Supplementary Material for "Collusion with Persistent Cost Shocks"

Susan Athey and Kyle Bagwell

July 13, 2006

Abstract

This document has three parts. The first part analyzes generalizations of our model to downward-sloping demand, imperfect substitutes, Cournot competition, and nonlinear cost functions. The second part describes a general dynamic programming approach to games with serially correlated private information. The third part provides a proof of our result about the existence of an equilibrium with productive efficiency for perfectly persistent types, and briefly describes a generalization to more than two firms.

---

*Athey: Harvard University (Department of Economics) and NBER. Bagwell: Columbia University (Kelvin J. Lancaster Professor of Economic Theory in the Department of Economics, and School of Business) and NBER.

1. Model Generalizations and Robustness of Results

This section discusses robustness of our results on rigid pricing to several modifications of the basic model, allowing for alternative specifications of demand, quantity competition, and nonlinear cost functions. In the last subsection, we briefly discuss how our results about first-best equilibria generalize.

1.1. Downward-Sloping Demand, Perfect Substitutes

Modify the model as follows. For simplicity of notation, eliminate announcements and quantity restrictions from the model, so that the only choice in the stage game is the price. (This does not affect our analysis.) Maintaining the assumption that goods are perfect substitutes, define the "profit-if-lowest-price" function as $\pi(p_{i,t}, \theta_{i,t}) \equiv (p_{i,t} - \theta_{i,t})D(p_{i,t})$, where D is a twice-continuously differentiable market demand function that satisfies $D > 0 > D'$ over the relevant range. We assume that π is strictly quasiconcave in ρ, with a unique maximizer, $\rho^m(\theta_{i,t})$, where $\rho^m(\bar{\theta}) \geq \bar{\theta}$. The monopoly price, $\rho^m(\theta_{i,t})$, is strictly increasing in $\theta_{i,t}$.

With these modifications, given period strategies, the expected market share for firm i in period t can be written

$$\bar{m}_i(\hat{\theta}_{i,t}; \mathbf{s}_t, \boldsymbol{\nu}_{-i,t}) = \mathbb{E}_{\boldsymbol{\theta}_{-i,t}} \left[\varphi_i \left(\boldsymbol{\rho}_t(\hat{\theta}_{i,t}, \boldsymbol{\theta}_{-i,t}) \right) \middle| \boldsymbol{\nu}_{-i,t} \right]$$

and the interim profit function can be written

$$\bar{\pi}_i(\hat{\theta}_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\nu}_{-i,t}) = \pi(\rho_{i,t}(\hat{\theta}_{i,t}), \theta_{i,t}) \bar{m}_i(\hat{\theta}_{i,t}; \mathbf{s}_t, \boldsymbol{\nu}_{-i,t}).$$

We next define $Q(p_{i,t}; \boldsymbol{\rho}_{-i,t}, \boldsymbol{\nu}_{-i,t})$ as the quantity a firm expects to sell when it sets price $p_{i,t}$ and opponents use pricing function $\boldsymbol{\rho}_{-i,t}$. In the present context, $Q_i(p_{i,t}; \boldsymbol{\rho}_{-i,t}, \boldsymbol{\nu}_{-i,t}) = D(p_{i,t}) \cdot \mathbb{E}_{\boldsymbol{\theta}_{-i,t}} \left[\varphi_i \left(p_{i,t}, \boldsymbol{\rho}_{-i,t}(\boldsymbol{\theta}_{-i,t}) \right) \middle| \boldsymbol{\nu}_{-i,t} \right]$, and so the function $\bar{\pi}_i$ may be alternatively expressed as

$$\bar{\pi}_i(\hat{\theta}_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\nu}_{-i,t}) = (p_{i,t} - \theta_{i,t}) Q_i(\rho_{i,t}(\hat{\theta}_{i,t}); \boldsymbol{\rho}_{-i,t}, \boldsymbol{\nu}_{-i,t}).$$

We ask three related questions about the optimality of rigid pricing in this model. First, can we bound the profits from an equilibrium with productive efficiency (if it is an equilibrium at all)? Second, is price rigidity optimal when demand is sufficiently inelastic? (That is, is our price rigidity result of Proposition 2 "knife-edge"?) Third, what can we say about optimal equilibria more generally?

Following the proof of Proposition 2, let $\breve{R}_i(\hat{\theta}_{i,1})$ and $\breve{Q}_i(\hat{\theta}_{i,1})$ denote the expected future discounted revenues and demand that firm i anticipates if it mimics type $\hat{\theta}_{i,1}$ throughout the game, and let $\breve{M}_i(\hat{\theta}_{i,1})$ be the associated market share. If firm i's type is $\theta_{i,1}$, then the present discounted value of profits for firm i can be represented as $U_i(\hat{\theta}_{i,1}, \theta_{i,1}) \equiv \breve{R}_i(\hat{\theta}_{i,1}) - \theta_{i,1}\breve{Q}_i(\hat{\theta}_{i,1})$.

Incentive compatibility implies the monotonicity constraint that $\breve{Q}_i(\theta_{i,1})$ is nonincreasing. Further, we note that local incentive compatibility together with the envelope theorem imply that

$$U_i(\theta_{i,1}, \theta_{i,1}) = U_i(\overline{\theta}, \overline{\theta}) + \int_{\theta_{i,1}}^{\overline{\theta}} \breve{Q}_i(\tilde{\theta}) d\tilde{\theta}. \tag{1.1}$$

Thus, using integration by parts, given prior F_0, ex ante profits are

$$\mathbb{E}_{\theta_{i,1}}[U_i(\theta_{i,1}, \theta_{i,1})] = U_i(\overline{\theta}, \overline{\theta}) + E_{\theta_{i,1}}\left[\breve{Q}_i(\theta_{i,1}) \frac{F_0}{f_0}(\theta_{i,1})\right]. \tag{1.2}$$

This expression shows that there are important similarities and differences between the downward-sloping demand case and the inelastic demand case. In both cases, monotonicity of $\frac{F_0}{f_0}(\theta_{i,1})$ implies that, all else equal, we would like to have $\breve{Q}_i$ nondecreasing as well, allocating more expected demand to higher-cost types. However, the force in favor of $\breve{Q}_i$ nondecreasing conflicts with incentive compatibility and creates a force for pooling.

In addition, in the case of inelastic demand, a rigid price at r maximizes both terms of expression (1.2). This is no longer true when demand is downward-sloping. Instead, the two terms are in conflict about the level of the price. To maximize the profit to the high-cost type, it would be optimal to have a rigid price, as the market share for type $\overline{\theta}$ is thereby made as large as possible; furthermore, the best rigid price would be the monopoly price for the highest-cost type. To maximize the second term (subject to monotonicity constraints), it would also be optimal to have a rigid price, following an argument analogous to the one we used for inelastic demand; however, the best rigid price would now be as low as possible, in order to make the quantity produced as large as possible and maximize the second term (recall that with inelastic demand, the expected market share was fixed at 1). The conflict between the two terms about the optimal price level implies that price rigidity is not necessarily optimal.

At the same time, (1.2) illustrates forces in favor of at least partial price rigidity. To see this, first consider a scheme that has productive efficiency. Even if deviations from a collusive agreement can be punished by giving a firm 0 profits forever (as in the belief threat equilibrium of Section 5.2), it is impossible to construct an equilibrium with productive efficiency in every period, where prices are greater than $\bar{\theta}$. To see why, note that in an equilibrium with productive efficiency, unless all firms have the highest possible cost, the high-cost type gets zero market share and zero profit in every period–so there can be no future reward to the high-cost type for pricing high. Thus, in any period where the market price was greater than $\bar{\theta}$, the high-cost type would have an incentive to deviate and undercut the market price. Thus, prices must be less than or equal to $\overline{\theta}$ in any productive efficiency equilibrium, and $U_i(\overline{\theta}, \overline{\theta}) = 0$.

Consider now a comparison between profits from a productive efficiency equilibria and those from the best rigid-pricing equilibrium. The best rigid-pricing equilibrium (when it exists) entails all firms selecting the price p^* that maximizes $E_{\theta_{i,1}}[D(p)(p - \theta_{i,1})]$. If p^* is greater than $\overline{\theta}$, and firms are sufficiently patient, there will be a PPBE where all firms set this price on the equilibrium path. Note that the shape of the demand curve for $p > \overline{\theta}$ does not affect profits in a productive efficiency equilibrium. Thus, starting from an initial demand curve $D(p)$, we can always find an alternative demand curve that generates the same profits from a productive efficiency equilibrium, but with sufficiently high p^* such that the best rigid-pricing equilibrium (if it exists) dominates any productive efficiency equilibrium that might exist for those parameter values.

