable-anticipating-stable equilibrium. Then each Q_k is of the Apt form. Explicitly if Q is Apt with parameter values $\lambda, \overline{\rho} \neq 1, \tilde{p}$, and α then Q_k has the Apt form*

$$\overline{\rho}_k = \overline{\rho} \tag{71}$$

$$\lambda_k = \lambda \left(\prod_{k'=1}^{k-1} \frac{\overline{\rho}}{N_{k'}(1-\overline{\rho}) + \overline{\rho}} \right)^{\frac{\overline{\rho}}{1-\overline{\rho}}} \tag{72}$$

$$\alpha_k = \alpha \prod_{k'=1}^{k-1} \frac{2\overline{\rho} + N_{k'}\,(1-\overline{\rho})}{2\sqrt{\overline{\rho}\,(\overline{\rho} + N_{k'}[1-\overline{\rho}])}} \tag{73}$$

$$\tilde{p}_k = \tilde{p} + sign(1-\overline{\rho}) \sum_{k'=1}^{k-1} \frac{N_{k'}^2 \alpha_{k'}^2 \overline{\rho}}{N_{k'}(1-\overline{\rho}) + \overline{\rho}} \tag{74}$$

When $\overline{\rho} = 1$, $\lambda_k = \lambda e^{-\sum_{k'=1}^{k-1} N_{k'}}$ and $\tilde{p}_k = \tilde{p} + sign(1-\rho)\alpha^2 \sum_{k'=1}^{k-1} N_{k'}^2$ and $sign(1-\rho)$ is evaluated at any p in the Apt range (is positive if demand is cost-absorbing, negative if it is cost-amplifying) .

Note that $\overline{\rho}_k$ is independent of k and that pass-through at level k, the ρ_k of Subsection II.B, is always on the same side of 1 as $\overline{\rho}_k$ by fact that Apt demand obeys SPAs. Furthermore the sign of α is preserved in α_k. Thus, as discussed in Subsection V.B (point 8), both the SPA-obedience and the flexibility of Apt demand are preserved at all levels of the GCS model, making Apt demand extremely flexible in analyzing such models, at least with respect to results like those established in Subsection II.A.

Also it is worth noting three special cases that further simplify the expressions for the parameters of Q_k. First, if demand is in the Bulow-Pfleiderer class (i.e. $\alpha = 0$) then it is easy to see that only the level shifter differs across Q_k and therefore each group of firms price as if they were simultaneous vertical monopolists facing the final demand, taking as part of their costs the mark-ups of all upstream firms. Second, if $\overline{\rho} = 1$ then α_k is constant and only the notional choke or explosive price shifts the firm away from the Bulow-Pfleiderer case. Finally in the case when $N_k = 1$ for all k, as considered by Anderson and Engers (1992), $\alpha_k = \alpha \left(\frac{\overline{\rho}+1}{2\sqrt{\overline{\rho}}} \right)^{k-1}$, $\lambda_k = \lambda \overline{\rho}^{\frac{(k-1)\overline{\rho}}{1-\overline{\rho}}}$ and $\tilde{p}_k = \tilde{p} + \alpha^2 \frac{1 - \left(\frac{[\overline{\rho}+1]^2}{4\overline{\rho}} \right)^{k-1}}{\left| 1 - \frac{[\overline{\rho}+1]^2}{4\overline{\rho}} \right|}$.

[6] In fact, they show it has a much stronger property, which we show below generalizes to the entire Bulow-Pfleiderer class and an arbitrary number of firms at each stage: precisely the same (inverse) demand function applies at each step.

Proof. We prove by induction. The base case is assumed, so we simply need to show that if Q_{k-1} is of the Apt form with parameters $(\overline{\rho}_{k-1}, \lambda_{k-1}, \tilde{p}_{k-1}, \alpha_{k-1})$ then Q_k takes the Apt form with parameters given in equations (71-73). By precisely the same reasoning as in Subsection II.C, at a stable equilibrium at level k it must be that

$$\sqrt{\text{sign}(\overline{\rho}-1)\left(\overline{m}_k + c_k - \tilde{p}_k\right)} =$$

$$\overline{\rho}\frac{\sqrt{N_k^2\alpha_k^2 + \text{sign}(\overline{\rho}-1)(c_k - \tilde{p}_k)\frac{N_k(1-\overline{\rho})+\overline{\rho}}{\overline{\rho}}} + \text{sign}(1-\overline{\rho})N_k\alpha_k}{N_k(1-\overline{\rho})+\overline{\rho}}$$

