Online Appendix to
Market Power Screens Willingness-to-Pay

E. Glen Weyl Jean Tirole

June 2011

In this appendix we treat more general policies than in the paper by invoking the revelation principle. This also allows an alternative means of deriving the results in the paper.

1 Incentive Compatibility

Ideally, the social planner would like every innovation yielding social value greater than its cost to be created. The natural solution to this problem, the principle of payment in accordance with product advocated by Pigou (1920), would be to give each innovation a reward equal to the social value it creates. However, given that the planner cannot observe σ and m, he is unable to perfectly implement payment in accordance with product. An entrepreneur of type (σ, m) can pretend to be of another type $(\hat{\sigma}, \hat{m})$ if she cannot be distinguished by observing the demand that entrepreneur generates. We assume *free disposal of demand*: an entrepreneur can freely reduce the demand for her product. Thus, she is able to imitate another entrepreneur if, at the price that other is asked to charge, she would generate at least as much demand.

Formally, type (σ, m) can successfully imitate type $(\hat{\sigma}, \hat{m})$ if and only if

$$\sigma Q\left(a\left(\hat{\sigma}, \widehat{m}\right)\frac{\widehat{m}}{m}\right) \geq \hat{\sigma} Q\left(a\left(\hat{\sigma}, \widehat{m}\right)\right)$$

We will say that the points $(\hat{\sigma}, \hat{m})$ satisfying this with equality lie on (σ, m)'s *imitation frontier given pricing policy* a. This is a sort of "production possibilities frontier" for the entrepreneur. To provide entrepreneurs with incentives to truthfully reveal their type, the social planner must provide at least as great a reward to each entrepreneur type as that she could earn at any other point along her imitation frontier.

Thus, the social planner's program is

$$\max_{\{\tau(\cdot,\cdot), a(\cdot,\cdot)\}} \int_{\{\theta: c<\tau(\sigma,m)\}} \left[\sigma m S(a(\sigma, m)) - c\right] f(\theta) d\theta \tag{1}$$

subject to

$$\sigma Q\left(a\left(\hat{\sigma},\widehat{m}\right)\frac{\widehat{m}}{m}\right)\geq\hat{\sigma}Q\left(a\left(\hat{\sigma},\widehat{m}\right)\right)\implies\tau(\sigma,m)\geq\tau(\hat{\sigma},\hat{m}). \tag{2}$$

1.1 From imitation frontiers to isoreward curves

A standard approach to mechanism design problems is to reduce their often unwieldy global incentive compatibility constraints to local constraints. This "Mirrlees-Rogerson first-order approach" must then be complemented by an ex-post verification of its validity. In this section we develop an "isoreward" approach that acts as an intuitive multi-dimensional extension of the first-order approach.[1] We begin by developing an analogous relaxed approach in this subsection, and then verify it validity in the next subsection.

In doing so we draw an analogy to a neoclassical production economy.[2] We can think of the entrepreneur as "producing" her report by choosing a point along her production possibility frontier (imitation frontier) to maximize her reward. Thus, at the optimal report, the marginal rate at which the entrepreneur can *transform* $\hat{\sigma}$ into $\hat{m}$ must be equated to the marginal rate at which the social planner rewards m relative to σ, as this is the entrepreneur's marginal rate of substitution between m and σ. This marginal rate of transformation is $-\frac{d\hat{\sigma}}{d\hat{m}}$, defined by implicit differentiation of the imitation frontier (her production possibility frontier) at $(\hat{\sigma},\hat{m})=(\sigma,m)$, which after some algebraic manipulations yields the simple formula,

$$-\frac{d\hat{\sigma}}{d\hat{m}}\frac{\hat{m}}{\hat{\sigma}}=\epsilon\left(a(\sigma,m)\right)\qquad\text{at }(\hat{\sigma},\hat{m})=(\sigma,m) \tag{3}$$

Thus, local to the truth, a one-percent increase in $\hat{m}$ requires a sacrifice of ϵ of a percent of $\hat{\sigma}$, as raising $\hat{m}$ by one percent forces the entrepreneur to raise prices (locally) by one percent. Thus, crucially, an increase in $\hat{m}$ requires a sacrifice of $\hat{\sigma}$ only to the extent that a is large, as ϵ increases from 0 to 1 as a does.[3]

[1] Our extension is, like Milgrom and Segal (2002)'s extension of the Mirrlees-Rogerson approach, robust to the possibility of non-differentiable mechanisms.

[2] The problem of incentive compatibility in our context can be seen as equivalent to that of market equilibrium in Rosen (1974)'s model of hedonic pricing in which every product exists (hence first-order conditions for its production by the most efficient producer, our incentive compatibility constraint, must be satisfied). However, our solution method via (in his context) isoprice curves, an alternative interpretation of his partial differential equations, has not, to our knowledge, been applied and might aid in the solution of such models.

[3] Finite demand and zero marginal distortion at the ex-post efficient price imply that $\epsilon(0)=0$ and monopoly optimization that $\epsilon(1)=1$. Furthermore,

$$MR=p-\frac{p}{\epsilon}\propto a-\frac{a}{\epsilon(a)}$$

so

$$Q'(MR)'>0\iff\left[a\left(1-\frac{1}{\epsilon}\right)\right]'<0\iff\epsilon'>-\frac{\epsilon}{a}\left(1-\frac{1}{\epsilon}\right)\iff\epsilon'>-\frac{\epsilon^2}{a}\left(1-\frac{1}{\epsilon}\right)$$

So, for marginal revenue to be decreasing, $\epsilon'>0$ whenever $\epsilon\leq1$.

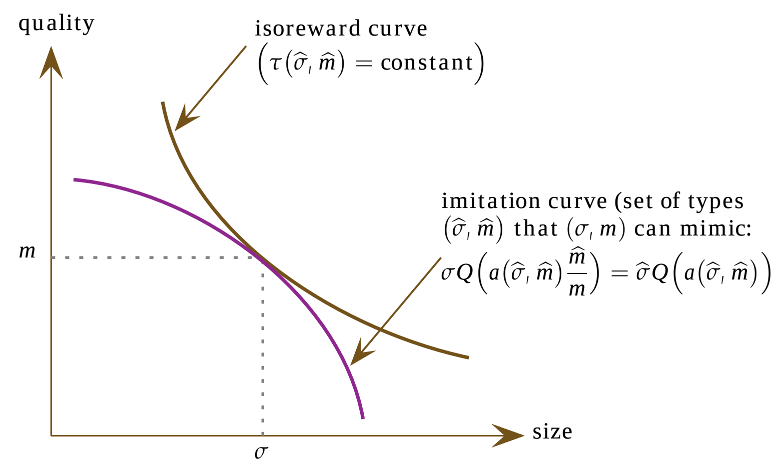

Figure 1: Isoreward curves must be tangent to imitation frontiers at the truthful point

As pictured in Figure 1, the relative marginal rate of substitution, $-\frac{\tau_m}{\tau_\sigma}$, must be equal to the marginal rate of transformation derived above:

$$\frac{\tau_m}{\tau_\sigma} = \frac{\sigma}{m}\epsilon\left(a(\sigma, m)\right). \tag{4}$$

That is, the local reward given to m relative to σ is proportional to the elasticity of the demand curve for the innovation at the prevailing price, which increases as price increases.

Because the tangency conditions above must hold *for every* (σ, m) *pair*, and thus the marginal rate of substitution is everywhere proportional to $\frac{\sigma}{m}$, Condition (4) uniquely traces out a series of *isoreward curves* along which rewards must be constant, as formalized in the corollary below.

Lemma 1: *Under an arbitrary differentiable pricing policy* $a(\cdot, \cdot)$, *incentive compatibility requires that* τ *be weakly monotone in both its arguments and that rewards be constant along any curve* $m = \mathcal{M}(\sigma)$ *obeying*

$$\frac{d\mathcal{M}}{d\sigma} = -\frac{\mathcal{M}(\sigma)}{\sigma\epsilon\left(a\left(\sigma, \mathcal{M}(\sigma)\right)\right)}$$

except at most on a (welfare irrelevant) countable set of such curves.

Proof. The next subsection is dedicated to the proof. □

We index isoreward curves by the point at which they intersect the $45°, \sigma = m$ line. We refer to this point as k. Given free disposal, the reward clearly must be increasing in k.

1.2 Necessary conditions for incentive compatibility

Nota bene: the proof in this section was written jointly by Weyl and Michal Fabinger. We are very grateful to Michal and these proofs, along with most of the rest of the material in this online appendix, will be spun off into a separate paper (Fabinger and Weyl, 2011). For this reason, this subsection provides more of a sketch of a proof of Lemma 1 than a fully detailed proof.

Our strategy for establishing Lemma 1, using a series of sublemmata, consists of five steps. In describing them, we use the terminology "conjectured isoreward curve" to refer to curves defined by the differential equations in the lemma statement and "actual isoreward curve" to describe a curve along which τ must in fact be constant. We also repeatedly rely on the fact that conjectured isoreward curves pass through each point in the (σ, m) plane, smoothly deform, do not intersect,and intersect the 45° line at exactly one point. These are classical results on the solutions to classes of first-order ordinary differential equations of the form $\mathcal{M}'(\sigma; \sigma, m) = -\frac{\mathcal{M}(\sigma)}{\sigma\epsilon(a(\sigma,\mathcal{M}(\sigma)))}$ where $\epsilon \geq 0$ and smooth and the curve passes through (σ, m).