Although we have now argued that neither rigid pricing nor productive efficiency will typically be optimal, it is useful to illustrate forces in favor of partial pooling (and in particular, intervals of pooling). Suppose that $\breve{Q}_i$ is strictly decreasing on $[x, y] \subset [\underline{\theta}, \bar{\theta}]$. Now define the series of pricing strategies implicitly (where pricing strategies need only be modified on $[x, y]$) so that $\breve{Q}_i'$ is equal to $\breve{Q}_i$ outside of $[x, y]$ but is constant on $[x, y]$, and

$$\mathbb{E}_{\theta_{i,t}}[\breve{Q}_i(\theta_{i,t})|\theta_{i,t} \in [x, y]] = \mathbb{E}_{\theta_{i,t}}[\breve{Q}_i'(\theta_{i,t})|\theta_{i,t} \in [x, y]] = \breve{Q}_i'(y). \tag{1.3}$$

Then, we can define a probability distribution

$$G(\theta_{i,1}; \breve{Q}_i) = \frac{1}{\breve{Q}_i'(y)} \int_x^{\theta_{i,1}} \breve{Q}_i(\theta_{i,1}) dF_0(\theta_{i,1}|\theta_{i,1} \in [x, y]).$$

Since $G(\cdot; \breve{Q}_i')$ dominates $G(\cdot; \breve{Q}_i)$ by FOSD, if F_0 is log-concave, then

$$E_{\theta_{i,1}}\left[\breve{Q}_i'(\theta_{i,1})\frac{F_0}{f_0}(\theta_{i,1})\right] \geq E_{\theta_{i,1}}\left[\breve{Q}_i(\theta_{i,1})\frac{F_0}{f_0}(\theta_{i,1})\right],$$

and a force in favor of rigidity is thus illustrated. However, the new scheme does not in general satisfy local on-schedule incentive constraints at x and y, and so this argument does not establish that partial rigidity is optimal. A full analysis would incorporate the additional modifications to the scheme that would restore on-schedule incentive compatibility and also satisfy off-schedule incentive compatibility, and establish conditions under which the modification increases expected profits. In general, we expect that partial pooling will be optimal, taking a form similar to the market-sharing two-step scheme analyzed in Section 3.2.3 (though there may be more than two steps).

Despite these complications, we report the following limiting result, establishing that rigid pricing is not a knife-edge result.

Proposition S1 Suppose that demand is downward-sloping. For δ sufficiently large, if F is log-concave and demand is sufficiently inelastic, then the best rigid-pricing scheme is the optimal PPBE.

Proof: Consider a family of demand functions normalized so that $D(0) = 1$. Then, expected profit is given by

$$\begin{aligned}\mathbb{E}_{\theta_{i,1}}[U_i(\theta_{i,1}, \theta_{i,1})] &= U_i(\overline{\theta}, \overline{\theta}) + \mathbb{E}_{\theta_{i,1}}\left[\breve{Q}_i(\theta_{i,1})\frac{F_0}{f_0}(\theta_{i,1})\right] \\ &= U_i(\overline{\theta}, \overline{\theta}) + \mathbb{E}_{\theta_{i,1}}\left[\breve{M}_i(\theta_{i,1})\frac{F_0}{f_0}(\theta_{i,1})\right] + \mathbb{E}_{\theta_{i,1}}\left[\left[\breve{Q}_i(\theta_{i,1}) - \breve{M}_i(\theta_{i,1})\right]\frac{F_0}{f_0}(\theta_{i,1})\right].\end{aligned}$$

for each member of the family. Assuming log-concavity and inelastic demand, we showed in Proposition 2 that a rigid price at the reservation value uniquely maximizes the first two terms. As demand becomes more inelastic, $\breve{Q}_i(\theta_{i,1})$ approaches $\breve{M}_i(\theta_{i,1})$ for all prices below the reservation value, and so the first two terms dominate. The level of the optimal rigid price approaches the reservation value as demand becomes inelastic; further, as in Proposition 2, this scheme is a PPBE sufficiently patient firms. ■

1.2. Imperfect Substitutes

We now consider briefly the possibility that firms sell imperfect substitutes. Note that there are several differences from the perfect substitutes case: with imperfect substitutes, the first-best allocation typically has all firms producing a positive amount, and typically demand is continuous in prices. However, these differences do not affect the analysis of the optimal collusive scheme very much. In this case, given strategies, we can represent demand in period t by

$$Q_i(p_{i,t};\boldsymbol{\rho}_{-i,t},\boldsymbol{\nu}_{-i,t}) = \mathbb{E}_{\boldsymbol{\theta}_{-i,t}}\left[D_i(p_{i,t},\boldsymbol{\rho}_{-i,t}(\boldsymbol{\theta}_{-i,t}))\middle|\,\boldsymbol{\nu}_{-i,t}\right],$$

and let $\breve{Q}_i(\theta_{i,1})$ be the expected discounted demand a firm expects over the course of the game from mimicking $\theta_{i,1}$. Then, (1.2) still characterizes profits.

In this model, expected demand for each cost type depends on the entire pricing function of opponents. If the demand function is linear in prices, however, players care only about the average price of opponents. Then, it is possible to modify pricing functions for each player only on the interval of types $[x,y]$, to change a pricing function from being strictly increasing to being constant on that interval, while leaving the expected price and the expected demand for opponents unchanged outside that interval. Then, the arguments of the last section can be applied: in particular, if $\breve{Q}_i$ is strictly decreasing on $[x,y]\subset[\underline{\theta},\bar{\theta}]$, we can find new pricing strategies such that $\breve{Q}_i'$ is equal to $\breve{Q}_i$ outside of $[x,y]$ but is constant on $[x,y]$, and expected demand on that region is unchanged. Then, as above, there will be a force in favor of pooling. From a formal perspective, this model is closely related to one analyzed by Athey, Atkeson, and Kehoe (2005). We conjecture that their arguments can be modified to show that if F and $1-F$ are log-concave, and demand is linear, then an optimal PPBE is characterized by intervals of pooling.

1.3. Cournot

The analysis is similar when firms compete in quantities rather than prices. To see this, let P be the inverse demand function, and let firm i's expected profit in period t be given by (where firm i's quantity strategy in period t is $\psi_{i,t}$)

$$\bar{\pi}_i(\hat{\theta}_{i,t},\theta_{i,t};\mathbf{s}_t,\boldsymbol{\nu}_{-i}) = \left(\mathbb{E}_{\theta_{i,1}}\left[P(\psi_{i,t}(\theta_{i,t}),\boldsymbol{\psi}_{i,t}(\boldsymbol{\theta}_{-i,t}))|\boldsymbol{\nu}_{-i,t}\right]-\theta_{i,t}\right)\psi_{i,t}(\theta_{i,t}).$$

Again, we can let $\breve{Q}_i(\theta_{i,1})$ denote the expected discounted demand firm i expects from mimicking $\theta_{i,1}$ throughout the game. Then, (1.2) characterizes profits, and a force in favor of pooling remains. Again, we expect optimal PPBE to be characterized by partial pooling.

1.4. Nonlinear Costs

So far we have assumed constant marginal costs. Suppose instead that total costs in period t are given by $h(q_{i,t}, \theta_{i,t}) = h^q(q_{i,t})h^\theta(\theta_{i,t})$ when firm with type $\theta_{i,t}$ produces $q_{i,t}$. Then, we provide sufficient conditions for rigid pricing to dominate alternative schemes that have the property that the highest-cost type serves less than $1/I$ of the market in each period. Any scheme with greater period-by-period productive efficiency than rigid pricing will have this feature. At first it might seem impossible to find an incentive-compatible scheme where the highest-cost type serves more than $1/I$; however, even though it might seem pathological, nonlinearities in cost do make it possible. Thus, we stop short of a full proof of the optimality of rigid pricing. Despite this, our analysis makes clear that nonlinear costs do not remove the incentive for pooling, nor do they invalidate our overall approach. In addition, our analysis establishes that rigid pricing dominates a wide range of schemes with greater productive efficiency.

Proposition S2 Suppose that (i) h^q and h^θ are differentiable, with $h^\theta > 0$, $h^q > 0$, $h^q_q > 0$, $h^\theta_\theta > 0$, $h^q_{qq} \leq 0$, and $h^\theta_{\theta\theta} \geq 0$, and (ii)

$$r > h_q(0, \bar{\theta}). \tag{1.4}$$

Then rigid pricing at r dominates any scheme satisfying the market share restriction that $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) \leq 1/I$ for each i, t and $\boldsymbol{\theta}_{-i,1}$.

Before proving the result, we pause to interpret the sufficient conditions. They require economies of scale: marginal costs are decreasing. It may seem somewhat surprising that with economies of scale, pooling could be optimal, since cost considerations favor shifting production to one firm even more than in the case of constant marginal costs. However, like in the case of constant marginal cost, allocating more market share to high-cost types relaxes incentive constraints for low-cost types. Thus, expected "information rents" to the firms are higher when more market share goes to high-cost firms. When r is higher than marginal cost for the high-cost type in the relevant range, and if rigid pricing increases the market share to the high-cost type, then rigid pricing increases profits to the high-cost type as well.

When the sufficient conditions of the proposition fail, there are generally competing effects, and it may be necessary to consider a parameterized model to fully characterize optimal PPBE.