Note that $\overline{M}_k(p) = \sum_{k'=1}^{k} \overline{m}_{k'}(p)$ so that $Q_{k+1}(p) = Q_k\left(p + \overline{m}_k[p]\right)$. Therefore, so long as we are in the Apt range which we are by the fact that the stable solution is always in the Apt range, then for $\overline{\rho} \neq 1$

$$Q_{k+1}(p) = \lambda_k \left(|\overline{\rho}-1|\overline{\rho}\frac{\sqrt{N_k^2\alpha_k^2 + \text{sign}(\overline{\rho}-1)(p - \tilde{p}_k)\frac{N_k(1-\overline{\rho})+\overline{\rho}}{\overline{\rho}}} + \text{sign}(1-\overline{\rho})N_k\alpha_k}{N_k(1-\overline{\rho})+\overline{\rho}} - 2\overline{\rho}\alpha_k \right)^{\frac{2\overline{\rho}}{1-\overline{\rho}}} =$$

$$\lambda_k \left(\frac{\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}\right)^{\frac{\overline{\rho}}{1-\overline{\rho}}} \left(|\overline{\rho}-1|\sqrt{\left|p - \tilde{p}_k - \text{sign}(1-\overline{\rho})\frac{N_k^2\alpha_k^2\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}\right|} - 2\overline{\rho}\alpha_k\frac{2\overline{\rho} + N_k\left(1-\overline{\rho}\right)}{2\sqrt{\overline{\rho}\left(\overline{\rho}+N_k[1-\overline{\rho}]\right)}} \right)^{\frac{2\overline{\rho}}{1-\overline{\rho}}} =$$

$$\lambda_{k+1}\left(|\overline{\rho}-1|\sqrt{p - \tilde{p}_{k+1}} - 2\overline{\rho}\overline{\alpha}_{k+1}\right)^{\frac{2\overline{\rho}}{1-\overline{\rho}}}$$

Thus

$$\lambda_{k+1} = \lambda_k\left(\frac{\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}\right)^{\frac{\overline{\rho}}{1-\overline{\rho}}}$$

$$\alpha_{k+1} = \alpha_k\frac{2\overline{\rho} + N_k\left(1-\overline{\rho}\right)}{2\sqrt{\overline{\rho}\left(\overline{\rho}+N_k[1-\overline{\rho}]\right)}}$$

$$\tilde{p}_{k+1} = \tilde{p}_k + \text{sign}(1-\overline{\rho})\frac{N_k^2\alpha_k^2\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}$$

The solution to this recursion is given in equations (71-74).

On the other hand if $\overline{\rho} = 1$, maintaining the convention from the proposition, then

$$Q_{k+1}(p) = \lambda_k e^{\text{sign}(1-\rho)\frac{\sqrt{\text{sign}(\rho-1)(p-\tilde{p}_k)+\alpha^2 N_k} + \text{sign}(\rho-1)\alpha N_k}{\alpha}} = \lambda_k e^{-N_k} e^{\text{sign}(1-\rho)\frac{\sqrt{|p-\tilde{p}_{k+1}|}}{\alpha}}$$

Thus α_k is constant and

$$\lambda_{k+1} = \lambda_k e^{-N_k}$$

$$\tilde{p}_{k+1} = \tilde{p}_k + \text{sign}(1-\rho)\alpha^2 N_k^2$$

Again this recursion clearly has the posited solution.

Thus the simple structure of Apt demand is maintained even as it is filtered through many layers of many double- (multi-) marginalizing firms. Unsurprisingly this gives rise to simple, explicit solutions for the mark-ups of all firms.

Proposition 7. *When $\overline{\rho} \neq 1$ the final price to consumers at a stable-anticipating-stable equilibrium is*

$$p^\star = \tilde{p} + sign(\overline{\rho}-1)\left(\sqrt{|c-\tilde{p}_{K+1}|}\left[\prod_{k=1}^{K}\sqrt{\frac{\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}}\right] + \frac{2\overline{\rho}\alpha}{|1-\overline{\rho}|}\left[1-\frac{1}{2^K}\left(\prod_{k=1}^{K}1+\frac{\overline{\rho}}{\overline{\rho}+N_k(1-\overline{\rho})}\right)\right]\right)^2 \tag{75}$$

