Prices, Profits, and Production*

Victor H. Aguiar   Roy Allen   Nail Kashaev [†]
vaguiar@uwo.ca   rallen46@uwo.ca   nkashaev@uwo.ca

This version: October 2018

Abstract
We study identification of multi-output production possibilities sets for heterogeneous firms that can be ranked in terms of productivity. Our setup applies with price-taking, profit-maximizing firms. We require observation of profits and either prices or proxies for prices. We characterize the identified set for production sets and provide conditions that ensure point identification of production sets. Our results extend classical duality results for the deterministic firm problem to a setup with rich heterogeneity, and with potentially limited variation in prices. We show that existing convergence results for quantile estimators may be directly converted to convergence results for production sets.
JEL classification numbers: C5.
Keywords: Production function, production set, nonseparable heterogeneity, shape restrictions.

*We are grateful to Lance Lochner, Salvador Navarro, David Rivers, and Al Slivinsky for useful comments and encouragement.
†Department of Economics, University of Western Ontario.

Introduction

This paper studies identification of production sets for firms with *multiple* outputs and inputs. We allow rich forms of complementarity and substitutability between inputs and outputs as well as rich heterogeneity across firms, but maintain the key assumption that firms can be ranked in terms of productivity. With this assumption, we characterize the most that can be said about production sets when one observes a cross-section of firm profits and both input and output prices.

The use of profits and prices to recover production sets has a long history in economics. It is now well-known that the profit function of a competitive firm fully characterizes its technological possibilities. This classical result applies, however, when there is no heterogeneity and when the analyst observes all possible (non-negative) prices. The main contribution of this paper is to extend this result to handle firm heterogeneity and to settings with potentially limited variation in prices.

Our key restriction on firm heterogeneity is that firms can be ranked in terms of productivity. We formalize this by assuming that a firm with higher productivity has access to all the production possibilities of a less productive firm, and possibly more. Our framework covers Hicks-neutral heterogeneity in productivity as a special case. With this assumption, the heterogeneous profit function satisfies a key monotonicity property in unobservable productivity similar to that used in Matzkin (2003).[1] We provide a novel identification result generalizing Matzkin (2003), showing how to recover the heterogeneous profit function from the joint distribution of prices and profits.

Once the heterogeneous profit function has been identified, we turn to our main result, which has two parts. First, we provide a sharp identification result characterizing the envelope of all production possibilities sets that can generate the data. This result applies regardless of the variation in prices, and is constructive. Next, we provide conditions under which the production possibilities sets may be uniquely recovered from data. This result does not require that all non-negative prices be observed, in contrast with textbook results (e.g. Kreps (2012)). Instead, we require that all possible "directions" of prices be observed. This condition can be satisfied even if all prices are bounded from above.

We next turn to estimation, providing an equality relating estimation error of

[1] We use a weak monotonicity condition rather than strict monotonicity in Matzkin (2003). This allows us to handle the important possibility that some firms earn zero profits – i.e. they shut down. In addition, it allows us to treat discrete and continuous heterogeneity in a common framework

profit functions and estimation error of production possibilities sets. This is a *generic* result allowing one to adapt consistency results for quantile estimators of the profit function, which is a well-understood problem (e.g. Matzkin (2003)), for the purpose of set estimation. The result is related to a classical result in convex analysis linking the (sup) distance of support functions with the (Hausdorff) distance of the corresponding sets. We generalize this result to our setting, requiring novel techniques because prices are restricted to be positive.

In some empirical settings the analyst may not observe all prices. We extend our analysis to allow observable proxy variables, which determine prices via an unknown, good-specific link function. Proxy variables can include observed product characteristics (e.g. observed quality). We show that with *either* profits and proxies, or quantities and proxies, it is possible to fully identify production sets. These additional identification results are established by using shape restrictions arising from the firm optimization problem. When prices and proxies are observed, we use a novel identification technique exploiting homogeneity, which may be of independent interest.[2]

Our analysis builds heavily on duality theory in the firm problem. Duality is a classical tool in producer theory for price-taking firms. Theoretical analysis includes the elegant and powerful contributions of Shephard (1953), Fuss & McFadden (1978), and Diewert (1982) among many others (see Kreps (2012) for a textbook treatment). Duality has also been heavily used to motivate parametric estimation of production functions (e.g. Lau (1972), Diewert (1973), Christensen et al. (1973)). This literature, however, examines a *representative agent* framework, either not including any random variables or including unobservable measurement error or heterogeneity as an additive error. We depart from this body of existing work by allowing rich nonseparable heterogeneity.[3]

There is little existing work concerning identification with limited (possibly finite) variation in prices. The closest paper appears to be Varian (1984), who works with different primitives and does not study unobservable heterogeneity. While we require observation of the joint distribution of profits and prices, Varian (1984) requires finite

[2]We provide a sufficient condition in terms of a rank condition. This condition arises by exploiting Euler's homogeneous function theorem to generate a system of linear equations that can be used for identification.

[3]Outside of the firm problem, duality has been used in the presence of heterogeneity in discrete choice (McFadden (1981)), matching models (Galichon & Salanié (2015)), hedonic models (Chernozhukov et al. (2017)), dynamic discrete choice (Chiong et al. (2016)), and the additively separable framework of Allen & Rehbeck (2017).

and deterministic datasets on prices and *quantities*.[4] While observation of prices and quantities implies observation of profits, the reverse is not true. We also note that it is not always necessary to observe all suitably normalized prices to uniquely recover the production set; thus, the variation in prices used to apply classical results (cf. Kreps (2012), Corollary 9.18) can be weakened.

A recent literature on the identification and practical estimation of a firm's technology has focused on outputs and inputs, sometimes not using prices or profits at all (e.g. Griliches & Mairesse (1995), Olley & Pakes (1996), Levinsohn & Petrin (2003), and Ackerberg et al. (2015)). A *pure quantities* approach has encountered an important stumbling block when modelling multi-output firms. With multiple outputs, there is no longer a production function. Instead, for a given level of inputs there is a *set* of possible outputs that can be produced. Without accounting for prices or placing more structure on the problem, the specific output may be indeterminate. The approach taken by e.g. De Loecker et al. (2016) completes the model by assuming separable technologies so the firm may be viewed as a composition of several single-output firms. Grieco & McDevitt (2016) does not assume separable technologies but imposes a linearity assumption. In contrast, the duality approach we take allows one to handle multi-output and single-output firms in a unified framework without such separability conditions or parametric restrictions.

Input price variation has recently been used by Gandhi et al. (2017) using a first order conditions approach.[5] While they focus on price variation in a single intermediate input, we study identification with variation in all prices. In contrast with their setup, our analysis requires only prices and profits, not quantities. We believe our analysis highlights the importance of price information to learn about the technology of a firm. In addition, it provides a complementary approach to methodologies that need to observe quantities, allowing practitioners to estimate the technology of firms in situations where the observability of outputs and inputs is problematic.[6]

The rest of this paper proceeds as follows. In Section 1 we present a model of production and several characterizations of it. Then we proceed with our main

[4]See also Cherchye et al. (2014) and Cherchye et al. (2018).

[5]See also Malikov (2017).

[6]Some examples are: (i) The housing market. the observability of output quantities is difficult because houses provide different services that are difficult to measure (Epple et al. (2010)). Profits and prices may be observable when quantities are not. (ii) Tax revenue agencies administrative datasets. In this type of dataset firms report profits but quantities of outputs and inputs are not observed. Price variability can be obtained from sales-tax datasets. (iii) Scanner wholesale price datasets, where prices are observed for each output and input in different geographic markets. These datasets can be matched with profits datasets to estimate production sets.

identification result in Section 2. In Section 3 we propose a consistent estimator of the production correspondence. In Section 4 we extend our methodology to environments where one observes proxies of unobserved prices. Section 5 is concerned with potential price endogeneity. We conclude in Section 6. All proofs can be found in Appendix A. Appendix B contains additional identification results for endogenous prices.

Notation

We use $\mathbb{R}^d_+$, $\mathbb{R}^d_-$, and $\mathbb{R}^d_{++}$, to denote component-wise nonnegative, nonpositive, and positive elements of the d-dimensional Euclidean space $\mathbb{R}^d$, respectively. The transpose of a vector is denoted y' and its Euclidean norm is denoted $\|y\|$. We denote the indicator function by $\mathbb{1}(\cdot)$ ($\mathbb{1}(A)$ is equal to 1 when the statement A is true, otherwise it is zero). We use boldface font (e.g. $\mathbf{p}$) to denote random objects and regular font (e.g. p) for deterministic ones. We denote the probability of an event A by the expression $\mathbb{P}(A)$. The supports of random vectors are usually denoted by capital letters; i.e. for the random vector $\mathbf{p}$, the support is denoted P, and is the smallest closed set such that $\mathbb{P}(\mathbf{p} \in P) = 1$. The cumulative distribution function (c.d.f.) of a random vector $\mathbf{p}$ is denoted by $F_{\mathbf{p}}$, and $F_{\boldsymbol{\pi}|\mathbf{p}}$ denotes the conditional c.d.f. of $\boldsymbol{\pi}|\mathbf{p}$.

1. Model

We study firms that may produce multiple outputs. Because we focus on technologies with multiple outputs, we work with production possibilities sets rather than production functions.[7] Specifically, every firm is characterized by a realization of $\mathbf{e} \in E$ and a correspondence $Y : E \rightrightarrows \mathbb{R}^{d_y}$, where $E \subseteq \mathbb{R}$ is a closed interval with nonempty interior.

The random variable $\mathbf{e}$ is interpreted as a scalar unobserved productivity term. The set $Y(e)$ is the production possibilities set for a firm with productivity level e. The set $Y(e)$ describes possible net output vectors. For each vector $y \in Y(e)$, a positive component indicates the firm is a net supplier of that good, and a negative

[7] An alternative approach is to use transformation functions. See Grieco & McDevitt (2016) for a recent application.

component indicates the firm is a net demander. For concreteness, for every $y \in Y(e)$ we may interpret the first d_{y_o} components as outputs and the last d_{y_i} components as inputs. The possible output/input vector is denoted $y = (y_o', y_i')'$, where $y_o \in \mathbb{R}^{d_{y_o}}$ and $y_i \in \mathbb{R}^{d_{y_i}}$ are the vectors of outputs and inputs, respectively.

Definition 1. A correspondence $Y : E \rightrightarrows \mathbb{R}^{d_y}$ is a production correspondence if, for every $e \in E$,

(i) $Y(e)$ is closed and convex;

(ii) $Y(e)$ satisfies *free disposal*: if y in $Y(e)$, then any y^* such that $y_l^* \leq y_l$ for all $l \in \{1, \cdots, d_y\}$ is also in $Y(e)$;

(iii) $Y(e)$ satisfies *the recession cone property*: if $\{y_k\}$ is a sequence of points in $Y(e)$ satisfying $\|y_k\| \to \infty$ as $k \to \infty$, then accumulation points of the set $\{y_k / \|y_k\|\}$ lie in the negative orthant of $\mathbb{R}^{d_y}$.

These conditions are standard. With closedness of $Y(e)$ maintained, condition (iii) is equivalent to the profit maximizing problem having a solution, and rules out constant or increasing returns to scale.[8] In particular, it implies that profits are finite.

We consider a setting in which, given a realization of $\mathbf{e}$ and market prices $\mathbf{p} \in P \subseteq \mathbb{R}^{d_y}_{++}$, each firm chooses a production plan $y \in Y(e)$ in order to maximize profits. We write the *profit maximization problem* for the firm as

$$\max_{y \in Y(e)} p'y.$$

Summarizing, we assume that firms are static profit maximizers, face no uncertainty, and are price takers.

In order to have a structural interpretation for unobserved productivity captured by $\mathbf{e}$, we impose that firms can be ranked according to productivity. We formalize this as follows.

Assumption 1. *If $\tilde{e} \geq e$, then $Y(e) \subseteq Y(\tilde{e})$.*

This assumption states that firms with higher values of e have access to weakly more possibilities than firms with lower values of e. Recall that the set E is a subset of the reals, so that the ranking $\tilde{e} \geq e$ is the usual linear order. One can think of e as unobserved one-dimensional input (e.g. managerial quality). That is, one can work with otherwise homogeneous firms that are different only in one unobserved input.

[8] See Kreps (2012), p. 199 for more details.

1.1. Production possibilities set and profit function

In this section we recall classical duality relationships between production sets and the profit function that will be used in our identification analysis. These results show how the profit function can be used to recover production possibilities sets. They are not immediately applicable when the analyst allows heterogeneity and observes only the distribution of profits and prices. Incorporating heterogeneity will be tackled in subsequent analysis.

Definition 2. The profit function of a price taking firm, denoted $\pi : \mathbb{R}^{d_y}_{++} \times E \to \mathbb{R}_+$, is given by

$$\pi(p, e) = \max_{y \in Y(e)} p'y.$$

The profit function is convex, i.e. for each $\alpha \in [0, 1]$ and possible prices $p, \tilde{p}$, $\pi(\alpha p + (1-\alpha)\tilde{p}, e) \geq \alpha\pi(p, e) + (1-\alpha)\pi(\tilde{p}, e)$. It is also homogeneous of degree 1 in prices, i.e. for each scalar $\lambda > 0$, $\pi(\lambda p, e) = \lambda\pi(p, e)$ for all e. These conditions are also sufficient for a function to be a profit function.