Proof: Following our notation above, let $\breve{R}_i(\hat{\theta}_{i,1})$ denote the expected future discounted revenue that firm i anticipates if it mimics type $\hat{\theta}_{i,1}$ throughout the game, and let $\breve{M}_i(\hat{\theta}_{i,1})$ be the associated market share. Let $\breve{H}_i(\hat{\theta}_{i,1}, \theta_{i,1})$ be the expected discounted total cost when type $\theta_{i,1}$

mimics $\hat{\theta}_{i,1}$ throughout the game, decomposed into two components, so that $\breve{H}_i(\hat{\theta}_{i,1}, \theta_{i,1}) = h^\theta(\theta_{i,1}) \cdot \breve{H}^q(\hat{\theta}_{i,1})$.

Let

$$U_i(\hat{\theta}_{i,1}, \theta_{i,1}) = \breve{R}_i(\hat{\theta}_{i,1}) - \breve{H}_i(\hat{\theta}_{i,1}, \theta_{i,1}).$$

Using the envelope theorem and on-schedule incentive compatibility,

$$U_i(\theta_{i,1}, \theta_{i,1}) = E_{\boldsymbol{\theta}_{-i,1}} \left[\breve{R}_i(\bar{\theta}) - \breve{H}_i(\bar{\theta}, \bar{\theta}) + \int_{\theta_{i,1}}^{\bar{\theta}} h^\theta_\theta(\tilde{\theta}_{i,1}) \cdot \breve{H}^q(\tilde{\theta}_{i,1}) d\tilde{\theta}_{i,1} \right].$$

So, using integration by parts,

$$\mathbb{E}[U_i(\theta_{i,1}, \theta_{i,1})] = \mathbb{E} \left[\breve{R}_i(\bar{\theta}) - \breve{H}_i(\bar{\theta}, \bar{\theta}) + h^\theta_\theta(\theta_{i,1}) \breve{H}^q(\theta_{i,1}) \frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})} \right]. \tag{1.5}$$

The on-schedule incentive constraints also imply that

$$\begin{aligned} U_i(\theta_{i,1}, \theta_{i,1}) &\geq U_i(\hat{\theta}_{i,1}, \theta_{i,1}) = U_i(\hat{\theta}_{i,1}, \hat{\theta}_{i,1}) + \breve{H}_i(\hat{\theta}_{i,1}, \hat{\theta}_{i,1}) - \breve{H}_i(\hat{\theta}_{i,1}, \theta_{i,1}) \\ U_i(\hat{\theta}_{i,1}, \hat{\theta}_{i,1}) &\geq U_i(\theta_{i,1}, \hat{\theta}_{i,1}) = U_i(\theta_{i,1}, \theta_{i,1}) + \breve{H}_i(\theta_{i,1}, \theta_{i,1}) - \breve{H}_i(\theta_{i,1}, \hat{\theta}_{i,1}), \end{aligned}$$

so combining,

$$\breve{H}_i(\hat{\theta}_{i,1}, \hat{\theta}_{i,1}) + \breve{H}_i(\theta_{i,1}, \theta_{i,1}) \leq \breve{H}_i(\hat{\theta}_{i,1}, \theta_{i,1}) + \breve{H}_i(\theta_{i,1}, \hat{\theta}_{i,1}).$$

In words, on-schedule incentive compatibility implies that $\breve{H}_i$ is submodular. Using our definitions and monotonicity restrictions, that in turn implies that $\breve{H}^q(\hat{\theta}_{i,1})$ is nonincreasing in $\widehat{\theta}_{i,1}$. This does not, however, imply that expected market shares are globally decreasing in $\hat{\theta}_{i,1}$, since h^q is nonlinear.

"Profit at the Top"

Fix a PPBE. Let $m_{i,t}(\boldsymbol{\theta}_1)$ be the market share allocation to firm i in period t as a function of firm types. For every i, t and $\boldsymbol{\theta}_{-i,1}$, we assume that $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) \leq 1/I$. The ex ante expected value of "profit-at-the-top" to firm i in period t is then

$$\mathbb{E}_{\boldsymbol{\theta}_{-i,1}}[\rho_{i,t}(\bar{\theta}) m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) - h^\theta(\bar{\theta}) h^q(m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}))].$$

We claim that this profit-at-the-top cannot exceed that which is achieved when a best rigid-pricing scheme is used in period t.

To establish this claim, we fix i and t and suppose that $\rho_{i,t}(\bar{\theta}) < r$ and/or $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) < 1/I$ for a positive measure of values for $\boldsymbol{\theta}_{-i,1}$. (Recall that a best rigid-pricing scheme has $\rho_{i,t}(\bar{\theta}) = r$ and $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) \equiv 1/I$.) Clearly, if $\rho_{i,t}(\bar{\theta}) < r$, then profit-at-the-top would be increased if the price were raised to r. Suppose then that $\rho_{i,t}(\bar{\theta}) = r$ and $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) < 1/I$ for a positive measure of values for $\boldsymbol{\theta}_{-i,1}$. For each such $\boldsymbol{\theta}_{-i,1}$, if we were to increase firm i's market share from $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1})$ to $1/I$, then firm i would enjoy a strict increase in profit when state $(\bar{\theta}, \boldsymbol{\theta}_{-i,1})$ occurs. This follows since $r - h^\theta(\bar{\theta}) h^q_q(q_{i,t}) \geq r - h^\theta(\bar{\theta}) h^q_q(0) = r - h_q(0, \bar{\theta}) > 0$ for $q_{i,t} \in [0, 1/I]$, where the first inequality uses $h^q_{qq}(q_{i,t}) \leq 0$ and the second inequality uses (1.4). Over a positive measure of values for $\boldsymbol{\theta}_{-i,1}$, these pointwise improvements imply a strict increase of profit-at-the-top to firm i in period t.

Thus, when a best rigid-pricing scheme is used, the effect on the profit-at-the-top in each period is positive, and so the aggregate effect must be positive. We have therefore established that any PPBE with the feature that $m_{i,t}(\bar{\theta}, \boldsymbol{\theta}_{-i,1}) \leq 1/I$ for each i, t and $\boldsymbol{\theta}_{-i,1}$ is dominated by using the best-rigid pricing scheme in each period. Since rigid pricing at r is a PPBE for sufficiently large δ, it provides greater expected profit at the top than any other PPBE that satisfies the market share restriction.

"Information Rents"

The last part of the ex ante expected profits expression (1.5), referred to as the "information rents," is

$$\mathbb{E}_{\theta_{i,1}}\left[h^{\theta}_{\theta}(\theta_{i,1})\breve{H}^{q}(\theta_{i,1})\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}\right].$$

For any $\breve{M}_i$ associated with a PPBE, consider an alternative market share allocation function derived from rigid pricing, so that market shares are equal for all types and all firms in each period, and $\breve{H}^q(\theta_i)$ is constant in θ_i. Rigid pricing dominates from the perspective of information rents if

$$\mathbb{E}_{\theta_{i,1}}\left[\left(h^{\theta}_{\theta}(\theta_{i,1})\frac{h^q(1/I)}{1-\delta} - h^{\theta}_{\theta}(\theta_{i,1})\breve{H}^{q}(\theta_{i,1})\right)\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}\right] \geq 0.$$

The expression can be rewritten as

$$\text{cov}\left(\left(\frac{h^q(1/I)}{1-\delta} - \breve{H}^q(\theta_{i,1})\right), h^{\theta}_{\theta}(\theta_{i,1})\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}\right) + \mathbb{E}_{\theta_{i,1}}\left[\left(\frac{h^q(1/I)}{1-\delta} - \breve{H}^q(\theta_{i,1})\right)\right] \cdot \mathbb{E}\left[h^{\theta}_{\theta}(\theta_{i,1})\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}\right]. \tag{1.6}$$

Note that $\mathbb{E}_{\theta_{i,1}}\left[h^{\theta}_{\theta}(\theta_{i,1})\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}\right] \geq 0$ since the integrand is positive everywhere. In addition, concavity of h^q and Jensen's inequality imply that for each t,

$$h^q(1/I) = h^q(\mathbb{E}_{\boldsymbol{\theta}_1}[m_{i,t}(\boldsymbol{\theta}_1)]) \geq \mathbb{E}_{\boldsymbol{\theta}_1}\left[h^q(m_{i,t}(\boldsymbol{\theta}_1))\right] = \mathbb{E}_{\theta_{i,1}}\left[\mathbb{E}_{\boldsymbol{\theta}_{-i,1}} h^q(m_{i,t}(\boldsymbol{\theta}_1))\right],$$

so that the second term of (1.6) is positive, using the definition of $\breve{H}^q(\theta_{i,1})$. On the first term, given log-concavity, $h^{\theta}_{\theta} > 0$, and $h^{\theta}_{\theta\theta} \geq 0$, $h^{\theta}_{\theta}(\theta_{i,1})\frac{F_0(\theta_{i,1})}{f_0(\theta_{i,1})}$ is nondecreasing. By incentive-compatibility, $\frac{h^q(1/I)}{1-\delta} - \breve{H}^q(\theta_{i,1})$ is nondecreasing. The covariance of two nondecreasing functions of a single random variable must be positive. Therefore, (1.6) must be positive, and the alternative with equal market shares must be preferred. ■

We conclude by noting that the restriction on market shares was used only in the analysis of "profit-at-the-top," so that there will still be a strong force for at least partial pooling (from the information rents term) even when the market share restriction fails.

