Nonlinear Pricing in Village Economies: Supplementary Appendix

Orazio Attanasio* Elena Pastorino†

A Summary of Appendix

We provide details of the model in Jullien (2000), which are useful to understand our derivations, in Section A.1. We present details of the examples mentioned in Section 3 in the paper in Section A.2. We discuss how the problem of a single seller that we focus on in the empirical analysis could be equivalently interpreted as the problem of an oligopolist in Section A.3. We further elaborate on our identification strategy in Section A.4. We present estimation results omitted from the paper in Section A.5.

A.1 Model with Heterogeneous Reservation Utilities

We provide here omitted details of the model in Jullien (2000) under the assumption that a seller's cost function is separable across consumers only for consistency with his formulation. Recall that in Jullien (2000) the seller's optimal menu is chosen to maximize expected profits subject to incentive compatibility and participation constraints, that is,

$$\text{(IR problem)} \qquad \max_{\{t(\theta),q(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} [t(\theta) - c(q(\theta))] f(\theta) d\theta \;\text{ s.t.}$$

$$\text{(IC)}\;\; v(\theta, q(\theta)) - t(\theta) \geq v(\theta, q(\theta')) - t(\theta') \text{ for any } \theta, \theta'$$

$$\text{(IR)}\;\; v(\theta, q(\theta)) - t(\theta) \geq \overline{u}(\theta) \text{ for any } \theta.$$

We refer to this model in which the seller's constraints are IC and IR as the *IR model*. We define an allocation $\{u(\theta), q(\theta)\}$ to be *implementable* if it satisfies the IC and IR constraints. The standard approach to solve such a problem is to reduce the IC constraints to a local version that is analytically more tractable. To do so, notice that the IC constraint is satisfied for a consumer of type θ whenever choosing $q(\theta)$ for the price $t(\theta)$ maximizes the left-side of the IC constraint. Taking first-order conditions, this requires that $v_q(\theta, q(\theta))q'(\theta) = t'(\theta)$. It will prove convenient to expression this condition as

$$u'(\theta) = v_\theta(\theta, q(\theta)), \tag{1}$$

by using the fact that differentiating $u(\theta)$ yields $u'(\theta) = v_\theta(\theta, q(\theta)) + [v_q(\theta, q(\theta))q'(\theta) - t'(\theta)]$. A standard result is that under the assumption that $v_{\theta q}(\theta, q) > 0$, an allocation is incentive compatible if, and only if, it is *locally* incentive compatible in that (1) holds, the quantity schedule $q(\theta)$ is weakly increasing (a.e.), and the associated utility $u(\theta)$ is absolutely continuous.

*University College London, Institute for Fiscal Studies, NBER, and CEPR.

†Stanford University, Hoover Institution, and Federal Reserve Bank of Minneapolis.

There are several steps to the solution of the resulting IR problem. First, we effectively "substitute out" the local incentive compatibility condition (1) by integrating both sides of it from $\underline{\theta}$ to θ to obtain

$$u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x, q(x))dx \tag{2}$$

and substituting it into the seller's objective function, using the fact that $t(\theta) - c(q(\theta)) = v(\theta, q(\theta)) - c(q(\theta)) - u(\theta) = s(\theta, q(\theta)) - u(\theta)$. Second, we rewrite the resulting problem as a Lagrangian problem with $d\gamma(\theta)$ representing the multiplier on the IR constraint for type θ. Third, after simple manipulations detailed in Result 2, we express the Lagrangian problem in the following simple form,

$$\text{(simple IR problem)} \qquad \max_{\{q(\theta)\}} \int_{\underline{\theta}}^{\bar{\theta}} \left\{ v(\theta, q(\theta)) - c(q(\theta)) + \left[\frac{F(\theta) - \gamma(\theta)}{f(\theta)}\right] v_\theta(\theta, q(\theta)) \right\} f(\theta)d\theta, \tag{3}$$

with $q(\theta)$ weakly increasing and $\gamma(\theta) = \int_{\underline{\theta}}^{\theta} d\gamma(x)$ defined to be the *cumulative multiplier* on the IR constraint for type θ. This cumulative multiplier has the properties of a cumulative distribution function, that is, it is nonnegative, weakly increasing, and $\gamma(\bar{\theta}) = 1$, as shown in Result 1. Note that the integral in the definition of $\gamma(\theta)$ is interpreted as accommodating not just discrete and continuous distributions but also mixed discrete-continuous ones. That is, this formulation covers the case in which the IR constraints bind at isolated points. In the standard nonlinear pricing model, consumers' reservation utilities are assumed to be independent of θ so that the IR constraints simplify to $u(\theta) \geq \overline{u}$ and bind only for the lowest type, which implies that $\gamma(\theta) = 1$ for all types. See Result 3.

Jullien (2000) shows that under three assumptions, referred to as *potential separation*, *homogeneity*, and *full participation*, there is a unique optimal allocation inducing full participation that is characterized by the first-order conditions to (3),

$$v_q(\theta, q(\theta)) - c'(q(\theta)) = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q(\theta)) \tag{4}$$

for each type, together with the complementary slackness condition on the IR constraints,

$$\int_{\underline{\theta}}^{\bar{\theta}} [u(\theta) - \bar{u}(\theta)]d\gamma(\theta) = 0. \tag{5}$$

Note that conditions (4) and (5) are actually those of a relaxed version of (3) in which the constraint that $q(\theta)$ is weakly increasing has been dropped.

The final step uses the potential separation assumption, which, as discussed in the paper, is a generalization of the standard monotone hazard rate condition, to show that the solution to these first-order conditions is increasing and, hence, a solution to the original IR problem. (See Result 4 for a precise statement of this result.) For later use, we find it convenient to let $l(\gamma, \theta)$ denote the solution to the first-order condition (4) for a consumer of type θ for a given value of the cumulative multiplier $\gamma \in [0, 1]$, as we do in the paper. Thus, $l(\gamma, \theta)$ should be interpreted as the quantity that would be chosen by the seller for some arbitrary cumulative multiplier, γ. We start with a preliminary result:

Result 1. *The cumulative multiplier $\gamma(\theta)$ satisfies $\gamma(\bar{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}} d\gamma(\theta) = 1$.*

Before we prove this result, note that the cumulative multipliers are measures over $\left[\underline{\theta}, \bar{\theta}\right]$ that may jump discretely at some points. Hence, we need to adopt a convention on what the integral symbol means for mixed discrete-continuous measures. An intuitive approach is as follows. A mixed discrete-continuous measure $\mu(\theta)$ on $\left[\underline{\theta}, \bar{\theta}\right]$ can be represented as the sum of a discrete measure $\mu_d(\theta)$, defined on the mass points $\{\theta_1, \ldots, \theta_K\}$ with generic element θ_k, and a continuous measure $\mu_c(\theta)$ on $\left[\underline{\theta}, \bar{\theta}\right]$. Then, the symbol $\int_{\theta'}^{\theta''} d\mu(\theta)$ is defined as

$$\int_{\theta'}^{\theta''} d\mu(\theta) \equiv \sum_{\theta' \leq \theta_k \leq \theta''} \mu_d(\theta_k) + \int_{\theta'}^{\theta''} d\mu_c(\theta), \tag{6}$$

where the integral on the right-side of (6) is the standard one for continuous measures. Critically, the integral over an interval $[\theta', \theta'']$ may contain discrete mass at both endpoints θ' and θ'' as well as discrete and continuous mass between these endpoints. Moreover, by this definition we have that

$$\mu(\bar{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}} d\mu(\theta). \tag{7}$$

Here and in the paper, we use the definition of integration given by (6) without further remark.
Proof of Result 1: We prove this result by considering a uniform marginal reduction in the participation constraint from $\bar{u}(\theta)$ to $\bar{u}(\theta) - \delta$ for all types by a given small amount, $\delta > 0$. The first part of the proof uses the standard envelope condition to derive an expression for the resulting change in the value of the IR problem, expressed in quasi-Lagrangian form,

$$\max_{\{u(\theta)\},\{q(\theta)\} \in Q} \left\{ \int_{\underline{\theta}}^{\bar{\theta}} \left[s(\theta, q(\theta)) - u(\theta)\right] f(\theta) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \left[u(\theta) - \overline{u}(\theta)\right] d\gamma(\theta) \right\} \text{ s.t. } u'(\theta) = v_\theta(\theta, q(\theta)),$$

where Q is the set of functions that are weakly increasing with θ. To this purpose, rewrite the value of this problem as

$$W\left(\bar{u}(\theta) - \delta\right) = \max_{\{u(\theta), q(\theta)\}} \left\{ \int_{\underline{\theta}}^{\bar{\theta}} \left[s(\theta, q(\theta)) - u(\theta)\right] f(\theta) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \left[u(\theta) - \bar{u}(\theta) + \delta\right] d\gamma(\theta) \right\}$$

ignoring the requirement that $q(\theta)$ be weakly increasing, so that

$$\frac{dW(\bar{u}(\theta) - \delta)}{d\delta} = \int_{\underline{\theta}}^{\bar{\theta}} d\gamma(\theta), \tag{8}$$

where the integral in (8) is defined as in (6).