Lemma 1 (Kreps (2012), Proposition 9.14)**.** *In order for a conjectured function $\pi : \mathbb{R}^{d_y}_{++} \times E \to \mathbb{R}_+$ to be the profit function for price-taking firms, indexed by e, it is necessary and sufficient that $\pi(\cdot, e)$ be convex and homogeneous of degree 1 for all $e \in E$.*

Both directions of this lemma may be used for empirical analysis of profit - maximizing firms. When one assumes a firm is profit maximizing, then convexity and homogeneity emerge as shape restrictions on the firm problem that can be used for extrapolation. Alternatively, homogeneity and convexity of a conjectured profit function are testable implications of the assumption of price-taking, profit-maximizing behavior. We discuss each of these aspects of the lemma when we introduce heterogeneity in the next section.

In our environment, the profit function provides a complete characterization of the production set for a given realization of $\mathbf{e}$.

Lemma 2 (Kreps (2012), Corollary 9.18)**.** *For all $e \in E$, the realized production set is described by*

$$Y(e) = \left\{ y \in \mathbb{R}^{d_y} : p'y \leq \pi(p, e),\ \forall p \in \mathbb{R}^{d_y}_{++} \right\}.$$

This result follows because we are assuming that Y is a production correspondence (recall Definition 1), not just a general correspondence. The result shows that if we can recover the profit function, then we can recover the production set.

Ranking firms according to productivity and according to profits are equivalent, as formalized below.

Lemma 3. *With the maintained assumption that $Y(\cdot)$ is a production correspondence, the following are equivalent:*

(i) Production Monotonicity: If $\tilde{e} \geq e$, then $Y(e) \subseteq Y(\tilde{e})$;

(ii) Profit Monotonicity: If $\tilde{e} \geq e$, then $\pi(p, \tilde{e}) \geq \pi(p, e)$ for every $p \in \mathbb{R}^{d_y}_{++}$.

Proof. See Rockafellar (1970), Corollary 13.1.1. ∎

Thus, Assumption 1 (condition (i) in this lemma) is equivalent to monotonicity of profits.

We require an additional technical continuity condition for our subsequent analysis. This condition will allow us to identify production sets under either a continuous or discrete distribution of heterogeneity.

Lemma 4 (Aliprantis & Border (2006), Lemma 17.29.)**.** *Suppose $Y(\cdot)$ is lower hemicontinuous, i.e., whenever $e_k \to e$ as $k \to \infty$, and $y \in Y(e)$, there is a sequence $y_k \in Y(e_k)$ such that $y_k \to y$ as $k \to \infty$. It follows that $\pi(p, \cdot)$ is lower semicontinuous.*

Lemmas 3 and 4 allow us to translate economically relevant assumptions on the primitives (production correspondences) into restrictions on the observable quantities (profits).

The following examples illustrate production monotonicity and lower hemicontinuity of $Y(\cdot)$.

Example 1. [Single output, Hicks-neutral production] Suppose a firm chooses capital (k) and labor (l) to produce a single output good. That is, $y_o \in \mathbb{R}$, $y_i = (-k, -l)'$, and $y = (y_o, -k, -l)'$. Negatives on capital and labor denote that these quantities are demanded rather than supplied. The production function is specified as $F(e, k, l) = A(e)f(k, l)$ with $f(\cdot, \cdot)$ a strictly concave, continuous, and weakly increasing function, and $A(\cdot) \geq 0$. The production possibilities set, $Y(e)$, is the set of all vectors y satisfying $y_o \leq F(e, k, l)$. Note that if $A(\cdot)$ is a weakly increasing and lower semicontinuous function, then $Y(e)$ satisfies production monotonicity and is lower hemicontinuous. For example, $A(e) = \exp(e)$ with $E = [0, M]$, $M > 0$, is both weakly increasing and lower semicontinuous. The function $A(e) = \mathbb{1}(0 \leq e \leq 1/2) + 2\mathbb{1}(1/2 < e \leq 1)$ with $E = [0, 1]$ is also weakly increasing and lower semicontinuous, yet has only two distinct types of firm (determined by whether $e > 1/2$). These two choices of A both imply lower hemicontinuity of $Y(\cdot)$ and production monotonicity, and so this example illustrates how we may treat discrete and continuous types in a common framework.

Example 2. [Multiple outputs] Suppose a firm operates two plants. Plant 1 produces $y_{o,1}$ that can be sold on the market or shipped to Plant 2. Plant 2 uses $y_{o,1}$ as an input in the production of $y_{o,2}$. Both plants use two types of labor: "skilled" (l_1^s and l_2^s) and "unskilled" (l_1^u and l_2^u). The firm is managed by a CEO with "quality" $e \in E = [0,1]$. Plant j is controlled by a manager whose quality is determined by e via an unknown weakly increasing and continuous function $A_j(\cdot)$. That is, the quality of plants' managers is a deterministic function of the quality of the CEO. The production function of Plant 1 is specified as $F_1(e, l^s, l^u) = A_1(e) f_1(l_1^s, l_1^u)$ with f_1 a strictly concave, continuous, and weakly increasing function. The production function of Plant 2 is specified as $F_2(e, y_{o,1}, l_2^s, l_2^u) = A_2(e) f_2(y_{o,1}, l_2^s, l_2^u)$ with f_2 a strictly concave, continuous, and weakly increasing function. Thus, $y_o = (y_{o,1}, y_{o,2})' \in \mathbb{R}^2$, $y_i = (-l_1^s - l_2^s, -l_1^u - l_2^u)'$, and $y = (y_o', y_i')'$. The production possibilities set is

$$Y(e) = \Big\{ y \in \mathbb{R}^2 \times \mathbb{R}^2_+ \ : y_{o,1} \leq F_1(e, l_1^s, l_1^u),\ y_{o,2} \leq F_2(e, y_{o,1}, l^s - l_1^s, l^u - l_1^u) \\ \text{for some } l_1^s \in [0, l^s],\ l_1^u \in [0, l^u] \Big\}.$$

Under our monotonicity conditions on functions f_1, f_2, A_1, and A_2, the correspondence Y satisfies production monotonicity and lower hemicontinuity.

2. Identification of the Production Correspondence

We now present our core identification results for production correspondences. We observe profits and prices, and so we identify the production correspondence by first identifying the profit function, and then using (and extending) duality results presented in the previous section. Recall that we use boldface font to denote random objects and regular font for deterministic ones.

In order to recover the profit function we impose the assumption that prices and unobservable heterogeneity are independent. In Section 5 we relax this assumption.

Assumption 2 (Independence)**.** *The unobserved shocks* $\mathbf{e}$ *are independent from prices* $\mathbf{p}$. *That is,* $F_{\mathbf{e}}(\cdot) = F_{\mathbf{e}|\mathbf{p}}(\cdot|p)$ *for all* $p \in P$.

The following result extends the results of Matzkin (2003) to weakly monotone functions. Allowing weak monotonicity of $\pi(p,\cdot)$ is empirically relevant since it accommodates discrete heterogeneity (see Example 1). It also accommodates the important possibility that firms may shut down, since then $\pi(p,\cdot)$ may be flat (at 0) for multiple values of e.

Theorem 1. *Let Assumption 2 hold and assume $\pi(p,\cdot)$ is lower semicontinuous and weakly increasing for every $p \in P$. It follows that $\pi(p,\cdot)$ is constructively identified from $F_{\boldsymbol{\pi}|\mathbf{p}}(\cdot|p)$ up to any strictly increasing $F_{\mathbf{e}}(\cdot)$ for all $p \in P$. In particular,*

$$\pi(p,e) = \inf\left\{k \ : \ e \leq F_{\mathbf{e}}^{-1}(F_{\boldsymbol{\pi}|\mathbf{p}}(k|p))\right\}.$$

We present the theorem with assumptions directly on $\pi(p,\cdot)$ to line up more cleanly with Matzkin (2003), and because our generalization may be of independent interest. We differ because we do not assume π is strictly increasing in its second argument, and we do not assume it is continuous.

The primitive in this paper is the production correspondence, however, and so we present a corollary with assumptions on this primitive. Production monotonicity is equivalent to profit monotonicity (Lemma 3), and lower hemicontinuity of $Y(\cdot)$ implies lower semicontinuity of $\pi(p,\cdot)$ (Lemma 4). Thus, we obtain the following result as a corollary.

Corollary 1. *Let Assumptions 1 and 2 hold and assume $Y(\cdot)$ is lower hemicontinuous. It follows that $\pi(p,\cdot)$ is constructively identified from $F_{\boldsymbol{\pi}|\mathbf{p}}(\cdot|p)$ up to any strictly increasing $F_{\mathbf{e}}(\cdot)$ for all $p \in P$. In particular,*

$$\pi(p,e) = \inf\left\{k \ : \ e \leq F_{\mathbf{e}}^{-1}(F_{\boldsymbol{\pi}|\mathbf{p}}(k|p))\right\}.$$

In light of Lemma 1, a testable implication is that for every e, the function

$$\inf\left\{k \ : \ e \leq F_{\mathbf{e}}^{-1}(F_{\boldsymbol{\pi}|\mathbf{p}}(k|p))\right\}$$

must be convex and homogeneous of degree 1 in prices.

2.1. From Profits to Production

Note that Theorem 1 identifies $\pi(p,\cdot)$ only over P (the support of prices). When prices have full positive support, i.e. $P = \mathbb{R}_{++}^{d_y}$, from Lemma 2 we immediately

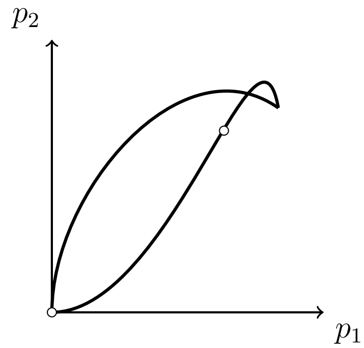

Figure 1 – The set P (depicted by black curve) satisfies Assumption 3 and has an empty interior. Dots represent "holes" in the support. Thus, P is not a connected set.

deduce that $Y(\cdot)$ is identified. We instead consider the possibility that P may have limited support. We characterize the sharp envelope of all production correspondences consistent with the data, as well as the support condition for prices that ensures point identification of $Y(\cdot)$.

Our results exploit homogeneity of $\pi(\cdot, e)$. By leveraging homogeneity, we know that if we identify $\pi(p, e)$ for some $p \in P$, then we also identify $\pi(\lambda p, e)$ for any positive λ. That is, we do not need to observe prices that are proportional to a price that we already observe. This simple property leads to a drastic shrinkage of the set of prices that we need to observe in the data in order to nonparametrically recover the profit function. Moreover, $\pi(\cdot, e)$ is convex and, hence, continuous.[9] These features lead to consideration of the following assumption, which ensures $Y(\cdot)$ may be recovered uniquely.

Assumption 3.

$$\operatorname{int}\left(\operatorname{cl}\left(\bigcup_{\lambda>0}\{\lambda p \,:\, p \in P\}\right)\right) = \mathbb{R}^{d_y}_{++},$$

where $\operatorname{cl}(A)$ *and* $\operatorname{int}(A)$ *are the closure and the interior of* A*, respectively.*

The set

$$\bigcup_{\lambda>0}\{\lambda p \,:\, p \in P\}$$

[9]Beyond continuity, the manner in which convexity affects the data requirements that ensure point identification is subtle, and depends on the shape of $Y(\cdot)$. We provide an illustrative example in Section 2.2.

consists of all prices where $\pi(\cdot, e)$ is known because of homogeneity. If that set has "holes," then we can fill them by taking the closure of the set since $\pi(\cdot, e)$ is convex, hence continuous. Assumption 3 means that after we consider the implications of homogeneity and continuity, it is as if we have full variation in prices. Figure 1 is an example of a set satisfying this assumption.

Note that Assumption 3 does not impose that the support of $\mathbf{p}$ contains an open ball. In particular, Assumption 3 can be satisfied if $\mathbf{p}$ is discrete but has a countable support. Assumption 3 is equivalent to

$$\operatorname{int}\left(\operatorname{cl}\left(\{p/\|p\| \; : \; p \in P\}\right)\right) = \mathbb{S}^{d_y-1} \cap \mathbb{R}^{d_y}_{++},$$

where $\mathbb{S}^{d_y-1}$ denotes the unit sphere in $\mathbb{R}^{d_y}$. This clarifies that the support condition involves directions of prices $p/\|p\|$. In particular, with one input and one output it requires that ratios of prices (e.g. p_1/p_2) can be made arbitrary close to 0 and ∞. In Figure 1, such extreme directions are obtained for vectors local to the origin.

Finally, we impose a normalization on the distribution of $\mathbf{e}$.

Assumption 4. *The distribution of* $\mathbf{e}$ *is uniform over* $[0, 1]$.

This assumption facilitates exposition; if it is dropped, subsequent identification results hold up to the distribution of $\mathbf{e}$. See Matzkin (2003) for a discussion of normalizations in related settings.

Theorem 2. *Let Assumption 4 and the assumptions of Theorem 1 hold. Moreover, let $\tilde{Y}$ be a correspondence such that*

$$\tilde{Y}(e) = \left\{ y \in \mathbb{R}^{d_y} \; : \; p'y \leq \pi(p, e), \; \forall p \in P \right\}$$

for all $e \in E$. Then

(i) $\tilde{Y}(\cdot)$ can generate the data and for each $e \in E$, $\tilde{Y}(e)$ is a closed, convex set that satisfies free disposal.

(ii) A production correspondence $Y'(\cdot)$ can generate the data if and only if

$$\max_{y \in Y'(e)} p'y = \max_{y \in \tilde{Y}(e)} p'y,$$

for every $e \in E$ and $p \in P$. It follows that for any such $Y'(\cdot)$, $Y'(e) \subseteq \tilde{Y}(e)$ for each $e \in E$.