1. We extend the classic theorem of Young and Young (1924) to show that the set of discontinuities of a monotone function of several variables can be placed along a countable set of non-increasing curves, which we call *extended discontinuity curves.*

2. We argue that any curve of discontinuities must lie entirely along a conjectured isoreward curve, since if it were to "cut through" an isoreward curve, it would offer a profitable opportunity for deviation.

3. Because each extended discontinuity curve can contain at most a countable number of curves (or almost-curves, curves from which a set of measure zero has been removed) of discontinuity, these lie along at most a countable number of conjectured isoreward curves.

4. We show that any conjectured isoreward curve not including discontinuities must be an actual isoreward curve.

5. Finally, we conclude that the set of conjectured isoreward curves failing to be actual isoreward curves is at most countable and, as such, can be disregarded.

We use a concept that, while in some sense common in economics, does not have a standard name we are aware of and thus we introduce the terminology that follow.

Definition 1: *A* non-increasing curve *is a curve in* $\mathbb{R}^2$ *with the property that if* (x, y) *belongs to the curve and* $x' > x, y' > y$ *then* (x', y') *do not belong to the curve. A* non-increasing almost-curve *is a non-increasing curve from which a set of one-dimensional Lebesgue measure zero has been removed.*

Note that a demand curve is really a non-increasing curve, not a non-increasing function, if it may be perfectly inelastic over some range. For verbal economy we will refer to a point at which τ fails to be continuous in (σ, m) (and not merely those in which it is discontinuous in some direction) as points of discontinuity of τ. Furthermore by countable we mean any set of cardinality less than or equal to that of the integers.

Sublemma 1: *The points of discontinuity of τ are a subset of a countable set of non-increasing curves, which we refer to as* extended discontinuity curves*, through the (σ, m) plane.*

Proof. The proof of this sublemma is fairly involved so we begin with a brief outline:

1. We begin by considering a finite box on the real plane and investigate the maximal number of non-increasing curves needed to accommodate all of its points of discontinuity of a size larger than a fixed amount.

2. Along any non-increasing curve through this box there must be no more than a finite number of discontinuities larger than some size in any direction. In the step that forms the heart of the proof, we show this implies that the number of non-increasing curves needed to accommodate all points of sufficiently large discontinuity is finite.

3. We then take the limit as the size of discontinuities grows small, obtaining a countable number of non-increasing curves for each box.

4. Finally we take the limit as the box grows to encompass the full plane, obtaining a countable union of countable sets of non-increasing curves.

Note that by the monotonicity of τ in σ and m, any discontinuity of τ in a direction from the $(-,-)$ quadrant to the $(+,+)$ quadrant of a point must be a jump discontinuity and therefore have a (supremal) size s (in some such direction). We refer to such discontinuities as *monotone discontinuities of size s*. Let us refer to the set of all monotone discontinuities of size great than δ contained in closed box $[0, N] \times [0, N]$ as $D(N, \delta)$ and to its closure as $\overline{D}(N, \delta)$.

We will construct a set of no more than $\left\lfloor \frac{\tau(N,N)}{\delta} \right\rfloor$ (where $\lfloor x \rfloor$ represents the greatest integer function) non-increasing curves, the union of which contains $\overline{D}(N, \delta)$. We do so inductively and thus define the base inductive case $\overline{D}_0(N, \delta) \equiv \overline{D}(N, \delta)$.

Let the *weakly undominating set* of $\overline{D}_i(N, \delta)$,

$$U\left(\overline{D}_i\left(N, \delta\right)\right) \equiv \left\{(\sigma, m) \in \overline{D}_i(N, \delta) : \nexists (\sigma', m') \in \overline{D}_i(N, \delta) s.t. (\sigma', m') \ll (\sigma, m)\right\}$$

Note that $U\left(\overline{D}_i\left(N, \delta\right)\right)$ is a closed subset of a non-increasing curve:

1. $U\left(\overline{D}_i\left(N, \delta\right)\right)$ is closed because it is defined by the failure of strict inequalities among elements of a closed set.

2. It is (at most) one-dimensional as any two-dimensional closed subset of $\mathbb{R}^2$ contains a interior point and thus dominates another point.

3. It may be chosen to be non-increasing because it is undominating.

Thus $U\left(\overline{D}_i\left(N,\delta\right)\right)$ is either connected, and thus forms a non-increasing curve itself, or there is a non-increasing curve, of which it is a subset, which connects it. While there may be many such curves, choose one and note that because the curve is non-increasing and every point in $U\left(\overline{D}_i\left(N,\delta\right)\right)$ is undominating, this curve may also be chosen to be undominating. Call this *extended* curve $\mathfrak{c}_{i+1}(N,\delta)$, define $\overline{D}_{i+1}(N,\delta)$ as the closure of $\overline{D}_i(N,\delta)\backslash\mathfrak{c}_{i+1}(N,\delta)$ and repeat this process unless $\overline{D}_{i+1}(N,\delta)=\emptyset$.

$\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor}(N,\delta)$ must be empty. To see this, suppose that this were not the case. Then there would be some point $(\sigma,m)\in\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor}(N,\delta)$ which is actually in $D(N,\delta)$, not merely its closure, as only closed sets are removed at each step and thus mere limit points will always be removed. Therefore $(\sigma,m)\notin U\left(\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor-1}(N,\delta)\right)$ and thus strictly dominates some point in $\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor-1}(N,\delta)$. This point again cannot be merely a limit point but must actually be an element of $D(N,\delta)$ as dominance is strict. But this point in turn must strictly dominate a point of $\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor-2}(N,\delta)\cap D(N,\delta)$ and so on. Thus, by the transitivity of strict dominance, (σ,m) lies atop a hierarchy of (at least) $\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor$ strict dominance relations among points in the box.

However, because the dominance is strict, it is possible to draw a non-decreasing curve (analogous to a non-increasing curve) between these points hitting each at from any direction from the $(-,-)$ to the $(+,+)$ quadrant. Thus, by the monotonicity of τ, $\tau(N,N)\geq\left(\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor+1\right)\delta>\tau(N,N)$ which is a contradiction. Thus $\overline{D}_{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor}(N,\delta)$ is in fact empty.

Thus $\overline{D}(N,\delta)\subset\bigcup_{i=1}^{\left\lfloor\frac{\tau(N,N)}{\delta}\right\rfloor}\mathfrak{c}_i(N,\delta)$ as desired. Thus clearly the set of all monotone discontinuities τ (of any size) is a subset of $\lim_{N\to\infty}\bigcup_{j=1}^{N}\bigcup_{i=1}^{\lfloor N\tau(N,N)\rfloor}\mathfrak{c}_i\left(N,\frac{1}{N}\right)$. But by Theorem 4 of Young and Young (1924), any point at which a monotone function of two variables is continuous in all directions from the $(-,-)$ to $(+,+)$ quadrants it is continuous from all directions. Thus the set of discontinuities of τ is a subset of $\lim_{N\to\infty}\bigcup_{j=1}^{N}\bigcup_{i=1}^{\lfloor N\tau(N,N)\rfloor}\mathfrak{c}_i\left(N,\frac{1}{N}\right)$, a countable union of non-increasing extended discontinuity curves. □

Sublemma 2: *Every non-increasing curve or almost-curve of discontinuities of τ is a subset of some conjectured isoreward curve.*

Proof. If the discontinuity curve consists of a point, it clearly lies along a conjectured isoreward curve. For non-point curves and almost-curves, our proof strategy for this sublemma is again quite intricate so we again outline it before diving in:

1. We begin by investigating discontinuity curves (rather than almost-curves) in a finite box consisting entirely of discontinuities of size at least δ. In particular we focus on a single curve lying along the highest extended discontinuity curve, to avoid any interference by other curves of discontinuity.

2. We use an inductive argument to show that no curve of discontinuity of at least size δ may "cut through" a conjectured isoreward curve as a series of local dominance relationships of the

smooth imitation frontiers might then be established which would force τ to take an infinite value in this finite region.

3. Because the discontinuity curve may not cut through, it must either be locally differentiable or kinked (in a Dini derivative sense). But if kinked a local value can be found which will cut through a sufficiently close isoreward curve, contradicting the prior step and establishing the differentiability of the discontinuity curve.

4. Any differentiable discontinuity curve which cannot cut through conjectured isoreward curves must be everywhere tangent to any such curve it intersects, determining a differential equation "pinning" the discontinuity curves to the conjectured isoreward curves.

5. The argument extends without any trouble to almost-curves, then to other extended discontinuity curves in the same finite region and then over all such curves as the region grows large and increment size small.

First we define the notion of tangency we are interested in, using the standard notation for Dini derivatives: D_-, D^-, D_+, D^+ represent respectively the lower left, upper left, lower right and upper right Dini derivatives. For a treatment of Dini derivatives, which are not used commonly in economics, see for example Royden (1988). We say that a non-increasing curve is *tangent* to a conjectured isoreward curve at point (σ, m) if

1. The discontinuity curve is a function $\check{m}(\check{\sigma})$ in a neighborhood about (σ, m).

2. The discontinuity curve does not cut the conjectured isoreward curve from above at (σ, m): $D_-\check{m}(\sigma) < \mathcal{M}'(\sigma; \sigma, m) \implies D_+\check{m}(\sigma) \geq \mathcal{M}'(\sigma; \sigma, m)$.

