Prices, Profits, and Production: Identification and Counterfactuals*

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

First version: October 10, 2018
This version: January 21, 2019

Abstract This paper studies nonparametric identification and counterfactual bounds for heterogeneous firms that can be ranked in terms of productivity. We require observation of profits or other optimizing-values such as costs or revenues, and either prices or attributes that determine prices. We extend classical duality results for price-taking firms to a setup with rich heterogeneity, and with limited variation in prices. We characterize the identified set for production sets, and provide conditions that ensure point identification. We present a general computationally-feasible framework for sharp counterfactual bounds, such as bounds on quantities at a counterfactual price. We show that existing convergence results for quantile estimators may be directly converted to convergence results for production sets, which facilitates statistical inference.

JEL classification: C5, D24.

Keywords: Counterfactual bounds, cost minimization, nonseparable heterogeneity, partial identification, profit maximization, production set, revenue maximization, shape restrictions.

*We are grateful to Paul Grieco, Lance Lochner, Salvador Navarro, David Rivers, Susanne Schennach, and Al Slivinsky for useful comments and encouragement.

†Department of Economics, University of Western Ontario.

Introduction

This paper studies identification of production sets and counterfactual bounds for firms with potentially multiple outputs and inputs. We assume an analyst has data on the values of an optimization problem, such as profits, costs, or revenues, as well as prices.[1] Our framework allows rich forms of complementarity and substitutability between outputs and inputs as well as rich heterogeneity across firms, but maintains 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 values (such as profits or costs) and prices of flexibly-chosen factors.

The use of values 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 prices. The main contribution of this paper is to study recoverability of production sets and sharp counterfactual bounds, both in the presence of heterogeneity and in settings with potentially limited variation in prices.

In order to obtain identification of firm-specific production possibility sets, we restrict firm heterogeneity by assuming 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 weak monotonicity property in unobservable productivity.[2] We exploit this monotonicity assumption to recover the heterogeneous profit function from the joint distribution of prices and profits.

Once the heterogeneous profit function has been identified, we study identification of production sets. First, we provide a sharp identification result characterizing the envelope of all production possibility sets that can generate the data. This result applies regardless of the variation in prices, and is constructive. Next, we provide

[1] Profit is the total revenue minus the total cost of flexible inputs. Our framework allows for fixed inputs.

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

conditions under which the production possibility sets may be uniquely recovered from data. This result does not require observability of all possible values of the price vector. Instead, we require that all possible "directions" of the price vector be observed. This condition can be satisfied if either all prices are bounded from above, or all prices are bounded from below, but not both.

Our sharp bounds on production sets apply with finite variation in prices, and can be adapted for the dual purpose of providing sharp counterfactual bounds. We present a general framework for sharp bounds on counterfactuals. For example, we describe sharp lower and upper bounds on profits at a new counterfactual price, as well as sharp bounds on outputs and inputs at a new counterfactual price. When prices take finitely many values, the bounds in these examples are described by linear programming problems and are computationally tractable. We emphasize that these sharp bounds on outputs and inputs require data on profits (or other values) and prices, not quantities.[3]

We next turn to estimation, providing an equality relating estimation error of profit functions and estimation error of production possibility sets. This result allows 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 a new argument because prices are restricted to be positive.

In some empirical settings the analyst may not observe prices for all inputs or outputs.[4] We extend our analysis to allow observable attribute variables, which determine prices via an unknown, good-specific link function. For example, Combes et al. (2017) use the location of a house as an attribute that is linked to price, since price itself is unobservable. We show that when an analyst observes profits and attributes, it is possible to fully identify production sets. We establish this by using a novel identification technique exploiting homogeneity, which may be of independent interest.[5]

[3]Our analysis requires observation of the values of optimization problems, such as profits, costs, or revenues. Observation of quantities and prices implies observation of these values but is not required.

[4]This problem may lead to *omitted price bias* (Zellner et al. (1966), Epple et al. (2010)), which can result in failure to identify the profit function.

[5]We exploit Euler's homogeneous function theorem to generate a system of linear equations. A rank condition on certain coefficients of this system provides a sufficient condition for identification for the function linking prices and attributes.