1.5. Robustness of Results about Efficient Collusion

Although we have not conducted a full numerical analysis of alternative models, the general approach we take to constructing first-best equilibria can be generalized to alternative models. The first step would be constructing a punishment equilibrium analogous to the carrot-stick pooling equilibrium (or some alternative). Although critical discount factors would clearly differ, and some details of the construction would differ, there is no reason to expect difficulties generalizing the carrot-stick equilibrium to alternative models. The dynamic programming approach we employ for analyzing first-best equilibria generalizes directly to other models of product market competition. Since first-best equilibrium payoffs are higher than payoffs from pooling equilibria, for sufficiently patient firms off-schedule incentive constraints should not bind.

The main challenge in eliciting truthful revelation alongside efficient market-share allocation is to provide future "rewards" and "punishments" that provide sufficient incentives for firms to give up market share, while still implementing efficient allocation in the reward and punishment continuation equilibria. With discrete types and perfect substitutes (either Bertrand or Cournot), states of the world arise with positive probability where market share can be shifted among firms. For particular functional forms, it is straightforward to write the system of equations and incentive constraints that must be satisfied to implement a first-best equilibrium, analogous to the system described in the paper. One could then describe the parameter values for which a first-best equilibrium exists. We note that if firms compete in quantities, communication plays a more substantive role in coordinating firms on the desired quantities as a function of cost.

With imperfect substitutes, there is typically a unique market-share allocation for any cost vector, even when costs are identical. Thus, some inefficiency should be expected for any discount factor strictly less than 1. Similarly, with nonlinear costs, diseconomies of scale imply that firms should share the market in a particular way. Thus, distorting production in order to provide rewards or punishments for past revelation will lead to some inefficiency. In both cases, the requisite inefficiency should decrease as patience increases.

2. Using Dynamic Programming to analyze the First-Best Scheme

This section provides more details about applying dynamic programming approaches in the spirit of Abreu, Pearce, and Stacchetti (1986, 1990) and Cole and Kocherlakota (2001) to this model. Although this is not necessary to establish that the first-best scheme is a PPBE, the techniques are useful more generally for analyzing self-generating sets of PPBE values either numerically or analytically.

We retain the notation from the paper, but also introduce some new notation. Let $\mathcal{V}$ be the set of functions $\mathbf{v} = (v_1, .., v_I)$ such that $v_i : \Theta_i \to \mathbb{R}$. This is the set of possible "type-contingent payoff functions." Let $\mathcal{W} = \mathcal{V} \times \Delta\Theta$. The set of PPBE then corresponds to a subset of $\mathcal{W}$. Each equilibrium is described by a set of initial beliefs about opponents, $\boldsymbol{\mu} \in \Delta\Theta$, and a function $\mathbf{v} \in \mathcal{V}$ that specifies the payoff each player expects to attain, conditional on the player's true type.

Consider "continuation value and belief functions" $(\mathbf{V}, \mathbf{M})$ mapping $\mathcal{Z}$ to $\mathcal{W}$. For every possible publicly observed outcome $\mathbf{z}_t$ from period t, these functions specify an associated belief and a type-contingent continuation payoff function. That is, $V_i(\mathbf{z}_t)$ is the type-contingent payoff function that will be realized following observed actions $\mathbf{z}_t$, and $V_i(\mathbf{z}_t)(\theta_{i,t+1})$ is the payoff firm i expects starting in period $t+1$ if $\mathbf{z}_t$ was the vector of observed actions in period t and its true type in period $t+1$ is $\theta_{i,t+1}$. Note that this structure makes it possible to compute firm i's expected future payoffs if firm i mimics another type in period t. The chosen actions affect which continuation payoff function is used through $\mathbf{z}_t$, but the firm's true type determines the firm's beliefs about $\theta_{i,t+1}$ and thus the firm's continuation value.

Let $\eta_j(a_{j,t}, s_{j,t}, \mu_{j,t})$ represent the belief that firm $i \neq j$ has about player j in period t after firm j has made announcement $a_{j,t}$, given that firm $i \neq j$ began the period with beliefs $\mu_{j,t}$ and that it posits that player j uses period strategy $s_{j,t}$. Then, for a continuation value function $\mathbf{V} = (V_1, .., V_I)$, define expected discounted payoffs for firm i in period t, when firm i's type is $\theta_{i,t}$, after announcements have been observed to be $\mathbf{a}_t$, and firm i chooses actions $p_{i,t}, q_{i,t}$ following these announcements:

$$
\begin{aligned}
&u_i(\mathbf{a}_t, p_{i,t}, q_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}, V_i) \\
&= (p_{i,t} - \theta_{i,t}) \cdot \mathbb{E}_{\boldsymbol{\theta}_{-i,t}} \left[\varphi_i \left(\left(p_{i,t}, \boldsymbol{\rho}_{-i,t}(\mathbf{a}_t, \boldsymbol{\theta}_{-i,t})\right); \left(q_{i,t}, \boldsymbol{\psi}_{-i,t}(\mathbf{a}_t, \boldsymbol{\theta}_{-i,t})\right)\right) \middle| \ \boldsymbol{\eta}_{-i}(\mathbf{a}_{-i,t}, \mathbf{s}_{-i,t}, \boldsymbol{\mu}_{-i,t}) \right] \\
&\quad + \delta \mathbb{E}_{\theta_{i,t+1}, \boldsymbol{\theta}_{-i,t}} \left[V_i \left((a_{i,t}, p_{i,t}, q_{i,t}), \mathbf{s}_{-i,t}(\mathbf{a}_t, \boldsymbol{\theta}_{-i,t}) \right) (\theta_{i,t+1}) \mid \boldsymbol{\eta}_{-i}(\mathbf{a}_{-i,t}, \mathbf{s}_{-i,t}, \boldsymbol{\mu}_{-i,t}), \theta_{i,t} \right].
\end{aligned}
$$

As in Section 4.2 of the paper, the following represents firm i's expected payoffs in period t, before announcements are made, when firm i has type $\theta_{i,t}$ and mimics type $\hat{\theta}_{i,t}$:

$$
\begin{aligned}
&\bar{u}_i(\hat{\theta}_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}, V_i) = \bar{\pi}_i(\hat{\theta}_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}) \\
&\quad + \delta \mathbb{E}_{\theta_{i,t+1}, \boldsymbol{\theta}_{-i,t}} \left[V_i \left(\mathbf{s}_t(\boldsymbol{\alpha}_t(\hat{\theta}_{i,t}, \boldsymbol{\theta}_{-i,t}), (\hat{\theta}_{i,t}, \boldsymbol{\theta}_{-i,t})) \right) (\theta_{i,t+1}) \ \middle| \ \boldsymbol{\mu}_{-i,t}, \theta_{i,t} \right].
\end{aligned}
$$

Then, following Cole and Kocherlakota (2001), we define a mapping $B : \mathcal{W} \to \mathcal{W}$, such that

$$
B(W) = \left\{ \begin{array}{c} (\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{W} : \ \forall i, \ \exists s_i^* \in \mathcal{S}_i \text{ and } (\mathbf{V}, \mathbf{M}) : \mathcal{Z} \to W, \\ \text{s.t. } \mathbf{M}(\mathbf{z}) \in \tilde{\mathbf{T}}(\boldsymbol{\mu}, \mathbf{s}^*, \mathbf{z}) \ \forall \mathbf{z} \in \mathcal{Z}, \\ v_i(\theta_i) = \bar{u}_i(\theta_i, \theta_i, \mathbf{s}^*, \boldsymbol{\mu}_{-i}, V_i) \ \forall \theta_i \in \Theta_i, \text{ and} \\ \text{(IC) } s_i^* \in \arg\max_{s_i \in \mathcal{S}_i} \bar{u}_i(\theta_i, \theta_i, (s_i, \mathbf{s}_{-i}^*), \boldsymbol{\mu}_{-i}, V_i). \end{array} \right\}. \qquad (2.1)
$$

Note that in this definition, the requirement that $(\mathbf{V}, \mathbf{M}) : \mathcal{Z} \to W$ is restrictive, since W is not a product set. The requirement ensures that the continuation value and belief functions are compatible: given the beliefs that arise given posited strategies and observations, $\mathbf{M}(\mathbf{z})$, only a subset of potential continuation value functions satisfy $(\mathbf{V}(\mathbf{z}), \mathbf{M}(\mathbf{z})) \in W$. Intuitively, today's actions reveal information about cost types that restrict the expected value of costs, and thus feasible payoffs, tomorrow.