3. The discontinuity curve does not cut the conjectured isoreward curve from below at (σ, m): $D^-\check{m}(\sigma) > \mathcal{M}'(\sigma; \sigma, m) \implies D^+\check{m}(\sigma) \leq \mathcal{M}'(\sigma; \sigma, m)$.

Consider $D\left(N, \frac{1}{N}\right)$ as defined above and its highest extended discontinuity curve and any non-point curve in this extended discontinuity curve, if such exists. Choose any interior point of the curve; this is some (σ, m) corresponding to an isoreward and imitation frontier. Suppose that the discontinuity curve fails to be tangent to this conjectured isoreward curve at (σ, m). There are two possibilities. Either the curve is locally a correspondence or it is a continuous function whose upper and lower Dini derivatives from each side exist, but they fail to obey the specified bounds. We focus on the second case to begin with, as the argument in the first case is essentially a special case of the second. Furthermore, we focus on showing the impossibility of a cut from above in the case when τ is continuous at (σ, m) from the $(+, +)$ quadrant rather than the $(-, -)$ quadrant (it must be from one by monotonicity), as the argument for a contradiction in all other cases is perfectly analogous. Thus we assume that $D_+\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m)$ and that $D_-\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m)$, seeking a contradiction.

By differentiability and tangency to the conjectured isoreward curve of the imitation frontier anchored at (σ,m), if $D_+\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m)$, there exists a region to the southeast of (σ,m) above $\check{m}$ and below the imitation frontier $\hat{m}(\sigma,m)$. By the arguments from the proof of Sublemma 1, $\tau \geq \frac{1}{N}$ in this region. But, by continuity of the imitation frontier in its parameters for any (σ',m') along the conjectured isoreward curve $\check{m}$ sufficiently close to (σ,m), the imitation frontier $\hat{m}(\sigma',m')$ contain points dominating points in his region. Thus by incentive compatibility, $\tau(\sigma',m') \geq \frac{1}{N}$.

However, again by continuity and monotonicity of the imitation frontiers in the (σ,m) at which they are anchored, if we chose any $(\sigma'',m'') > (\sigma,m)$ interior to $[0,N]\times[0,N]$, which is possible because (σ,m) was constructed as an interior point, there is some point of the form (σ',m') described in the preceding paragraph dominated by a neighborhood of (σ'',m''). Furthermore because $D_-\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m)$, by the exact argument of the preceding paragraph, some such point lies below $\check{m}$ and that there is some region, in the dominating neighborhood of (σ'',m'') lying above $\check{m}$ and below $\mathcal{M}(\sigma'',m'')$. By the argument above, τ in this region is at least $\frac{1}{N}$ greater than $\tau(\sigma',m')$ which it dominates through $\check{m}$. Thus τ must be at least $\frac{2}{N}$ in this region. Thus we can iterate the argument to show that there is some point, interior to $[0,N]\times[0,N]$ on which τ takes arbitrarily high values, a contradiction.

To prove that $\check{m}$ may not be vertical at (σ,m) follows precisely the same logic and to prove it cannot cut below follows the same logic in reverse (starting from the left, then moving to the right, moving down rather than up across conjectured isoreward curves).

Therefore $\check{m}$ must be tangent to any isoreward curve it intersects. Thus either $D_-\check{m}(\sigma) \geq \mathcal{M}'(\sigma;\sigma,m)$ or $D_+\check{m}(\sigma) \geq \mathcal{M}'(\sigma;\sigma,m)$ and either $D^-\check{m}(\sigma) \leq \mathcal{M}'(\sigma;\sigma,m)$ or $D^+\check{m}(\sigma) \leq \mathcal{M}'(\sigma;\sigma,m)$. We now wish to establish that at all σ, $\check{m}$ is differentiable and $\check{m}'(\sigma) = \mathcal{M}'(\sigma;\sigma,m)$. This is equivalent to

$$D_-\check{m}(\sigma) = D^-\check{m}(\sigma) = D_+\check{m}(\sigma) = D^+\check{m}(\sigma) = \mathcal{M}'(\sigma;\sigma,m)$$

This may fail, for example, if $D_-\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m)$. If so, $\exists\delta > 0 : D_-\check{m}(\sigma) < \mathcal{M}'(\sigma;\sigma,m) - \delta$. A well-known property of Dini derivatives[4] is their *potential continuity*: $\forall\eta > 0, \exists\sigma^\star \in [\sigma-\eta,\sigma]$: $D_-\check{m}(\sigma^\star), D_+\check{m}(\sigma^\star) < D_-\check{m}(\sigma) + \eta$. But by the continuous differentiability of ϵ and a and continuity of $\check{m}$, for η sufficiently small $\mathcal{M}'(\sigma^\star;\sigma^\star,\check{m}(\sigma^\star))$ is arbitrarily close to $\mathcal{M}'(\sigma;\sigma,m)$. Thus $D_-\check{m}(\sigma^\star), D_+\check{m}(\sigma^\star) < \mathcal{M}'(\sigma^\star;\sigma^\star,\check{m}(\sigma^\star))$ and $\check{m}$ is not tangent to the isoreward curve passing through $(\sigma^\star,\check{m}(\sigma^\star))$. But this contradicts our above reasoning. All other ways in which $\check{m}$ may fail to be differentiable can be ruled out similarly. Thus $\check{m}(\sigma)$ is differentiable at all σ in its support and

$$\check{m}'(\sigma) = \mathcal{M}'(\sigma;\sigma,m)$$

[4]See, for example, Hagood and Thomson (2006).

an ordinary differential equation with a unique solution passing through any point, corresponding to the conjectured isoreward curve passing through that point. Thus $\check{m}$ lies entirely along a single conjectured isoreward curve.

This argument may be repeated first for all other curves lying along the extended discontinuity curve in consideration, showing each of these lies along a conjectured isoreward curve. Furthermore it is straightforward to demonstrate that the same argument applies to almost-curves as to curves: removing a set of measure zero does not interfere with any of the steps as proper closure was applied in all of our arguments above.

Moving to the next extended discontinuity curve down, we may now ignore all points in $D\left(N, \frac{1}{N}\right)$, as these dominate all points in this curve and we can thus consider the argument over a sufficiently small box so as to exclude these points. The same argument can then be repeated to show all curves along this, and inductively all other, extended discontinuity curves containing $D\left(N, \frac{1}{N}\right)$ lie entirely along conjectured isoreward curves.

But of course this argument hold for any N and thus, by the same limiting argument as in the proof of Sublemma 1, establishes that all curves and almost-curves of discontinuities lie each entirely along a single conjectured isoreward curve.

□

Sublemma 3: *At most a countable number of conjectured isoreward curves contain discontinuities of τ.*

Proof. It is a classic real analysis result that every subset of the real line is dense in a countable union of closed intervals. Thus, because any non-increasing curve is uni-dimensional and therefore isomorphic to (a subset of) the real line, the points of discontinuity along every extended discontinuity curve from Sublemma (1) is a countable union of non-increasing curves or almost-curves. But by Sublemma 2 each of these intersects at most a single isoreward curve. Thus the set of isoreward curves containing discontinuities has the cardinality of non-increasing curves of discontinuity, which are a countable union of countable sets and therefore countable. □

Sublemma 4: *Suppose that τ is continuous in (σ, m) at every point along a conjectured isoreward curve. Then this is an actual isoreward curve.*

Proof. Our proof proceeds in three steps:

1. We argue that along any conjectured isoreward curve where τ is continuous, the upper Dini derivative from the direction below and perpendicular to the conjectured isoreward curve is continuous along that curve.

2. We show that if this Dini derivative is bounded from above at a point, then the conjectured isoreward curve is constant in the neighborhood of that point.

3. We show that arbitrarily close to any conjectured isoreward curve from below there must exist another conjectured isoreward curve along which all points have finite upper Dini derivatives and, therefore, that this other conjectured isoreward curve is an actual isoreward curve.

4. Finally, by continuity, we conclude that the conjectured isoreward curve is, in fact, an isoreward curve.

We begin by establishing the continuity of Dini derivatives along conjectured isoreward curves at which τ is continuous. Consider a point (σ, m). We would like to show that the upper Dini derivative of τ at (σ, m) from below in the direction $\left(1, -\frac{1}{\mathcal{M}'(\sigma;\sigma,m)}\right)$ is continuous along the conjectured isoreward curve. We will denote this derivative by $\tilde{D}^{-}$. To do this we begin by establishing bounds on these Dini derivatives.

To place a lower bound on $D^{-}\tau\left(\sigma', \mathcal{M}\left(\sigma';\sigma,m\right)\right)$ we must, for any distance we are challenged with, find a point closer than that distance to $\left(\sigma', \mathcal{M}\left(\sigma';\sigma,m\right)\right)$ in the lower perpendicular direction to the isoreward curve which is sufficiently lower than $\tau\left(\sigma', \mathcal{M}\left(\sigma';\sigma,m\right)\right)$. To see that this is possible, note that points very close in the negative perpendicular direction to (σ, m) will be dominated by, and therefore with a lower value than, $\left(\sigma', \mathcal{M}\left(\sigma';\sigma,m\right)\right)$; by continuity, such a point, jointly with (σ', m') may be chosen so that the value of τ is arbitrarily close to $\tau(\sigma, m)$. Furthermore the definition of the lower perpendicular Dini derivative at (σ, m) implies that a close point to on the lower perpendicular of (σ, m) may be chosen to be dominated by a slightly lower point on the lower perpendicular of $(\sigma', \mathcal{M}(\sigma';\sigma,m))$ and still have a sufficiently low value of τ. This ensures that the desired point on the lower perpendicular of $(\sigma', \mathcal{M}(\sigma';\sigma,m))$ does in fact exist and establishes the lower bound.