and more generally the total price of all firms weakly above group k, $c_{k-1}^\star \equiv c + \sum_{k'=k}^{K}\overline{m}_{k'}^\star =$ is

$$c_{k-1}^\star = \tilde{p}_k + sign(\overline{\rho}-1)\left(\sqrt{|c-\tilde{p}_{K+1}|}\left[\prod_{k'=k}^{K}\sqrt{\frac{\overline{\rho}}{N_{k'}(1-\overline{\rho})+\overline{\rho}}}\right] + \frac{2\overline{\rho}\alpha_k}{|1-\overline{\rho}|}\left[1-\frac{1}{2^{K-k+1}}\left(\prod_{k'=k}^{K}1+\frac{\overline{\rho}}{\overline{\rho}+N_{k'}(1-\overline{\rho})}\right)\right]\right)^2 \tag{76}$$

When $\overline{\rho} = 1$, maintaining the convention from Proposition 6,

$$c_{k-1}^\star = \tilde{p} + sign(\rho-1)\left(\left[\sqrt{|c-\tilde{p}| + \alpha^2\sum_{k'=1}^{K}N_{k'}^2} - \alpha\sum_{k'=k}^{K}N_{k'}\right]^2 - \alpha^2\sum_{k'=1}^{k-1}N_{k'}^2\right) \tag{77}$$

Note that these are explicit expressions for all mark-ups as for all $k = 1,\dots,K$ as

$$\overline{m}_k^\star = c_{k-1}^\star - c_k^\star$$

and

$$m_{i_k}^\star = \frac{\overline{m}_k^\star}{N_k} \quad \forall i = 1,\dots,N_k$$

where $c_K^\star \equiv c$ and $c_0^\star \equiv p^\star$.

In the special cases mentioned above, the solution obviously takes an even simpler form. For example, in the case of $\alpha = 0$ we have

$$m_{i_k}^\star = \frac{(c-\tilde{p})(\overline{\rho}-1)}{N_k(1-\overline{\rho})+\overline{\rho}}\prod_{k'=k+1}^{K}\frac{\overline{\rho}}{N_{k'}(1-\overline{\rho})+\overline{\rho}} \tag{78}$$

This immediately dramatically generalizes Anderson and Engers's results: they argue that the earlier one is in a chain, the better off one is in one very special case of the Bulow-

Pfleiderer class. Equation (78) shows that this result is driven by their assumption that demand is cost-absorbing: for any size of firms in a group, it is better to be earlier in the chain if and only if $\overline{\rho} < 1$ and better to be later in the chain if $\overline{\rho} > 1$. This follows from the fact that $\frac{\overline{\rho}}{N_k(1-\overline{\rho})+\overline{\rho}}$ is on the same side of 1 as $\overline{\rho}$ for any N_k.

Proof. By definition $Q_{K+1}(c) = Q_K(\overline{m}_K^\star + c) = Q_k(p_k^\star) = Q(p^\star)$ for all $k = 1, \dots, K$. A bit of algebra, which we omit here as it is very similar to that in the proof of Proposition 6, shows that $c_{k-1}^\star = Q_k^{-1}\left(Q_{K+1}[c]\right)$ is given by equation (76).

Note that, immediately by duality, all of the results here hold for the case of quantity competition if Apt inverse demand is substituted for Apt demand, quantities and mark-ups are switched through out and exogenous quantity $\tilde{q}$ is substituted[7] for c. Thus Apt inverse demand gives simple, closed-form solutions to the entire GCS class.

B. Comparing industrial organizations

Now, to show how these formulae can be used to establish substantive results, we confirm two general theoretical results using the formulae from Apt demand. These results are not intended to be novel, though as far as we know one is, nor of particular substantive interest. Nor are the derivations of them meant to be insightful or general. Quite the contrary: what we hope to demonstrate is that that the tractability and flexibility of Apt demand allow insight into general properties of various industrial organization models in a straightforward mechanical way, without requiring sophisticated, general analysis of aspired to by the main text.

In particular we study two- and three-firm vertical monopoly games. In the first game, we confirm one of the general result of Subsection II.A. We show that the comparison between the Spengler-Stackelberg (SS) sequential equilibrium final price and that arising at Cournot-Nash (CN) equilibrium is determined by whether demand is cost-absorbing or cost-amplifying. In the second game we use an alternative characterization of prices to final consumers in the general to show that the comparison of these prices between the case of $K = 2$, $N_1 = 1$ and $N_2 = 2$ and that when $K = 3$ and $N_k = 1 \forall k = 1, 2, 3$ is determined by a combination of the sign of the slope of pass-through and whether demand is cost-absorbing or cost-amplifying; we then confirm this result in the case of Apt demand.