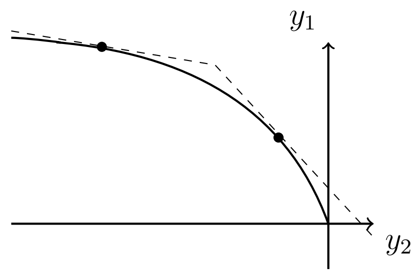

Figure 2 – $\tilde{Y}(e)$ and $Y'(e)$ for $d_y = 2$ and $P = \{p^*, p^{**}\}$. $\tilde{Y}(e)$ is the area under the dashed lines. $Y'(e)$ is the area under the solid curve. Dashed lines correspond to two hyperplanes $p_1^* y_1 + p_2^* y_2 = \pi(p^*, e)$ and $p_1^{**} y_1 + p_2^{**} y_2 = \pi(p^{**}, e)$. They are tangental to the solid curve.

(iii) If Assumption 3 holds, then $\tilde{Y}$ is the only production correspondence that can generate the data.

Parts (i) and (ii) of Theorem 2 are a sharp identification result, stating the most that can be said about production correspondences under our assumptions. These results are related to Varian (1984), Theorem 15.[10] However, Varian (1984) works only with finite datasets, which are comparable to finite support of prices in our setting. In addition, Varian (1984) observes prices and quantities while we observe prices and profits.

Theorem 2(ii) establishes that $\tilde{Y}(\cdot)$ is the envelope of all production correspondences that can generate the data (see Figure 2). We note, however, that $\tilde{Y}(\cdot)$ may not be a production correspondence because it need not satisfy the recession cone property (recall Definition 1(iii)). To see this, suppose that a firm of type $e \in E$ has a single input and output, prices are a constant vector $P = \{(1,1)'\}$, and profits at that price are given by $\pi((1,1)', e) = 0$. Then the set $\tilde{Y}(e)$ is $\{y \in \mathbb{R}^2 : y_1 + y_2 \leq 0\}$. This set induces infinite profits for a price taking firm whenever $p_1 > p_2$. Hence, this set violates the recession cone property, which is necessary for the firm problem to have a maximizer since $\tilde{Y}(e)$ is closed and nonempty.[11]

Theorem 2(iii) is a refinement of Lemma 2, which is the textbook version of recovering production sets from the profit function. In econometric settings, it is not always natural to assume that all prices are observed ($P = \mathbb{R}_{++}^{d_y}$). Theorem 2 clarifies

[10]The set $\tilde{Y}(e)$ is essentially the "outer" set considered in Varian (1984), Section 7. It is constructed from price and profit information, however, rather than price and quantity information as in Varian (1984).

[11]See e.g. Kreps (2012), Proposition 9.7. Note from part (iii), when Assumption 3 holds it follows that $\tilde{Y}$ is a production correspondence, and thus satisfies the recession cone property.

the variation in prices sufficient for nonparametric identification of production sets.

We emphasize that the full strength of Assumption 3 may be needed only if one wants to fully identify $Y(\cdot)$. If one is only interested in identification of some economically relevant region of the production possibilities frontier, then it suffices to observe only those prices that are tangential to that region of interest as the following example demonstrates.

Example 3. Fix some $e \in E$ and suppose that

$$Y''(e) = \{y \in \mathbb{R} \times \mathbb{R}_- \,:\, y_1 \leq \sqrt{-y_2}\}.$$

That is, the production possibilities frontier is $\{y \in \mathbb{R} \times \mathbb{R}_- \,:\, y_1 = \sqrt{-y_2}\}$. Suppose that one is only interested in identification of the production possibilities frontier when $y_1 \in [\underline{y}_1, \overline{y}_1]$ with $0 < \underline{y}_1 \leq \overline{y}_1 < \infty$. Then Theorem 2 implies that it suffices to observe prices only in the set $\{p \in \mathbb{R}^2_{++} \,:\, 2\underline{y}_1 \leq p_1/p_2 \leq 2\overline{y}_1\}$.

2.2. Point Identification and Assumption 3

It is natural to wonder when Assumption 3 is necessary and sufficient for point identification of $Y(\cdot)$. Unfortunately, this question is technical. It is essentially equivalent to asking when the function π_P defined as π restricted to $P \times E$, has a *unique* extension $\tilde{\pi} : \mathbb{R}^{d_y}_{++} \times E \to \mathbb{R}^{d_y}$ such that $\tilde{\pi}$ is homogeneous of degree 1, convex, and satisfies $\tilde{\pi}(p, e) = \pi(p, e)$ for every $(p', e)' \in P \times E$. More formally, the extension $\tilde{\pi}(p, \cdot)$ must also be increasing in its second argument and lower semicontinuous in e for each p.

First, we note that by exploiting continuity and homogeneity of degree 1, we know that there is a unique extension of π_P to the set

$$\operatorname{int}\left(\operatorname{cl}\left(\bigcup_{\lambda>0} \{\lambda p \,:\, p \in P\}\right)\right) \times E$$

that satisfies the desired properties. It is, however, possible that this set is strictly nested in $\mathbb{R}^{d_y}_{++} \times E$, and yet there is a unique extension of π_P to all of $\mathbb{R}^{d_y}_{++} \times E$.

Example 4 (Unique Extension without Assumption 3)**.** Consider $\pi(p, e) = e \sum_{j=1}^{d_y} |p_j|$ with $E = [0, M)$. This functions is homogeneous of degree 1 and convex, and hence the profit function for a price taking firm (Lemma 1). Let $\Delta^{d_y - 1} = \{y \in \mathbb{R}^{d_y}_{++} \,:\, \sum_{j=1}^{d_y} y_j = 1\}$ denote the relative interior of the probability simplex, and let

$S = \{y \in \Delta^{d_y - 1} : |y_j - 1/d_y| \leq 1/d_y \text{ for each } j\}$ denote a convex set centered at the midpoint of the simplex. Let P be the probability simplex with the region S removed, i.e. $P = \Delta^{d_y-1} \cap S^C$, where S^C denotes the complement of S. Note that $\pi_P(\cdot, e)$ is a subset of the affine space $\{y \in \mathbb{R}^{d_y} : \sum_{j=1}^{d_y} y_j = e\}$ for each e, and hence any convex extension $\tilde{\pi}$ must also be affine over the convex hull of P, which is Δ^{d_y-1}. There is a unique such extension because Δ^{d_y-1} has dimension $d_y - 1$ (i.e. the smallest affine space containing this set has dimension $d_y - 1$). Thus $\tilde{\pi}$ is unique. In particular, by Lemma 2 the production correspondence is identified even though Assumption 3 fails to hold.

We are only able to construct "knife edge" examples in which the support restriction of Assumption 3 is *not* equivalent to point identification of $Y(\cdot)$. We note that strict convexity of $\pi(\cdot, e)$ rules out this sort of example.

2.3. Supply Function

Some interesting objects are directly identifiable from the profit function, without (directly) working with inequalities as described in Theorem 2. One such object is the supply function, which exists whenever the profit maximization problem has a unique solution.

Definition 3. The supply function of a price taking firm, denoted $y : \mathbb{R}^{d_y}_{++} \times E \to \mathbb{R}^{d_y}$, is given by

$$y(p, e) = \arg\max_{y \in Y(e)} p'y.$$

A direct consequence of our main result is that the supply function is also identified from only observing profits and prices, without needing to observe quantities directly. We let $\nabla_p \pi(p, e)$ denote the gradient with respect to prices of the profit function $\pi(\cdot, \cdot)$ at the point (p, e). This derivative exists provided the supply function $y(p, e)$ exists (e.g. Mas-Colell et al. (1995), Proposition 5.C.1). The following result is Hotelling's lemma.

Corollary 2. *Let $p \in P$, $e \in E$, and suppose $y(p, e)$ is the unique maximizer. Then if $\pi(\cdot, \cdot)$ is identified, the supply function is identified via the formula*

$$y(p, e) = \nabla_p \pi(p, e).$$

Note that this formula may be used with continuous quantities, since then we may take a derivative of the profit function. If quantities are observed in addition to prices and profits, we note that this equality may be used as an overidentifying restriction. We note that when the maximizer is not unique, identification of $Y(\cdot)$ instead identifies the set of profit-maximizing quantities for each p and e.

3. Estimation and Consistency

In this section, we describe how an estimator $\hat{\pi}(\cdot, e)$ of the profit function may be used to construct an estimator $\hat{Y}(e)$ of the production function for a firm with productivity level e. The main result in this section provides a *generic* result relating the estimation error of $\hat{\pi}$ (for π) and that of the constructed set $\hat{Y}$ (for Y). Consistency and rates of convergence results for $\hat{\pi}$ thus have analogous statements for $\hat{Y}$.

As setup, we now formalize our notions of distance both for functions and sets. We present our result for a fixed $e \in E$. We assume that $\pi(\cdot, e)$ is identified over $\mathbb{R}^{d_y}_{++}$ (we assume Assumption 3). Given a fixed $e \in E$ and $\hat{\pi}(\cdot, e)$, a natural estimator for $Y(e)$ is the following random convex set:

$$\hat{Y}(e) = \left\{ y \in \mathbb{R}^{d_y} \ : \ p'y \leq \hat{\pi}(p, e), \forall p \in P \right\}.$$

This set is a plug-in estimator motivated by Theorem 2. A commonly used notion of a distance between convex sets is the Hausdorff distance. The Hausdorff distance between two convex sets $A, B \subseteq \mathbb{R}^{d_y}$ is given by,

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|a - b\| \right\}.$$

Unfortunately, the Hausdorff distance between $Y(e)$ and $\hat{Y}(e)$ can be infinite. For this reason we will consider the Hausdorff distance between certain extensions of these sets. The following example illustrates why the original distance may be infinite.

Example 5. Suppose that $d_y = 2$ and for some $e \in E$

$$Y(e) = \left\{ y \in \mathbb{R} \times \mathbb{R}_- \ : \ y_1 \leq \sqrt{-y_2} \right\},$$
$$\hat{Y}_n(e) = \left\{ y \in \mathbb{R} \times \mathbb{R}_- \ : \ y_1 \leq (1 - 1/n)\sqrt{-y_2} \right\}, \quad n \in \mathbb{N}.$$

Note that although $\lim_{n \to \infty} (1 - 1/n)\sqrt{-y_2} = \sqrt{-y_2}$ for every finite $y_2 \leq 0$, the

Hausdorff distance is equal to $\sup_{y_2 \in \mathbb{R}_-} \sqrt{-y_2}/n = \infty$ for every finite $n \in \mathbb{N}$.

We consider the Hausdorff distance between extensions such as

$$Y_{\bar{P}}(e) = \left\{ y \in \mathbb{R}^{d_y} \ : \ p'y \leq \pi(p,e), \ \forall p \in \bar{P} \right\},$$
$$\hat{Y}_{\bar{P}}(e) = \left\{ y \in \mathbb{R}^{d_y} \ : \ p'y \leq \hat{\pi}(p,e), \ \forall p \in \bar{P} \right\},$$

where $\bar{P} \subseteq P$ is convex and compact. These sets nest the original sets because the inequalities hold only for $p \in \overline{P}$, not for every $p \in P$. Let $\bar{\mathcal{P}}$ be a collection of all compact and convex subsets of P. The next lemma formalizes that these sets are extensions (i.e. bigger) and that they can approximate the original sets.

Lemma 5. *(i) For every $\bar{P}, \bar{P}' \in \bar{\mathcal{P}}$, if $\bar{P} \subseteq \bar{P}'$, then $Y_{\bar{P}'}(e) \subseteq Y_{\bar{P}}(e)$ and $\hat{Y}_{\bar{P}'}(e) \subseteq \hat{Y}_{\bar{P}}(e)$;*

(ii) $Y(e) \subseteq Y_{\bar{P}}(e)$ and $\hat{Y}(e) \subseteq \hat{Y}_{\bar{P}}(e)$ for all $\bar{P} \in \bar{\mathcal{P}}$;

(iii) For any convex, compact K and $\hat{K}$ such that $Y(e) \bigcap K \neq \emptyset$ and $\hat{Y}(e) \bigcap \hat{K} \neq \emptyset$, there exist $\bar{P} \in \bar{\mathcal{P}}$ such that $Y_{\bar{P}}(e) \bigcap K = Y(e) \bigcap K$ and $\hat{Y}_{\bar{P}}(e) \bigcap \hat{K} = \hat{Y}(e) \bigcap \hat{K}$.

We now turn to the main result in this section, which establishes an equality relating the distance between $\hat{\pi}$ and π, and the distance between extensions of $\hat{Y}$ and Y. Our distance for these profit functions is given by

$$\eta_{\bar{P}}(e) = \sup_{p \in \bar{P}} \left\| \frac{\hat{\pi}(p,e) - \pi(p,e)}{\|p\|} \right\|.$$

Theorem 3. *Suppose that the assumptions of Theorem 2 hold. Suppose, moreover, that for every $e \in E$, $\hat{\pi}(\cdot, e)$ is an estimator of $\pi(\cdot, e)$ that is homogeneous of degree 1 and continuous. If $\hat{\pi}(\cdot, e)$ is convex, then*

$$d_H(Y_{\bar{P}}(e), \hat{Y}_{\bar{P}}(e)) = \eta_{\bar{P}}(e) \quad \text{a.s.},$$

for every $\bar{P} \in \bar{\mathcal{P}}$.

Theorem 3 is a non-trivial extension of a well-known relation between the Hausdorff distance and the support functions of convex *compact* sets to convex, closed, and *unbounded* sets.[12]

[12] See Kaido & Santos (2014) for a recent application of this result for convex compact sets.