By repeating the same argument, but reversing the choice of point dominance we obtain the opposite inequality: $\forall\delta > 0$, $\exists\eta : \forall\sigma' \in [\sigma, \sigma+\eta]$, $\tilde{D}^{-}\tau(\sigma, m) - \delta \geq \tilde{D}^{-}\tau\left(\sigma', \mathcal{M}\left(\sigma';\sigma,m\right)\right)$. Combining these two implies the continuity of the lower Dini derivatives along the conjectured isoreward curve.

Next we argue that Dini derivatives being bounded above at a point from below along the local perpendicular implies the conjectured isoreward curve being constant in that neighborhood. Note that this hypothesis is equivalent, by the continuity of the Dini derivatives shown above to the Dini derivative being bounded in a neighborhood about this point.

Suppose this were false. Then there must be two points, (σ, m) and (σ', m') lying on the same conjectured isoreward curve in a neighborhood of bounded upper Dini derivative; take this upper bound to be M. Suppose we consider another conjectured isoreward curve that is sufficiently close (δ^2 along the local perpendicular) to the original one in question. Then by differentiability of the conjectured isoreward curves and tangency of the imitation frontier, the local imitation frontier will intersect the slightly lower conjectured isoreward curve at a distance of order δ along the isoreward curve (if a local perpendicular is drawn). The value of τ at this point must be no greater than $\tau(\sigma, m)$ by incentive compatibility. Furthermore, the value at the point along the original conjectured isoreward curve must not exceed $\tau(\sigma, m)$ by more than $M\delta^2$, by the upper bound on the upper Dini derivative. Iterating this argument one finds that if the distance between (σ, m) and (σ', m') is less than d, then $\tau(\sigma', m') < \tau(\sigma, m) + d\delta$. Because δ maybe chosen arbitrarily small we obtain that $\tau(\sigma', m') \leq \tau(\sigma, m)$.

The argument may be repeated in the opposite direction to show that $\tau(\sigma', m') \geq \tau(\sigma, m)$ and thus it must be that $\tau(\sigma', m') = \tau(\sigma, m)$. Thus any neighborhood of bounded upper Dini derivatives from below along perpendicular must have a locally constant τ value along the conjectured isoreward curve.

Now consider the actual value of these upper Dini derivatives from below along the local perpendicular. Note that for the perpendicular emanating from any point along the conjectured isoreward curve, the upper Dini derivative from below along this curve cannot be infinite, in any neighborhood of the curve, over any set of positive measure; otherwise the value of the curve would, by monotonicity, be infinite at some finite point. Thus along each perpendicular at most a set of measure zero has points where the upper Dini derivative from below is infinite.

Choose a countable, dense set of points in a neighborhood about (σ, m) and consider the perpendiculars emanating from these points. The set of conjectured isoreward curves passing through as point of infinite Dini derivative from below along any of these perpendiculars must be of measure zero, as a countable union of sets of measure zero is of measure zero. So must be the set of these conjectured isoreward curves that contain discontinuities of τ by Sublemma 3. Therefore we can always find a "normative" conjectured isoreward curve avoiding both of these sins and arbitrarily close to the original conjectured isoreward curve of interest.

This normative curve must have uniformly bounded upper Dini derivatives along the diagonal in the neighborhood of each point in the countable dense set. It must therefore be constant over these sets. Furthermore because it may be chosen arbitrarily close to the original conjectured isoreward curve, at which τ is continuous, this original curve must have constant τ over this neighborhood as well.

Putting this all together, we have established that each conjectured isoreward curve, if it contains no discontinuities of τ, must be constant in the neighborhood of each point along the curve. But clearly this establishes constancy over the whole curve. □

Proof of Lemma 1. By Sublemma 3 and Sublemma 4, the set of conjectured isoreward curves which fail to be actual isoreward curves is countable. Now suppose some $\tau^{\star\star}$ was incentive compatible, but violated the isoreward property on some countable set of pathological conjectured isoreward curves. Then consider another $\tau^\star$ which matches the values of $\tau^{\star\star}$ on the non-pathological conjectured isoreward curves but assigned to all point of the pathological isoreward curves any value taken by $\tau^{\star\star}$ along these pathological isoreward curves. Because f is continuous and has all finite moments social welfare is exactly the same under $\tau^{\star\star}$ as under $\tau^\star$ as changing the value of a function along a countable set does not alter its Riemann-Stieljes or Lebesgue integral. And clearly $\tau^\star$ satisfies the isoreward property along all conjectured isoreward curves. Thus *some* optimal τ always satisfies the isoreward property of the Lemma. $\square$

1.3 Sufficient conditions for incentive compatibility

Suppose that for every (q, p) pair there is *some* entrepreneur whom we would like to produce q and charge p.[5] By free disposal, t must be increasing in (q, p) for this to be feasible. If entrepreneur (σ, m) is to choose a quantity price pair $\left(\sigma Q\left(\frac{p}{m}\right), p\right)$ then the isoquant (isoreward curve) of $t(q, p)$ must be tangent to (σ, m)'s demand curve at $\left(\sigma Q\left(\frac{p}{m}\right), p\right)$.

Suppose we wish to implement some continuous mapping $(q, p) : \mathbb{R}^2_+ \to \mathbb{R}^2_+$ which prescribes the quantity $q(\sigma, m)$ and price $p(\sigma, m)$, which any entrepreneur is instructed to produce and charge, respectively. Clearly to be implementable the price-quantity pair must lie along the appropriate demand curve:

$$q(\sigma, m) = \sigma Q\left(\frac{p(\sigma, m)}{m}\right)$$

What else is required of (q, p) to be implementable? Suppose that q were to be decreasing in σ or p decreasing in m. This would effectively ask types with a comparative advantage in producing demand or prices to produce less of these than another type with a comparative disadvantage. The classic logic of Spence (1973) and Mirrlees shows this may not be incentive compatible. Moreover, in two variables individual (weak) monotonicity in each argument is not sufficient (McAfee and McMillan, 1988); if (q, p) is differentiable the Jacobian of the transformation must be positive semidefinite (it cannot flip points across quadrants in coordinate systems based at any point). Such a differentiable mapping is known as a monotone orientation-preserving weak self-diffeomorphism (weak MOPSD) of $\mathbb{R}^2_+$; a strict MOPSD has a positive definite Jacobian.

While this condition is quite imposing as a formal statement, it is really just the most natural two-dimensional generalization of the standard monotonicity condition for one-dimensional implementation. Calculating the Jacobian of the (σ, m) to (q, p) transformation in logs yield

$$\begin{bmatrix} 1 - \epsilon\epsilon_{a_\sigma} & \epsilon_{a_\sigma} \\ -\epsilon\epsilon_{a_m} & 1 + \epsilon_{a_m} \end{bmatrix}.$$

Thus a MOPSD simply requires that $\epsilon\epsilon_{a_\sigma} < 1$ (where ϵ_{a_σ} is the elasticity of a with respect to σ), $\epsilon_{a_m} > -1$ and that $\epsilon\epsilon_{a_\sigma} - \epsilon_{a_m} < 1$. The first condition is that increasing σ should not increase prices so far as to offset the direct increase in quantity this causes, so that q remains monotone in σ. The second condition posits that a should not fall so rapidly in m that p is actually declining in m. The final condition is equivalent to the condition that moving along the local demand function towards

[5]Note that even if this fails it is irrelevant and we can always restore monotonicity, as any time there *is* an entrepreneur at (q, p) the reward there must be greater than at any point dominated by this.

higher prices also moves towards higher m and lower σ. This last is the requirement, in addition to monotonicity, that ensures the preservation of orientation.

Lemma 2: *Suppose that* (q, p) *is a strict MOPSD of* $\mathbb{R}^2_+$ *with* $q(\sigma, m) = \sigma Q\left(\frac{p(\sigma,m)}{m}\right)$*. Then the conditions in Lemma 1 are necessary and sufficient for incentive compatibility. Almost conversely if* (q, p) *is differentiable then any incentive compatible* t *implementing* (q, p) *is constant over any neighborhood where* (q, p) *fails to be MOPSD.*

Thus, if we are willing to assume (q, p) is differentiable, we may use the isoreward approach except when there is bunching (a weak MOPSD or no MOPSD at all with flat t).[6] However, any weak MOPSD is the limit of a series of strict MOPSD and thus little is lost by restricting attention to the latter.

Proof. In the forward direction, necessity follows from the fact that a MOPSD is clearly equivalent to a differentiable pricing policy a: by the nature of a MOPSD is isoreward curves in the (σ, m) space may easily be transformed into those in the (q, p) space.