Our simplest identification result for profit functions assumes that prices and productivity are independent. In an extension, we relax this assumption to allow endogeneity, applying the results of Chernozhukov & Hansen (2005). Allowing endogeneity also facilitates application of our identification techniques to cost minimization and revenue maximization. The main difference between these and the unconstrained profit maximization problem is that the firm fixes some variables such as output quantities in the cost minimization problem. Endogeneity may arise if these fixed variables are choice variables. Once the heterogeneous cost or revenue function is identified using instruments, our previous analysis, including counterfactual bounds, applies.

Our analysis uses duality theory and shape restrictions arising 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. Duality has also been used to motivate parametric estimators (e.g. Lau (1972), Diewert (1973), Christensen et al. (1973)). This literature focuses on a *representative agent* framework in which all prices are observed. In contrast, we allow rich nonseparable heterogeneity and focus on the important case of limited variation in prices.[6] There is little existing work concerning identification with limited (possibly finite) variation in prices. One such paper is Hanoch & Rothschild (1972), which focuses on finite deterministic datasets of individual firms' profits or costs, and prices. Hanoch & Rothschild (1972) does not study identification of the production set or the profit function, but focuses on providing necessary and sufficient conditions under which an observed production function is consistent with profit maximization or cost minimization.[7] Another paper studying limited price variation is Varian (1984), which works with quantities and prices and does not study unobservable heterogeneity.[8] While observation of prices and quantities implies observation of profits, the reverse is not true.

A recent literature on the identification and practical estimation of a firm's technology has focused on output and input quantities, sometimes not using prices at all (e.g. Griliches & Mairesse (1995), Olley & Pakes (1996), Levinsohn & Petrin (2003),

[6]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 (2018).

[7]Cherchye et al. (2016) studies the identification of profits and production sets with a finite deterministic dataset on prices and quantities.

[8]See also Cherchye et al. (2014) and Cherchye et al. (2018). Cherchye et al. (2018) differs from us because they assume observed input quantities in the context of cost minimization. Also, they do not study a cross section of firms.

and Ackerberg et al. (2015)). Our analysis of revenue maximization complements a production function approach, showing that when inputs, revenues, and output prices are observed, it is possible to learn about heterogeneous firms' production sets. This analysis applies even when there are multiple outputs. In contrast, a pure quantities approach that does not use output prices faces the challenge that 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 heterogeneous 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.[9] 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 prices and profits (or other values).

Complementing this recent work, our analysis further 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 some outputs and inputs is problematic. For instance, in the housing market the observability of output quantities is difficult because houses provide different services that are hard to measure. However, prices may be observed (Combes et al. (2017) and Albouy & Ehrlich (2018)). In the health industry, an analyst may find it difficult to measure inputs such as drugs since they vary widely in their physical characteristics. However, prices and total costs may be observable (Bilodeau et al. (2000)). In the banking industry, outputs such as business loans and consumers loans are difficult to measure because a loan is a financial service that entails many unobservable goods and services. However, the price of a loan is observed as well as profits in some settings (Berger et al. (1993)).

The rest of this paper proceeds as follows. In Section 1 we present a model of heterogeneous production in which firms are rankable in terms of productivity. Then we proceed with our main identification result for production possibility sets in Section 2. Section 3 provides a general framework to conduct sharp counterfactual

[9] See also Doraszelski & Jaumandreu (2013) for an application with a parametric structure, and Malikov (2017).

analysis in production environments. In Section 4 we propose a consistent estimator of the production correspondence. In Section 5 we extend our methodology to environments where one observes attributes that determine unobservable prices. Section 6 is concerned with potential correlation between price (or quantities) and productivity. Section 7 extends the previous results to a general class of constrained maximization problems such as cost minimization and revenue maximization. We conclude in Section 8. All proofs can be found in Appendix A.

1. Profit Maximization

This paper studies the question of recoverability of the technology of heterogeneous firms given data on the value function of their maximization problems, as well as data on prices or attributes that alter the maximization problems. The simplest example is profit maximization for a price-taking firm. Profits (revenue minus costs) are the value function of the problem, and prices are shifters that alter the maximization problem. Other examples include cost minimization given a fixed level of outputs and revenue maximization given a fixed level of inputs.