Following Cole and Kocherlakota (2001), standard arguments can be adapted to show that the operator B is monotone (where a set A is larger than B if $A \supseteq B$). Showing that the set of PPBE is the largest fixed point of the operator B involves more work, because our model differs in a few respects from that of Cole and Kocherlakota. Cole and Kocherlakota's (2001) assumptions about the monitoring technology imply that $\tilde{\mathbf{T}}$ is single-valued, since it is impossible to observe outcomes $\mathbf{z}$ that are inconsistent with strategies $\mathbf{s}^*$. Our definition of B has an additional degree of freedom, since $B(W)$ may include different elements that are supported using different off-equilibrium-path beliefs. Although this in itself does not pose a difficulty, the fact that we consider a hidden-information game does raise additional issues. In particular, beliefs are not continuous in strategies: in the limit as a separating strategy approaches pooling (for example, announcements are uninformative, and the prices chosen by different types approach one another), the sequence of separating strategies induce very different beliefs than those induced by the limiting pooling strategy.

Two additional technical differences from the literature arise. First, the strategy space in each period is a compact set but is not finite. Second, there is a discontinuity in payoffs due to the Bertrand stage game: a firm can receive discretely higher payoffs by selecting a slightly different price. We believe that these differences can be addressed, but a full treatment is beyond the scope of this paper, and it is not necessary for our purposes.

The most useful insight for analyzing a particular class of PPBE is that if we can explicitly construct a set W such that $W \subseteq B(W)$, this set must be a PPBE. Formally:

Lemma S1 Let W^* be the set of PPBE type-contingent payoff functions and beliefs. For any compact set $W \subseteq \mathcal{W}$ such that $W \subseteq B(W)$, W is a self-generating equilibrium set: $W \subseteq W^*$.

This lemma follows by adapting the findings of Abreu, Pearce, and Stacchetti (1986) and Cole and Kocherlakota (2001) to our game; the extension is straightforward given the definitions.

Each PPBE is described by a $\mathbf{w} = (\mathbf{v}, \boldsymbol{\mu}) \in W^*$, which is the type-contingent payoff function and the belief. Since each $\mathbf{w} \in W^*$ corresponds to a PPBE outcome, we will simply refer to $\mathbf{w} \in W^*$ as a PPBE. Further, since $\mathbf{w} \in W^*$ implies $\mathbf{w} \in B(W^*)$, each such equilibrium can be "decomposed" into the period strategies, $\mathbf{s}^*$, and the continuation belief and payoff mappings $(\mathbf{V}, \mathbf{M}) : \mathcal{Z} \to W$ that are guaranteed to exist by the definition of B. We rely heavily on this way of describing and analyzing equilibria below.

The incentive constraint in the definition of B includes all deviations. We wish to separate the types of deviations into "on-schedule" deviations, whereby one type mimics another, and "off-schedule" deviations, where a type chooses an action that was not assigned to any type.

Unlike the case of perfect persistence, here there is always a chance that types change from period to period, and so any kind of "mimicking" behavior constitutes an on-schedule deviation, even if the firm mimics different types at different points in time. The on-schedule constraint for firm i can be written as follows:

$$\bar{u}_i(\theta_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_t, V_i) \geq \bar{u}_i(\hat{\theta}_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}, V_i)\ \forall \hat{\theta}_{i,t}, \theta_{i,t}. \tag{2.2}$$

The off-schedule constraint is written:

$$\begin{aligned} & u_i(\boldsymbol{\alpha}_t(\boldsymbol{\theta}_t), \rho_{i,t}(\boldsymbol{\alpha}_t(\boldsymbol{\theta}_t), \theta_{i,t}), \psi_{i,t}(\boldsymbol{\alpha}_t(\boldsymbol{\theta}_t), \theta_{i,t}), \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}, V_i) \\ & \geq u_i((a_{i,t}, \boldsymbol{\alpha}_{-i,t}(\boldsymbol{\theta}_{-i,t})), p_{i,t}, q_{i,t}, \theta_{i,t}; \mathbf{s}_t, \boldsymbol{\mu}_{-i,t}, V_i) \text{ for all } \theta_{i,t},\ \boldsymbol{\theta}_{-i,t}, \\ & \text{and all } (a_{i,t}, p_{i,t}, q_{i,t}) \notin \left\{ \begin{array}{c} (a'_{i,t}, p'_{i,t}, q'_{i,t}) : \exists \hat{\theta}_{i,t} \in \Theta_i \text{ s.t. } a'_{i,t} = \alpha_{i,t}(\hat{\theta}_{i,t}), \\ p'_{i,t} = \rho_{i,t}(\boldsymbol{\alpha}_t(\hat{\theta}_{i,t}, \boldsymbol{\theta}_{-i,t}), \hat{\theta}_{i,t}),\ q'_{i,t} = \psi_{i,t}(\boldsymbol{\alpha}_t(\hat{\theta}_{i,t}, \boldsymbol{\theta}_{-i,t}), \hat{\theta}_{i,t}) \end{array} \right\}. \end{aligned} \tag{2.3}$$

If, for all i, both of these constraints are satisfied for $\mathbf{s} = \mathbf{s}^*$, then $s_i^* \in \arg\max_{s_i \in \mathcal{S}_i} \bar{u}_i(\theta_i, \theta_i, (s_i, \mathbf{s}^*_{-i}), \boldsymbol{\mu}_{-i}, V_i)$.

2.1. Applying Dynamic Programming Tools to the First-Best Scheme

In order to verify that a particular first-best scheme is a PBBE, there are two steps. First, we construct a set $W^{fb} \subset \mathcal{W}$, which requires constructing the (state-contingent payoff functions, belief) pairs induced by the first-best scheme, as well as associated "continuation value functions." Using this development, the second step is to define a set of incentive constraints and verify that they are satisfied, thus ensuring that strategies of the first-best scheme are indeed a PPBE.

The paper defines period strategies $\tilde{s}_i$ and continuation value functions $\tilde{\mathbf{V}}$ for each state, and it states a system of equations that yields the state-contingent payoff functions $\tilde{\mathbf{v}}(\cdot; j, \boldsymbol{\theta}_{t-1}) \in \mathcal{V}$.

Recal that v^{cs} and v^r are the type-contingent payoff functions from the worst carrot-stick and best rigid pricing equilibria, respectively. We consider a modified version of a carrot-stick scheme where players announce their types in each period. Since behavior does not depend on beliefs about opponents on or off of the equilibrium path in this scheme, firms are indifferent about their reports, and so the carrot-stick scheme with truthful reporting is a PPBE so long as the original carrot-stick scheme was a PPBE. Let the set of possible beliefs in a fully separating equilibrium be $\mathcal{M}^{FS} = \{\boldsymbol{\mu} \in \Delta\Theta^2 : \boldsymbol{\mu} = \mathbf{F}(\cdot; \boldsymbol{\theta}_{t-1}) \text{ for some } \boldsymbol{\theta}_{t-1} \in \Theta^2\}$. Then, define the following sets of (state-contingent payoff function, belief) pairs:

$$\begin{aligned} W^e &= \{(\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{W} : \exists \omega_j(\boldsymbol{\theta}_{t-1}) \in \Omega^e \text{ s.t. } \boldsymbol{\mu} = \mathbf{F}(\cdot; \boldsymbol{\theta}_{t-1}) \text{ and } \mathbf{v} = \tilde{\mathbf{v}}(\cdot; j, \boldsymbol{\theta}_{t-1})\}, \\ W^p &= \{(\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{W} :\ \mathbf{v} = (v^{cs}, v^{cs}),\ \boldsymbol{\mu} \in \mathcal{M}^{FS}\} \cup \{(\mathbf{v}, \boldsymbol{\mu}) \in \mathcal{W} :\ \mathbf{v} = (v^r, v^r),\ \boldsymbol{\mu} \in \mathcal{M}^{FS}\}. \end{aligned}$$

We wish to show that $W^{fb} = W^e \cup W^p$ is a self-generating set in the sense of Abreu, Pearce, and Stacchetti (1986), and as applied to our problem in Lemma S1. Informally, we require that for each element w of W^{fb}, we can find strategies $\mathbf{s}^*$ and a continuation value function $\mathbf{V}: \mathcal{Z} \to \mathcal{V}$, such that three conditions hold. First, we require feasibility: $(\mathbf{V}(\mathbf{z}_t), \mathbf{F}(\cdot;\mathbf{a}_t)) \in W^{fb}$, so that the continuation valuation function gives a state-contingent payoff function that is in W^{fb} given the beliefs induced by the period's public outcomes. Second, we require "promise-keeping": the strategies and continuation valuation functions deliver the promised state-contingent payoffs. Third, the strategies $\mathbf{s}^*$ must be best responses.

The following result formalizes the sufficient conditions required to verify that a first-best scheme is a PPBE. It provides an alternative to the approach outlined in the paper.

Proposition S3 Fix $I = 2$ and consider the two-type model with imperfect persistence, with primitives δ, r, L, H, and $\mathbf{F}$, with $\delta \geq \delta_c$. Fix the specification of a first-best scheme $\tilde{g}$, $\tilde{q}_1$ and $\tilde{T}$, and consider the corresponding $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{V}}$. If for each $(i, j, \boldsymbol{\theta}_{t-1}) \in \{1,2\}^2 \times \Theta^2$, the on-schedule and off-schedule constraints, (2.2) and (2.3), hold when $\boldsymbol{\mu}_t = \mathbf{F}(\cdot;\boldsymbol{\theta}_{t-1})$, $\mathbf{V} = \tilde{\mathbf{V}}(\cdot; j, \boldsymbol{\theta}_{t-1})$, and $\mathbf{s}^* = \tilde{\mathbf{s}}(\cdot; j, \boldsymbol{\theta}_{t-1})$, then, $W^{fb} \subseteq B(W^{fb})$, and W^{fb} is a self-generating PPBE set that yields first-best profits in every period.