For the second part of the proof, we argue that it is immediate that the solution to the problem obtained from this proposed change in the participation constraints implies the same quantities as in the original problem, with the price schedule shifted up by the constant δ and consumers' utilities shifted down by δ. Of course, such a change in the participation constraints will just shift up the value of the program by δ. Specifically, if $\{u(\theta), q(\theta)\}$ with associated $t(\theta)$ is the solution to the original problem, then $\{u(\theta) - \delta, q(\theta)\}$ with associated $t(\theta) + \delta$ is the solution to the new problem,

$$W(\bar{u}(\theta) - \delta) = \int_{\underline{\theta}}^{\bar{\theta}} \left[t(\theta) + \delta - c(q(\theta))\right] f(\theta) d\theta = \delta + \int_{\underline{\theta}}^{\bar{\theta}} \left[t(\theta) - c(q(\theta))\right] f(\theta) d\theta = \delta + W(\bar{u}(\theta)).$$

Hence,

$$\frac{dW(\bar{u}(\theta)-\delta)}{d\delta}=1. \tag{9}$$

Thus, we have

$$\gamma(\bar{\theta})=\int_{\underline{\theta}}^{\bar{\theta}} d\gamma(\theta)=\frac{dW(\bar{u}\,(\theta)-\delta)}{d\delta}=1,$$

where these equalities follow from the definition in (6), (8), and (9). Hence, $\gamma(\bar{\theta})=1$. □

We now show how the IR problem can be reduced to the simple IR problem.

Result 2. *The IR problem can be reduced to the simple IR problem.*

Proof of Result 2: The first step is to rewrite the IC constraints in their local form and express the resulting problem in quasi-Lagrangian form by letting $d\gamma(\theta)$ denote the multiplier on the participation constraint of a consumer of type θ so that the seller's problem becomes

$$\max_{\{u(\theta)\},\{q(\theta)\}\in Q}\left\{\int_{\underline{\theta}}^{\bar{\theta}}\left[s(\theta,q(\theta))-u(\theta)\right]f(\theta)d\theta+\int_{\underline{\theta}}^{\bar{\theta}}\left[u(\theta)-\overline{u}(\theta)\right]d\gamma(\theta)\right\}\text{ s.t. }u'(\theta)=v_{\theta}(\theta,q(\theta)), \tag{10}$$

where Q is the set of functions that are weakly increasing with θ. The second step is to establish two simple results, that is,

$$\int_{\underline{\theta}}^{\bar{\theta}} u(\theta)dF(\theta)=u(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))d\theta-\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))F(\theta)d\theta, \tag{11}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} u(\theta)d\gamma(\theta)=u(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))d\theta-\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))\gamma(\theta)d\theta. \tag{12}$$

Here we establish (12), and note that the proof of (11) is analogous. To do so, recall that the constraint $u'(\theta)=v_{\theta}(\theta,q(\theta))$ is equivalent to (2). Integrating this condition from $\underline{\theta}$ to $\bar{\theta}$ with respect to $\gamma(\theta)$ gives

$$\int_{\underline{\theta}}^{\bar{\theta}} u(\theta)d\gamma(\theta)=\int_{\underline{\theta}}^{\bar{\theta}}\left[u(\underline{\theta})+\int_{\underline{\theta}}^{\theta} v_{\theta}(x,q(x))dx\right]d\gamma(\theta)=u(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\underline{\theta}}^{\theta} v_{\theta}(x,q(x))dx\right)d\gamma\left(\theta\right). \tag{13}$$

We then use integration by parts to simplify the second term on the right-most side of (13). Using $\int AdB=AB|-\int BdA$, with $A=\int_{\underline{\theta}}^{\theta} v_{\theta}(x,q(x))dx$, $B=\gamma(\theta)$, $\gamma(\underline{\theta})=0$, and $\gamma(\bar{\theta})=1$, we obtain

$$\begin{aligned}\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\underline{\theta}}^{\theta} v_{\theta}(x,q(x))dx\right)d\gamma(\theta)&=\left(\int_{\underline{\theta}}^{\theta} v_{\theta}(x,q(x))dx\right)\gamma(\theta)\Big|_{\underline{\theta}}^{\bar{\theta}}-\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))\gamma(\theta)d\theta\\&=\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))d\theta-\int_{\underline{\theta}}^{\bar{\theta}} v_{\theta}(\theta,q(\theta))\gamma(\theta)d\theta.\end{aligned} \tag{14}$$

The third step consists in substituting (11) and (12) into the quasi-Lagrangian form of the objective function in (10) to obtain

$$\int_{\underline{\theta}}^{\bar{\theta}} s(\theta,q(\theta))f(\theta)d\theta-\int_{\underline{\theta}}^{\bar{\theta}} u(\theta)f(\theta)d\theta+\int_{\underline{\theta}}^{\bar{\theta}} u(\theta)d\gamma(\theta)-\int_{\underline{\theta}}^{\bar{\theta}}\overline{u}(\theta)d\gamma(\theta)$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} s(\theta, q(\theta))f(\theta)d\theta - u(\underline{\theta}) - \int_{\underline{\theta}}^{\bar{\theta}} v_\theta(\theta, q(\theta))d\theta + \int_{\underline{\theta}}^{\bar{\theta}} v_\theta(\theta, q(\theta))F(\theta)d\theta$$

$$+u(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} v_\theta(\theta, q(\theta))d\theta - \int_{\underline{\theta}}^{\bar{\theta}} v_\theta(\theta, q(\theta))\gamma(\theta)d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \bar{u}(\theta)d\gamma(\theta),$$

so that the seller's objective function can be expressed as

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[s(\theta, q(\theta)) + \frac{F(\theta) - \gamma(\theta)}{f(\theta)} v_\theta(\theta, q(\theta)) \right] f(\theta)d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \bar{u}(\theta)d\gamma(\theta), \tag{15}$$

where $\int_{\underline{\theta}}^{\bar{\theta}} \bar{u}(\theta)d\gamma(\theta)$ is an irrelevant constant that we can drop without affecting the solution to the problem. □

Consider the standard nonlinear pricing model in which the IR constraints reduce to $u(\theta) \geq \bar{u}$, where $\bar{u}$ is consumers' reservation utility. The following result is immediate.

Result 3. *In the standard nonlinear pricing model, $\bar{u}(\theta) = \bar{u}$ and $\gamma(\theta) = 1$ for all θ.*

Proof of Result 3: Note that any incentive compatible allocation implies that $u(\theta)$ is strictly increasing in θ, since $u'(\theta) = v_\theta(\theta, q(\theta)) > 0$. Thus, if the IR constraints are to bind, they must bind just for the lowest type and be slack for higher types, that is, $d\gamma(\theta) = 0$ for $\theta > \underline{\theta}$. Note that if the IR constraint did not bind for the lowest type, $\underline{\theta}$, then the seller could increase profits by increasing $t(\underline{\theta})$ until the constraint binds. Next, since $\gamma(\theta)$ is weakly increasing and has the properties of a cumulative distribution function, it follows that $\gamma(\bar{\theta}) = 1 = \int_{\underline{\theta}}^{\bar{\theta}} d\gamma(\theta)$. But if $d\gamma(\theta) = 0$ for all $\theta > \underline{\theta}$, then $\gamma(\theta)$ must have a mass point at $\underline{\theta}$ with $\gamma(\underline{\theta}) = 1$, since $\gamma(\bar{\theta}) = 1 = \int_{\underline{\theta}}^{\bar{\theta}} d\gamma(\theta)$. □

Consider now the three main assumptions of Jullien (2000). The first assumption, potential separation, requires $l(\gamma, \theta)$ to be a weakly increasing function of θ for all $\gamma \in [0, 1]$, sufficient conditions for which are

$$\frac{\partial}{\partial\theta}\left(\frac{s_q(\theta, q)}{v_{\theta q}(\theta, q)}\right) \geq 0 \text{ and } \frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0 \geq \frac{d}{d\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right).$$

The second assumption, homogeneity, is a critical one. The easiest way to understand it is to imagine that the reservation utility $\bar{u}(\theta)$ is generated by consumers trading with an "outside" seller who offers an incentive-compatible menu $\{\bar{t}(\theta), \bar{q}(\theta)\}$ such that $\bar{u}(\theta) = v(\theta, \bar{q}(\theta)) - \bar{t}(\theta)$. Then, homogeneity amounts to assuming that the outside seller offers a locally incentive compatible menu that achieves $\bar{u}(\theta)$ for each consumer type in that

$$\bar{u}'(\theta) = v_\theta(\theta, \bar{q}(\theta)) \text{ and } \bar{q}(\theta) \text{ is weakly increasing.} \tag{16}$$