Clearly, if we take a supremum with respect to e or $\bar{P}$ on both sides, then the analogous equality holds. If $\hat{\pi}$ is homogeneous of degree 1 in prices, then

$$\eta_{\bar{P}}(e) = \sup_{p \in \bar{P}} \left\| \hat{\pi}\left(\frac{p}{\|p\|}, e\right) - \pi\left(\frac{p}{\|p\|}, e\right) \right\| \leq \sup_{p \in P} \left\| \hat{\pi}\left(\frac{p}{\|p\|}, e\right) - \pi\left(\frac{p}{\|p\|}, e\right) \right\|.$$

The final equality replaces $\overline{P}$ by the larger set P. Thus, consistency of $\hat{\pi}$ for π in this appropriately-scaled sup-norm over P implies consistency of the extensions in Hausdorff distance.

Homogeneity of the estimator can be imposed by rescaling the data by dividing by one of the prices. Unfortunately, convexity is more challenging to impose and so we turn to a related result that covers cases in which $\hat{\pi}$ is not convex. To formalize our result, we introduce two additional parameters:

$$R_{\bar{P}}(e) = \sup_{p \in \bar{P}} \frac{\pi(p, e)}{\|p\|}, \quad r_{\bar{P}}(e) = \inf_{p \in \bar{P}} \frac{\pi(p, e)}{\|p\|}.$$

Proposition 1. *Suppose that the assumptions of Theorem 2 hold. Suppose, moreover, that for every $e \in E$, $\hat{\pi}(\cdot, e)$ is an estimator of $\pi(\cdot, e)$ that is homogeneous of degree 1 and continuous. If $\eta_{\bar{P}}(e) = o_p(1)$ and $0 < r_{\bar{P}}(e) < R_{\bar{P}}(e) < \infty$, then*

$$d_H(Y_{\bar{P}}(e), \hat{Y}_{\bar{P}}(e)) \leq \eta_{\bar{P}}(e) \frac{R_{\bar{P}}(e)}{r_{\bar{P}}(e)} \frac{1 + \eta_{\bar{P}}(e)/R_{\bar{P}}(e)}{1 - \eta_{\bar{P}}(e)/r_{\bar{P}}(e)}$$

with probability approaching 1, for every $\bar{P} \in \bar{\mathcal{P}}$. In particular,

$$d_H(Y_{\bar{P}}(e), \hat{Y}_{\bar{P}}(e)) = o_p(1).$$

The equivalence result in Theorem 3, together with the ensuing consistency result, may be used for any convex estimator of π. We outline a specific approach to estimating π by adapting the flexible functional form of Diewert (1973) to our setting. This class of functions applies with multiple outputs and inputs. Consider a profit function of the form

$$\pi(p, e) = \sum_{i=1}^{d_y} \sum_{j=1}^{d_y} b_{i,j}(e) p_i^{1/2} p_j^{1/2},$$

where $b_{i,j}(e) = b_{j,i}(e)$ for all i, j. The original class of Diewert (1973) considers a deterministic model or representative agent model, in which each $b_{i,j}(\cdot)$ is a constant function. We allow unobservable heterogeneity by allowing $b_{i,j}(e)$ to be a function of e. This functional form exhibits several desirable properties: (i) it is linear in the coefficients $b_{i,j}(e)$; (ii) monotonicity of $\pi(p,\cdot)$ can be imposed by assuming that each $b_{i,j}(\cdot)$ is weakly increasing;[13] (iii) convexity can be also imposed using linear inequalities on the coefficients;[14] (iv) homogeneity of degree 1 in p is built-in. These features facilitate its estimation using constrained linear quantile regression (Koenker & Ng (2005)). The supply function for good k is described by the formula:

$$y_k(p,e) = \sum_{j=1}^{d_y} b_{k,j}(e)(p_j/p_k)^{1/2}.$$

Thus, if quantities are observed in addition to prices and profits, then this equation provides overidentifying information. Finally, we note that our identification results in Section 4 show that if the price of a good j is not observed, we may replace p_j by a parametric function $g_j(x_j;\beta_j)$ such as $\beta_j x_j$, where x_j is a price proxy.

4. Extensions: Price proxies

In many empirical applications not all prices are observed, but the researcher has access to some observables (proxies) that are informative about unobserved prices. For example, the rental rate of capital and wages may be proxied as unknown functions of different market specific observables (e.g. short-term and long-term interest rates, unemployment level and fraction of high-skilled workers in the market). One can also use observed characteristics such as quality as a proxy for unobserved input prices.[15] This section provides identification results if unobserved prices are unknown functions of these proxies.

To formalize this, suppose that for each price p_j we have an observable proxy x_j that satisfies

$$p_j = g_j(x_j)$$

for an unknown function $g_j : X_j \to \mathbb{R}$, where X_j denotes the support of $\mathbf{x}_j$. Note that we assume that every price is a function of only one proxy variable to simplify the notation. We can also allow for existence of additional covariates that enter every g_j.

[13] Recall our identification arguments require only that $\pi(p,\cdot)$ be weakly increasing, not strictly increasing as in Matzkin (2003).

[14] A sufficient condition for convexity in prices is that $b_{i,j}(e) \leq 0$ for all $i \neq j$ and $b_{i,i}(e) \geq 0$.

[15] De Loecker et al. (2016) uses output price, market shares, product dummies, firm location, and export status as proxies for unobserved input prices in a related framework.

In this case the analysis below is performed as if we condition on a fixed value of those common covariates.[16] We denote $x = (x_j)_{j=1,\ldots,d_y} \in X$ and $g(x) = (g_j(x_j))_{j=1,\ldots,d_y}$. Profits are then given by $\pi(g(x), e)$. If the function g were known, we could calculate these profits directly and then apply Theorem 2.

We do not assume g is known, and instead provide two approaches to identify g. The first approach requires observation of profits and price proxies. The second approach requires observation of price proxies and quantities chosen. Each strategy identifies each function g_j up to location and scale. We place the following additional conditions to identify g_j uniquely.

Assumption 5. *(i) $g_{d_y}(x_{d_y}) = x_{d_y}$, i.e. the price of one input or output is observed;*

(ii) The value of g is known at one point, i.e. there exist known $x_0 \in X$ and p_0 such that $g(x_0) = p_0$;

(iii) g is differentiable on the interior of X, and the set $\left\{x_j \in X_j \,:\, \partial_{x_j} g(x_j) = 0\right\}$ has Lebesgue measure zero for every j.

(iv) (Rectangular Support) $X = \prod_{j=1}^{d_y} X_j$ where each set $X_j \subseteq \mathbb{R}$ is an interval with nonempty interior.

Assumptions 5(i)-(ii) allow us to identify the scale and the location, respectively, of the multivariate function g. Since we can always relabel both inputs and outputs, Assumption 5(i) is equivalent to assuming that at least one price (not necessary p_{d_y}) is observed.

First, we establish identification of $\tilde{\pi}(\cdot,\cdot) = \pi(g(\cdot),\cdot)$. This is a preliminary step to identify g, which in turn may be used to identify π. The function $\tilde{\pi}$ is a composition of two objects of interest: g and π. In order to identify $\tilde{\pi}$ we impose an independence restriction that implies Assumption 2, and is implied by Assumption 2 if g is invertible.

Assumption 6. *The unobserved shocks $\mathbf{e}$ are independent from proxies $\mathbf{x}$. That is, $F_{\mathbf{e}}(\cdot) = F_{\mathbf{e}|\mathbf{x}}(\cdot|x)$ for all $x \in X$.*

The following lemma is an analog of Theorem 1.

Lemma 6. *Suppose that Assumptions 4 and 6 are satisfied. If $\tilde{\pi}(x,\cdot) = \pi(g(x),\cdot)$ is lower semicontinuous and weakly increasing for every $x \in X$, then $\tilde{\pi}(x,\cdot)$ is identified for every $x \in X$ up to $F_{\boldsymbol{\pi}|\mathbf{x}}$.*

[16] If for some $\tilde{x}$ we have that $p_j = \tilde{g}_j(x_j, \tilde{x})$ for all j, then for every $\tilde{x}$ there exists $g_j(\cdot) = \tilde{g}_j(\cdot, \tilde{x})$ such that $p_j = g_j(x_j)$ for all j.

Proof. The proof follows from the proof of Theorem 1 with "p" replaced by "x". ∎

Lemma 6 implies that if one knows or can consistently estimate $F_{\boldsymbol{\pi}|\mathbf{x}}$, then $\tilde{\pi}$ can be uniquely recovered. Thus, π is identified up to $F_{\boldsymbol{\pi}|\mathbf{x}}$ and g, since $F_{\boldsymbol{\pi}|\mathbf{x}}$ is identified. If profits and proxies are observed, then we only need to identify g. If only proxies and quantities chosen ($\mathbf{y}$) are observed, then, since $\boldsymbol{\pi} = g(\mathbf{x})'\mathbf{y}$, we again only need to identify g. In the next two section we provide identification of g either from profits-proxies data or from proxies-quantities data.

4.1. Price Proxies and Profits Observed

In this section, we study identification if price proxies and profits are observed. We show that by using homogeneity of $\pi(\cdot, e)$ for each e, we can recover g. Once g is identified, it is straightforward to identify $\pi(\cdot, e)$, and, hence, the production possibilities sets $Y(\cdot)$ by our previous arguments.

We leverage the identifying power of homogeneity via Euler's homogeneous function theorem. This identification technique appears to be novel and may be of independent interest, and so we present an informal outline before presenting our formal results. Recall that the profit function $\pi(\cdot, e)$ is homogeneous of degree 1, which from Euler's homogeneous function theorem yields the system of equations

$$\sum_{i=1}^{d_y} \partial_{p_i} \pi(p, e) p_i = \pi(p, e).$$

Replacing prices with price proxies, we obtain

$$\sum_{i=1}^{d_y} \partial_{p_i} \pi(g(x), e) g_i(x_i) = \pi(g(x), e). \tag{1}$$

Since $\tilde{\pi}(x, e) = \pi(g(x), e)$, we have

$$\partial_{p_i} \pi(g(x), e) \partial_{x_i} g_i(x_i) = \partial_{x_i} \tilde{\pi}(x, e).$$

Plugging this in to (1) we obtain

$$\sum_{i=1}^{d_y} \partial_{p_i} \tilde{\pi}(x, e) \frac{g_i(x_i)}{\partial_{x_i} g_i(x_i)} = \tilde{\pi}(x, e). \tag{2}$$

Recall that $\tilde{\pi}(\cdot, e)$ is identified by Lemma 6. Thus the only unknowns involve g. We focus on cases in which *only* one price is observed. We note, however, that if all prices except one are observed, one may directly use (2) to identify the fraction $\frac{g_i(x_i)}{\partial_{x_i} g_i(x_i)}$ for the component i whose price is not observed. This is because that fraction is the only unknown in the linear equality. By varying x_i, holding everyone else fixed, one readily identifies the entire function g_i using an appropriate scale/location normalization.

We now turn to the more challenging setting in which we only observe the price of the d_y-th good. Identification of the terms $g_i(x_i)/\partial_{x_i} g_i(x_i)$ for $i \neq d_y$ is possible by varying x_{d_y} to generate a system of equations from (2), and then inverting a matrix. Our formal result makes use a rank condition to facilitate this inversion step.

Definition 4. We say that $f : X \to \mathbb{R}$ satisfies the rank condition at a point $x_{-d_y} \in \mathbb{R}^{d_y - 1}$ if there exists a collection of $\{x_{d_y,l}\}_{l=1}^{d_y-1}$ such that

(i) $x_l^* = (x'_{-d_y}, x_{d_y,l})' \in X$;

(ii) The square matrix

$$A(f, x^*) = \begin{bmatrix} \partial_{x_1} f(x_1^*) & \dots & \partial_{x_{d_y-1}} f(x_1^*) \\ \partial_{x_1} f(x_2^*) & \dots & \partial_{x_{d_y-1}} f(x_2^*) \\ \dots & \dots & \dots \\ \partial_{x_1} f(x_{d_y-1}^*) & \dots & \partial_{x_{d_y-1}} f(x_{d_y-1}^*) \end{bmatrix}$$

is nonsingular.

We will apply this rank condition to $\tilde{\pi}$ in place of f. It is helpful to recall that by Hotelling's lemma, partial derivatives of $\tilde{\pi}$ take the following form

$$\partial_{x_j} \tilde{\pi}(x, e) = \partial_{p_j} \pi(p, e)|_{p=g(x)} = y_j(g(x), e) \partial_{x_j} g_j(x_j),$$

where $y_j(g(x), e)$ is the supply function for good j. Thus, this rank condition applied to π may equivalently be interpreted as a rank condition involving the supply function for the goods as well as certain derivatives of g.

The following result provides conditions under which *either* quantiles or the conditional mean of $\boldsymbol{\pi}$ given $\mathbf{x}$ is sufficient to recover the price proxy function g.

Theorem 4. *Suppose that $\pi(\cdot, e)$ is differentiable for every $e \in E$ and Assumptions 4, 5, and 6 are satisfied. Then g is identified from the observed distribution of $F_{\boldsymbol{\pi}|\mathbf{x}}$ if one of the following testable conditions holds:*

(i) The assumptions of Lemma 6 are satisfied, and for every x_{-d_y} there exists $e^* \in [0,1]$ such that $\tilde{\pi}(\cdot, e^*)$ satisfies the rank condition at x_{-d_y};

(ii) Assumption 6 holds and $\mathbb{E}[\boldsymbol{\pi}|\mathbf{x} = \cdot]$ satisfies the rank condition at every x_{-d_y}.

This result states that the rank condition need only hold at some level of productivity e^* *or* the representative agent profit function $\mathbb{E}[\boldsymbol{\pi}|\mathbf{x} = \cdot]$. While our core analysis focuses on identification with heterogeneity, because the firm problem aggregates this result is also of interest for a representative agent analysis.

The rank condition is testable and can be satisfied, for instance, for the Diewert (1973) profit function presented in Section 3.