For sufficiency consider some type (σ, m). Suppose that rather than choosing $(q(\sigma, m), p(\sigma, m))$, (σ, m) strictly prefers $\left(\sigma Q\left(\frac{p'}{m}\right), p'\right)$ where $p' > p(\sigma, m)$; in particular consider the point $\left(\sigma Q\left(\frac{p'}{m}\right), p(\sigma, m)\right)$ and let (σ', m') be its inverse under (q, p) which exists as it is a strict MOPSD. Also, let (σ'', m'') be the inverse of $\left(\sigma Q\left(\frac{p'}{m}\right), p'\right)$. It is well-known that the inverse of a strict MOPSD is itself a strict MOPSD so by orientation-preservation (property 3),

$$(m'' - m')(\sigma - \sigma') > (m - m')(\sigma'' - \sigma') \tag{5}$$

while by monotonicity we have that $m' > m''$ and $\sigma > \sigma'$. However we also know that $\left(\sigma Q\left(\frac{p'}{m}\right), p'\right)$ lies on (σ, m) *and* (σ'', m'')'s (downward sloping) demand curves. So either $\sigma > \sigma''$ and $m < m''$ or $\sigma < \sigma''$ and $m > m''$. Suppose the second were the case; then clearly $\sigma'' > \sigma > \sigma'$ and $m > m'' > m'$ so

$$(m - m')(\sigma'' - \sigma') > (m'' - m')(\sigma - \sigma')$$

in contradiction of inequality (5). Thus we must have $\sigma > \sigma''$ and $m < m''$. But then clearly the elasticity demand for (σ'', m'') at p' is great smaller than that of (σ, m). Thus if type (σ'', m'') is locally indifferent to raising q to increase k, type (σ, m) must be able to strictly raise k (which is strictly monotone by construction). Thus imitating (σ'', m'') cannot be the k-maximizing choice for (σ, m). But this argument may be repeated for any point on the frontier $(\hat{\sigma}, m) \neq (\sigma, m)$ (for points to the southeast the argument is analogous but reversed) proving that the k-maximizing point for (σ, m) is (q, p).

For the partial converse, suppose that (q, p) is not a weak MOPSD. Then, by continuity, there exists a neighborhood where either monotonicity in some direction or orientation-preservation in

[6] Fabinger and Weyl (2011) are working to relax the differentiability and other technical assumptions.

some direction is violated for each pair or triple of relevant points in the neighborhood. This can easily be shown to give rise to the exactly the opposite of the argument above, establishing the local convexity of the objective function of the entrepreneur at a point in the neighborhood and thus establishing that (q, p) is not in fact an optimal choice for her. If t is constant over that range, however, the argument fails and t may clearly still implement the desired function. □

Note that this implies that over regions where t is endogenously optimally flat or near flat it may be optimal to adopt very low prices as anything may be implemented. This is intuitive: if t is not optimally increasing over a range there is little incentive to sort, as which isoreward curve innovations are assigned to is irrelevant.

2 Optimal pricing

Building on the preceding section we may now investigate the properties of more general, incentive compatible pricing rules. Our analysis is less complete than under proportional pricing, but we provide four results. First, we present a general first-order derivative with respect to adjusting prices at any point and use it to provide a generalization of Theorem 1 from the text. Second, we discuss a version of our analysis that may apply even absent proportional pricing: a first-order derivative with respect to the overall level of market power. Third, we use a similar strategy to prove a generalized version of Theorems 2 and 3 from the text. Finally, we discuss which direction, starting from proportional pricing, may be optimal to move the price schedule. This provides some limited justification for proportional pricing. For simplicity we restrict attention to the case when $\tau^\star$ is differentiable.

Under general pricing we can again change variables from (σ, m) to (k, x). Here k again represents the isoreward curve measured by the point at which it intersects the $45°, \sigma = m$ line and x is again $\frac{m}{\sigma}$. However, now that pricing is not proportional, demand elasticity is not the same at all points. It is thus necessary to consider, for any point (k, x), the elasticity of demand at that point $\epsilon(k, x)$; we also use this convention for other quantities such as S and Q. However, not only the elasticity at a point is relevant; the value of k corresponding to any (σ, m) is determined by the elasticity *at every point along the isoreward curve passing through* (σ, m) between x and the 45° line. Thus, one must consider the *average elasticity of demand*:[7]

$$\tilde{\epsilon}(k, x) \equiv \frac{1}{\overline{\frac{1}{1+\epsilon}}(k, x)} - 1$$

[7]As described in Online Appendix Subsection 5.2, this is a (transformed) log-average elasticity along the isoreward curve from point (k, x) to point $(k, 1)$.

where

$$\overline{\frac{1}{1+\epsilon}}(k,x) \equiv \int_{z=1}^{x} \frac{1}{z\log(x)\left[1+\epsilon(k,z)\right]}dz$$

We begin be establishing the validity of our change of variables. We want to show that point $(\sigma, m) = \left(kx^{-\frac{\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}, kx^{\frac{1}{1+\tilde{\epsilon}(k,x)}}\right)$ lies along an isoreward curve intersecting the $45°, \sigma = m$ line at $m = \sigma = \sqrt{\sigma m x^{-\frac{1-\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}}$ where

$$\tilde{\epsilon}(k,x) \equiv \frac{1}{\overline{\frac{1}{1+\epsilon}}(k,z)} - 1$$

$$\overline{\frac{1}{1+\epsilon}}(k,x) \equiv \int_{z=1}^{x} \frac{1}{z\log(x)\left[1+\epsilon(k,z)\right]}dz$$

where $\epsilon(k,x) \equiv \epsilon\left(a\left(kx^{-\frac{\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}, kx^{\frac{1}{1+\tilde{\epsilon}(k,x)}}\right)\right)$. By Lemma 1, along an isoreward curve

$$\frac{\partial \sigma}{\partial \mathcal{M}} = -\frac{\epsilon(k,x)}{x}$$

Thus moving along an isoreward curve while adjusting x until one reaches the 45° line makes

$$\log(m) + \int_{l=\log(x)}^{0} \frac{1}{1+\epsilon\left(k, e^{l}\right)}dl = \log(k) \iff \log(m) = \log(k) + \log(x)\overline{\frac{1}{1+\epsilon}}(k,x) \iff m = kx^{\frac{1}{1+\tilde{\epsilon}(k,x)}}$$

A similarly derivation applies to σ.

Next we define the notion of a variation of $a(\sigma, m)$ in the direction of another policy $\hat{a}(\sigma, m)$ in keeping with the classical calculus of variations. In particular letting maximized social welfare under policy a be $W(a)$ the first variation of welfare in the direction of $\hat{a}$ is

$$\delta W(\hat{a}) \equiv \lim_{\delta \to 0} \frac{W\left[(1-\delta)a + \delta\hat{a}\right] - W\left(a\right)}{\delta}$$

We now seek to calculate $\delta W(\hat{a})$ for arbitrary $\hat{a}$. To do this we exploit some terminology we now define.

Let $\Delta a = \hat{a} - a$ and generalize $\bar{\cdot}(x,k)$ to potentially run over a range other than x to 1. For example, $\left.\overline{\frac{\Delta a \epsilon'}{(1+\epsilon)^2}}(k,x)\right|_{\tilde{x}}$ denotes a case when the lower bound of integration in the relevant equations is replaced by $\tilde{x}$ and $\log(x)$ is replaced by $\log\left(\frac{x}{\tilde{x}}\right)$ so that $\overline{\frac{\Delta a \epsilon'}{(1+\epsilon)^2}}(k,x) = \left.\overline{\frac{\Delta a \epsilon'}{(1+\epsilon)^2}}(k,x)\right|_{1}$.

First note that, as before, the points lying along the isoreward curve of a point which is, under a, assigned to (k,x) take the form (where we drop the arguments of σ and $\mathcal{M}$ inside the large expressions and assume $1 < \tilde{x} < x$):

$$\sigma(\delta; \tilde{x}, k, x) = k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}} e^{\int_{z=\log(\tilde{x})}^{\log(x)} \frac{1}{z}\left(\frac{\epsilon(k,z)}{1+\epsilon(k,z)} - \frac{\epsilon([1-\delta]a(\sigma,\mathcal{M}) + \delta\hat{a}(\sigma,\mathcal{M}))}{1+\epsilon([1-\delta]a(\sigma,\mathcal{M}) + \delta\hat{a}(\sigma,\mathcal{M}))}\right)dz} \equiv k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}} e^{\delta_\sigma(\tilde{x};k,x)}$$

$$\mathcal{M}(\delta;\tilde{x},k,x)=k\tilde{x}^{\frac{1}{1+\tilde{\epsilon}(k,\tilde{x})}}e^{\int_{z=\log(\tilde{x})}^{\log(x)}\frac{1}{z}\left(\frac{1}{1+\epsilon(k,z)}-\frac{1}{1+\epsilon([1-\delta]a(\sigma,\mathcal{M})+\delta\hat{a}(\sigma,\mathcal{M}))}\right)dz}\equiv k\tilde{x}^{\frac{1}{1+\tilde{\epsilon}(k,\tilde{x})}}e^{\delta_m(\tilde{x};k,x)}$$