Formally, the assumption requires that the reservation utility be implementable through an incentive compatible schedule, $\{\bar{q}(\theta)\}$. Technically, given the assumption that $v_{\theta q}(\theta, q) > 0$, condition (16) requires the reservation utility $\bar{u}(\theta)$ be sufficiently convex.[1]

The third assumption, full participation, simply ensures that the seller can make nonnegative profits when trading with each consumer type so that all consumers participate. A sufficient condition for this assumption to hold is that homogeneity holds and for each type θ, the seller can make weakly positive profits by supplying the reservation quantity $\bar{q}(\theta)$ at price $\bar{t}(\theta)$ to each type so that a consumer of type θ

[1]Under the assumption that $v_{\theta\theta}(\cdot, \cdot) \geq 0$, which is typically made in the literature, it follows that $0 \leq \bar{u}''(\theta) = v_{\theta\theta}(\theta, \bar{q}(\theta)) + v_{\theta q}(\theta, \bar{q}(\theta))\bar{q}'(\theta)$, since $v_{\theta q}(\cdot, \cdot) > 0$ and $\bar{q}'(\theta) \geq 0$ by assumption. When, for instance, $v(\theta, q) = \theta\nu(q)$, the homogeneity assumption just requires convex reservation utilities since $\bar{u}''(\theta) = \nu'(\bar{q}(\theta))\bar{q}'(\theta) \geq 0$.

obtains the utility $\bar{u}(\theta) = v(\theta, \bar{q}(\theta)) - \bar{t}(\theta)$, that is,

$$\bar{t}(\theta) - c(\bar{q}(\theta)) = v(\theta, \bar{q}(\theta)) - c(\bar{q}(\theta)) - \bar{u}(\theta) \geq 0. \tag{17}$$

Result 4 (Jullien's Theorem 1)**.** *Under the assumptions of potential separation, homogeneity, and full participation, there exists a unique optimal allocation with full participation. An implementable allocation $\{u(\theta), q(\theta)\}$ is optimal if, and only if, there exists a cumulative distribution function $\gamma(\theta)$ on $[\underline{\theta}, \overline{\theta}]$ such that the first-order conditions (4) and the complementary slackness condition (5) are satisfied. Moreover, $q(\theta)$ is continuous.*

We now state Jullien's characterization of the optimal menu for the highly convex and weakly convex cases, respectively. To do so, let $\Theta_B = \{\theta : l(1, \theta) \leq \overline{q}(\theta) \leq l(0, \theta)\}$ denote the set of types such that for each such type θ, there exists a reservation multiplier $\bar{\gamma}(\theta)$ between zero and one that can support the reservation quantity as a solution to the seller's first-order condition. Thus, if we restrict attention to the set of types in Θ_B, then the only remaining condition that needs to be met for the reservation multiplier $\bar{\gamma}(\theta)$ to be a legitimate cumulative multiplier is that $\bar{\gamma}(\theta)$ be increasing on Θ_B, where

$$\bar{\gamma}(\theta) = F(\theta) + \frac{v_q(\theta, \bar{q}(\theta)) - c'(\bar{q}(\theta))}{v_{\theta q}(\theta, \bar{q}(\theta))} f(\theta).$$

For such a set, if

$$d\overline{q}(\theta)/d\theta \geq l_\theta(\overline{\gamma}(\theta), \theta) \text{ on } \Theta_B, \tag{18}$$

then the highly convex case applies, whereas if

$$d\overline{q}(\theta)/d\theta \leq l_\theta(\overline{\gamma}(\theta), \theta) \text{ on } \Theta_B, \tag{19}$$

then the weakly convex case applies. To interpret these conditions, note that differentiating $\bar{q}(\theta) = l(\bar{\gamma}(\theta), \theta)$ yields

$$\bar{\gamma}'(\theta) = \frac{d\overline{q}(\theta)/d\theta - l_\theta(\overline{\gamma}(\theta), \theta)}{l_\gamma(\overline{\gamma}(\theta), \theta)}. \tag{20}$$

Since $l_\gamma(\gamma, \theta) < 0$ whenever $l(\gamma, \theta) > 0$, condition (18) implies that $\bar{\gamma}(\theta)$ is decreasing on Θ_B so that the reservation multiplier cannot be legitimate for any interior type. Condition (19) implies that $\bar{\gamma}(\theta)$ is increasing on Θ_B so that the reservation multiplier is legitimate for all types in Θ_B. Hence, under (18), participation constraints cannot bind for any interior type whereas under (19), they bind for all types in Θ_B. Under the assumptions of potential separation, homogeneity, and full participation, Jullien's Propositions 2 and 3 imply the following result:

Result 5 (Jullien's Propositions 2 and 3)**.** *Under (18), the highly convex case applies so that there exists a constant γ such that $q(\theta) = l(\gamma, \theta)$. Under (19), if $\overline{q}(\cdot)$ is continuous and Θ_B is nonempty, then Θ_B is an interval $[\theta_1, \theta_2]$ and the weakly convex case applies so that $q(\theta) = l(0, \theta)$ for $\theta < \theta_1$, $q(\theta) = \bar{q}(\theta)$ for $\theta_1 \leq \theta \leq \theta_2$, and $q(\theta) = l(1, \theta)$ for $\theta > \theta_2$.*

A.2 Details of Examples in the Paper

Setup of Example 1. We provide here details about Example 1 in the paper. Suppose utility is HARA and given by $\nu(q) = (1 - d)[aq/(1 - d) + b]^d/d$, with $a > 0$, $aq/(1 - d) + b > 0$, and $d < 1$, so that $\nu'(q) = a[aq/(1-d)+b]^{d-1}$ and $\nu''(q) = -a^2[aq/(1-d)+b]^{d-2}$. From the seller's first-order condition, it follows that $\nu'(q) = cf(\theta)/[\theta f(\theta) + F(\theta) - \gamma(\theta)]$. So, the quantity $q(\theta) = l(\gamma(\theta), \theta)$ implied by the augmented model is

$$q(\theta) = l(\gamma(\theta), \theta) = \frac{(1-d)}{a} \left[\frac{a\theta f(\theta) + aF(\theta) - a\gamma(\theta)}{cf(\theta)} \right]^{\frac{1}{1-d}} - \frac{b(1-d)}{a}.$$

The quantity $q_s(\theta) = l(1, \theta)$ implied by the standard nonlinear pricing model, instead, satisfies $\nu'(q) = cf(\theta)/[\theta f(\theta) + F(\theta) - 1]$. The first-best quantity $q_{FB}(\theta)$ solves $\theta\nu'(q_{FB}(\theta)) = c$.

As for the linear monopoly quantity and price, from $\theta\nu'(q_m(\theta)) = p_m$ it follows

$$q_m(\theta) = \frac{(1-d)}{a} \left[\left(\frac{a\theta}{p_m} \right)^{\frac{1}{1-d}} - b \right]$$

and $|\varepsilon_{PQ}| = E_\theta[A(q_m(\theta))^{-1}]/E_\theta[q_m(\theta)] = 1/(1-d) + b/\{aE_\theta[q_m(\theta)]\}$, where $A(\cdot)$ is the coefficient of absolute risk aversion. Assume that $b = 0$ so that $|\varepsilon_{PQ}| = (1-d)^{-1}$, $p_m = c/d$, and

$$q_m(\theta) = \frac{(1-d)}{a} \left(\frac{ad\theta}{c} \right)^{\frac{1}{1-d}}.$$

Using the fact that $u_m(\theta) = \theta\nu(q_m(\theta)) - p_m q_m(\theta)$, we obtain

$$u_m(\theta) = \frac{(1-d)^2}{d^{\frac{1-2d}{1-d}}} \left(\frac{a}{c} \right)^{\frac{d}{1-d}} \theta^{\frac{1}{1-d}}.$$

From $u_s(\theta) = \overline{u}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \nu(q_s(x))dx$, where $\overline{u}(\underline{\theta})$ is intended to be consumers' reservation utility for all types under the standard model, it also follows

$$u_s(\theta) = \overline{u}(\underline{\theta}) + \frac{a^{\frac{d}{1-d}}(1-d)}{d} \int_{\underline{\theta}}^{\theta} \left[\frac{xf(x) + F(x) - 1}{cf(x)} \right]^{\frac{d}{1-d}} dx.$$

Note that $q_m(\theta) \geq l(\gamma(\theta), \theta)$ when $b = 0$ if, and only if, $[\gamma(\theta) - F(\theta)]/f(\theta) \geq (1-d)\theta$.