We begin with the two-player game. By Proposition 7 the total price to consumers under

[7]Note that by duality an exogenous quantity $\tilde{q}$ could also be added to our model of vertical monopolies, so long as it is homogenous across firms.

SS ($K = 2, N_1, N_2 = 1$) is

$$\tilde{p} + \text{sign}(\overline{\rho} - 1)\left(\overline{\rho}\sqrt{|c - \tilde{p}| + \alpha^2\left(\overline{\rho} + \frac{[\overline{\rho}+1]^2}{4}\right)} + \frac{2\overline{\rho}\alpha}{|1-\overline{\rho}|}\left[1 - \frac{(\overline{\rho}+1)^2}{4}\right]\right)^2$$

On the other hand at CN equilibrium ($K = 1, N_1 = 2$)

$$\tilde{p} + \text{sign}(\overline{\rho} - 1)\left(\sqrt{\frac{\overline{\rho}}{2-\overline{\rho}}}\sqrt{|c - \tilde{p}| + \alpha^2\frac{4\overline{\rho}}{2-\overline{\rho}}} + \frac{2\overline{\rho}\alpha}{|1-\overline{\rho}|}\left[1 - \frac{1}{2}\left(1 + \frac{\overline{\rho}}{2-\overline{\rho}}\right)\right]\right)^2$$

Simplifying the difference between these expressions yields

$$\sqrt{\frac{1}{2-\overline{\rho}}}\sqrt{|c - \tilde{p}| + \alpha^2\frac{4\overline{\rho}}{2-\overline{\rho}}} - \sqrt{\overline{\rho}}\sqrt{|c - \tilde{p}| + \alpha^2\left(\overline{\rho} + \frac{[\overline{\rho}+1]^2}{4}\right)} - \frac{\sqrt{\overline{\rho}}\alpha|1-\overline{\rho}|(\overline{\rho}+2)}{2(2-\overline{\rho})} > 0 \quad (79)$$

We want to show that this holds for all $\alpha, \overline{\rho}, |c - \tilde{p}|$ obeying $|c - \tilde{p}| > \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2}1_{\alpha>0}$ and $\overline{\rho} \in (0, 2)\backslash\{1\}$. Note that we can normalize $|c - \tilde{p}| = 1$, because our goal is to show these properties hold globally. Furthermore note that

$$\sqrt{\frac{1}{2-\overline{\rho}}}\sqrt{1 + \alpha^2\frac{4\overline{\rho}}{2-\overline{\rho}}} > \sqrt{\overline{\rho}}\sqrt{1 + \alpha^2\left(\overline{\rho} + \frac{[\overline{\rho}+1]^2}{4}\right)}$$

as

$$\overline{\rho}(2 - \overline{\rho}) < 1$$

for $\rho \in (0, 2)\backslash\{1\}$ and

$$\frac{4\overline{\rho}}{(2-\overline{\rho})^2} > \overline{\rho}^2 + \frac{\overline{\rho}(\overline{\rho}+1)^2}{4}$$

which can be seen easily by graphing the relevant functions on $\overline{\rho} \in (0, 2)$. Thus for $\alpha \leq 0$ the result follows immediately.

On the other hand for $\alpha > 0$, it must be that $\alpha < \frac{|1-\overline{\rho}|}{2\overline{\rho}}$. Thus we can rewrite condition (79) as

$$\sqrt{\frac{1}{2-\overline{\rho}}}\sqrt{1 + a^2\frac{(1-\overline{\rho})^2}{\overline{\rho}(2-\overline{\rho})}} - \sqrt{\overline{\rho}}\sqrt{1 + a^2\frac{(1-\overline{\rho})^2}{4\overline{\rho}^2}\left(\overline{\rho} + \frac{[\overline{\rho}+1]^2}{4}\right)} - \frac{a(1-\overline{\rho})^2(\overline{\rho}+2)}{4\sqrt{\overline{\rho}}(2-\overline{\rho})} > 0 \quad (80)$$

where $a \equiv \frac{2\alpha\overline{\rho}}{|1-\overline{\rho}|}$. We need to show this holds for any $\overline{\rho} \in (0, 2)\backslash\{1\}$ and for any $a \in (0, 1)$. But this is a compact set, so graphical comparison using mathematical software suffices, especially for such simple smooth functions. The right hand side of the inequality is shown in Figure 6 on the range of $(\alpha, \rho) \in (0, 2) \times (0, 1)$; clearly it is always strictly positive. This figure is generated using the toolkit[8] attached to this paper, which thus allows the graphical

[8]The exact graph does not exactly match what the toolkit produces as it is the difference between the

comparison of arbitrary equilibrium outcomes within and across industrial organizations in the GCS model without any work by the analyst.