Proof: We established in the paper that W^p is a self-generating equilibrium set. By Lemma 2, it remains to show that $\mathbf{w} \in W^e$ implies $\mathbf{w} \in B(W^{fb})$. By construction, each $\mathbf{w} \in W^e$ is associated with a $\omega_j(\boldsymbol{\theta}_{t-1}) \in \Omega^e$. Let $\mathbf{s}^* = \tilde{\mathbf{s}}(\cdot; j, \boldsymbol{\theta}_{t-1})$ and $\mathbf{V} = \tilde{\mathbf{V}}(\cdot; j, \boldsymbol{\theta}_{t-1})$, and note that $\mathbf{s}^* \in \mathcal{S}$ and, for all $\mathbf{z}$, $(\mathbf{V}(\mathbf{z}), \mathbf{F}(\cdot;\mathbf{a})) \in W^{fb}$, so that feasibility is satisfied. Further, letting $v_i = \tilde{v}_i(\cdot; j, \boldsymbol{\theta}_{t-1})$ for each i, it follows by definition of $\tilde{v}_i$ that promise-keeping holds: $v_i(\theta_i) = \bar{u}_i(\theta_i, \theta_i, \mathbf{s}, \boldsymbol{\mu}_{-i}, V_i)$. Finally, if (2.2) and (2.3) hold with these definitions, $\mathbf{s}_i^*$ is a best response to $\mathbf{s}_{-i}^*$ for each i. Thus, $\mathbf{w} \in B(W^{fb})$, as desired. ■

3. Productive Efficiency with Perfect Persistence

3.1. Proof of Proposition 6

The paper stated two sufficient conditions for the existence of the desired equilibrium:

$$\inf_{\theta'_{i,1} > \theta''_{i,1}} \frac{f_0(\theta'_{i,1})}{f_0(\theta''_{i,1})} > \frac{\delta}{1 - \delta(1-\delta)} \text{ and } \inf_{\theta'_{i,1} < \theta''_{i,1}} \frac{f_0(\theta'_{i,1})}{f_0(\theta''_{i,1})} > \delta, \tag{3.1}$$

and δ is sufficiently small that, for all $\theta_{i,1}$,

$$\frac{2 f_0(\theta_{i,1})}{(1 - F_0(\theta_{i,1}))^2} \int_{\widetilde{\theta}_i = \theta_{i,1}}^{\overline{\theta}} (1 - F_0(\widetilde{\theta}_i)) d\widetilde{\theta}_i > \delta. \tag{3.2}$$

In a separating equilibrium with productive efficiency, $\rho_{i,1}(\theta_{i,1})$ is strictly increasing and symmetric across firms, and $\beta(\theta_w, \theta_l) \in [\theta_w, \theta_l]$ and thus $\beta(\theta_w, \theta_w) = \theta_w$. Let $\rho(\theta_{i,1})$ denote the symmetric first-period pricing function. Suppose further that $\beta(\theta_w, \theta_l)$ is monotonic, in that it is strictly increasing in each argument.

Fix $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$. First, suppose firm i engages in a downward deviation, by mimicking the first-period price of $\widehat{\theta}_{i,1}$ slightly below $\theta_{i,1}$. Consider the types $\widetilde{\theta}_{j,1}$ for the rival firm j such that the rival loses (i.e., $\widetilde{\theta}_{j,1} > \widehat{\theta}_{i,1}$) and the winner chooses a future price that exceeds $\theta_{i,1}$ (i.e., $\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1}) > \theta_{i,1}$). Observe that $\beta(\widehat{\theta}_{i,1}, \theta_{i,1}) < \theta_{i,1}$; further, $\beta(\theta_{i,1}, \overline{\theta}) > \beta(\theta_{i,1}, \theta_{i,1}) = \theta_{i,1}$, and so for $\widehat{\theta}_{i,1}$ slightly below $\theta_{i,1}$, we have that $\beta(\widehat{\theta}_{i,1}, \overline{\theta}) > \theta_{i,1}$. We conclude that there exists a unique value $\theta_c(\widehat{\theta}_{i,1}, \theta_{i,1}) \in (\theta_{i,1}, \overline{\theta})$ that satisfies $\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1}) = \theta_{i,1}$. Second, suppose firm i engages in an upward deviation, by mimicking the first-period price of $\widehat{\theta}_{i,1}$ slightly above $\theta_{i,1}$. Consider the types $\widetilde{\theta}_{j,1}$ for the rival such that the rival wins (i.e., $\widetilde{\theta}_{j,1} < \widehat{\theta}_{i,1}$) and chooses a future price that exceeds $\theta_{i,1}$ (i.e., $\beta(\widetilde{\theta}_{j,1}, \widehat{\theta}_{i,1}) > \theta_{i,1}$). Observe that $\beta(\theta_{i,1}, \widehat{\theta}_{i,1}) > \theta_{i,1}$; further, $\beta(\underline{\theta}, \widehat{\theta}_{i,1}) < \theta_{i,1}$ for $\widehat{\theta}_{i,1}$ sufficiently little above $\theta_{i,1}$. We conclude that there exists a unique value $\theta_b(\widehat{\theta}_{i,1}, \theta_{i,1}) \in (\underline{\theta}, \theta_{i,1})$ that satisfies $\beta(\widetilde{\theta}_{j,1}, \widehat{\theta}_{i,1}) = \theta_{i,1}$.

Consider the following downward deviation: Firm i with type $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$ mimics $\widehat{\theta}_{i,1}$ slightly below $\theta_{i,1}$ (i.e., chooses $\rho(\widehat{\theta}_{i,1}) < \rho(\theta_{i,1})$), and then (i) if $\widetilde{\theta}_{j,1} < \widehat{\theta}_{i,1}$, firm i makes no first-period sale and exits (e.g., prices above r) in all future periods; (ii). if $\widetilde{\theta}_{j,1} \in (\widehat{\theta}_{i,1}, \theta_c)$, firm i makes the first-period sale and exits (e.g., prices above r) in all future periods; and (iii). if $\widetilde{\theta}_{j,1} > \theta_c$, firm i makes the first-period sale and mimics thereafter the equilibrium pricing behavior of type $\widehat{\theta}_{i,1}$ (i.e, sets $\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1})$ in all future periods). As $\widehat{\theta}_{i,1} \uparrow \theta_{i,1}$, the deviating firm i's payoff approaches that which it earns in the putative equilibrium. Thus, a necessary feature of a productive efficient equilibrium is that firm i does better by announcing $\widehat{\theta}_{i,1} = \theta_{i,1}$ than any other $\widehat{\theta}_{i,1} < \theta_{i,1}$, given the associated strategies described in (i)-(iii) above.

Under (i)-(iii), the profit from a downward deviation is defined as

$$\pi^D(\widehat{\theta}_{i,1}, \theta_{i,1}) = [\rho(\widehat{\theta}_{i,1}) - \theta_{i,1}][1 - F_0(\widehat{\theta}_{i,1})] + \frac{\delta}{1-\delta} \int_{\theta_c}^{\overline{\theta}} [\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1}) - \theta_{i,1}] dF_0(\widetilde{\theta}_{j,1}). \tag{3.3}$$

Lemma S2 For any $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$, if derivatives are evaluated as $\widehat{\theta}_{i,1} \uparrow \theta_{i,1}$,

$$\pi^D_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = [\rho(\theta_{i,1}) - \theta_{i,1}][-F_0'(\theta_{i,1})] + [1 - F_0(\theta_{i,1})]\rho'(\theta_{i,1}) + \frac{\delta}{1-\delta} \int_{\theta_{i,1}}^{\overline{\theta}} \beta_{\theta_w}(\theta_{i,1}, \widetilde{\theta}_{j,1}) dF_0(\widetilde{\theta}_{j,1}) \tag{3.4}$$

$$\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\theta_{i,1}, \theta_{i,1}) = F_0'(\theta_{i,1})[1 - \frac{\delta}{1-\delta} \frac{\beta_{\theta_w}(\theta_{i,1}, \theta_{i,1})}{\beta_{\theta_l}(\theta_{i,1}, \theta_{i,1})}] \tag{3.5}$$

Proof: Using (3.3) and the definition of θ_c, we find that

$$\pi^D_{\widehat{\theta}_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) = [\rho(\widehat{\theta}_{i,1}) - \theta_{i,1}][-F_0'(\widehat{\theta}_{i,1})] + [1 - F_0(\widehat{\theta}_{i,1})]\rho'(\widehat{\theta}_{i,1}) + \frac{\delta}{1-\delta} \int_{\theta_c}^{\overline{\theta}} \beta_{\theta_w}(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1}) dF_0(\widetilde{\theta}_{j,1}).$$