Note that $\lim_{\delta\to 0}\delta_m(\tilde{x};k,x),\delta_\sigma(\tilde{x};k,x)=0\ \forall(\tilde{x};k,x)$. Furthermore we can obtain

$$d\sigma(\tilde{x},k,x)(\hat{a})\equiv\lim_{\delta\to 0}\frac{k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}}e^{\delta_\sigma(\tilde{x};k,x)}-k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}}}{\delta}=k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}}\lim_{\delta\to 0}\frac{\delta_\sigma(\tilde{x};k,x)}{\delta}=$$

$$-k\tilde{x}^{-\frac{\tilde{\epsilon}(k,\tilde{x})}{1+\tilde{\epsilon}(k,\tilde{x})}}\log\left(\frac{x}{\tilde{x}}\right)\overline{\frac{\Delta a\epsilon'}{(1+\epsilon)^2}(k,x)}\Bigg|_{\hat{x}}$$

$$d\mathcal{M}(\tilde{x},k,x)(\hat{a})=k\tilde{x}^{\frac{1}{1+\tilde{\epsilon}(k,\tilde{x})}}\log\left(\frac{x}{\tilde{x}}\right)\overline{\frac{\Delta a\epsilon'}{(1+\epsilon)^2}(k,x)}\Bigg|_{\hat{x}}$$

We can then calculate that the isoreward curve $\hat{k}$ assigned to (k,x) under the new policy is

$$\hat{k}(k,x;\delta,\hat{a})=ke^{\int_{z=1}^{\log(x)}\frac{1}{z}\left(\frac{1}{1+\epsilon(k,z)}-\frac{1}{1+\epsilon(\delta\hat{a}(\sigma(z;k,x),\mathcal{M}(z;k,x))+[1-\delta]a(\sigma(z;k,x),\mathcal{M}(z;k,x)))}\right)dz}$$

Let $dk(k,x)(\hat{a})\equiv\lim_{\delta\to 0}\frac{\hat{k}(k,x;\delta,\hat{a})-k}{\delta}$. It is straightforward, but tedious, to show that a number of second-order effects drop out because a small move towards $\hat{a}$ only causes a small change in σ; thus we are left with two effects:

$$\frac{dk(k,x)(\hat{a})}{k}=\log(x)\left(\underbrace{\overline{\frac{\Delta a\epsilon'}{(1+\epsilon)^2}}(k,x)}_{\text{direct sorting}}+\underbrace{\overline{\frac{\log\left(\frac{x}{z}\right)\overline{\frac{\Delta a\epsilon'}{(1+\epsilon)^2}}(k,x)\Big|_z\left(\epsilon_{a_m}-\epsilon_{a_\sigma}\right)\epsilon'}{(1+\epsilon)^2}}(k,x)}_{\text{indirect sorting}}\right)$$

The source of the first effect is both familiar from before and discussed more extensively below: raising a locally causes all innovations with a higher x on the same isoreward curve to have a higher k changing the local direction of the isoreward curve. The second effect is a bit more subtle, but not fundamentally different. To the extent a is not constant, changing a and thus the path of the (k,x) isoreward curve not only directly changes the path of the isoreward curve, but also does so indirectly by changing the set of elasticities "encountered" by the isoreward curve on its way to the 45° line.

We can now return to $\delta W(\hat{a})$. As usual we have two effects: one on the boundary and one on the interior of the integral. If we skip over steps now familiar from the calculation of numerous first-order derivatives of W above, we obtain

$$\delta W(\hat{a})=\int_k\int_{x\geq 1}k\left(\tau^{\star\prime}(k)\frac{dk(k,x)}{k}\left[S(k,x)k^2x^{\frac{1-\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}-\tau^\star(k)\right]\tilde{f}(k,x,\tau^\star(k))-\Delta a(k,x)Q(k,x)\epsilon(k,x)x^{\frac{1-\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}\int_{c=0}^{\tau^\star(k)}\tilde{f}(k,x,c)\right)dx+\ldots dk=$$

where ... represents the corresponding opposite term for $x < 1$. Dropping arguments where possible

$$\int_k \int_{x \geq 1} k^2 \left(\frac{\tau^{\star\prime}}{k} \log(x) \frac{\overline{\left[\log\left(\frac{x}{z}\right) \overline{\frac{\Delta a \epsilon'}{(1+\epsilon)^2}}\Big|_z (\epsilon_{a_m} - \epsilon_{a_\sigma}) + \Delta a \right] \epsilon'}}{(1+\epsilon)^2} \left[S k^2 x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} - \tau^\star(k) \right] \tilde{f}\left(x|k,\tau^\star\right) \tilde{f}\left(\tau^\star|k\right) - \Delta a Q \epsilon x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} \tilde{f}\left(x|k, c<\tau^\star\right) \tilde{F}\left(\tau^\star|k\right) \right) \tilde{f}(k) dx + \ldots dk \propto$$

$$\int_k k^4 \int_{x \geq 1} \left(\frac{\tau^{\star\prime}}{k} \log(x) \frac{\overline{\left[\log\left(\frac{x}{z}\right) \overline{\frac{\Delta a \epsilon'}{(1+\epsilon)^2}}\Big|_z (\epsilon_{a_m} - \epsilon_{a_\sigma}) + \Delta a \right] \epsilon'}}{(1+\epsilon)^2} \left(S x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} - E_{\tilde{f},x} \left[S x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} \right] \right) \eta(\tau^\star(k)|k) - \Delta a Q \epsilon x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} E_{\tilde{f},x} \left[S x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} \right] \tilde{f}(\hat{x}|\hat{k},\tau^\star) \frac{E\left[\tilde{f}(\hat{x}|k,c) \middle| c<\tau^\star(k), k \right]}{\tilde{f}(x|c=\tau^\star(k),k)} \right) \tilde{f}(k) dx + \ldots d \tag{6}$$

First-order conditions for maximization over all differentiable schedules a require that $\delta W(\hat{a})$ be 0 for all differentiable $\hat{a}$. It is well known that this condition[8] is equivalent to this condition holding for the class of differentiable functions

$$\hat{a}(k, x; \hat{k}, \hat{x}, \eta) = a(k, x) + \frac{e^{-\frac{(k-\hat{k})^2 + (x-\hat{x})^2}{\eta^2}}}{2\pi\eta}$$

for all choices of $(\hat{k}, \hat{x})$ and all $\eta < \overline{\eta}$ for an arbitrarily small η, as such functions form a basis for the set of all differentiable functions. If we allow $\eta \to 0$ we obtain a point mass difference between a and $\hat{a}$ and the notion of a *perturbation* of W at $(\hat{k}, \hat{x})$:

$$W'(\hat{k}, \hat{x}) \equiv \lim_{\eta \to 0} \frac{W\left(\hat{a}\left(\hat{k}, \hat{x}, \eta\right)\right)}{\eta}$$

As the proof of our first result in this section shows, these perturbations can be computed as limits of formula (6) when Δa converges to 0 everywhere by $(\hat{k}, \hat{x})$ and to ∞ at that point.

An innovation's social value is, after some algebraic manipulations, $S(k, x) k^2 x^{\frac{1-\tilde{\epsilon}(k,x)}{1+\tilde{\epsilon}(k,x)}}$. The benefits of raising a again arise from sorting and those of lowering it from reducing ex-post inefficiency, both locally holding fixed rewards given to each k by the envelope theorem; the optimum balances these two incentives. However for general pricing we must consider this trade-off at each (σ, m), or (k, x), pair.

Proposition 1: *If $\tau^\star$ is differentiable in k, the first-order net benefit of increasing a at $\left(\hat{k}, \hat{x}\right)$ beginning from a strict MOPSD pricing policy $a(\cdot, \cdot)$ is, if $x \geq 1$, proportional to*

$$\underbrace{\frac{\tau^{\star\prime}}{\hat{k}} \frac{\epsilon'(\hat{k},\hat{x})}{\hat{x}\left[1+\epsilon(\hat{k},\hat{x})\right]^2} \left[1 + \log(\hat{x}) \overline{\frac{\epsilon'(\epsilon_{a_m} - \epsilon_{a_\sigma})}{(1+\epsilon)^2}}\left(\hat{k},\hat{x}\right)\right] \left(\frac{E_{\tilde{f},x>\hat{x}}\left[S x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}{E_{\tilde{f},x}\left[S x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]} - 1 \right) \eta\left(\tau^\star|\hat{k}\right)}_{\textit{sorting}} - \underbrace{Q \epsilon \hat{x}^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} H\left(\hat{x}|\hat{k},\tau^\star\right) \frac{E\left[\tilde{f}\left(\hat{x}|\hat{k},c\right)|c<\tau^\star,\hat{k}\right]}{\tilde{f}\left(x|c=\tau^\star,\hat{k}\right)}}_{\textit{ex-post distortion}}$$

[8] Note that π in this condition is the geometric constant, not a variable for profits.

where H is the (conditional) hazard rate of x under $\tilde{f}$ and if $x < 1$, proportional to

$$\frac{\tau^{\star\prime}}{\hat{k}} \frac{\epsilon'(\hat{k},\hat{x})}{\hat{x}[1+\epsilon(\hat{k},\hat{x})]^2} \left[1+\log(\hat{x})\overline{\frac{\epsilon'(\epsilon_{a_m}-\epsilon_{a_\sigma})}{(1+\epsilon)^2}}(\hat{k},\hat{x})\right] \left(1-\frac{E_{\tilde{f},x<\hat{x}}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}{E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}\right) \eta(\tau^\star|\hat{k}) - Q\epsilon\hat{x}^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}} R(\hat{x}|\hat{k},\tau^\star) \frac{E[\tilde{f}(\hat{x}|\hat{k},c)|c<\tau^\star,\hat{k}]}{\tilde{f}(x|c=\tau^\star,\hat{k})}$$

where R is the reversed hazard rate. A necessary condition for a strict MOPSD (no-bunching) solution is that these equal zero at every point in the plane.