Example 6 (Diewert function, $d_y = 3$)**.** Let

$$\pi(p,e) = \sum_{i=1}^{3}\sum_{j=1}^{3} b_{i,j}(e) p_i^{1/2} p_j^{1/2}.$$

Suppose that p_3 is observed, and $p_1 = g_1(x_1)$ and $p_2 = g_2(x_2)$. Assume, moreover, that $\partial_{x_i} g_i(x_i) \neq 0$, for all x_i and $i = 1,2$. Fix any x_1 and x_2. Then the rank condition is satisfied if and only if there exists e^* such that

$$\frac{b_{1,1}(e^*)\sqrt{g_1(x_1)} + b_{1,2}(e^*)\sqrt{g_2(x_2)}}{b_{2,2}(e^*)\sqrt{g_2(x_2)} + b_{1,2}(e^*)\sqrt{g_1(x_1)}} \neq \frac{b_{1,3}(e^*)}{b_{2,3}(e^*)}.$$

In particular, if $g_1(\cdot) = g_2(\cdot)$, then the rank condition is satisfied if and only if

$$\frac{b_{1,1}(e^*) + b_{1,2}(e^*)}{b_{2,2}(e^*) + b_{1,2}(e^*)} \neq \frac{b_{1,3}(e^*)}{b_{2,3}(e^*)}.$$

In Example 6 the rank condition is satisfied except for a set of parameter values with Lebesgue measure zero. However, as the following example demonstrates, the rank condition may fail to hold for all possible values of parameters.

Example 7 (Cobb-Douglas)**.** For a fixed e, let $y_o \leq k^\alpha l^\beta$ such that $\alpha + \beta < 1$ and $\alpha, \beta > 0$. Then

$$\pi(p,e) = (1-\alpha-\beta)\left[\frac{p_k}{\alpha}\right]^{\frac{\alpha}{\alpha+\beta-1}}\left[\frac{p_l}{\beta}\right]^{\frac{\beta}{\alpha+\beta-1}}(p_o)^{-\frac{1}{\alpha+\beta-1}},$$

where $p = (p_o, p_k, p_l)'$. Suppose that only p_o is perfectly observed. Fix some $p_k = g_k(x_k)$ and $p_l = g_l(x_l)$. Then for any two p'_o and p''_o let $p' = (p'_o, p_k, p_l)'$ and $p'' =$

$(p_o'', p_k, p_l)'$. The matrix $A(\tilde{\pi}, x^*)$ is singular since it is equal to

$$\begin{bmatrix} \frac{\alpha\pi(p',e)}{(\alpha+\beta-1)g_k(x_k)}\partial_{x_k}g_k(x_k) & \frac{\beta\pi(p',e)}{(\alpha+\beta-1)g_l(x_l)}\partial_{x_l}g_l(x_l) \\ \frac{\alpha\pi(p'',e)}{(\alpha+\beta-1)g_k(x_k)}\partial_{x_k}g_k(x_k) & \frac{\beta\pi(p'',e)}{(\alpha+\beta-1)g_l(x_l)}\partial_{x_l}g_l(x_l) \end{bmatrix}$$

It can be shown that the rank condition is never satisfied for Cobb-Douglas production function if only one of the prices is perfectly observed.

The rank condition is not satisfied for the Cobb-Douglas production function because the ratios of any two different quantities chosen (e.g. l/k, or y_o/l) do not depend on the price of the last quantity. Indeed, recall that

$$\partial_{x_j}\tilde{\pi}(x,e) = y_j(g(x),e)\partial_{x_j}g_j(x_j).$$

Thus, if $y_j(g(x),e)/y_i(g(x),e)$ does not depend on observed price p_{d_y}, then the i-th column of $A(\tilde{\pi}, x^*)$ is a scaled version of the j-th column of $A(\tilde{\pi}, x^*)$. Hence, $A(\tilde{\pi}, x^*)$ is singular.

4.2. Price Proxies and Quantities Observed

In this section we show that if prices and profits are not observed, but price proxies and the net production vector are, then we may recover the distribution of what we term *pseudo-profits*. This distribution has the same distribution of profits conditional on prices, up to a scale parameter. Using the fact that one price is observed and with a location normalization on g (recall Definition 5(ii)), we recover the location and scale of profits, and thus we identify the distribution of profits conditional on prices, even though we have only observed a single price. Using this distribution we can identify the production possibilities sets by our previous arguments.

Assumption 7. *(i) The random variables* $\mathbf{x}, \mathbf{y}$*, and* $\mathbf{e}$ *satisfy*

$$\mathbf{y} = \arg\max_{y \in Y(\mathbf{e})} g(\mathbf{x})'y \quad \text{a.s.}$$

(ii) $\mathbb{E}\left[y(g(x),\mathbf{e})\right]$ *exists for each* $x \in X$ *and satisfies*

$$\mathbb{E}\left[y(g(x),\mathbf{e})\right] = \arg\max_{y \in \overline{Y}} g(x)'y$$

for some $\overline{Y}$, where the expectation is over the marginal distribution of $\mathbf{e}$.

(iii) For each $x \in X$,

$$\mathbb{E}\left[y(g(x), \mathbf{e})\right] = \mathbb{E}\left[\mathbf{y}|\mathbf{x} = x\right],$$

where

$$\mathbb{E}\left[\mathbf{y}|\mathbf{x} = x\right] = \lim_{\delta \to 0} \mathbb{E}\left[\mathbf{y}|\mathbf{x} \in B(\delta, x)\right]$$

and $B(\delta, x)$ is the closed ball of radius δ around x.

Part (i) states that y maximizes profits and is the unique maximizer. Parts (ii) and (iii) essentially state that a representative agent exists, and the conditional mean of $\mathbf{y}$ given $\mathbf{x}$ identifies the average supply function $\mathbb{E}\left[y(g(x), \mathbf{e})\right]$. To elaborate, part (ii) states that the average supply function $\mathbb{E}\left[y(g(x), \mathbf{e})\right]$ maximizes profits with a representative agent production possibilities set $\overline{Y}$. If $\mathbf{e}$ has finite support, this is a standard representative agent result for the firm problem (e.g. Kreps (2012), Proposition 13.1; Allen & Rehbeck (2017) provide an aggregation result that applies when $\mathbf{e}$ does not have finite support). Given the other assumptions, part (iii) is implied if $g(\cdot)$ is continuous and $\mathbf{x}$ and $\mathbf{e}$ are independent.[17]

By exploiting a symmetry feature that arises due to optimization (cf. Allen & Rehbeck (2017)), we obtain the following constructive identification result. To state the result, first define the representative agent profit function $\overline{\pi}(p) = \mathbb{E}\left[\pi(p, \mathbf{e})\right]$, where the expectation is taken over the marginal distribution of $\mathbf{e}$.

Theorem 5. *Let Assumptions 5, 6, and 7 hold and assume* $\mathbf{x}$ *and* $\mathbf{y}$ *are observed. If $\overline{\pi}$ is twice continuously differentiable and the mixed partial derivatives satisfy $\nabla_{j,d_y}\overline{\pi} \neq 0$ everywhere, then g is identified. In particular,*

$$g_j(t) - g_j(x_{0j}) = \int_{x_{0j}}^{t} \frac{\partial_{x_j}\mathbb{E}\left[\mathbf{y}_{d_y}|\mathbf{x} = x\right]}{\partial_{d_y}\mathbb{E}\left[\mathbf{y}_k|\mathbf{x} = x\right]} dx_j.$$

Recall that by Hotelling's lemma, twice differentiability of the aggregate profit function $\overline{\pi}(\cdot)$ amounts to differentiability of the aggregate supply function $\mathbb{E}\left[y(\cdot, \mathbf{e})\right]$. Assuming that the mixed partial derivatives of $\overline{\pi}$ are nonzero thus requires that there is some complementarity/substitutability between the components of the output/input vector. Formally, the aggregate supply function for each good j must have a nonzero derivative with respect to the price of good d_y. This rules out cases in which the representative firm production possibilities set $\overline{Y}$ can be written as a Cartesian product

[17] See Allen & Rehbeck (2017) for a rigorous statement.

of two nonempty sets, e.g. $\overline{Y} = \overline{Y}^1 \times \overline{Y}^2$.[18]

Once g is identified, profits are identified from the relation $\boldsymbol{\pi} = g(\mathbf{x})'\mathbf{y}$ whenever we observe price proxies and the netput vector $\mathbf{y}$. Thus, we may identify the conditional distribution of profits given prices from the conditional distribution of netputs given price proxies. This extends the applicability of our earlier analysis to settings in which profits and prices may not be observable. Recall that we assume at least one price is identified for this analysis. We note that if we drop this assumption (i.e. we drop the assumption that $g_{d_y}(x_{d_y}) = x_{d_y}$ for all x_{d_y}), it is possible to identify the function g up to location and scale by adapting arguments in Allen & Rehbeck (2017). Such an approach can be used to identify the distribution of profits given prices up to scale.

5. Extensions: Endogeneity

In this section we consider the possibility of endogeneity in prices. In particular, we study cases in which the independence condition that we have been using so far is violated (i.e., $F_{e|p}(\cdot|p) = F_e(\cdot)$ fails).

Our benchmark model considers perfectly competitive firms that face different prices. Price variation may arise because firms operate in different markets. In a general equilibrium setup, variation in market endowments can then drive variation in prices. Market endowments can be understood as the market characteristics that determine the initial distribution of inputs and outputs in each market before production and consumption take place. Price endogeneity may arise if productivity depends on some market characteristics. In this case, our setup will require some other market characteristics (instruments) that are independent of unobserved productivity. These instruments have to affect prices but must not be related to productivity.[19]

In order to address endogeneity, we describe how an instrumental variable can be used to identify the profit function π. In particular, assume that the analyst observes

[18] Such structure means that the supply function for components corresponding to $\overline{Y}^1$ does not depend on the prices for components corresponding to $\overline{Y}^2$. This in turn means that certain mixed partials of $\overline{\pi}$ must be zero. This does not pose a conceptual problem, since one could conduct analysis just for the components corresponding to $\overline{Y}^1$ separately from those corresponding to $\overline{Y}^2$.

[19] We note endogenous market entry/exit is less of a concern as a source of endogeneity due to the static nature of our exercise. In addition, our framework can accommodate zero profits. In a cross section, we can consider, as part of our population, those firms that have exited (temporarily) from the marketplace and are obtaining zero profits as a result.

$(\boldsymbol{\pi}, \mathbf{p}', \mathbf{w}')'$, where the instrumental variable $\mathbf{w}$ is supported on W.

The following assumption is an independence condition that requires the instrumental variable to be independent of the unobserved heterogeneity $\mathbf{e}$.

Assumption 8. *$F_{\mathbf{e}|\mathbf{w}}(\cdot|w) = F_{\mathbf{e}}(\cdot)$ for all $w \in W$.*

Assumption 8 together with the requirement that the profit function $\pi(p, \cdot)$ is strictly monotone imply the following integral equation familiar from the literature on nonparametric quantile instrumental variable models:

Lemma 7. *If $\pi(p, \cdot)$ is strictly increasing for all $p \in P$ and Assumptions 4 and 8 hold, then the following holds:*

$$\mathbb{P}\left(\mathbb{1}\left(\, \boldsymbol{\pi} \leq \pi(\mathbf{p}, e)\,\right) | \mathbf{w} = w\right) = e \tag{3}$$

for all $e \in E$ and $w \in W$.

The previous result says that in the presence of endogeneity, we can still rank firms conditional on the instrumental variable. Note that Equation (3) is an integral equation that connects the unknown profit function and distribution of the observables and normalized distribution of $\mathbf{e}$. Indeed, Equation (3) can be rewritten as

$$\int_{P_w} F_{\boldsymbol{\pi}|\mathbf{p},\mathbf{w}}(\pi(p, e)|p, w) f_{\boldsymbol{p}|\mathbf{w}}(p|w) dp = e$$

for all $w \in W$ and $e \in E$, where P_w denotes the support of $\mathbf{p}|\mathbf{w} = w$ and we assume the p.d.f. of $\mathbf{p}$ given $\mathbf{w}$ exists. The above integral equation has a unique solution in

$$\mathcal{L}^2(P) = \left\{ m(\cdot) \ : \ \int_P |m(x)|^2 dx < \infty \right\}$$

for every $e \in E$, if the operator $T_e : \mathcal{L}^2(P) \to \mathcal{L}^2(W)$ defined by

$$(T_e m)(w) = \int_{P_w} F_{\boldsymbol{\pi}|\mathbf{p},\mathbf{w}}(m(p)|p, w) f_{\boldsymbol{p}|\mathbf{w}}(p|w) dp,$$

is injective for every $e \in E$. Injectivity of integral operators is closely related to the notion of completeness. Numerous sufficient conditions for injectivity of integral operators are available in the literature.[20] For brevity we provide a simple identification result for our model when only one price is endogenous based on Hu et al. (2017). In

[20]See for example Newey & Powell (2003), Chernozhukov & Hansen (2005), DH́aultfoeuille et al. (2010), Andrews (2011), DH́aultfoeuille (2011), and Hu et al. (2017).

Appendix B we establish identification under high-level conditions using the results of Chernozhukov & Hansen (2005), and we derive a generalization of the identification result of Hu et al. (2017) to the multidimensional case.

Let $p_{-1} = (p_2, p_3, \ldots, p_{d_y})'$ (i.e., $p = (p_1, p_{-1}')'$). Suppose that $w = (w_1, p_{-1}')'$ with $w_1 \in W_1$ (i.e., there is one endogenous price and one non-price instrument for it). Assume that the support of $\mathbf{p}_1|(\boldsymbol{w_1} = w_1, \mathbf{p}_{-1} = p_{-1})$ is $[0, \overline{p}(w_1, p_{-1})]$. Fix some p_{-1}, and denote $\overline{p}^*(w_1) = \overline{p}(w_1, p_{-1})$, $f_1(p_1, w_1) = f_{\mathbf{p}_1|\mathbf{w}}(p_1|(w_1, p_{-1}')')$, $G(\pi, p_1) = F_{\boldsymbol{\pi}|\mathbf{p}}(\pi|(p_1, p_{-1}')')$, and $G_w(\pi, p_1, w_1) = F_{\boldsymbol{\pi}|\mathbf{p},\mathbf{w}_1}(\pi|(p_1, p_{-1}')', w_1)$.