So that when $\overline{\rho} > 1$, CN equilibrium leads to a higher price, when $\overline{\rho} < 1$ SS equilibrium leads to a higher price. The same can easily be seen to extend to the division between cost-amplifying and cost-absorbing constant limiting mark-up. Thus we confirm the general theoretical prediction that the CN price is above (below) the SS price if demand is cost-amplifying (absorbing). Thus Apt demand takes a general problem that would require a fortuitous insight to solve and reduces them to comparing algebraic expressions, a mechanical task made trivial by the toolkit attached to this paper. This dramatically reduces the burden on an analyst, while maintaining a large degree of flexibility, allowing fairly general results to be easily derived.

To further highlight this reduced burden, we now turn to the three-firm vertical monopolies problem. We establish a comparison between final prices under two industrial organization first using a general theoretical approach and then show how this same result can be shown entirely mechanically using he toolkit, even in this more complex three-player game, using Apt demand.

For the general theory, we use Lemma 2 established in appendix *Generalized Cournot-Stackelberg Models*. This "lucky" result provides a deceptively simple ranking of $p^\star$ under $K = 3, N_k = 1$ (which we refer to as 3SS equilibrium for three-player Spengler-Stackelberg) and $K = 2, N_1 = 2, N_2 = 1$ (which we refer to as SCN equilibrium for Stackelberg-Cournot-Nash). In particular, directly from Lemma 2 we have, that under 3SS equilibrium

$$p^\star_{3SS} = \mu(p^\star_{3SS}) \left(1 + \frac{1}{\rho(p^\star_{3SS})} + \frac{1 + \mu(p^\star_{3SS})\rho'(p^\star_{3SS})}{\rho(p^\star_{3SS})^2}\right)$$

and

$$p^\star_{SCN} = \mu(p^\star_{SCN}) \left(1 + \frac{2}{\rho(p^\star_{SCN})}\right)$$

Therefore if globally

$$\frac{1 + \mu\rho'}{\rho^2} > (<) \frac{1}{\rho}$$

then $p^\star_{3SS} > (<) p^\star_{SCN}$. If demand has increasing pass-through and is cost-absorbing, $p^\star_{3SS} > p^\star_{SCN}$; if demand has decreasing pass-through and is cost-amplifying $p^\star_{3SS} < p^\star_{SCN}$. We focus on these two results for simplicity. Note that while these results easily follow, by simple algebra, from Lemma 2, they required a conceptual leap such as that proposition, and a

square root of raw prices rather than the difference between the raw prices themselves. However the two outputs are essentially indistinguishable and prove the same point; we included the square root here because this is simpler to present in the text.

strategy of ranking effect of market powers, to make them possible. On the other hand, as above, these results can be shown mechanically using the formulae for prices under the two industrial organizations. $p^\star_{3SS}$ is given by

$$\tilde{p} + \mathrm{sign}(\overline{\rho} - 1)\left(\overline{\rho}^{\frac{3}{2}}\sqrt{|c - \tilde{p}| + \alpha^2 \frac{1 - \left(\frac{(\overline{\rho}+1)^2}{4\overline{\rho}}\right)^3}{1 - \frac{(\overline{\rho}+1)^2}{4\overline{\rho}}}} + \frac{\overline{\rho}\alpha\left(8 - [1+\overline{\rho}]^3\right)}{4|1-\overline{\rho}|}\right)^2$$

while $p^\star_{SCN}$ is

$$\tilde{p} + \mathrm{sign}(\overline{\rho} - 1)\left(\frac{\overline{\rho}}{\sqrt{2-\overline{\rho}}}\sqrt{|c - \tilde{p}| - \alpha^2\frac{\overline{\rho}^3 - 3\overline{\rho}^2 - 2\overline{\rho} - 1}{\overline{\rho}(2-\overline{\rho})}} + \mathrm{sign}(1-\overline{\rho})\frac{3\overline{\rho}\alpha}{2-\overline{\rho}}\right)^2$$