If the type distribution is uniform with $f(\theta) = 1/(\overline{\theta} - \underline{\theta})$ and $F(\theta) = (\theta - \underline{\theta})/(\overline{\theta} - \underline{\theta})$, then

$$u_s(\theta) = \overline{u}(\underline{\theta}) + \frac{a^{\frac{d}{1-d}}(1-d)^2}{2c^{\frac{d}{1-d}}d} \left[\left(2\theta - \overline{\theta}\right)^{\frac{1}{1-d}} - \left(2\underline{\theta} - \overline{\theta}\right)^{\frac{1}{1-d}} \right].$$

Similarly, since $u(\theta) = \overline{u}(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \nu(q(x))dx$, in the highly convex case of our model we obtain

$$u(\theta) = \overline{u}(\underline{\theta}) + \frac{a^{\frac{d}{1-d}}(1-d)}{c^{\frac{d}{1-d}}d} \int_{\underline{\theta}}^{\theta} \left[\frac{xf(x) + F(x) - \gamma}{f(x)} \right]^{\frac{d}{1-d}} dx = \overline{u}(\underline{\theta}) + \frac{a^{\frac{d}{1-d}}(1-d)^2}{2c^{\frac{d}{1-d}}d}$$

$$\cdot \left\{ \left[2\theta - \underline{\theta} - \gamma(\overline{\theta} - \underline{\theta})\right]^{\frac{1}{1-d}} - \left[\underline{\theta} - \gamma(\overline{\theta} - \underline{\theta})\right]^{\frac{1}{1-d}} \right\}.$$

Assume that $\overline{u}(\underline{\theta}) = 0$. Then $u_s(\theta) \geq u_m(\theta)$ if, and only if, $\left(2\theta - \overline{\theta}\right)^{\frac{1}{1-d}} \geq \left(2\underline{\theta} - \overline{\theta}\right)^{\frac{1}{1-d}} + 2d^{\frac{d}{1-d}}\theta^{\frac{1}{1-d}}$. Instead, $u(\theta) \geq u_m(\theta)$ if, and only if,

$$\left[2\theta - \underline{\theta} - \gamma(\overline{\theta} - \underline{\theta})\right]^{\frac{1}{1-d}} - \left[\underline{\theta} - \gamma(\overline{\theta} - \underline{\theta})\right]^{\frac{1}{1-d}} \geq 2d^{\frac{d}{1-d}}\theta^{\frac{1}{1-d}}.$$

These are the calculations behind Example 1 in the paper.

Example of Impact of Cash Transfers on Unit Prices with HARA Utility Function. We provide here omitted details of the example discussed in the paper after Corollary 2. Since $T(q(\theta)) = \theta\nu(q(\theta)) - u(\theta)$, it follows that

$$T(q(\theta)) = \frac{(1-d)}{d}\left(\frac{a}{c}\right)^{\frac{d}{1-d}}\left[2\theta-\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{d}{1-d}}\left\{\theta-\frac{(1-d)}{2}\left[2\theta-\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]\right\}$$

$$+\frac{(1-d)^2}{2d}\left(\frac{a}{c}\right)^{\frac{d}{1-d}}\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{1}{1-d}}$$

in the highly convex case, with the price per unit $p(q(\theta)) = T(q(\theta))/q(\theta)$ given by

$$p(q(\theta)) = \frac{c}{2d\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]}\left\{2\theta d+(1-d)\left[\underline{\theta}+\gamma(\overline{\theta}-\underline{\theta})\right]+\frac{(1-d)\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{1}{1-d}}}{\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{d}{1-d}}}\right\}.$$

By straightforward algebraic manipulations, it is easy to show that

$$\frac{\partial p(q(\theta))}{\partial\gamma} = \frac{c\left(\overline{\theta}-\underline{\theta}\right)}{2d\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^2}\left\{2\theta d+(1-d)\left[\underline{\theta}+\gamma(\overline{\theta}-\underline{\theta})\right]+\frac{(1-d)\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{1}{1-d}}}{\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{d}{1-d}}}\right\}$$

$$+\frac{c(\overline{\theta}-\underline{\theta})}{2d\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]}\left\{1-d+\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{d}{1-d}}\frac{\left[(1+d)\underline{\theta}-2\theta+(1-d)\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]}{\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{1}{1-d}}}\right\}.$$

Moreover, using the expression for $q(\theta)$, that is,

$$q(\theta) = \frac{(1-d)}{a}\left(\frac{a}{c}\right)^{\frac{1}{1-d}}\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{1}{1-d}},$$

to obtain the inverse function $\theta(q)$, it follows that

$$p(q) = \frac{c}{2}+\frac{(1-d)^{1-d}}{2d}\left\{\frac{a^{\frac{d}{1-d}}(1-d)^{1+d}\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{1}{1-d}}}{c^{\frac{d}{1-d}}q}+\frac{a^d\left[\underline{\theta}+\gamma(\overline{\theta}-\underline{\theta})\right]}{q^{1-d}}\right\}.$$

It is immediate that $p(q) > 0$ and $p(q)$ decreases with quantity when $d < 0$. Now, observe that

$$\frac{\partial p(q(\underline{\theta}))}{\partial\gamma} = \frac{c\underline{\theta}\left(\overline{\theta}-\underline{\theta}\right)}{d\left[\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^2},$$

which is negative when $d < 0$, whereas

$$\frac{\partial p(q(\overline{\theta}))}{\partial\gamma} = \frac{c\left(\overline{\theta}-\underline{\theta}\right)}{2d\left[2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^2}\left\{2\overline{\theta}d+(1-d)\left[\underline{\theta}+\gamma(\overline{\theta}-\underline{\theta})\right]+\frac{(1-d)\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{1}{1-d}}}{\left[2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{d}{1-d}}}\right\}$$

$$+\frac{c(\overline{\theta}-\underline{\theta})}{2d\left[2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]}\left\{1-d+\left[\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right]^{\frac{d}{1-d}}\frac{\left[(1+d)\underline{\theta}-2\overline{\theta}+(1-d)\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]}{\left[2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{\frac{1}{1-d}}}\right\},$$

which is positive when $d<0$ as long as

$$\frac{\overline{\theta}}{\overline{\theta}-\underline{\theta}}<\left[\frac{2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)}{\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})}\right]^{\frac{\widetilde{d}}{1+\widetilde{d}}}, \tag{21}$$

$\widetilde{d}=-d>0$, which is satisfied if, for instance, $\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})$ is chosen small enough. Since, as argued in the paper, γ decreases after an increase in income, and $\partial p(q(\underline{\theta}))/\partial\gamma<0$ whereas $\partial p(q(\overline{\theta}))/\partial\gamma>0$, then such an increase implies an increase in unit prices at *low* quantities and a decrease in unit prices at *high* quantities. Note that (21) is equivalent to

$$\frac{\widetilde{d}}{1+\widetilde{d}}>\frac{\log\left(\frac{\overline{\theta}}{\overline{\theta}-\underline{\theta}}\right)}{\log\left[\frac{2\overline{\theta}-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)}{\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})}\right]},$$

so a sufficient condition for (21) is $\widetilde{d}/(1+\widetilde{d})>\log\left(\frac{\overline{\theta}}{\overline{\theta}-\underline{\theta}}\right)/\log\left(\frac{\overline{\theta}}{\underline{\theta}}\right)$, which is easily satisfied for d small enough. Observe, finally, that

$$\frac{\partial p(q(\theta))}{\partial\gamma\partial\theta}=\frac{c\left(\overline{\theta}-\underline{\theta}\right)}{d\left[2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]^{3}}\left\{-2\theta-\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})\right.$$

$$\left.+\left[\frac{\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})}{2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)}\right]^{\frac{d}{1-d}}\left[\frac{2\theta+(d-3)\underline{\theta}+(1-d)\gamma\left(\overline{\theta}-\underline{\theta}\right)}{1-d}\right]\right\}.$$

Thus, $\partial p(q(\theta))/\partial\gamma\partial\theta\geq 0$ is equivalent to

$$\left[\frac{2\theta-\underline{\theta}-\gamma\left(\overline{\theta}-\underline{\theta}\right)}{\underline{\theta}-\gamma(\overline{\theta}-\underline{\theta})}\right]^{\frac{\widetilde{d}}{1+\widetilde{d}}}\left[2\theta-(\widetilde{d}+3)\underline{\theta}+(1+\widetilde{d})\gamma\left(\overline{\theta}-\underline{\theta}\right)\right]\leq(1+\widetilde{d})[2\theta+\underline{\theta}+\gamma\left(\overline{\theta}-\underline{\theta}\right)], \tag{22}$$

which is always satisfied if

$$2\theta+(1+\widetilde{d})\gamma\left(\overline{\theta}-\underline{\theta}\right)-(3+\widetilde{d})\underline{\theta}\leq 0\Leftrightarrow 2\theta\leq(3+\widetilde{d})\underline{\theta}-(1+\widetilde{d})\gamma(\overline{\theta}-\underline{\theta}).$$

With $\underline{\theta}=1$ and $\overline{\theta}=\underline{\theta}+1$, this last displayed inequality becomes $2\theta\leq 3+\widetilde{d}-(1+\widetilde{d})\gamma$. If $\gamma=1/2$, this latter inequality reduces to $2\theta\leq 5/2+\widetilde{d}/2$ and a sufficient condition is $4\leq 5/2+\widetilde{d}/2$ or $\widetilde{d}\geq 3$. When $\underline{\theta}=1$, $\overline{\theta}=\underline{\theta}+1$, and $\gamma=1/2$, it also follows that (21) reduces to

$$2<\left(\frac{3-\gamma}{1-\gamma}\right)^{\frac{\widetilde{d}}{1+\widetilde{d}}}\Leftrightarrow\frac{\widetilde{d}}{1+\widetilde{d}}>\frac{\log(2)}{\log\left(\frac{3-\gamma}{1-\gamma}\right)}=\frac{\log(2)}{\log\left(5\right)}.$$

Hence, (21) and (22) can both be easily satisfied.