Theorem 6. *Under the assumptions of Lemma 7, if for any fixed e and p_{-1}*

(i) $W_1 = [\underline{w}_1, \overline{w}_1]$;

(ii) $\overline{p}^*(\cdot)$ *is a continuously differentiable and strictly increasing function such that* $\overline{p}^*(\underline{w}_1) = 0$ *and* $\overline{p}^*(\overline{w}_1) < \infty$;

(iii) $f_1(p_1, \cdot)$ *is continuously differentiable for all p_1 such that*

$$f_1(\overline{p}^*(w_1), w_1) > 0,$$
$$\partial_{w_1} f_1(0, w_1) = 0,$$

for every $w_1 \in W_1$;

(iv) $G_w(\cdot, p_1, w_1) = G(\cdot, p_1)$ *for all $w_1 \in W_1$ and $p_1 \in [0, \overline{p}^*(w_1)]$;*

(v) $G(\cdot, p_1)$ *is invertible for every* $p_1 \in [0, \overline{p}^*(\overline{w}_1)]$,

Then for any $\pi(\cdot)$ and $\pi'(\cdot)$ that solve Equation (3) it must hold that $\pi(\mathbf{p}, \boldsymbol{e}) = \pi'(\mathbf{p}, \boldsymbol{e})$ a.s.

Condition (i) requires the instrument to have bounded support. Condition (ii) is a support restriction on the endogeneous price after conditioning on all other prices and the instrument. For instance, it is satisfied if $\overline{p}^*(w_1) = w_1$ (there is an exogeneous bound on the output price that is binding with positive probability). The assumption that the lower bound of the support of the endogeneous price is 0 can be relaxed (see Theorem 8 in Appendix B). Condition (iv) is an exclusion restriction requiring profits to be independent of the instrument conditional on all prices.

Note that if the profit function is identified and firms are price-takers and profit maximizers, then all the results of Theorem 2, including point identification of $Y(\cdot)$, hold since Assumption 3 can be satisfied even if prices have bounded support.

6. Conclusion

Classical analysis of the firm problem has demonstrated the power of duality. This paper extends existing work focused on deterministic settings to settings with rich heterogeneity. Our key assumption is that firms can be ranked in terms of productivity. We use a weak monotonicity condition on profits, which generalizes a strict monotonicity condition in Matzkin (2003). With this assumption, we show how to identify firm production possibilities sets under rich variation in prices, and also describe the most that can be learned about such sets with limited price variation. To facilitate estimation, we present a generic result relating estimation error in profit functions and estimation error of production sets. This parallels a classical result in convex analysis but is novel because it applies when one only observes nonnegative prices. We also present novel results showing how to work with price proxies instead of prices. We provide a new identification technique exploiting homogeneity, which may be of independent interest. Finally, we describe how the independence conditions in our main analysis may be relaxed in the presence of endogeneity.

References

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411–2451.

Aliprantis, C. D. & Border, K. (2006). *Infinite Dimensional Analysis*. Springer-Verlag.

Allen, R. & Rehbeck, J. (2017). Identification with additively separable heterogeneity. *Working Paper*.

Andrews, D. W. (2011). Examples of l2-complete and boundedly-complete distributions. *Mimeo*.

Brunel, V.-E. (2016). Concentration of the empirical level sets of Tukey's halfspace depth. *Probability Theory and Related Fields*, (pp. 1–32).

Cherchye, L., Demuynck, T., De Rock, B., & De Witte, K. (2014). Non-parametric analysis of multi-output production with joint inputs. *The Economic Journal*, 124(577), 735–775.

Cherchye, L., Demuynck, T., De Rock, B., & Verschelde, M. (2018). *Nonparametric identification of unobserved technological heterogeneity in production*. Working Paper Research 335, National Bank of Belgium.

Chernozhukov, V., Galichon, A., Henry, M., & Pass, B. (2017). Single market nonparametric identification of multi-attribute hedonic equilibrium models. *arXiv preprint arXiv:1709.09570*.

Chernozhukov, V. & Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1), 245–261.

Chiong, K. X., Galichon, A., & Shum, M. (2016). Duality in dynamic discrete-choice models. *Quantitative Economics*, 7(1), 83–115.

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The review of economics and statistics*, (pp. 28–45).

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., & Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2), 445–510.

DH́aultfoeuille, X. (2011). On the completeness condition in nonparametric instrumental problems. *Econometric Theory*, 27(3), 460–471.

DH́aultfoeuille, X., Février, P., et al. (2010). Identification of mixture models using support variations. *Manuscript, ENSAE.*

Diewert, W. (1973). Functional forms for profit and transformation functions. *Journal of Economic Theory*, 6(3), 284 – 316.

Diewert, W. E. (1982). Duality approaches to microeconomic theory. *Handbook of mathematical economics*, 2, 535–599.

Epple, D., Gordon, B., & Sieg, H. (2010). A new approach to estimating the production function for housing. *American Economic Review*, 100(3), 905–24.

Fuss, M. & McFadden, D., Eds. (1978). *Production Economics: A Dual Approach to Theory and Applications: Applications of the Theory of Production*. Elsevier.

Galichon, A. & Salanié, B. (2015). Cupid's invisible hand: Social surplus and identification in matching models.

Gandhi, A., Navarro, S., & Rivers, D. (2017). On the identification of gross output production functions. *Mimeo.*

Grieco, P. L. & McDevitt, R. C. (2016). Productivity and quality in health care: Evidence from the dialysis industry. *The Review of Economic Studies*, 84(3), 1071–1105.

Griliches, Z. & Mairesse, J. (1995). *Production functions: the search for identification*. Technical report, National Bureau of Economic Research.

Hu, Y., Schennach, S. M., & Shiu, J.-L. (2017). Injectivity of a class of integral operators with compactly supported kernels. *Journal of Econometrics*, 200(1), 48–58.

Kaido, H. & Santos, A. (2014). Asymptotically efficient estimation of models defined by convex moment inequalities. *Econometrica*, 82(1), 387–413.

Koenker, R. & Ng, P. (2005). Inequality constrained quantile regression. *Sankhyā: The Indian Journal of Statistics*, (pp. 418–440).

Kreps, D. M. (2012). *Microeconomic foundations I: choice and competitive markets*, volume 1. Princeton university press.

Lau, L. J. (1972). Profit functions of technologies with multiple inputs and outputs. *The Review of Economics and Statistics*, (pp. 281–289).

Levinsohn, J. & Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2), 317–341.

López, M. & Still, G. (2007). Semi-infinite programming. *European Journal of Operational Research*, 180(2), 491–518.

Malikov, E. (2017). Structural estimation of multi-product production functions and productivity. *Mimeo*.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). *Microeconomic theory*, volume 1. Oxford university press New York.

Matzkin, R. L. (2003). Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5), 1339–1375.

McFadden, D. (1981). Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272.

Newey, W. K. & Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5), 1565–1578.

Olley, G. S. & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263–1297.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.

Shephard, R. W. (1953). *Cost and production functions*. Princeton University Press.

Varian, H. R. (1984). The nonparametric approach to production analysis. *Econometrica: Journal of the Econometric Society*, (pp. 579–597).

A. Proofs of Main Results

A.1. Proof of Theorem 1

Fix some p. For every k define

$$E(\pi, k) = \{e \in E \mid \pi(p, e) \leq k\}.$$

Note that this set is compact because $\pi(p, \cdot)$ is lower semicontinuous and E is closed. Since $E(\pi, k) \subseteq E$ and E is bounded, the set $E(\pi, k)$ is bounded, hence compact. Define also

$$e^*(\pi, k) = \max_{e \in E(\pi,k)} e,$$

where the maximum exists because $E(\pi, k)$ is compact. Note that by weak monotonicity of $\pi(p, \cdot)$, $e \in E(\pi, k)$ if and only if $e \leq e^*(\pi, k)$. Hence,

$$F_{\boldsymbol{\pi}|\mathbf{p}}(k|p) = \mathbb{P}\left(\pi(\mathbf{p}, \mathbf{e}) \leq k|\mathbf{p} = p\right) = \mathbb{P}\left(\mathbf{e} \leq e^*(\pi, k)|\mathbf{p} = p\right) = F_{\mathbf{e}}(e^*(\pi, k)),$$

where the last equality follows from Assumption 2. Thus, for any conjectured $F_{\mathbf{e}}$ that is strictly monotone, we identify $e^*(\pi, k)$ via

$$F_{\mathbf{e}}^{-1}(F_{\boldsymbol{\pi}|\mathbf{p}}(k|p)) = e^*(\pi, k).$$

To identify $\pi(p,\cdot)$, first note that for each k, $\pi(p, e^*(\pi, k)) = k$ because $\pi(p,\cdot)$ is lower semicontinuous. For arbitrary e, we have

$$\pi(p,e) = \inf \{k \,:\, e \leq e^*(\pi, k)\}.$$

by weak monotonicity of $\pi(p,\cdot)$. Thus, $\pi(p,\cdot)$ is identified.

A.2. Proof of Theorem 2

It is immediate that $\tilde{Y}(e)$ is closed, convex, and satisfies free disposal for every $e \in E$. Moreover, $\sup_{y\in\tilde{Y}(e)} p'y = \pi(p,e)$ for every $p \in P$ and $e \in E$. Thus, conclusion (i) follows from the fact that $\pi(p,\cdot)$ is identified for each $p \in P$ by Theorem 1.

To establish conclusion (ii), recall that under the assumptions of Theorem 1 and Assumption 4, $Y'(\cdot)$ can generate the data if and only if $\max_{y\in\tilde{Y}'(e)} p'y = \pi(p,e)$ for every $p \in P$. The set $\tilde{Y}(e)$ is constructed as the largest set consistent with profit maximization, and since it is closed, convex, and satisfies free disposal, it is the largest such set consistent with profit maximization. Since production correspondences *also* must satisfy the recession cone property, we obtain that $Y'(e) \subseteq \tilde{Y}(e)$.

To prove (iii), note that since $\pi(\cdot, e)$ is homogeneous of degree 1 for every $e \in E$ we can identify $\pi(\cdot, e)$ over

$$\bigcup_{\lambda>0} \{\lambda p \,:\, p \in P\}.$$

Next, since $\pi(\cdot, e)$ is convex it is continuous, hence it is identified over

$$\operatorname{int}\left(\operatorname{cl}\left(\bigcup_{\lambda>0} \{\lambda p \,:\, p \in P\}\right)\right).$$

When Assumption 3 holds, identification of $Y(\cdot)$ follows from Lemma 1.

A.3. Proof of Theorem 3 and Proposition 1

The Hausdorff distance between two convex sets $A, B \subseteq \mathbb{R}^{d_y}$ is given by

$$d_H(A,B) = \max\left\{\sup_{a\in A}\inf_{b\in B}\|a-b\|, \sup_{b\in B}\inf_{a\in A}\|a-b\|\right\}.$$

Alternatively, the Hausdorff distance can be defined as

$$d_H(A,B) = \min\{\rho \geq 0 \ : \ A \subseteq B + \rho \mathbb{S}^{d_y-1}, B \subseteq A + \rho \mathbb{S}^{d_y-1}\},$$

where $\mathbb{S}^{d_y-1} = \{y \in \mathbb{R}^{d_y} \ : \ \|y\| \leq 1\}$ is the unit ball. The support function of a closed convex set A is defined for $u \in \mathbb{R}^{d_y}$ via $h_A(u) = \sup_{w \in A} u'w$.

We now provide a key lemma. This result generalizes a classical result that holds for $\bar{P} \subseteq \mathbb{S}^{d_y-1}$, and may be of independent interest.

Lemma 8. *Let $a, b : \mathbb{R}^{d_y}_{++} \to \mathbb{R}$, $d_y \geq 2$, be convex and homogeneous of degree 1. Define*

$$A = \left\{ y \in \mathbb{R}^{d_y} \ : \ p'y \leq a(p), \ \forall p \in \bar{P} \right\}$$
$$B = \left\{ y \in \mathbb{R}^{d_y} \ : \ p'y \leq b(p), \ \forall p \in \bar{P} \right\},$$

where $\bar{P} \subseteq \mathbb{S}^{d_y-1} \bigcap \mathbb{R}^{d_y}_{++}$ is convex and compact. Then

$$d_H(A,B) = \sup_{p \in \bar{P}} \|a(p) - b(p)\|.$$

Proof. First note that $a(p)$ and $b(p)$ are values of the support functions of A and B evaluated at $p \in \bar{P}$, respectively, since $a(\cdot)$ and $b(\cdot)$ are homogeneous of degree 1 and convex.