Rather than compare these analytically, we highlight the tractability of by supplying, in Figures 3 and 4, graphs generated by mathematical software demonstrating the results. We show the difference between the term inside the square in the expression for $p^\star_{SCN}$ and that for $p^\star_{3SS}$; this should be positive whenever $(1-\overline{\rho})\alpha > 0$. In the graphs again we normalize $|c - \tilde{p}|$ to 1. Figure 3 shows the case of $\overline{\rho} > 1$ and $\alpha < 0$ over the domain of $\alpha \in (-1000, 0)$ and $\overline{\rho} \in (1, 2)$. For the case, shown in Figure 4, of $\alpha > 0$ and $1 > \overline{\rho}$ we let $a \equiv \frac{2\alpha\overline{\rho}}{|1-\overline{\rho}|}$ and graph over the range $(a, \overline{\rho}) \in (0,1)^2$; this covers all the relevant range as we have assumed (see Subsection II.B) that $\overline{p} - \frac{4\alpha^2\overline{\rho}^2}{(1-\overline{\rho})^2} > c$ and $\overline{p} - c = 1$. Clearly over both of these ranges the difference is strictly positive. Conversely, when $(1-\overline{\rho})\alpha < 0$, mathematical can immediately identifies cases when the expression is either strictly positive or strictly negative. For example when $\overline{\rho} = 1.002$, $c = \underline{p} + 1$ and $\alpha = .00002$ the expression is positive, when when $\overline{\rho} = 1.11$, $c = \underline{p} + 1$ and $\alpha = .001$ it is negative.

Thus a general comparison of these industrial organizations can be uncovered using Apt demand with essentially no analytic work using the toolkit accompanying this paper and simultaneously the ambiguity, in the case where ambiguity arises, is also immediately evident.

References

Anderson, Simon P. and Maxim Engers, "Stackelberg Versus Cournot Oligopoly Equilibrium," *International Journal of Industrial Organization*, 1992, *10* (1), 127–135.

Bulow, Jeremy I. and Paul Pfleiderer, "A Note on the Effect of Cost Changes on Prices," *Journal of Political Economy*, 1983, *91* (1), 182–185.

Rochet, Jean-Charles and Jean Tirole, “Platform Competition in Two-Sided Markets,” *Journal of the European Economic Association*, 2003, *1* (4), 990–1029.

Seade, Jesus, “The Stability of Cournot Revisited,” *Journal of Economic Theory*, 1980, *23* (1), 15–27.