A.3 An Oligopoly Model with Price Discrimination

Consider a market (village) in which two or more identical firms with the same cost functions compete in nonlinear price-quantity menus to exclusively serve any given consumer. We assume that over the relevant period of time—a week, in light of the frequency of our data—a consumer only purchases from one seller. Consumers' preferences and characteristics are as described in the paper in Section 3. Here we prove the following result: regardless of the pattern of individual consumers' purchases across sellers, equilibrium prices and quantities can be characterized as the solution to the problem of a single seller that we focused on in the paper. Hence, in this precise sense, our construction entails no loss of generality.

Formally, the strategy of each firm $j = 1, \ldots, J$ consists of the offer of an incentive compatible menu $\{t_j(\theta), q_j(\theta)\}$ for all consumer types. Let

$$u_j(\theta) = v(\theta, q_j(\theta)) - t_j(\theta)$$

denote the utility of a consumer of type θ from choosing to purchase from firm j. The strategy of each consumer simply consists of choosing which firm to visit. Conditional on consumers' visits, each firm offers incentive compatible menus that encode the optimal purchasing strategies of consumers, so there is no need to separately specify a consumer's purchasing strategy conditional on a visit.

A strategy for a consumer is a vector $x(\theta) = (x_1(\theta), \ldots, x_J(\theta))$ of probabilities of visiting firms with $\sum_j x_j(\theta) = 1$ and $x_j(\theta) \geq 0$. Clearly, given a vector of utilities associated with purchasing from any of the firms, $u(\theta) = (u_1(\theta), \ldots, u_J(\theta))$, the best response of a consumer of type θ satisfies

$$x_j(\theta, u(\theta)) = \begin{cases} 0 \text{ if } u_j(\theta) < \overline{u}_j(\theta) = \max_{i \neq j} u_i(\theta) \\ \geq 0 \text{ otherwise.} \end{cases} \tag{23}$$

Here $\overline{u}_j(\theta)$ is the highest utility of a consumer of type θ from purchasing from some firm *other than* firm j. This best response takes into account the fact that a consumer will never visit a firm that offers a contract yielding a strictly lower utility than some other firm. For simplicity, if the consumer is indifferent between visiting, say, k firms, then we posit that the consumer visits each firm with probability $1/k$.

Given the best response of each consumer, denoted by $x = \{x_j(\theta, u(\theta))\}_{j,\theta}$, and the offered utilities of all other firms, $u_{-j} = \{u_i(\theta)\}_{i \neq j, \theta}$, the problem of firm j consists of choosing an incentive compatible menu to maximize expected profits,

$$\max_{\{t_j(\theta), q_j(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} x_j(\theta, u(\theta))[t_j(\theta) - c(q_j(\theta))] f(\theta) d\theta, \tag{24}$$

subject to the incentive constraints

$$\text{(IC)} \;\; v(\theta, q_j(\theta)) - t_j(\theta) \geq v(\theta, q_j(\theta')) - t_j(\theta') \text{ for any } \theta, \theta', \text{ if } x_j(\theta, u(\theta)) > 0.$$

An equilibrium is a set of best responses for each type of consumer, $\{x_j(\theta, u(\theta))\}_{j,\theta}$, together with incentive-compatible menus and offered utilities for each firm, $\{t_j(\theta), q_j(\theta), u_j(\theta)\}_{j,\theta}$, which satisfy (23) and (24).

Clearly, we can use (23) to write out a firm's profit maximization problem directly as

$$\text{(O problem)} \quad \max_{\{t_j(\theta),q_j(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} x_j(\theta, u(\theta))[t_j(\theta) - c(q_j(\theta))]f(\theta)d\theta \ \text{ s.t.}$$
$$\text{(IC)} \ v(\theta, q_j(\theta)) - t_j(\theta) \geq v(\theta, q_j(\theta')) - t_j(\theta') \text{ for any } \theta, \theta', \text{ if } x_j(\theta, u(\theta)) > 0$$
$$\text{(IR')} \ v(\theta, q_j(\theta)) - t_j(\theta) \geq \overline{u}_j(\theta) \text{ for any } \theta, \text{ if } x_j(\theta, u(\theta)) > 0.$$

In an equilibrium with identical sellers, it follows that $\overline{u}_j(\theta)$ is independent of j and can be written as $\overline{u}(\theta)$. Also, $x_j(\theta, u(\theta)) = 1/J$, that is, consumers of each type θ randomize equally among all J firms in their visits. With $\overline{u}_j(\theta) = \overline{u}(\theta)$ and omitting the multiplicative constant $1/J$ from a seller's objective function, this oligopoly problem reduces to the IR problem in the paper. We formally state this result in the next proposition.

Proposition 1. *Suppose the market is populated by a given number of sellers with identical cost function, $c(q)$. Then, equilibrium prices and quantities are solutions to the IR problem in the paper.*

The argument for Proposition 1 relies on sellers in a village competing to exclusively serve consumers. Competition in non-exclusionary nonlinear price schedules is beyond the scope of our paper. First, this type of competition does not conform to anecdotal evidence on consumption patterns in our data: typically, households purchase weekly from one seller only. Second, common approaches to characterize these problems rely on assumptions, like the ability of a seller to condition its prices on consumers' purchasing behavior with other sellers, that are counterfactual in our setting. See, for instance, Stole (2007).

A.4 Identification

A.4.1 Comparison with Existing Literature

Our identification argument builds on arguments common in the literature on the nonparametric identification of auctions and nonlinear pricing models (see Campo et al. (2011), Guerre et al. (2000) and Perrigne and Vuong (2010), cited in the paper). In auction models, the key object of interest, the distribution of bidders' valuations, is nonparametrically identified from the observed distribution of bids, based on the monotone relationship between bidders' valuations and actual bids that auction models usually imply—up to some knowledge of bidders' utility function if bidders are risk averse. Similarly, one of the key objects of interest in our analysis, the distribution of consumers' marginal willingness to pay, is nonparametrically identified from the observed distribution of quantity purchases (and the equilibrium price schedule), based on the monotone relationship between consumers' marginal willingness to pay and purchased quantities that our nonlinear pricing model gives rise to.

Unlike the estimation of common auction models, however, the estimation of nonlinear pricing models often involves the recovery of consumers' valuation of quantity. Like Perrigne and Vuong (2010), under the separability assumption $v(\theta, q) = \theta\nu(q)$ for consumers' utility, we identify both the distribution of θ, which describes a consumer's marginal willingness to pay, and consumers' "base" utility function, $\nu(q)$, just using information on the price schedule and the distribution of quantity purchases in a market, up to the coefficient of absolute risk aversion. We do so by exploiting the relationship between prices and quantities implied by a seller's first-order condition for the choice of quantities to offer and each consumer's first-order condition for the choice of quantity to buy.

A seller's problem, though, is more involved in our model than in the model of Perrigne and Vuong (2010). Unlike in the nonlinear pricing model that Perrigne and Vuong (2010) consider, the equilibrium price and quantity menu in our model depends not just on the distribution of θ, on $\nu(q)$, and on a seller's cost function, but also on the *endogenous* distribution of consumers who are indifferent between purchasing and not purchasing in a market, due to binding subsistence or participation constraints. As explained in the paper, consumers indifferent between purchasing and not purchasing need not be just consumers with the lowest possible marginal willingness to pay, as in the standard nonlinear pricing model. Since the characteristics of this group of "marginal" consumers affect a seller's choice of prices and quantities, the recovery of the primitives of interest requires the identification of the distribution of such consumers. We show that the empirical content of our model is rich enough to allow us to identify (and estimate) the distribution of "marginal consumers" and so to identify key primitives of our model.

Like in Perrigne and Vuong (2010), in our setup, the identification of more general preference structures than the separable one just described is made difficult by the need to recover the dependence of utility on *both* quantity and marginal willingness to pay, just based on first-order conditions for the optimal choice of quantity. In particular, only the first derivative of the utility function with respect to quantity, $v_q(\theta, q)$, and the cross-partial derivative of the utility function with respect to type and quantity, $v_{\theta q}(\theta, q)$, appear in the equilibrium conditions on which identification relies—respectively, in consumers' and sellers' first-order conditions ($v_q(\theta, q)$) or just in sellers' first-order conditions ($v_{\theta q}(\theta, q)$). Moreover, the cross-partial derivative $v_{\theta q}(\theta, q)$ only appears interacted with other unobserved primitives or unobserved endogenous variables. Hence, the identification of more general preferences typically requires additional information. Naturally, information on consumers' reservation utility in a market, $\overline{u}(\theta)$, would be sufficient to nonparametrically identify $v_\theta(\theta, q)$ under our homogeneity assumption, since $\overline{u}'(\theta) = v_\theta(\theta, q)$ under this assumption. By the same argument as in the paper, $v_q(\theta, q)$ is identified from the marginal price schedule, $T'(q)$. Alternatively, knowledge of marginal cost in a market allows to nonparametrically set- and point-identify some of the primitives of interest in this more general case, even without knowledge of $\overline{u}(\theta)$. We establish this point in the next subsection.