Differentiating with respect to $\theta_{i,1}$ and using $\partial\theta_c/\partial\theta_{i,1} = 1/\beta_{\theta_l}(\widehat{\theta}_{i,1}, \theta_c)$, we obtain

$$\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) = F_0'(\widehat{\theta}_{i,1}) - \frac{\delta}{1-\delta}\frac{\beta_{\theta_w}(\widehat{\theta}_{i,1}, \theta_c)}{\beta_{\theta_l}(\widehat{\theta}_{i,1}, \theta_c)}F_0'(\theta_c). \tag{3.6}$$

Finally, as $\widehat{\theta}_{i,1} \uparrow \theta_{i,1}$, we observe that $\theta_c \downarrow \theta_{i,1}$, and so we obtain the desired expressions. ■

Consider now the following upward deviation: Firm i with type $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$ mimics $\widehat{\theta}_{i,1}$ slightly above $\theta_{i,1}$ (i.e., chooses $\rho(\widehat{\theta}_{i,1}) > \rho(\theta_{i,1})$), and then (i) if $\widetilde{\theta}_{j,1} < \theta_b$, firm i makes no first-period sale and exits (e.g., prices above r) in all future periods; (ii). if $\widetilde{\theta}_{j,1} \in (\theta_b, \widehat{\theta}_{i,1})$, firm i makes no first-period sale, undercuts the rival's price $\beta(\widetilde{\theta}_{j,1}, \widehat{\theta}_{i,1})$ in the second period, and then exits (e.g., prices above r) in all future periods; and (iii). if $\widetilde{\theta}_{j,1} > \widehat{\theta}_{i,1}$, firm i makes the first-period sale and mimics thereafter the equilibrium pricing behavior of type $\widehat{\theta}_{i,1}$ (i.e, sets $\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1})$ in all future periods). As $\widehat{\theta}_{i,1} \downarrow \theta_{i,1}$, the deviating firm i's payoff approaches that which it earns in the putative equilibrium. Thus, a necessary feature of a productive efficient equilibrium is that firm i does better by announcing $\widehat{\theta}_{i,1} = \theta_{i,1}$ than any other $\widehat{\theta}_{i,1} > \theta_{i,1}$, given the associated strategies described in (i)-(iii) above.

Under (i)-(iii), the profit from an upward deviation is defined as $\pi^U(\widehat{\theta}_{i,1}, \theta_{i,1}) =$

$$[\rho(\widehat{\theta}_{i,1}) - \theta_{i,1}][1 - F_0(\widehat{\theta}_{i,1})] + \delta\int_{\theta_b}^{\widehat{\theta}_{i,1}}[\beta(\widetilde{\theta}_{j,1}, \widehat{\theta}_{i,1}) - \theta_{i,1}]dF_0(\widetilde{\theta}_{j,1}) + \frac{\delta}{1-\delta}\int_{\widehat{\theta}_{i,1}}^{\overline{\theta}}[\beta(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1}) - \theta_{i,1}]dF_0(\widetilde{\theta}_{j,1}) \tag{3.7}$$

Lemma S3 For any $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$, if derivatives are evaluated as $\widehat{\theta}_{i,1} \downarrow \theta_{i,1}$,

$$\pi^U_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = [\rho(\theta_{i,1}) - \theta_{i,1}][-F_0'(\theta_{i,1})] + [1 - F_0(\theta_{i,1})]\rho'(\theta_{i,1}) + \frac{\delta}{1-\delta}\int_{\theta_{i,1}}^{\overline{\theta}}\beta_{\theta_w}(\theta_{i,1}, \widetilde{\theta}_{j,1})dF_0(\widetilde{\theta}_{j,1}) \tag{3.8}$$

$$\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\theta_{i,1}, \theta_{i,1}) = F_0'(\theta_{i,1})[1 + \delta(\frac{\delta}{1-\delta} - \frac{\beta_{\theta_l}(\theta_{i,1}, \theta_{i,1})}{\beta_{\theta_w}(\theta_{i,1}, \theta_{i,1})})] \tag{3.9}$$

Proof: Using (3.7), the definition of θ_b, and $\beta(\widehat{\theta}_{i,1}, \widehat{\theta}_{i,1}) = \widehat{\theta}_{i,1}$, we find that

$$\begin{aligned}\pi^U_{\widehat{\theta}_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) \quad = \quad & [\rho(\widehat{\theta}_{i,1}) - \theta_{i,1}][-F_0'(\widehat{\theta}_{i,1})] + [1 - F_0(\widehat{\theta}_{i,1})]\rho'(\widehat{\theta}_{i,1}) - [\widehat{\theta}_{i,1} - \theta_{i,1}]F_0'(\widehat{\theta}_{i,1})\frac{\delta^2}{1-\delta} \\ & +\delta\int_{\theta_b}^{\widehat{\theta}_{i,1}}\beta_{\theta_l}(\widetilde{\theta}_{j,1}, \widehat{\theta}_{i,1})dF_0(\widetilde{\theta}_{j,1}) + \frac{\delta}{1-\delta}\int_{\widehat{\theta}_{i,1}}^{\overline{\theta}}\beta_{\theta_w}(\widehat{\theta}_{i,1}, \widetilde{\theta}_{j,1})dF_0(\widetilde{\theta}_{j,1}).\end{aligned}$$

Differentiating with respect to $\theta_{i,1}$ and using $\partial\theta_b/\partial\theta_{i,1} = 1/\beta_{\theta_w}(\theta_b, \widehat{\theta}_{i,1})$, we obtain

$$\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) = F_0'(\widehat{\theta}_{i,1})[1 + \frac{\delta^2}{1-\delta}] - \delta\frac{\beta_{\theta_l}(\theta_b, \widehat{\theta}_{i,1})}{\beta_{\theta_w}(\theta_b, \widehat{\theta}_{i,1})}F_0'(\theta_b). \tag{3.10}$$

Finally, as $\widehat{\theta}_{i,1} \downarrow \theta_{i,1}$, we observe that $\theta_b \uparrow \theta_{i,1}$, and so we obtain the desired expressions. ■

We now report two corollaries:

Corollary S1 For any $\theta_{i,1} \in (\underline{\theta}, \overline{\theta})$, $\pi^D_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = \pi^U_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) =$

$$[\rho(\theta_{i,1}) - \theta_{i,1}][-F_0'(\theta_{i,1})] + [1 - F_0(\theta_{i,1})]\rho'(\theta_{i,1}) + \frac{\delta}{1-\delta} \int_{\theta_{i,1}}^{\overline{\theta}} \beta_{\theta_w}(\theta_{i,1}, \widetilde{\theta}_{j,1}) dF_0(\widetilde{\theta}_{j,1}).$$

Corollary S2 Suppose that

$$\frac{\beta_{\theta_w}(\theta_w, \theta_l)}{\beta_{\theta_l}(\theta_w, \theta_l)} = 1 - \delta. \tag{3.11}$$

Then

$$\pi^D_{\widehat{\theta}_{i,1}\theta i,1}(\theta_{i,1}, \theta_{i,1}) = \pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\theta_{i,1}, \theta_{i,1}) = F_0'(\theta_{i,1})[1 - \delta] > 0, \tag{3.12}$$

$$\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) = F_0'(\widehat{\theta}_{i,1}) - \delta F_0'(\theta_c), \tag{3.13}$$

$$\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1}) = F_0'(\widehat{\theta}_{i,1})[1 + \frac{\delta^2}{1-\delta}] - \frac{\delta}{1-\delta} F_0'(\theta_b) \tag{3.14}$$

The corollaries follow directly from Lemmas S2 and S3 and expressions (3.6) and (3.10). The latter corollary motivates the specification for $\beta(\theta_w, \theta_l)$ in Proposition 6, which satisfies monotonicity and (3.11).

We now confirm that the pricing functions specified in Proposition 6 satisfy local incentive compatibility, with respect to our two deviation candidates. Define

$$\pi(\widehat{\theta}_{i,1}, \theta_{i,1}) = 1_{\{\widehat{\theta}_{i,1} \leq \theta_{i,1}\}} \pi^D(\widehat{\theta}_{i,1}, \theta_{i,1}) + 1_{\{\widehat{\theta}_{i,1} > \theta_{i,1}\}} \pi^U(\widehat{\theta}_{i,1}, \theta_{i,1}). \tag{3.15}$$

Since $\pi^D_{\widehat{\theta}_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1})$ and $\pi^U_{\widehat{\theta}_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1})$ exist everywhere, and $\pi^D_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = \pi^U_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1})$ (as shown in Corollary S1), it follows that $\pi_{\widehat{\theta}_{i,1}}(\widehat{\theta}_{i,1}, \theta_{i,1})$ exists everywhere. Imposing the specification for $\beta(\theta_w, \theta_l)$ in Proposition 6, we may use Corollary S1 to find that local incentive compatibility holds if and only if $\pi^D_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = \pi^U_{\widehat{\theta}_{i,1}}(\theta_{i,1}, \theta_{i,1}) = 0$, or equivalently

$$[\rho(\theta_{i,1}) - \theta_{i,1}][-F_0'(\theta_{i,1})] + [1 - F_0(\theta_{i,1})]\rho'(\theta_{i,1}) + \frac{\delta}{2-\delta}[1 - F_0(\theta_{i,1})] = 0 \tag{3.16}$$

Thus, we can characterize the first-period pricing function that achieves local incentive compatibility by

$$\rho(\overline{\theta}) = \overline{\theta} \tag{3.17}$$

$$\rho'(\theta_{i,1}) = \frac{F_0'(\theta_{i,1})}{1 - F_0(\theta_{i,1})}(\rho(\theta_{i,1}) - \theta_{i,1}) - \frac{\delta}{2-\delta} \tag{3.18}$$

It is now straightforward to verify that the first-period pricing function specified in Proposition 6 solves (3.17) and (3.18).