These latter examples involve additional constraints relative to a profit maximization problem, e.g. with cost minimization a firm is constrained to produce a given level of output. For notational simplicity and to obtain sharper results in some cases, the core of this paper focuses on the unconstrained profit maximization problem. In Section 7 we then describe how our analysis applies to important constrained problems such as cost minimization and revenue maximization.

Our analysis applies to heterogeneous firms that may produce multiple outputs. Because we allow multiple outputs, we work with production possibility sets rather than production functions.[10] 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.[11]

The random variable $\mathbf{e}$ is interpreted as a scalar unobservable productivity term.[12]

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

[11]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.

[12]We use boldface font (e.g. $\mathbf{p}$) to denote random objects and regular font (e.g. p) for deterministic ones. The supports of random vectors are usually denoted by capital letters; i.e. for the random

The set $Y(e)$ is the production possibility set for a firm with productivity level e. It 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 component indicates the firm is a net demander. The possible output/input vector is denoted y.[13]

We formalize our assumptions on production correspondences below.

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 $y_j^* \leq y_j$ for all $j \in \{1, \cdots, d_y\}$ is also in $Y(e)$;

(iii) $Y(e)$ satisfies *the recession cone property*: if $\{y^m\}$ is a sequence of points in $Y(e)$ satisfying $\|y^m\| \to \infty$ as $m \to \infty$, then accumulation points of the set $\{y^m / \|y^m\|\}_{m=1}^{\infty}$ 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 maximization problem having a solution, and rules out constant or increasing returns to scale.[14] 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.[15] 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 Section 7 we consider closely related problems that impose constraints on the feasible quantities. For example, with cost minimization we may require that a given level of output be produced.

vector $\mathbf{p}$, the support is denoted P, and is the smallest closed set such that $\mathbb{P}(\mathbf{p} \in P) = 1$, where $\mathbb{P}(A)$ denotes the probability of an event A.

[13]The transpose of a vector is denoted y' and its Euclidean norm is denoted $\|y\|$.

[14]See Kreps (2012), p. 199 for more details. Our approach can still be applied to the case of constant returns to scale because we can always transform this technology to a decreasing returns one by normalizing the quantity of one output or input to one.

[15]We implicitly assume that all components of $\mathbf{p}$ are strictly positive with probability one. Since formally P is a closed set we will abuse notation and associate P with $P \cap \mathbb{R}_{++}^{d_y}$.

In order to have a structural interpretation for unobservable productivity captured by **e**, we impose that firms can be ranked according to productivity. We formalize this as follows.

Assumption 1. *If $e \leq \tilde{e}$, 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 $e \leq \tilde{e}$ is the usual order. One can think of e as an unobservable one-dimensional input (e.g. managerial quality) that is fixed. Thus, one may interpret this setup as studying otherwise homogeneous firms that are different only in one unobservable input.

1.1. Production Possibility Sets and Profit Functions

In this section we recall classical duality relationships between production sets and profit functions that will be used in our identification analysis. These results show how the profit function can be used to recover production possibility 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, p^*, $\pi(\alpha p + (1 - \alpha)p^*, e) \geq \alpha\pi(p, e) + (1 - \alpha)\pi(p^*, e)$. 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 (Kreps (2012), Proposition 9.14).

When one assumes a firm maximizes profits taking prices as given, then convexity and homogeneity emerge as shape restrictions on the firm problem that can be used for counterfactual bounds. 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 profit function in the following sections.

In our environment, the profit function provides a complete characterization of the production set for a given realization of **e**.

Lemma 1 (E.g. Kreps (2012), Corollary 9.18)**.** *For all $e \in E$, the realized production set is described by*

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

The result shows that if we can recover the profit function for all prices, then we can fully recover the production set. In Section 2.1 we generalize this result by providing a sharp characterization of the production set when observability of prices is limited. Section 3 also provides sharp bounds for profits and production at prices outside the support of the data.