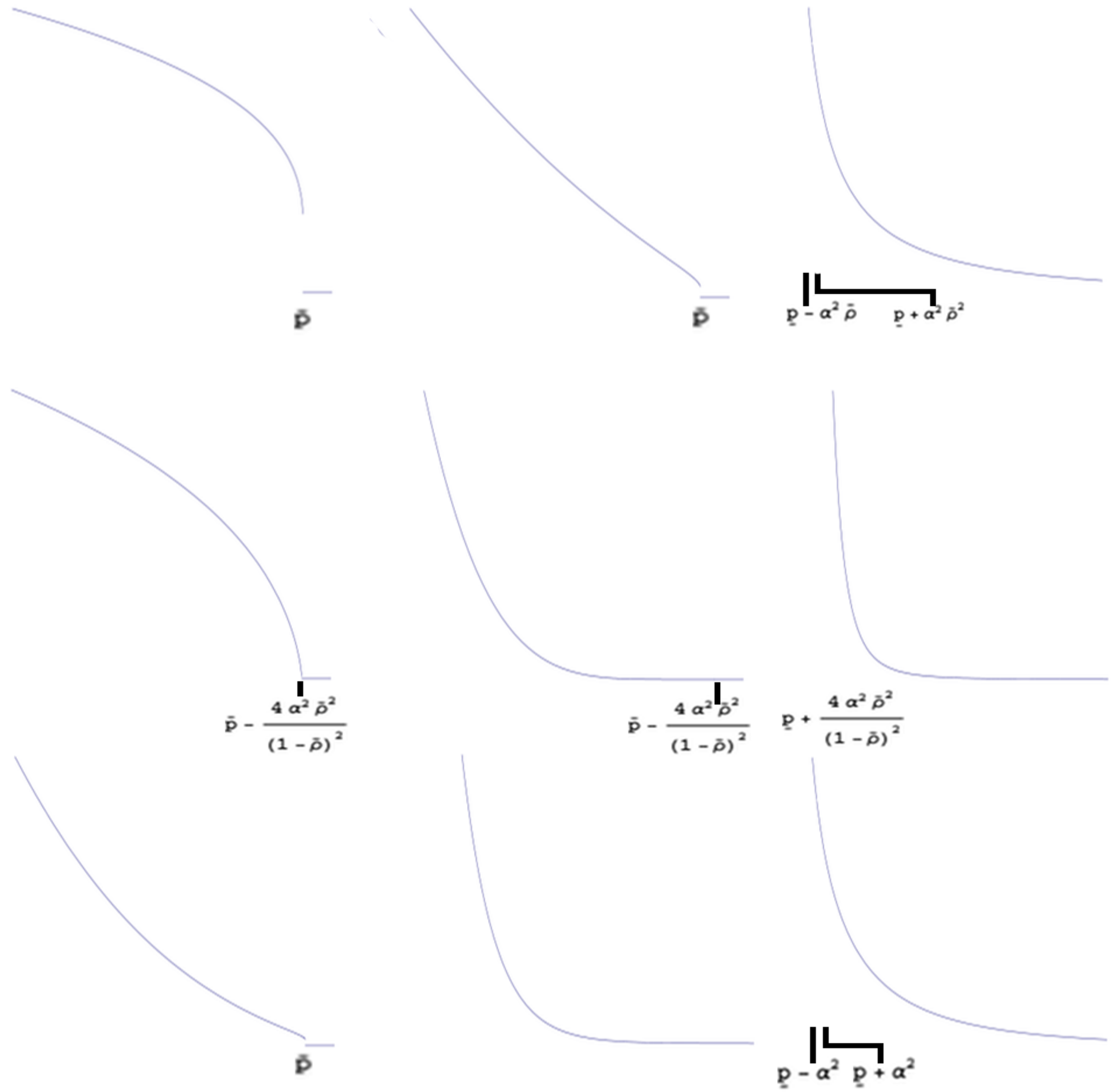

Figure 7: Apt demand $Q(p)$ for various parameter values. The first row has $\alpha = -1$ with $\overline{\rho} = .25, .75$ and 2 in the three columns. Row 2 has $\alpha = 1$ with the same column $\overline{\rho}$ values as row 1. Row 3 has $\overline{\rho} = 1$ with the cost-absorbing, constant mark-up and cost-amplifying forms in the three column; in the first and last case, $\alpha = -1$.

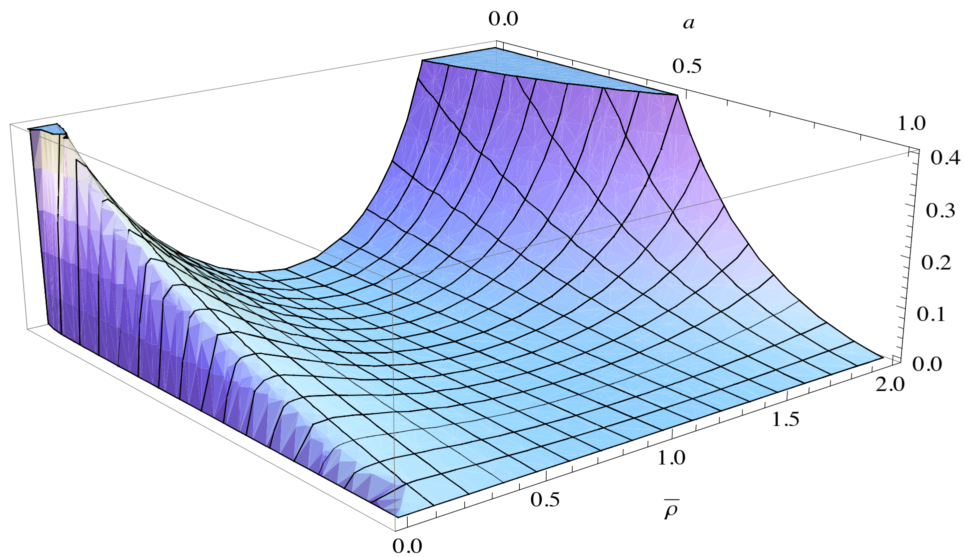

Figure 8: Expression (80) on $(\overline{\rho}, a) \in (0,2) \times (0,1)$; the expression is always strictly positive