We next confirm that the specified pricing functions satisfy global incentive compatibility, with respect to our two deviation candidates. As established in Corollary S2, $\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ and $\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ exist everywhere and $\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\theta_{i,1},\theta_{i,1}) = \pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\theta_{i,1},\theta_{i,1})$ for the $\beta(\theta_w,\theta_l)$ function that we specify. It follows that $\pi_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ exists everywhere as well. Now consider the sign of $\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ for $\widehat{\theta}_{i,1} < \theta_{i,1}$. Using (3.13), we see that $\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ is positive if $f_0(\widehat{\theta}_{i,1})/f_0(\theta_c) > \delta$. Since $\theta_c > \widehat{\theta}_{i,1}$, we may draw the following conclusion: given $\beta(\theta_w,\theta_l)$ is specified as in Proposition 6, for every $\widehat{\theta}_{i,1} < \theta_{i,1}$, $\pi^D_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1}) > 0$ if the second inequality in (3.1) holds. Next, consider the sign of $\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ for $\widehat{\theta}_{i,1} > \theta_{i,1}$. Using (3.14), we see that $\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ is positive if $f_0(\widehat{\theta}_{i,1})/f_0(\theta_b) > \delta/[1-\delta(1-\delta)]$. Since $\theta_b < \widehat{\theta}_{i,1}$, we may draw the following conclusion: given $\beta(\theta_w,\theta_l)$ is specified as in Proposition 6, for every $\widehat{\theta}_{i,1} > \theta_{i,1}$, $\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1}) > 0$ if the first inequality in (3.1) holds. Thus, under (3.1), $\pi_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1})$ is positive everywhere. Then, standard arguments can be used to show that local incentive compatibility implies global incentive compatibility.[1]

Next, we determine conditions under which the first-period pricing function is strictly increasing. Differentiating the first-period pricing function specified in Proposition 6, we may confirm that $\rho'(\theta_{i,1}) > 0$ if (3.2) holds.

Guided by the foregoing, we may specify a separating equilibrium with productive efficiency, when (3.1) and (3.2) hold. Along the equilibrium path, firms use the pricing strategies specified in Proposition 6. Following any history where an off-schedule deviation has been observed, the carrot-stick belief threat punishment is induced. This punishment is characterized in Proposition 7, and it ensures that a firm that undertakes an off-schedule deviation expects to make approximately zero expected profit over the subsequent periods. In the event that firm i undertakes an on-schedule deviation in period 1, we specify that firm i's subsequent behavior is determined as specified in the downward and upward deviation candidates discussed above.

We may now confirm that no deviation is attractive. Clearly, no firm would gain by taking an off-schedule deviation in the first period (i.e., by deviating outside of the range of the first-period pricing function). Likewise, if a firm did not deviate in the first period, then it would not gain by taking an off-schedule deviation in a later period. A losing firm would clearly not gain from undercutting $\beta(\theta_w,\theta_l)$; and a winning firm would not gain from raising price above $\beta(\theta_w,\theta_l)$, since the immediate gain is approximately zero (the future price of the losing firm is $\beta(\theta_w,\theta_l)+\epsilon$) and the induced subsequent profits are also approximately zero. Next, suppose that firm i took an on-schedule deviation in the first period and consider its optimal play in

[1] For $\widehat{\theta}_{i,1} < \theta_{i,1}$, observe that $\pi(\widehat{\theta}_{i,1},\theta_{i,1}) - \pi(\theta_{i,1},\theta_{i,1}) = \pi^D(\widehat{\theta}_{i,1},\theta_{i,1}) - \pi^D(\theta_{i,1},\theta_{i,1}) < 0$, where the inequality follows from standard arguments, given that $\pi^D(\widehat{\theta}_{i,1},\theta_{i,1})$ satisfies local incentive compatibility and positive cross partials. For $\widehat{\theta}_{i,1} > \theta_{i,1}$, the same argument applies, with π^U replacing π^D.

subsequent periods. Under our specification, if firm i takes an off-schedule deviation in a later period, then firm j is induced to follow the carrot-stick belief punishment thereafter. Thus, if firm i takes an on-schedule deviation in period 1, then it can do no better than to follow the behavior prescribed by the downward and upward deviation candidates discussed above in periods 2 and later. This observation, combined with our work above, ensures as well that firm i does not gain from taking an on-schedule deviation in period 1. ■

3.2. More Than Two Firms

It is straightforward to generalize the description of the productive efficiency scheme to more than 2 firms. However, stating sufficient conditions for the scheme to be a PPBE becomes more complex. The pricing strategies can be written:

$$\rho_{i,1}(\theta_{i,1}) = \theta_{i,1} + \frac{2}{2-\delta}\frac{1}{(1-F_0(\theta_{i,1}))^{I-1}}\int_{\widetilde{\theta}_i=\theta_{i,1}}^{\overline{\theta}}(1-F_0(\widetilde{\theta}_i))^{I-1}d\widetilde{\theta}_i.$$

Let $\theta_w = \min_{i\in I}\theta_{i,1}$, while $\theta_l = \max_{A\in 2^I,\ |A|=I-1}\min_{i\in A}\theta_{i,1}$. If firm i is the low-cost firm in period 1, then for all $t > 1$, firm i sets price

$$p_{i,t} = \beta(\theta_w,\theta_l) = \frac{1-\delta}{2-\delta}\theta_w + \frac{1}{2-\delta}\theta_l,$$

while firm $j \neq i$ sets price $p_{j,t} = \beta(\theta_w,\theta_l)+\varepsilon$ for $\varepsilon > 0$.

The sufficient conditions for global incentive compatibility from the $I = 2$ case generalize as follows:

$$\inf_{\theta'_{i,1}:\theta'_{i,1}>\theta''_{i,1}}\frac{\left(1-F_0\left(\theta'_{i,1}\right)\right)^{I-2}f_0(\theta'_{i,1})}{\left(1-F_0\left(\theta''_{i,1}\right)\right)^{I-2}f_0(\theta''_{i,1})} > \frac{\delta}{1-\delta(1-\delta)},\ \text{and}\ \inf_{\theta'_{i,1}:\theta'_{i,1}<\theta''_{i,1}}\frac{\left(1-F_0\left(\theta'_{i,1}\right)\right)^{I-2}f_0(\theta'_{i,1})}{\left(1-F_0\left(\theta''_{i,1}\right)\right)^{I-2}f_0(\theta''_{i,1})} > \delta. \tag{3.19}$$

The second condition is that δ is small enough that, for all $\theta_{i,1}$,

$$\frac{2\,(I-1)\,f_0(\theta_{i,1})}{(1-F_0(\theta_{i,1}))^I}\int_{\widetilde{\theta}_i=\theta_{i,1}}^{\overline{\theta}}(1-F_0(\widetilde{\theta}_i))^{I-1}d\widetilde{\theta}_i > \delta. \tag{3.20}$$

However, the first condition in (3.19) will fail when $I > 2$ (let $\theta'_{i,1} = \bar{\theta}$). In the proof of Proposition 6, the condition was used in establishing global incentive compatibility for upward deviations: that is, we used it to show that type $\theta_{i,1}$ does not want to mimic type $\hat{\theta}_{i,1} > \theta_{i,1}$. However, the condition is stronger than what is needed: if $\pi^U(\widehat{\theta}_{i,1}\theta_{i,1})$ is the payoff from such mimicry, the condition guarantees that $\pi^U_{\widehat{\theta}_{i,1}\theta_{i,1}}(\widehat{\theta}_{i,1},\theta_{i,1}) \geq 0$ for all $\hat{\theta}_{i,1} > \theta_{i,1}$. What is necessary is that $\pi^U(\widehat{\theta}_{i,1},\theta_{i,1}) \leq \pi^U(\theta_{i,1},\theta_{i,1})$ for all $\hat{\theta}_{i,1} > \theta_{i,1}$. This can be verified directly for a particular distribution of types; numerical calculations indicate that for F_0 uniform and $I > 2$, the critical discount factor *below* which global incentive compatibility holds is greater than zero, but diminishes rapidly with I.

4. References

Athey, S., Atkeson, A. and P. Kehoe (2005), "The Optimal Degree of Monetary Policy Discretion," *Econometrica* 73 (5), September,1431-1476.