Next note that for convex sets $C, D \subseteq \mathbb{R}^{d_y}$ the following is true: $C \subseteq D$ if and only if $h_C(u) \leq h_D(u)$ for all $u \in \mathbb{S}^{d_y-1}$. Hence,

$$\{\rho \in \mathbb{R}_+ \ : \ A \subseteq B + \rho \mathbb{S}^{d_y-1}, B \subseteq A + \rho \mathbb{S}^{d_y-1}\} \iff$$
$$\{\rho \in \mathbb{R}_+ \ : \ h_A(u) \leq h_{B+\rho\mathbb{S}^{d_y-1}}(u), h_B(u) \leq h_{A+\rho\mathbb{S}^{d_y-1}}(u), \forall u \in \mathbb{S}^{d_y-1}\}.$$

By construction, $h_A(u) = h_{B+\rho\mathbb{S}^{d_y-1}}(u) = h_B(u) = h_{A+\rho\mathbb{S}^{d_y-1}}(u) = \infty$ for all $u \in \mathbb{S}^{d_y-1} \setminus \bar{P}$. Hence,

$$\{\rho \in \mathbb{R}_+ \ : \ A \subseteq B + \rho \mathbb{S}^{d_y-1}, B \subseteq A + \rho \mathbb{S}^{d_y-1}\} \iff$$
$$\{\rho \in \mathbb{R}_+ \ : \ h_A(u) \leq h_{B+\rho\mathbb{S}^{d_y-1}}(u), h_B(u) \leq h_{A+\rho\mathbb{S}^{d_y-1}}(u), \forall u \in \bar{P}\} \iff$$
$$\{\rho \in \mathbb{R}_+ \ : \ h_A(u) \leq h_B(u) + h_{\rho\mathbb{S}^{d_y-1}}(u), h_B(u) \leq h_A(u) + h_{\rho\mathbb{S}^{d_y-1}}(u), \forall u \in \bar{P}\} \iff$$
$$\{\rho \in \mathbb{R}_+ \ : \ h_A(u) \leq h_B(u) + \rho, h_B(u) \leq h_A(u) + \rho, \forall u \in \bar{P}\} \iff$$
$$\{\rho \in \mathbb{R}_+ \ : \ \sup_{u \in \bar{P}} \|h_A(u) - h_B(u)\| \leq \rho\}.$$

Hence, $d_H(A,B) = \sup_{p\in\bar{P}} \|a(p) - b(p)\|$. ■

To prove Theorem 3 note that since $\pi(\cdot, e)$ and $\hat{\pi}(\cdot, e)$ are homogeneous of degree 1, we have

$$\pi(p,e)/\|p\| = \pi\left(p/\|p\|, e\right)$$
$$\hat{\pi}(p,e)/\|p\| = \hat{\pi}\left(p/\|p\|, e\right)$$

for all $p \in \bar{P}$ and $e \in E$. Thus, Theorem 3 is obtained as corollary.

We now turn to the proof of Proposition 1. We first present two lemmas, which are modifications of Lemmas 6 and 7 in Brunel (2016).

Lemma 9. *Assume that $\bar{P} \subseteq \mathbb{S}^{d_y-1} \bigcap P$ is compact and convex. Let $a : \bar{P} \to \mathbb{R}$ be a continuous function. Let $A = \{y \in \mathbb{R}^{d_y} \;:\; p'y \le a(p), p \in \bar{P}\}$. For all $p^* \in \bar{P}$ there exists $y^* \in A$ such that $h_A(p^*) = p^{*\prime} y^*$. Moreover, there exists $P^* \subseteq \bar{P}$ such that*

(i) The cardinality of P^* is less than or equal to d_y;

(ii) $p'y^* = a(p)$ for all $p \in P^*$;

(iii) $p^* = \sum_{p\in P^*} \lambda_p p$ for some nonnegative numbers λ_p.

Proof. Fix some $p \in \bar{P}$. Since A is closed and bounded from above, y^* always exists. Since $a(\cdot)$ is continuous on $\bar{P}$ the set $P' = \{p \in \bar{P} : p'y^* = a(p)\}$ is non-empty. Indeed, by way of contradiction assume that P' is empty. Hence, for all $p \in \bar{P}$ $p'y^* < a(p)$. Since the function $a(\cdot) - \cdot' y^*$ is strictly positive on a compact $\bar{P}$, there exists $\nu > 0$ that bounds $a(\cdot) - \cdot' y^*$ from below. Hence, for every $p \in \bar{P}$

$$p'(y^* + \nu p^*) = p'y^* + \nu p' p^* \le a(p) - \nu + \nu p' p^* \le a(p).$$

Thus, $(y^* + \nu p^*) \in A$. But the later is not possible since $p^*(y^* + \nu p^*) = a(p^*) + \nu > a(p^*)$ implies that y^* is not a maximizer.

The rest of the lemma follows from Theorem 2(b) in López & Still (2007). ■

Lemma 10. *Assume that $\bar{P} \subseteq \mathbb{S}^{d_y-1} \bigcap P$ is compact and convex. Let $a : \bar{P} \to \mathbb{R}$ be continuous convex homogeneous of degree 1 function and $\{b_n : \bar{P} \to \mathbb{R}\}$ be a sequence of continuous homogeneous of degree 1 functions such that*

$$A = \left\{ y \in \mathbb{R}^{d_y} \;:\; p'y \le a(p),\; \forall p \in \bar{P} \right\},$$
$$B_n = \left\{ y \in \mathbb{R}^{d_y} \;:\; p'y \le b_n(p),\; \forall p \in \bar{P} \right\}$$

are non-empty for all $n \in \mathbb{N}$. *Assume that* $\eta_n = \sup_{p\in\bar{P}} \|a(p) - b_n(p)\| = o(1)$ *and* $0 < r = \inf_{p\in\bar{P}} a(p) < R = \sup_{p\in\bar{P}} a(p) < \infty$, *then there exists* $N > 0$ *such that*

$$\sup_{p\in\bar{P}} \|a(p) - h_{B_n}(p)\| \leq \eta_n \frac{R}{r} \frac{1 + \eta_n/R}{1 - \eta_n/r},$$

for all $n > N$.

Proof. Fix some $p^* \in \bar{P}$ and some n such that $\eta_n < r$. By Lemma 9 there exists a finite set P_n^*, a collection of nonnegative numbers $\{\lambda_{p,n}\}_{p\in P_n^*}$ and $y_n^* \in B_n$ such that $h_{B_n} = p^{*\prime} y_n^*$, $p^* = \sum_{p\in P_n^*} \lambda_{p,n} p$, and $p' y_n^* = b_n(p)$ for all $p \in P_n^*$. Note that for all $p \in p_n^*$ we have that $b_n(p) = h_{B_n}(p)$. Then

$$\begin{aligned} a(p^*) = h_A(p^*) = h_A\left(\sum_{p\in P_n^*} \lambda_{p,n} p\right) &\leq \sum_{p\in P_n^*} \lambda_{p,n} h_A(p) = \sum_{p\in P_n^*} \lambda_{p,n} a(p) \leq \sum_{p\in P_n^*} \lambda_{p,n}(b_n(p) + \eta_n) \\ &= \sum_{p\in P_n^*} \lambda_{p,n} p' y_n^* + \eta_n \sum_{p\in P_n^*} \lambda_{p,n} = p^{*\prime} y_n^* + \eta_n \sum_{p\in P_n^*} \lambda_{p,n} = h_{B_n}(p^*) + \eta_n \sum_{p\in P_n^*} \lambda_{p,n} \end{aligned} \tag{4}$$

Moreover,

$$h_{B_n}(p^*) \leq b_n(p^*) \leq a(p^*) + \eta_n. \tag{5}$$

Hence, $\|a(p^*) - h_{B_n}(p^*)\| \leq \eta_n \max\{1, \sum_{p\in P_n^*} \lambda_{p,n}\}$.

Next note that the inequality in (5) implies that

$$\sum_{p\in P_n^*} \lambda_{p,n} p' y_n^* = p^{*\prime} y_n^* = h_{B_n}(p^*) \leq a(p^*) + \eta \leq R + \eta_n.$$

In addition,

$$\sum_{p\in P_n^*} \lambda_{p,n} p' y_n^* = \sum_{p\in P_n^*} \lambda_{p,n} b_n(p) \geq \sum_{p\in P_n^*} \lambda_{p,n}(a(p) - \eta_n) \geq \sum_{p\in P_n^*} \lambda_{p,n}(r - \eta_n).$$

Hence,

$$\sum_{p\in P_n^*} \lambda_{p,n} \leq \frac{R + \eta_n}{r - \eta_n}.$$

As a result,

$$\|a(p^*) - h_{B_n}(p^*)\| \leq \eta_n \max\{1, \sum_{p\in P_n^*} \lambda_{p,n}\} = \eta_n \max\left\{1, \frac{R + \eta_n}{r - \eta_n}\right\} = \eta_n \frac{R}{r} \frac{1 + \eta_n/R}{1 - \eta_n/r}.$$

$\blacksquare$

To prove Theorem 3 note that since $\pi(\cdot, e)$ and $\hat{\pi}(\cdot, e)$ are homogeneous of degree 1, we have

$$\pi(p,e)/\|p\| = \pi\left(p/\|p\|, e\right)$$
$$\hat{\pi}(p,e)/\|p\| = \hat{\pi}\left(p/\|p\|, e\right)$$

To prove Proposition 1, note that by Lemma 8, with probability 1,

$$d_H(Y_{\bar{P}}(e), \hat{Y}_{\bar{P}}(e)) = \sup_{p\in\bar{P}} \left\| \pi(p/\|p\|, e) - h_{\hat{Y}_{\bar{P}}(e)}(p/\|p\|) \right\|.$$

The conclusion then follows from applying Lemma 10 to the right hand side of the equality above.

A.4. Proof of Theorem 4

To prove sufficiency of (i), note that $\tilde{\pi}(x, \cdot)$ is identified for every $x \in X$ by Lemma 6.

Fix some x_{-d_y} and take $e^* \in E$ from condition (i). We abuse notation and drop e^*. By homogeneity of degree 1 of $\pi(\cdot)$ we have that for every $x \in X$

$$\sum_{i=1}^{d_y} \partial_{g_i}\pi(g(x))g_i(x_i) = \pi(g(x)). \tag{6}$$

Moreover, since $\tilde{\pi}(x) = \pi(g(x))$, we have that

$$\partial_{g_i}\pi(g(x))\partial_{x_i}g_i(x_i) = \partial_{x_i}\tilde{\pi}(x), \tag{7}$$

for every $i = 1, \ldots, d_y$. Combining (6) and (7) we get that

$$\sum_{i=1}^{d_y} \partial_{x_i}\tilde{\pi}(x)\frac{1}{\partial_{x_i}(\log(g_i(x_i)))} = \tilde{\pi}(x). \tag{8}$$

as long as $0 < \left\|\dfrac{\partial_{x_i}g_i(x_i)}{g_i(x_i)}\right\| < \infty$ for every $i = 1, \ldots, d_y$.

Let $t = \left(\dfrac{1}{\partial_{x_i}(\log(g_i(x_i)))}\right)_{1,\ldots,d_y-1}$. Note that t does not depend on x_{d_y}. Since $\tilde{\pi}$

satisfies the rank condition there exists non-singular $A(\tilde{\pi}(x^*))$ such that equation (8) can be rewritten as

$$At = b, \tag{9}$$

where $b = (b_l)_{l=1,\dots,d_y-1}$ and $b_l = \tilde{\pi}(x_l^*) - \partial_{x_{d_y}}\tilde{\pi}(x_l^*)x_{d_y,l}$. Since $A(\tilde{\pi}(x^*))$ is of full rank, t is identified. Since the choice of x_{-d_y} was arbitrary and we know the location (Assumption 5.(ii)) we identify $g_i(\cdot)$ for every $i = 1, \dots, d_y - 1$.

Sufficiency of (ii) follows from applying the same arguments as in the proof of sufficiency of (i) to the function $\mathbb{E}[\boldsymbol{\pi}|\mathbf{x} = \cdot]$.

A.5. Proof of Theorem 5

This follows by adapting arguments in Allen & Rehbeck (2017). The envelope theorem applied to the representative firm problem yields Hotelling's Lemma,

$$\mathbb{E}[y(g(x), \mathbf{e})] = \nabla\overline{\pi}(g(x)).$$

Differentiating, we obtain

$$\partial_{x_k}\mathbb{E}[y(g(x), \mathbf{e})] = \nabla_{j,k}\overline{\pi}(g(x))\partial_{x_k}g_k(x_k). \tag{10}$$

Because $\overline{\pi}$ is twice continuously differentiable, its Hessian is a positive semi-definite matrix. In particular, $\nabla_{j,k}\overline{\pi} = \nabla_{k,j}\overline{\pi}$. When this mixed cross-partial is nonzero, we can divide (10) and its counterpart with j, k interchanged to obtain,

$$\frac{\partial_{x_j}\mathbb{E}[y_k(g(x), \mathbf{e})]}{\partial_{x_k}\mathbb{E}[y_j(g(x), \mathbf{e})]} = \frac{\partial_{x_j}g_j(x_j)}{\partial_{x_k}g_k(x_k)}. \tag{11}$$

Now set $k = d_y$. Then (11) is valid because we have assumed the global restriction $\nabla_{j,d_y}\overline{\pi} \neq 0$ for each j. Since $\mathbb{E}[\mathbf{y}|\mathbf{x} = x] = \mathbb{E}[y(g(x), \mathbf{e})]$, and $\partial_{x_{d_y}}g(x_{d_y}) = 1$ by Assumption 5(i), we identify differences in $\partial_{x_j}g_j(\cdot)$ for all j by integrating (11). By Assumption 5(ii), we have $g(x_0) = p_0$ for some known x_0 and p_0, which identifies the levels, and hence g_j is identified for each j.