Our arguments also bear similarities with those in the hedonic pricing literature. See, in particular, Ekeland et al. (2004) and Heckman et al. (2010), cited in the paper. In these papers too, marginal payoff functions are not identified without further restrictions, in addition to the equilibrium conditions on the behavior of both sides of the market. The first paper proves identification of marginal payoff functions and the distribution of unobserved heterogeneity up to location and scale, under the assumption of an additively separable marginal payoff structure. Specifically, they consider nonparametric hedonic models with additive marginal utility and additive marginal product functions, and show that hedonic models with these additivity restrictions are nonparametrically identified based on single market data. No heterogeneity in the curvature of preference or production functions is allowed.

The second paper examines alternative identifying assumptions on the functional form of marginal payoff functions, combined with exogenous variables, for data from single and multiple markets. This second paper proves the nonparametric identification of structural functions and distributions in general hedonic models without imposing additivity. For instance, the authors allow the curvature of the marginal utility for a product attribute, and the distribution of marginal utilities, to vary in general ways across agents with different observed characteristics.

In analogy to these papers, we also investigate the identifiability of nonparametric structural relationships with nonadditive heterogeneity and assess which features of nonlinear pricing models can be identified from observations on equilibrium outcomes in a single market, under relatively mild assumptions that are common in the empirical auction and nonlinear pricing literature. Like these authors, we rely on separability conditions for the identification of consumer marginal utility. Note that the strategy of relying on multi-market data to achieve identification is less appealing in our setup, given the potential variation of consumers' marginal willingness to pay and reservation utility across markets, as our estimates confirm.

Our identification problem, however, differs from the one in those two papers in three important respects. First, to establish identification, those papers exploit the existence of exogenous covariates independent of the unobserved heterogeneity term of interest—here, consumers' marginal willingness to pay. The existence of such exogenous variables is unlikely in our case, as most household characteristics, such as the demographic composition of a household, are highly correlated with consumers' purchasing behavior in the data.

Second, in our case not just the cumulative distribution or probability density function of the heterogeneity distribution is unknown but also its support, which compounds the identification problem since knowledge of this support is crucial in identifying marginal utility.

Third, consumers' reservation utilities depend on their unobserved characteristics, which makes the participation constraint non-redundant for consumers with potentially *any* value of the unobserved characteristic, θ. Hence, a consumer's problem in our model is a mixed discrete-continuous choice problem of whether to participate and, conditional on participation, which quantity to choose, whose solution is not just characterized by the first-order conditions for optimal consumption. Similarly, a seller's problem consists in deciding whether to induce a consumer to trade and, if so, for which price and quantity combination. As shown, the interaction between participation (or budget) and incentive constraints in a seller's problem is crucial for the characteristics of the equilibrium price schedule and for its dependence on primitives. Importantly, as argued in the paper, this feature of our model is key to the distributional implications of nonlinear pricing, which are the focus of our analysis.

A.4.2 Identification of General Preference Structures

We show here which primitives of our model in a market can be identified with knowledge of a seller's marginal cost of the total quantity provided of a good, $c'(Q)$, once the separability assumption $v(\theta, q) = \theta\nu(q)$ is relaxed. Without loss of generality, consider the IR problem. Then, the primitives left to identify are $v(\theta, q)$, $F(\theta)$, the support of θ, $f(\theta)$, and $\overline{u}(\theta)$. Recall that the general form of a seller's and a consumer's first-order conditions is, respectively,

$$v_q(\theta, q) - c'(Q) = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q) \text{ and } T'(q) = v_q(\theta, q),$$

which imply

$$T'(q) - c'(Q) = \frac{\gamma(\theta) - F(\theta)}{f(\theta)} v_{\theta q}(\theta, q). \tag{25}$$

From $T'(q) = v_q(\theta, q)$, it follows that $v_q(\theta, q)$ is nonparametrically identified without the need of any further restriction. Since $\gamma(\theta(q)) \geq G(q)$ if, and only if, $T(q) \geq c'(Q)$, it also follows

$$\gamma(\theta(q)) \in \begin{cases} [0, G(q)) \text{ at any } q \text{ s.t. } T'(q) < c'(Q) \\ (G(q), 1] \text{ at any } q \text{ s.t. } T'(q) > c'(Q) \end{cases} \text{ and } \gamma(\theta(q)) = G(q) \text{ at any } q \text{ s.t. } T'(q) = c'(Q). \tag{26}$$

Note that these bounds are tight in that the set of values of $\gamma(\theta(q))$ they imply covers the identified set. As discussed in the paper, $F(\theta)$ is identified by $G(q)$.

Result 6. *Suppose that $c'(Q)$ is known. Then, $v_q(\theta, q)$ is identified from the marginal price schedule and $\gamma(\theta)$ is identified from the marginal price schedule and the distribution of quantity purchases by* (26). *If $\gamma(\theta)$ is identified, then the support of θ (up to $\underline{\theta}$), $f(\theta)$, and $v_{\theta q}(\theta, q)$ are also identified.*

Proof of Result 6: The first part of the claim is immediate from the discussion preceding it. As for the rest of the claim, observe that (25) implies that $v_{\theta q}(\theta, q)/f(\theta)$ is identified once $\gamma(\theta)$ is identified since $F(\theta) = G(q)$. If $\gamma(\theta)$ is identified, then $\theta(q)$ is identified up to $\underline{\theta}$ by the same argument as in the paper. Hence, $f(\theta)$ is identified too from the observed distribution of quantities, since $f(\theta) = g(q)/\theta'(q)$. Finally, $v_{\theta q}(\theta, q)$ is identified from $v_{\theta q}(\theta, q)/f(\theta)$. □

The following result is immediate by the homogeneity assumption, which implies $\overline{u}'(\theta) = v_\theta(\theta, q)$, and Result 6.

Result 7. *Suppose that $\overline{u}'(\theta)$ is known. Then, $v_q(\theta, q)$ is identified from the marginal price schedule and $v_\theta(\theta, q)$ is identified from $\overline{u}'(\theta)$.*

A.5 Omitted Estimation Results

Here we graph the estimates of the probability density function of consumer types for each good, namely, rice, kidney beans, and sugar, and in each village, defined as a Mexican municipality or as a Mexican locality. In the tables that follow, we report the t-statistics of the estimates of the model's primitives and the cumulative multiplier associated with consumers' participation (or budget) constraints obtained from villages defined as localities. We note that the estimates of the probability density function of consumer types are very similar across the two specifications of the multiplier function and the two definitions of villages.

A.5.1 Estimates of the Probability Density Functions of Consumer Types

In Figure 1, we plot the estimates of the probability density function of consumer types for each good and village estimated from villages defined at the level of the Mexican *municipality*. We graph the estimates obtained for the linear specification of the index of the multiplier function in the top panels and for the quadratic specification of the index of the multiplier function in the bottom panels. See Section 4 in the paper for details. In Figure 2, we plot the corresponding estimates from villages defined at the level of the Mexican *locality*.

Figure 1: Estimated Density Function of Types from Municipalities (Linear and Quadratic Specification)