This is just the standard calculus-of-variations first-order derivative at a point in this context, using the envelope theorem.

Proof. As usual with Dirac-convergent weighting functions, the integral converges to the value of the density at the limit mass point. The value of the second term of (6) is easy to evaluate at this mass point so we focus on the first term and begin by analyzing:

$$\log(x)\overline{\frac{\left[\log\left(\frac{x}{z}\right)\overline{\frac{\Delta a\epsilon'}{(1+\epsilon)^2}}\Big|_z(\epsilon_{a_m}-\epsilon_{a_\sigma})+\Delta a\right]\epsilon'}{(1+\epsilon)^2}} = \int_{z=1}^{x} \frac{1}{z}\frac{\epsilon'(k,z)}{[1+\epsilon(k,z)]^2}\left(\Delta a(k,z)+[\epsilon_{a_m}(k,z)-\epsilon_{a_\sigma}(k,z)]\int_{\alpha=z}^{x}\frac{1}{\alpha}\frac{\epsilon'(k,\alpha)\Delta a(k,\alpha)}{[1+\epsilon(k,\alpha)]^2}d\alpha\right)dz$$

Evaluating this in the limit as Δa becomes a point mass of 1 on $(\hat{k},\hat{x})$ yields

$$\frac{1_{x\geq\hat{x}}}{\hat{x}}\frac{\epsilon'(\hat{k},\hat{x})}{\left[1+\epsilon(\hat{k},\hat{x})\right]^2} + \int_{z=1}^{x}\frac{1_{z<\hat{x}<x}\epsilon'(\hat{k},\hat{x})}{\left[1+\epsilon(\hat{k},\hat{x})\right]^2}\left[\epsilon_{a_m}(\hat{k},z)-\epsilon_{a_\sigma}(\hat{k},z)\right]dz$$

$$\frac{1_{x\geq\hat{x}}}{\hat{x}}\frac{\epsilon'(\hat{k},\hat{x})}{\left[1+\epsilon(\hat{k},\hat{x})\right]^2}\left[1+\log(\hat{x})\overline{\frac{\epsilon'(\epsilon_{a_m}-\epsilon_{a_\sigma})}{(1+\epsilon)^2}}(\hat{k},\hat{x})\right]$$

So by equation (6) for $\hat{x} > 1$, $W'(\hat{k},\hat{x}) \propto$

$$\int_{x=\hat{x}}^{\infty}\frac{\tau^{\star\prime}\epsilon'(\hat{k},\hat{x})}{\hat{k}\hat{x}\left[1+\epsilon(\hat{k},\hat{x})\right]^2}\left[1+\log(\hat{x})\overline{\frac{\epsilon'(\epsilon_{a_m}-\epsilon_{a_\sigma})}{(1+\epsilon)^2}}(\hat{k},\hat{x})\right]\left(Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}-E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)\eta(\tau^\star(k)|k)-Q\epsilon\hat{x}^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\tilde{f}(\hat{x}|\hat{k},\tau^\star)\frac{E[\tilde{f}(\hat{x}|k,c)|c<\tau^\star(k),k]}{\tilde{f}(x|c=\tau^\star(k),k)}=$$

$$\frac{\tau^{\star\prime}}{\hat{k}}\frac{\epsilon'(\hat{k},\hat{x})}{\hat{x}\left[1+\epsilon(\hat{k},\hat{x})\right]^2}\left[1+\log(\hat{x})\overline{\frac{\epsilon'(\epsilon_{a_m}-\epsilon_{a_\sigma})}{(1+\epsilon)^2}}(\hat{k},\hat{x})\right]\left(\frac{E_{\tilde{f},x>\hat{x}}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}{E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}-1\right)\eta(\tau^\star(k)|k)-Q\epsilon\hat{x}^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}H\left(\hat{x}|\hat{k},\tau^\star\right)\frac{E\left[\tilde{f}(\hat{x}|k,c)\Big|c<\tau^\star(k),k\right]}{\tilde{f}(x|c=\tau^\star(k),k)} \tag{7}$$

as in the text. For $\hat{x} < 1$ the reasoning is analogous and thus omitted. □

Three things may be gleaned from these relatively dense expressions. First, despite their complexity, the same basic forces are at work as with proportional pricing. The first term is the product of materialism and the degree to which raising a is able to sort for the best innovations. This is quite naturally measured by the ratio of the social value of average innovations with a higher value of x than $\hat{x}$, along isoreward $\hat{k}$, to the social value of an overall-average innovation along that isoreward curve. Thus, the basic logic of our analysis carries through more generally. Second, there seems be a strong indication, discussed more extensively later, that along an isoreward curve a will optimally decline in x (decline in m and/or increase in σ), at least over a significant range. We will discuss the reasoning behind this more extensively below.

Finally, supposing this is the case, it is worth noting that this creates both greater, and lesser, selective pressure in favor of high x innovations. On the one hand, it raises selective pressure as their S values are higher as well. On the other hand in this case ϵ_{a_m} is likely smaller than ϵ_{a_σ} (so that a declines in x along k) so that the rather odd term

$$\log(\hat{x})\overline{\frac{\epsilon'\left(\epsilon_{a_m}-\epsilon_{a_\sigma}\right)}{(1+\epsilon)^2}}\left(\hat{k},\hat{x}\right)$$

becomes negative and thus depresses the incentives for market power. The basic source of this term, explained extensively in Online Appendix Subsection 5.2, is that changing pricing alters the set of prices through which (σ, m)'s isoreward curve passes as it approaches the 45° line and thus indirectly affects the rewards given to (σ, m) to the extent that a is not constant.

This analysis may be used to generalize Theorem 1 from the text.

Theorem 1: *At global ex-post efficient pricing there is a local incentive at all points to raise prices at any (σ, m) for which $\tau^\star(\sigma)$ is not constant in the neighborhood of σ. At global monopoly pricing there is a local incentive at all points to lower prices.*

Proof. At globally ex-post efficient prices $\epsilon = \epsilon_{a_m} = \epsilon_{a_\sigma} = 0$ and $S = 1$ for all innovations. Thus expression (7) becomes

$$\frac{\tau^{\star'}}{k}\epsilon'(0)\left(\frac{E_{\tilde{f},x>\hat{x}}[x]}{E_{\tilde{f},x}[x]}-1\right)\eta\left(\tau^\star(k)|k,x\right)$$

which is strictly positive whenever $\tau^{\star'} \neq 0$. By similar tricks at global monopoly pricing

$$\frac{E_{\tilde{f},x>\hat{x}}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}{E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]}-1=0$$

as $\tilde{\epsilon} = 1$ everywhere while the second term is

$$-Q\epsilon H(\hat{x}|k) < 0$$

A similar logic holds for $x < 1$. □

Another way of recovering some broader results of our analysis is to focus on its primary goal: determining the optimal "level", rather than structure, of market power. One of the necessary conditions for optimality is, of course, choosing this level correctly. In particular, we can consider the first-order costs and benefits of uniformly lowering $\frac{1}{1+\epsilon}$ at every point by a small amount; this particular direction is chosen for the analytic simplifications it allows. This gives a similar expression to our baseline proportional pricing first-order condition, as shown in the following proposition. However the additional forces identified above still play a role.

Proposition 2: *Starting from any strict MOPSD policy a with $a < 1$ everywhere and assuming $\tau^\star$ is differentiable at this policy, the first variation of W in the direction $a + \frac{(1+\epsilon)^2}{\epsilon'}$ (uniform decrease in $\frac{1}{1+\epsilon}$) is proportional to*

$$E_{k,\tilde{f}}\left[k^4\left(\frac{\eta\tau^{\star\prime}}{k} Cov_{x,\tilde{f}}\left[\frac{\epsilon'}{(1+\epsilon)^2}\left(1+\overline{\log\left(\frac{x}{z}\right)(\epsilon_{a_m}-\epsilon_{a_\sigma})}\right)\log(x), Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)-E_{\tilde{f},x}\left[Q\epsilon x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\middle|k,c<\tau^\star(k)\right]E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right]$$

Proof. Lowering $\frac{1}{1+\epsilon}$ uniformly by one unit corresponds to $\Delta a = \frac{[1+\epsilon]^2}{\epsilon'}$. Plugging this into expression (6) yields

$$\delta W\left(a+\frac{(1+\epsilon)^2}{\epsilon'}\right)=\int_k k^4\int_{x\geq 1}\left(\frac{\tau^{\star\prime}}{k}\log(x)\left[1+\overline{\log\left(\frac{x}{z}\right)(\epsilon_{a_m}-\epsilon_{a_\sigma})}\right]\left(Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}-E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)\eta(\tau^\star(k)|k,x)-\frac{Q\epsilon[1+\epsilon]^2}{\epsilon'}x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)\tilde{f}(k,x)dx+\ldots dk$$

The corresponding term for $x < 1$ is essentially identical so we obtain

$$\delta W\left(a+\frac{(1+\epsilon)^2}{\epsilon'}\right)=\int_k k^4\int_x\left(\frac{\tau^{\star\prime}}{k}\log(x)\left[1+\overline{\log\left(\frac{x}{z}\right)(\epsilon_{a_m}-\epsilon_{a_\sigma})}\right]\left(Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}-E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)\eta(\tau^\star(k)|k,x)-\frac{Q\epsilon[1+\epsilon]^2}{\epsilon'}x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)\tilde{f}(k,x)dxdk\propto$$

$$E_{k,\tilde{f}}\left[k^4\left(\frac{\eta\tau^{\star\prime}}{k}\text{Cov}_{x,\tilde{f}}\left[\frac{\epsilon'}{(1+\epsilon)^2}\left(1+\overline{\log\left(\frac{x}{z}\right)(\epsilon_{a_m}-\epsilon_{a_\sigma})}\right)\log(x), Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right)-E_x\left[Q\epsilon x^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\middle|k,c<\tau^\star(k)\right]E_{\tilde{f},x}\left[Sx^{\frac{1-\tilde{\epsilon}}{1+\tilde{\epsilon}}}\right]\right]$$

□

Similarly, we may consider the benefits of moving towards monopoly pricing from *any pricing policy* arbitrarily close to it, or towards ex-post efficiency from any pricing policy arbitrarily close to it, to derive a general version of Friedman's Conjecture, or its converse.