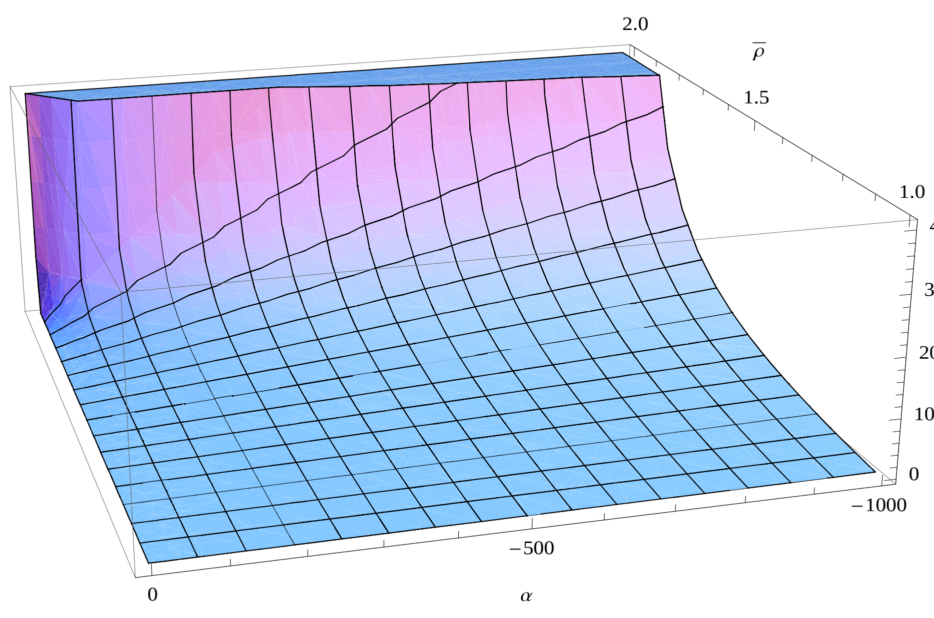

Figure 9: Re-scaled difference (in thousands) between $p^{\star}_{SCN}$ and $p^{\star}_{3SS}$ with $(\overline{\rho}, \alpha) \in (1,2) \times (-1000, 0)$ and c normalized to $\underline{p}+1$; the expression is always strictly positive

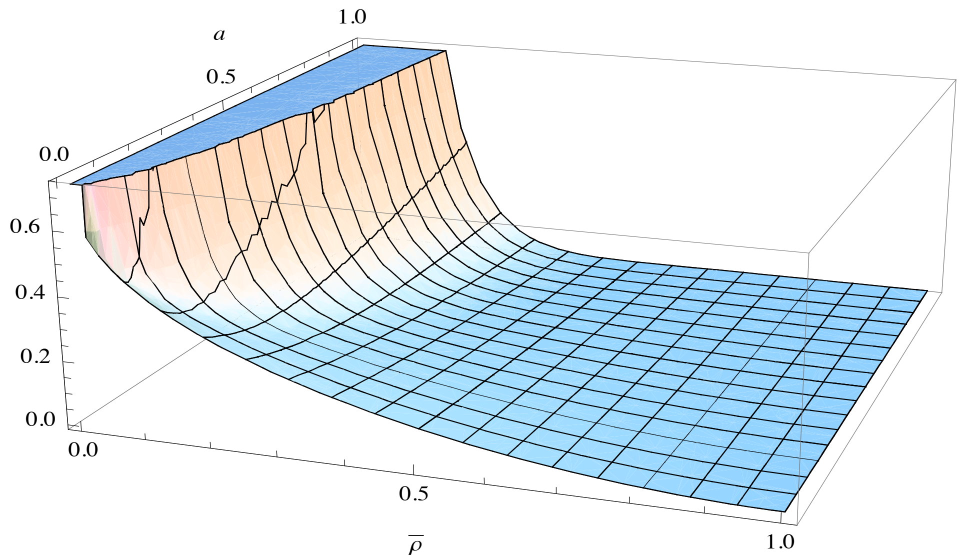

Figure 10: Re-scaled difference between $p^{\star}_{3SS}$ and $p^{\star}_{SCN}$ with $(\overline{\rho}, a) \in (0,1)^2$, $\alpha \equiv \frac{a|1-\overline{\rho}|}{2\overline{\rho}}$, c normalized to $\underline{p}+1$; the expression is always strictly positive