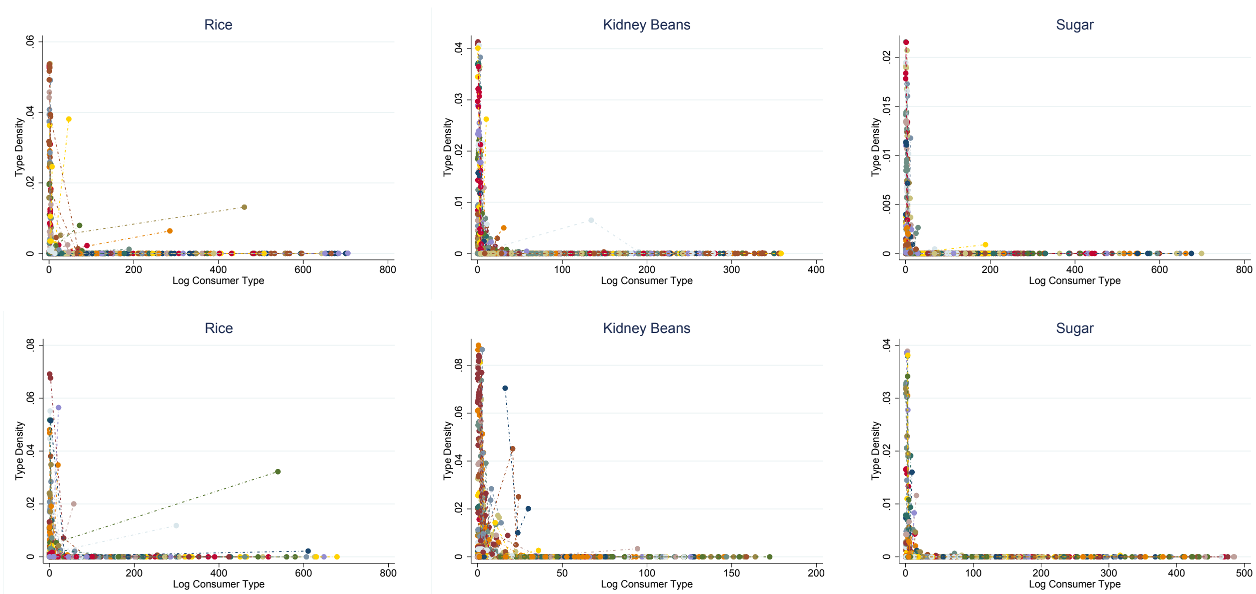
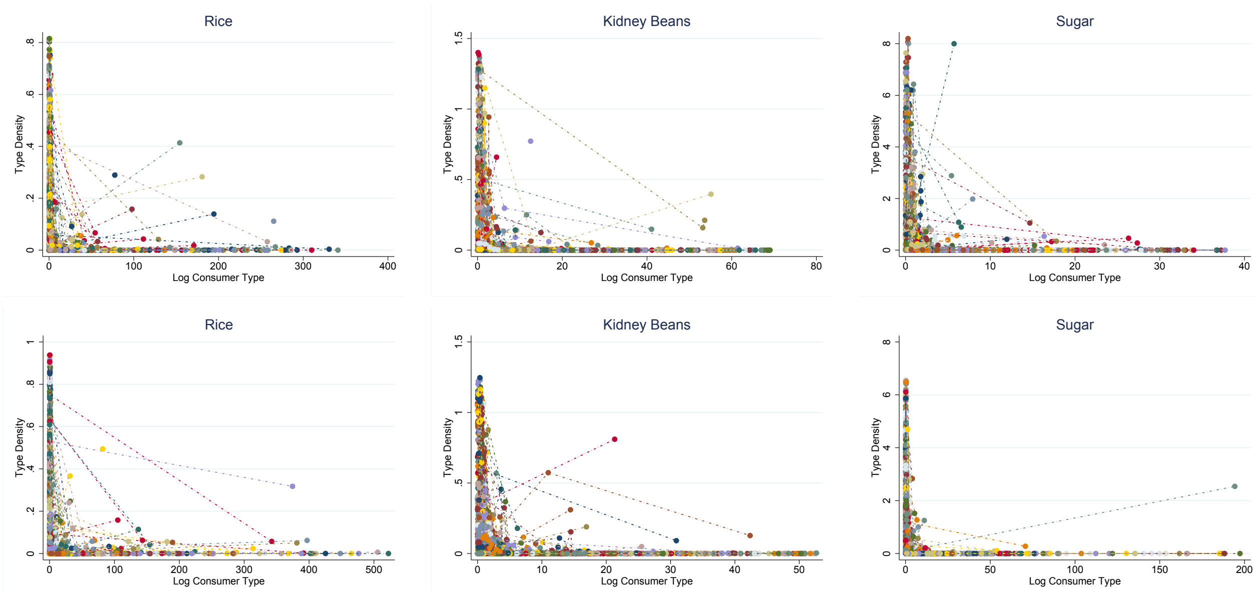


Figure 2: Estimated Density Function of Types from Localities (Linear and Quadratic Specification)



-

A.5.2 Estimates from Localities: Linear Specification of Multiplier

The following three tables report selected percentiles of the distribution of the t-statistics of the estimates of $c'(Q)$, $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ across villages. These statistics are meant to illustrate the overall precision of our estimates. The next three tables report the quartiles of the distribution across villages of selected percentiles of the distribution across village-level quantities of the t-statistics of the estimates of $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$. These statistics are meant to show the variability across villages of the precision of the estimates of $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ at the different quantities of a good in a village. All estimates have been obtained assuming that the cumulative multiplier for each good in each village is a logistic function of quantity with a *linear* index.

Table 1: Percentiles of t-Statistics across Quantities and Villages for Rice (Linear)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.046	2.078	6.137	17.464	40.573	67.370	116.908	157.698	219.232
$\gamma(\theta(q))$	1.492	4.919	8.687	32.465	423.474	2.1×10^4	6.4×10^5	7.8×10^6	7.6×10^8
$\theta(q)$	0.018	0.199	0.477	1.421	4.305	11.729	28.350	46.836	168.314
$\nu'(q)$	-82.665	-24.026	-12.946	-3.655	-0.651	2.092	23.368	53.139	186.895
$f(\theta)$	1.118	1.118	1.118	2.739	6.801	9.487	12.698	14.433	17.783

Table 2: Percentiles of t-Statistics across Quantities and Villages for Beans (Linear)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.006	0.082	0.531	4.252	20.117	52.698	102.618	145.407	253.494
$\gamma(\theta(q))$	1.348	4.068	6.851	19.377	63.942	298.560	2266.634	1.2×10^4	1.4×10^5
$\theta(q)$	0.017	0.121	0.233	0.760	1.979	4.897	9.943	14.400	36.705
$\nu'(q)$	-9.354	-4.244	-2.207	-0.801	-0.044	6.205	35.734	80.199	288.988
$f(\theta)$	1.118	1.118	1.581	3.868	7.259	10.124	12.829	14.371	19.074

Table 3: Percentiles of t-Statistics across Quantities and Villages for Sugar (Linear)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.028	0.458	2.593	14.432	81.714	168.106	310.166	383.611	1079.864
$\gamma(\theta(q))$	0.918	2.967	5.710	20.727	115.921	5521.459	1.6×10^5	1.3×10^6	1.2×10^8
$\theta(q)$	0.028	0.191	0.366	1.182	2.622	5.918	12.318	20.804	68.186
$\nu'(q)$	-9.528	-4.086	-2.061	-0.269	2.538	24.648	117.050	238.115	593.421
$f(\theta)$	1.118	1.118	1.581	4.330	7.583	10.308	13.399	15.890	18.873

Table 4: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Rice (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.868	2.238	4.915	10.943	42.727	1288.063	3.9×10^4	1.6×10^5	3.3×10^6
	p_{50}	2.981	9.293	12.284	34.977	219.328	5564.058	1.9×10^5	1.0×10^6	1.0×10^7
	p_{75}	8.288	38.873	71.004	239.436	2347.055	4.1×10^4	5.3×10^6	5.7×10^7	4.3×10^9
$\theta(q)$	p_{25}	0.012	0.173	0.285	0.921	2.707	7.235	15.518	23.228	77.506
	p_{50}	0.018	0.345	0.650	1.567	4.539	12.344	25.603	39.546	139.740
	p_{75}	0.024	0.536	0.984	2.383	6.674	16.967	37.531	69.579	178.023
$\nu'(q)$	p_{25}	−102.280	−30.619	−19.462	−6.528	−1.591	−0.434	0.246	4.267	16.539
	p_{50}	−46.631	−20.421	−12.514	−3.381	−0.546	1.501	12.688	27.224	90.916
	p_{75}	−26.391	−12.352	−7.253	−1.479	0.066	11.390	39.017	75.014	291.182
$f(\theta)$	p_{25}	1.118	1.118	1.118	1.704	2.958	5.000	7.746	9.354	13.334
	p_{50}	2.236	3.133	3.716	5.181	7.086	9.403	12.361	14.186	16.956
	p_{75}	5.000	5.700	6.134	7.200	8.488	11.307	13.509	15.108	19.133

Table 5: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Beans (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.985	3.168	4.303	10.108	23.698	47.761	127.314	244.325	1.8×10^4
	p_{50}	2.309	8.971	15.326	29.391	58.219	148.088	398.864	1565.996	4.3×10^4
	p_{75}	6.627	26.890	47.324	95.732	272.696	1185.432	5181.648	2.2×10^4	8.8×10^4
$\theta(q)$	p_{25}	0.009	0.088	0.162	0.535	1.380	3.507	6.741	9.957	19.674
	p_{50}	0.023	0.165	0.325	0.946	2.117	4.857	9.461	12.323	32.044
	p_{75}	0.023	0.233	0.500	1.321	3.102	6.460	12.790	16.810	46.251
$\nu'(q)$	p_{25}	−11.395	−6.792	−4.108	−1.709	−0.538	−0.018	4.480	10.778	39.869
	p_{50}	−6.629	−3.462	−2.085	−0.757	−0.084	3.662	20.573	36.142	120.305
	p_{75}	−3.579	−1.742	−1.193	−0.220	1.259	12.897	64.953	149.110	348.020
$f(\theta)$	p_{25}	1.118	1.581	1.581	2.500	4.047	5.659	7.071	8.062	10.488
	p_{50}	2.456	4.031	4.748	6.124	7.972	9.848	12.443	14.457	21.215
	p_{75}	4.464	6.088	7.020	8.062	9.782	11.478	14.048	16.880	23.555