A.6. Proof of Lemma 7

Fix some $w \in W$ and $e \in E$. First, note that by the law of iterated expectations

$$\mathbb{P}\left(\boldsymbol{\pi} - \pi(\mathbf{p}, e) \leq 0 | \mathbf{w} = w\right) = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}\left(\, \pi(p, \mathbf{e}) - \pi(p, e) \leq 0\, \right) | \mathbf{p} = p, \mathbf{w} = w\right] | \mathbf{w} = w\right].$$

By strict monotonicity of $\pi(p, \cdot)$ it follows that:

$$\mathbb{E}\left[\mathbb{1}\left(\, \pi(p, \mathbf{e}) - \pi(p, e) \leq 0\, \right) | \mathbf{p} = p, \mathbf{w} = w\right] = \mathbb{E}\left[\mathbb{1}\left(\, \mathbf{e} \leq e\, \right) | \mathbf{p} = p, \mathbf{w} = w\right]$$

The law of iterated expectations together with Assumptions 4 and 8 then imply that

$$\mathbb{P}\left(\boldsymbol{\pi} - \pi(\mathbf{p}, e) \leq 0 | \mathbf{w} = w\right) = e.$$

A.7. Proof of Theorem 6

Note that equation (3) holds by Lemma 7. Fix some p_{-1} and e. Under condition (iv), (3) can be rewritten as

$$\int_0^{\overline{p}^*(w_1)} G(\pi((p_1, p_{-1}')', e), p_1) f_1(p_1, w_1) dp_1 = e$$

for all $w_1 \in W_1$. Let $v(p_1) = G(\pi((p_1, p_{-1}')', e), p_1)$. Since $G(\cdot, p_1)$ is observed and invertible, if we identify $v(\cdot)$, then we identify $\pi((\cdot, p_{-1}')', e)$ as well. Consider the following integral equation:

$$\int_0^{\overline{p}^*(w_1)} v(p_1) f_1(p_1, w_1) dp_1 = e,$$

for all $w_1 \in W_1$. It has a unique solution since all the assumptions of Theorem 2.1 in Hu et al. (2017) are satisfied. The conclusion of the theorem then follows from the fact that e and p_{-1} were chosen arbitrary.

B. Endogeneity

Note that equation (3) is equivalent to the IV model of quantile treatment effect of Chernozhukov and Hansen (2005). Thus we can directly invoke their identification result. For some fixed $\delta, \underline{f} > 0$ define the relevant parameter space $\mathcal{P}$ as the convex hull of functions $\pi'(\cdot, e)$ that satisfy: (i) for every $w \in W$, $\mathbb{P}\left(\boldsymbol{\pi} \leq \pi(\mathbf{p}, e) | \mathbf{w} = w\right) \in [e - \delta, e + \delta]$, and (ii) for each $p \in P$,

$$\pi'(p, e) \in s_p = \left\{ \pi \ : \ f_{\boldsymbol{\pi}|\mathbf{p},\mathbf{w}}(\pi | p, w) \geq \underline{f} \text{ for all } w \text{ with } f_{\mathbf{w}|\mathbf{p}}(w|p) > 0 \right\}.$$

Moreover, let $f_{\boldsymbol{\epsilon}|\mathbf{p},\mathbf{w}}(\cdot|p, w; e)$ denote the density of $\boldsymbol{\epsilon} = \boldsymbol{\pi} - \pi(\mathbf{p}, e)$ conditional on $\mathbf{p}$ and $\mathbf{w}$.

Theorem 7. *Suppose that*

(i) $\pi(p, \cdot)$ *is strictly increasing for every* $p \in P$*;*

(ii) Assumptions 4 and 8 hold;

(iii) $\boldsymbol{\pi}$ *and* $\mathbf{w}$ *have bounded support;*

(iv) $f_{\boldsymbol{\epsilon}|\mathbf{p},\mathbf{w}}(\cdot|p, w; e)$ *is continuous and bounded over* $\mathbb{R}$ *for all* $p \in P$, $w \in W$, *and* $e \in E$*;*

(v) $\pi(p, e) \in s_p$ *for all* $p \in P$ *and* $e \in E$*;*

(vi) For every $e \in E$, $\mathbb{E}\left[(\pi'(\mathbf{p}, e) - \pi(\mathbf{p}, e))\omega(\mathbf{p}, \mathbf{w}; e)|\mathbf{w}\right] = 0$ a.s. *implies that* $\pi'(\mathbf{p}, e) - \pi(\mathbf{p}, e)$ a.s., *for* $\omega(p, w; e) = \int_0^1 f_{\boldsymbol{\epsilon}|\mathbf{p},\mathbf{w}}(\delta(\pi'(p, e) - \pi(p, e))|p, w; e)d\delta > 0$*;*

Then for any other $\pi'(\cdot, e) \in \mathcal{P}$ *such that*

$$\mathbb{P}\left(\mathbb{1}\left(\, \boldsymbol{\pi} \leq \pi'(\mathbf{p}, e) \,\right) | \mathbf{w} = w\right) = e$$

for all $w \in W$, $\pi'(\mathbf{p}, e) = \pi(\mathbf{p}, e)$ a.s..

Alternatively, we can use the identification strategy presented in Hu et al. (2017). For any reordering of p and w, $(\tilde{p}_1, \ldots, \tilde{p}_{d_y})'$ and $(\tilde{w}_1, \ldots, \tilde{w}_{d_y})'$, and for every $k = 1, \ldots, d_y$, let the support of

$$\tilde{\boldsymbol{p}}_{\boldsymbol{k}} | \tilde{\boldsymbol{p}}_{\mathbf{1}} = \tilde{p}_1, \ldots, \tilde{\boldsymbol{p}}_{k-1} = \tilde{p}_{k-1}, \tilde{\boldsymbol{w}}_k = \tilde{w}_k, \ldots, \tilde{\boldsymbol{w}}_{y_d} = \tilde{w}_{y_d}$$

be

$$[\underline{p}_k(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d}), \overline{p}_k(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d})].$$

Theorem 8. *Suppose that*

(i) $\pi(p, \cdot)$ *is strictly increasing for every* $p \in P$*;*

(ii) Assumptions 4 and 8 hold;

(iii) $F_{\boldsymbol{\pi}|\mathbf{p},\mathbf{w}}(\cdot|p, w) = F_{\boldsymbol{\pi}|\mathbf{p}}(\cdot|p)$ *for all* $p \in P$ *and* $w \in W$*;*

(iv) $F_{\boldsymbol{\pi}|\mathbf{p}}(\cdot|p)$ *is invertible for every* $p \in P$*;*

(v) $W = \times_{k=1}^{d_y} [\underline{w}_k, \overline{w}_k]$*;*

(vi) There exists reordering of p *and* w*,* $(\tilde{p}_1, \ldots, \tilde{p}_{d_y})'$ *and* $(\tilde{w}_1, \ldots, \tilde{w}_{d_y})'$*, such that*

$$f_{\tilde{\boldsymbol{p}}_k|\tilde{\boldsymbol{p}}_1,\ldots,\tilde{\boldsymbol{p}}_{k-1},\tilde{\boldsymbol{w}}}(\cdot|\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}) = f_{\tilde{\boldsymbol{p}}_k|\tilde{\boldsymbol{p}}_1,\ldots,\tilde{\boldsymbol{p}}_{k-1},\tilde{\boldsymbol{w}}_k,\ldots,\tilde{\boldsymbol{w}}_{y_d}}(\cdot|\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d})$$

for every $k = 1, \ldots, d_y$*;*

(vii) For every $k = 1, \ldots, d_y$ *and every* $(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_{k+1}, \ldots, \tilde{w}_{y_d})'$ *there exists* $\underline{p}_k$ *and* $\overline{p}_k$ *such that*

$$\underline{p}_k \le \underline{p}_k^*(\tilde{w}_k) \le \overline{p}_k^*(\tilde{w}_k) \le \overline{p}_k,$$

and

$$\underline{p}_k = \underline{p}_k^*(\underline{\tilde{w}}_k) = \overline{p}_k^*(\underline{\tilde{w}}_k),$$

where $\underline{p}_k^*(\tilde{w}_k) = \underline{p}_k(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d})$ *and* $\overline{p}_k^*(\tilde{w}_k) = \overline{p}_k(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d})$*;*

(viii) For every $k = 1, \ldots, d_y$ *and every* $(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_{k+1}, \ldots, \tilde{w}_{y_d})'$*,* $\underline{p}_k^*(\cdot)$ *and* $\overline{p}_k^*(\cdot)$ *a continuously differentiable with* $\overline{p}_k^*(\cdot)$ *being strictly increasing;*

(ix) For every $k = 1, \ldots, d_y$*, every* $(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_{k+1}, \ldots, \tilde{w}_{y_d})'$*, every* $\tilde{w}_k$ *either*

$$f_{\tilde{\boldsymbol{p}}_k|\tilde{\boldsymbol{p}}_1,\ldots,\tilde{\boldsymbol{p}}_{k-1},\tilde{\boldsymbol{w}}_k,\ldots,\tilde{\boldsymbol{w}}_{y_d}}(\underline{p}_k^*(\tilde{w}_k)|\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d}) = 0$$

or $\partial_{\tilde{w}_k} \underline{p}_k^*(\tilde{w}_k) = 0$*;*

(x) For every $k = 1, \ldots, d_y$*, every* $(\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_{k+1}, \ldots, \tilde{w}_{y_d})'$*, every* $\tilde{w}_k$*,*

$$\partial_{p_k} f_{\tilde{\boldsymbol{p}}_k|\tilde{\boldsymbol{p}}_1,\ldots,\tilde{\boldsymbol{p}}_{k-1},\tilde{\boldsymbol{w}}_k,\ldots,\tilde{\boldsymbol{w}}_{y_d}}(p_k|\tilde{p}_1, \ldots, \tilde{p}_{k-1}, \tilde{w}_k, \ldots, \tilde{w}_{y_d})$$

exists and continuous;

(xi) *For every* $k = 1, \dots, d_y$, *every* $(\tilde{p}_1, \dots, \tilde{p}_{k-1}, \tilde{w}_{k+1}, \dots, \tilde{w}_{y_d})'$, *every* $\tilde{w}_k$

$$f_{\tilde{\boldsymbol{p}}_k|\tilde{\boldsymbol{p}}_1,\dots,\tilde{\boldsymbol{p}}_{k-1},\tilde{\boldsymbol{w}}_k,\dots,\tilde{\boldsymbol{w}}_{y_d}}(\overline{p}_k^*(\tilde{w}_k)|\tilde{p}_1, \dots, \tilde{p}_{k-1}, \tilde{w}_k, \dots, \tilde{w}_{y_d}) > 0.$$

Then there is a unique continuous $\pi(\cdot, e)$ *that solves Equation (3) for every* $e \in E$.

Proof. Note that under conditions (i) and (ii) we can invoke Lemma 7. Thus, equation (3) holds. Moreover, under condition (iii), (3) can be rewritten as

$$\int_{P(w)} F_{\boldsymbol{\pi}|\mathbf{p}}(\pi(p, e)|p) f_{\mathbf{p}|\mathbf{w}}(p|w) dp = e,$$

where $P(w)$ is the support of $\mathbf{p}|\mathbf{w} = w$. Fix some e. Let $v(p) = F_{\boldsymbol{\pi}|\mathbf{p}}(\pi(p, e)|p)$. Since $F_{\boldsymbol{\pi}|\mathbf{p}}(\cdot|p)$ is observed and invertible, if we pointidentify $v(\cdot)$, then we identify $\pi(\cdot, e)$ as well. Consider the following integral equation:

$$\int_{P(w)} v(p) f_{\mathbf{p}|\mathbf{w}}(p|w) dp = e, \tag{12}$$

where $P(w)$ and $f_{\mathbf{p}|\mathbf{w}}$ are given. Without loss of generality assume that the reordering from condition (vi) is $(1, 2, \dots, d_y)$. Next note that under condition (vi)

$$\begin{aligned} f_{p|w} &= f_{p_{d_y}|p_1,p_2,\dots,p_{d_y-1},w} f_{p_{d_y-1}|p_1,p_2,\dots,p_{d_y-2},w} \cdots f_{p_2|p_1,w} f_{p_1|w} = \\ &= f_{p_{d_y}|p_1,p_2,\dots,p_{d_y-1},w_{d_y}} f_{p_{d_y-1}|p_1,p_2,\dots,p_{d_y-2},w_{d_y-1},w_{d_y}} \cdots f_{p_2|p_1,w_2,\dots,w_{d_y}} f_{p_1|w}. \end{aligned} \tag{13}$$

Let

$$g_{d_y}(p_1, \dots, p_{d_y-1}, w_{d_y}) = \int v(p) f_{p_{d_y}|p_1,p_2,\dots,p_{d_y-1},w_{d_y}}(p_{d_y}|p_1, p_2, \dots, p_{d_y-1}, w_{d_y}) dp_{d_y}$$

Recursively for $k = 2, \dots, d_y - 1$ define

$$\begin{aligned} &g_k(p_1, \dots, p_{k-1}, w_k, \dots, w_{d_y}) = \\ &= \int g_{k+1}(p_1, \dots, p_k, w_{k+1}, \dots, w_{d_y}) f_{p_k|p_1,p_2,\dots,p_{k-1},w_k,\dots,w_{d_y}}(p_k|p_1, p_2, \dots, p_{k-1}, w_k, \dots, w_{d_y}) dp_k \end{aligned}$$

Next note that (13) implies that (12) can be rewritten as

$$\int g_2(p_1, w_2, \dots, w_{d_y}) f_{p_1|w}(p_1|w) dp_1 = e$$

Since all the assumptions of Theorem 2.1 in Hu et al. (2017) are satisfied, we identify

g_2 using variation in w_1. Next note that

$$\int g_3(p_1, p_2, w_3, \ldots, w_{d_y}) f_{p_2|p_1,w_2,\ldots,w_{d_y}}(p_2|p_1, w_2, \ldots, w_{d_y}) dp_2 = g_2(p_1, w_2, \ldots, w_{d_y})$$

Since again all the assumptions of Theorem 2.1 in Hu et al. (2017) are satisfied, we identify g_3 using variation in w_2. Repeating this argument finitely many times we identify g_{d_y}, and, hence, we identify $v(\cdot)$. ■