Theorem 2: *Beginning from any a sufficiently, uniformly close to uniform monopoly pricing ($a = 1$ everywhere) but with $a < 1$ everywhere, if V_1 (of Theorem 2 from the text) is sufficiently large there are first-order benefits from moving a small amount (uniformly) towards uniform monopoly pricing. Beginning from any strict MOPSD a sufficiently, log-uniformly close to ex-post efficiency ($a = 0$ everywhere) but with $a > 0$ everywhere, if V_0 (of Theorem 3 from the text) is sufficiently small there are first-order benefits from moving a small amount (uniformly) towards ex-post efficiency.*

The proof is effectively identical to that of Theorems 2 and 3 from the text, with only slight complexities in simplifying the more general first-order condition local to the proportional policies of ex-post efficiency and monopoly pricing.

Proof. If a is sufficiently, uniformly close to 1 or 0 then, because $\frac{\epsilon'}{(1+\epsilon)^2}$ is approximately constant, moving towards monopoly pricing is equivalent to uniformly increasing $\frac{1}{1+\epsilon}$. Furthermore all the simplifications from the proof of Theorem 1 apply and the first-order derivative from Proposition 1 simplify to those in the proofs of Theorem 2 and 3 from the text and the results therefore follow by the same reasoning as there. □

Finally, to provide at least some notion of what the optimal structure of market power may look like, we can consider evaluating the first-order condition for optimal pricing at each point, beginning from optimal proportional pricing. This provides at least some indication of the optimal structure of market power.

Corollary 1: *If $\tau^\star$ is differentiable in k, the first-order net benefit of increasing a at $\left(\hat{k},\hat{x}\right)$ given beginning from a proportional pricing policy a is, if $x \geq 1$, proportional to*

$$\frac{\epsilon' \tau^{\star\prime}\left(\hat{k}\right)}{\left(1+\epsilon\right)^2 \hat{k}\hat{x}^{\frac{2}{1+\epsilon}}}\left(\frac{E_{\tilde{f},x>\hat{x}}\left[x^{\frac{1-\epsilon}{1+\epsilon}}\middle|\hat{k},\tau^\star\right]}{E_{\tilde{f},x}\left[x^{\frac{1-\epsilon}{1+\epsilon}}\middle|\hat{k},\tau^\star\right]}-1\right)\eta\left(\tau^\star|\hat{k}\right)-Q\epsilon H\left(\hat{x}|\hat{k},\tau^\star\right)\frac{E\left[\tilde{f}\left(\hat{x}|\,\hat{k},c\right)\middle|c<\tau^\star,\hat{k}\right]}{\tilde{f}\left(x|\,\tau^\star,\hat{k}\right)}$$

and if $x < 1$

$$\frac{\epsilon' \tau^{\star\prime}\left(\hat{k}\right)}{\left(1+\epsilon\right)^2 \hat{k}\hat{x}^{\frac{2}{1+\epsilon}}}\left(1-\frac{E_{\tilde{f},x<\hat{x}}\left[x^{\frac{1-\epsilon}{1+\epsilon}}\middle|\hat{k},\tau^\star\right]}{E_{\tilde{f},x}\left[x^{\frac{1-\epsilon}{1+\epsilon}}\middle|\hat{k},\tau^\star\right]}\right)\eta\left(\tau^\star|\hat{k}\right)-Q\epsilon R\left(\hat{x}|\hat{k},\tau^\star\right)\frac{E\left[\tilde{f}\left(\hat{x}|\,\hat{k},c\right)\middle|c<\tau^\star,\hat{k}\right]}{\tilde{f}\left(x|\,\tau^\star,\hat{k}\right)}$$

With some additional simplifying assumptions these formula may give some insight. Assume that k, x and c are all independent, which implies that $\tau^\star$ is quadratic in k. Furthermore, suppose that $\frac{E_{\tilde{f},\hat{x}}\left[\hat{x}^{\frac{1-\epsilon}{1+\epsilon}}\middle|\hat{x}>x,k,c=\tau^\star(k)\right]}{E_{\tilde{f},\hat{x}}\left[\hat{x}^{\frac{1-\epsilon}{1+\epsilon}}\middle|k,c=\tau^\star(k)\right]}-1=\gamma x^{\frac{1-\epsilon}{1+\epsilon}}$ as in a generalized Pareto distribution. Then the first formula simplifies to

$$\frac{\gamma\epsilon'\eta}{\left(1+\epsilon\right)^2 x}-Q\epsilon H\left(x\right)$$

This setting implies that the local incentives for distortion are independent of k. However it also gives a sense that, after some point at least, the local incentive for raising prices is likely to decline in x along an isoreward curve. This occurs for two reasons. First, if we have the standard increasing hazard rate condition, H grows in x. This is just the logic behind the classic "no distortion at the top" result of Mirrlees (1971), though based on sorting rather than rent extraction. The sorting benefits of higher prices only affect innovations with higher x than that at which prices are distorted; if the mass of such upper-tail innovations shrinks relative to that of those being distorted as hazard rates are increasing, there will be little incentive to create such distortion. However there is also an additional motive for lower distortions for higher x here: the multiplicative, log-linear nature of the stretch parameterization implies that the upper tail is less affected by shifts in elasticity at high x values than those at low values (above unity).[9] In some sense the closer x is to 1 the more

[9]It would be interesting to know if this is a more general property of multidimensional screening problems.

dramatically a change in the elasticity at that point shifts the isoreward curve and therefore the greater its sorting value. While for $x < 1$ things are a bit subtler, there seems to be some weak, but general, indication that optimal policy calls for a values declining along isoreward curves.[10]

References

Armstrong, Mark, "Multiproduct Nonlinear Pricing," *Econometrica*, 1996, *64* (1), 51–75.

Bulow, Jeremy I. and John Roberts, "The Simple Economics of Optimal Auctions," *Journal of Political Economy*, 1989, *97* (5), 1060–1090.

Fabinger, Michal and E. Glen Weyl, "A Multidimensional Envelope Theorem with Endogenous Choice Sets," 2011. This paper is in preparation.

Hagood, John W. and Brian S. Thomson, "Recovering a Function from a Dini Derivative," *American Mathematical Monthly*, 2006, *113* (1), 34–46.

Harrod, R. F., "The Law of Decreasing Costs," *The Economic Journal*, 1931, *41* (164), 566–576.

McAfee, R. Preston and John McMillan, "Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory*, 1988, *46* (2), 335–354.

Milgrom, Paul R., *Putting Auction Theory to Work*, Cambridge, UK: Cambridge University Press, 2004.

_ **and Ilya Segal**, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 2002, *70* (2), 583–601.

Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Pigou, Arthur C., *The Economics of Welfare*, London: MacMillan, 1920.

Rochet, Jean-Charles and Philippe Choné, "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 1998, *66* (4), 783–826.

[10]This extremely tentative conclusion merits two comments. First, if this is the case it might point towards bunching (a weak MOPSD) being optimal as a declining with rising x pushes up against order-preservation and/or monotonicity. This would be consistent with the arguments of Armstrong (1996) and Rochet and Choné (1998) that bunching is quite common in multidimensional screening and is an interesting observation from an economic perspective, as it would imply the easy-to-implement institution of price controls potentially forming part of optimal policy. Second, it provides at least some reassurance that proportional pricing is not too wildly off as a policy prescription. There are limits to how rapidly a may increase with x and still be a MOPSD and therefore incentive compatible; had it been optimal for a to decrease in x there would have been no limit to how far off proportional pricing might have been. Obviously all of this analysis is exceptionally preliminary and the more complete analysis of our model with general pricing remains an important topic for future research.

Rogerson, William P., "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 1985, *53* (6), 1357–1367.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 1974, *82* (1), 34–55.

Royden, H. L., *Real Analysis*, Vol. Third Edition, Englewood Cliffs, NJ: Prentice-Hall, 1988.

Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Viner, Jacob, "Cost Curves and Supply Curves," *Zietschrift für Nationalökonomie*, 1931, *3*, 23–46.

Wilson, Robert B., *Nonlinear Pricing*, Oxford: Oxford University Press, 1993.

Young, W. H. and Grace Chisolm Young, "On the Discontinuities of Monotone Functions of Several Variables," *Proceedings of the London Mathematical Society*, 1924, *s2-22* (124–142).