Table 6: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Sugar (Linear)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.418	2.249	3.338	8.486	24.133	476.983	2.8×10^4	1.7×10^5	3.3×10^6
	p_{50}	1.209	5.690	12.683	29.163	82.548	1569.923	6.3×10^4	5.2×10^5	5.3×10^6
	p_{75}	4.162	18.620	32.856	75.538	309.142	7259.133	2.2×10^5	3.3×10^6	2.5×10^8
$\theta(q)$	p_{25}	0.025	0.123	0.301	0.930	2.015	3.924	7.518	10.447	19.237
	p_{50}	0.028	0.218	0.466	1.334	2.832	5.879	9.549	15.821	34.383
	p_{75}	0.046	0.276	0.717	2.039	4.113	9.088	15.896	25.188	69.761
$\nu'(q)$	p_{25}	−10.931	−5.816	−4.288	−1.279	0.107	6.205	36.551	83.493	194.541
	p_{50}	−7.377	−2.711	−1.701	−0.240	2.677	25.290	78.036	139.464	306.498
	p_{75}	−3.424	−1.037	−0.314	0.516	10.199	64.145	199.016	257.552	628.068
$f(\theta)$	p_{25}	1.118	1.118	1.402	2.958	4.743	6.971	8.545	9.487	12.748
	p_{50}	2.927	3.987	5.054	6.296	7.906	10.092	13.555	15.969	18.337
	p_{75}	5.831	6.444	7.071	8.048	9.764	12.078	14.847	17.630	19.969

A.5.3 Estimates from Localities: Quadratic Specification of Multiplier

The following three tables report selected percentiles of the distribution of the t-statistics of the estimates of $c'(Q)$, $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$ across villages. The next three tables report the quartiles of the distribution across villages of selected percentiles of the distribution across village-level quantities of the t-statistics of the estimates of $\gamma(\theta(q))$, $\theta(q)$, $\nu'(q)$, and $f(\theta)$. All estimates have been obtained assuming that the cumulative multiplier for each good in each village is a logistic function of quantity with a *quadratic* index.

Table 7: Percentiles of t-Statistics across Quantities and Villages for Rice (Quadratic)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.197	1.918	5.199	16.728	37.022	65.785	111.986	156.545	225.224
$\gamma(\theta(q))$	1.035	3.547	6.640	24.053	232.322	6921.356	2.2×10^5	1.7×10^6	1.5×10^8
$\theta(q)$	0.030	0.256	0.544	1.591	4.553	14.492	41.178	74.819	191.568
$\nu'(q)$	−108.437	−28.931	−13.165	−3.345	−0.455	5.553	35.716	67.995	234.168
$f(\theta)$	1.118	1.118	1.118	2.727	6.794	9.487	12.550	14.287	17.655

Table 8: Percentiles of t-Statistics across Quantities and Villages for Beans (Quadratic)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.004	0.044	0.172	2.615	12.770	53.810	102.240	136.358	239.336
$\gamma(\theta(q))$	1.281	3.359	5.236	17.060	75.993	386.926	3984.676	2.5×10^4	2.0×10^5
$\theta(q)$	0.024	0.157	0.313	0.966	2.460	5.446	9.243	13.201	28.556
$\nu'(q)$	−9.701	−4.706	−2.639	−0.744	0.086	6.785	37.882	94.091	332.620
$f(\theta)$	1.118	1.118	1.581	3.873	7.364	10.062	13.078	15.810	20.321

Table 9: Percentiles of t-Statistics across Quantities and Villages for Sugar (Quadratic)

	p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$c'(Q)$	0.019	0.167	1.781	8.346	82.833	167.661	288.498	380.583	650.754
$\gamma(\theta(q))$	1.430	3.532	6.013	21.947	96.086	2019.979	6.5×10^4	3.2×10^5	2.4×10^8
$\theta(q)$	0.016	0.147	0.305	1.092	2.494	5.124	12.396	20.509	122.292
$\nu'(q)$	−25.074	−4.306	−2.269	−0.391	1.872	20.848	80.387	148.411	591.023
$f(\theta)$	1.118	1.118	1.581	4.330	7.665	10.460	13.463	15.572	18.884

Table 10: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Rice (Quadratic)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.310	2.528	4.166	9.665	51.009	790.278	1.9×10^4	7.1×10^4	1.3×10^6
	p_{50}	0.683	6.347	14.073	29.170	152.315	2299.945	6.0×10^4	5.3×10^5	3.0×10^6
	p_{75}	5.978	25.989	38.259	116.680	777.837	1.9×10^4	1.8×10^6	1.2×10^7	1.6×10^8
$\theta(q)$	p_{25}	0.025	0.172	0.346	0.984	2.972	7.786	19.299	29.881	66.676
	p_{50}	0.064	0.354	0.661	1.945	4.735	15.125	44.963	70.168	132.236
	p_{75}	0.183	0.821	1.229	2.970	8.701	24.310	64.569	100.735	215.639
$\nu'(q)$	p_{25}	-162.027	-50.724	-24.138	-5.239	-1.574	-0.360	0.156	3.172	13.915
	p_{50}	-95.212	-21.296	-13.165	-3.249	-0.364	2.376	13.803	23.804	70.276
	p_{75}	-46.973	-9.720	-4.752	-0.728	2.825	20.226	51.351	88.045	182.750
$f(\theta)$	p_{25}	1.118	1.118	1.118	1.759	2.958	5.375	8.211	12.247	14.185
	p_{50}	2.236	2.962	3.536	5.313	6.792	8.867	11.727	14.073	17.476
	p_{75}	5.000	5.590	6.058	7.245	8.637	10.949	13.155	14.958	20.456

Table 11: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Beans (Quadratic)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.822	2.250	3.422	7.292	18.897	60.567	149.344	268.732	2573.039
	p_{50}	1.820	4.227	9.847	24.617	67.114	169.289	414.829	1132.442	1.5×10^4
	p_{75}	4.032	18.955	38.337	105.138	422.276	1382.032	5278.679	1.6×10^4	2.0×10^5
$\theta(q)$	p_{25}	0.016	0.110	0.228	0.718	1.858	3.796	6.830	9.632	17.124
	p_{50}	0.017	0.160	0.330	1.078	2.588	5.411	8.229	11.303	25.619
	p_{75}	0.130	0.323	0.645	1.569	3.663	6.985	11.180	15.948	28.997
$\nu'(q)$	p_{25}	-13.748	-6.468	-4.595	-2.156	-0.518	0.064	3.390	6.724	30.630
	p_{50}	-6.826	-4.054	-2.356	-0.853	-0.058	3.184	14.733	26.128	96.637
	p_{75}	-3.527	-1.668	-0.857	0.017	3.121	16.823	51.626	128.570	250.859
$f(\theta)$	p_{25}	1.118	1.118	1.581	2.739	4.100	5.924	7.583	9.552	15.407
	p_{50}	2.427	3.708	4.450	6.124	7.982	9.937	12.924	16.074	20.282
	p_{75}	4.031	5.666	6.614	8.063	9.552	11.837	15.112	17.534	20.793

Table 12: Between-Village Quartiles of Percentiles of t-Statistics across Village Quantities for Sugar (Quadratic)

		p_1	p_5	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}	p_{95}	p_{99}
$\gamma(\theta(q))$	p_{25}	0.940	2.412	3.857	8.572	27.432	147.725	1.7×10^4	8.1×10^4	1.3×10^6
	p_{50}	3.166	7.223	10.878	29.042	77.897	530.951	2.8×10^4	1.3×10^5	2.3×10^6
	p_{75}	7.822	20.626	36.080	85.640	307.809	3935.793	1.0×10^5	6.7×10^5	3.2×10^9
$\theta(q)$	p_{25}	0.007	0.121	0.229	0.765	1.867	3.413	6.157	9.802	44.665
	p_{50}	0.007	0.165	0.384	1.283	2.747	5.033	9.987	15.675	72.555
	p_{75}	0.056	0.344	0.794	1.909	3.541	8.033	17.366	36.451	3885.028
$\nu'(q)$	p_{25}	-2.5×10^3	-6.658	-3.531	-1.381	-0.053	4.027	21.737	34.101	120.574
	p_{50}	-22.933	-3.397	-1.794	-0.325	2.064	19.568	47.653	87.399	222.200
	p_{75}	-4.993	-1.425	-0.691	0.390	9.453	40.488	115.893	202.440	622.084
$f(\theta)$	p_{25}	1.118	1.118	1.350	2.739	4.743	7.004	8.934	10.607	13.460
	p_{50}	2.739	3.783	4.815	6.347	8.129	10.457	13.239	15.572	18.895
	p_{75}	4.330	6.222	6.982	8.094	9.805	12.298	15.059	17.636	18.936