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

Lemma 2. *With the maintained assumption that $Y(\cdot)$ is a production correspondence, the following are equivalent:*

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

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

Thus, Assumption 1 (condition (i) in this lemma) is equivalent to monotonicity of profits. Lemma 2 allows us to translate an economically relevant assumption on the primitives (the production correspondence) into a restriction on the observable quantities (profits). The following example illustrates production monotonicity 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 while k and l are positive, 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 possibility 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 function, then $Y(\cdot)$ satisfies production monotonicity. For example, $A(e) = \exp(e)$ with $E = [0, M]$, $M > 0$, is weakly increasing. The function $A(e) = \mathbb{1}\left(0 \leq e \leq 1/2\right) + 2\mathbb{1}\left(1/2 < e \leq 1\right)$ with $E = [0,1]$ is also weakly increasing, yet has only two distinct types of firms (determined by whether $e > 1/2$).[16] These two choices of A both imply production monotonicity, and so this example illustrates how we may treat discrete and continuous types in a common framework.

[16] We denote the indicator function by $\mathbb{1}\left(\cdot\right)$. $\mathbb{1}\left(A\right)$ is equal to 1 when the statement A is true and 0 otherwise.

2. Identification of the Production Correspondence

We now present our core identification results for the production correspondence. 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. 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 $\boldsymbol{\pi}$ conditional on $\mathbf{p} = p$.

For concreteness, in order to identify the profit function we assume we observe a cross section of firms that operate in different markets. Price may vary across markets due to different market characteristics or endowments (Brown & Matzkin (1996)). Market endowments can be understood as the market characteristics that determine the initial distribution of outputs and inputs in each market before production and consumption take place.

In this section, in order to recover the profit function we impose the assumption that prices and unobservable heterogeneity are independent. In Section 6 we relax this assumption. We note that starting in Section 2.1, our analysis applies provided one has (somehow) identified the profit function $\pi(\cdot)$. Thus, if one takes the profit function as a primitive, our analysis of identification and counterfactual bounds still applies.

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

It is helpful to relate this independence condition with concern over *transmission bias*, which is a known problem in analysis of identification of production functions from inputs and outputs. This bias arises due to the endogeneity of some outputs/inputs that are determined partly by the productivity term (Marschak & Andrews (1944)). For analysis of identification of the profit function using profits and prices, transmission bias may be irrelevant because identification does not condition on choice variables (such as inputs in a production function setting).

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,*

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

for all $p \in P$ and $e \in E$.

We present this 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 $\pi(p,\cdot)$ is continuous or is strictly increasing in e.

Because the primitive of the paper is the production correspondence, we note that a version of Theorem 1 applies with assumptions placed directly on the profit function. Most notable, we may make use of the fact that production monotonicity is equivalent to profit monotonicity (Lemma 2). Moreover, lower hemicontinuity of the production correspondence is sufficient for lower semicontinuity of the profit function (Aliprantis & Border (2006), Lemma 17.29).[17] 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 $p \in P$.*

We note that Theorem 1 makes use of a shape restriction for each structural function $\pi(p,\cdot)$. The result does not make use of any shape restriction as p varies; i.e. this result does not formally require that $\pi(\cdot,e)$ be the profit function for a firm. Given properties of the profit function, a testable implication is that for every e, the function

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

must be convex and homogeneous of degree 1. We make use of these shape restrictions when we present counterfactual bounds in Section 3.

[17] $Y(\cdot)$ is lower hemicontinuous, if whenever $e^m \to e$ as $m \to \infty$, and $y \in Y(e)$, there is a sequence $y^m \in Y(e^m)$ such that $y^m \to y$ as $m \to \infty$.

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 1 we immediately 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 can lead 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 therefore continuous.[18] 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

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

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. Another example is the Cartesian product of all integers, $P = \{1, 2, \dots\}^{d_y}$.

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

[18] 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 Appendix B.

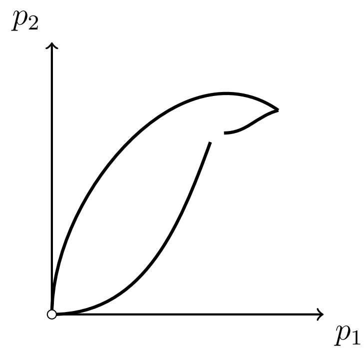

Figure 1 – The set P (depicted by black curve) satisfies Assumption 3 and has an empty interior. Dots represent "holes" in the support. Thus, P is not a connected set.

support. Assumption 3 is equivalent to

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

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 $p/\left\|p\right\|$. In particular, in two dimensions this condition 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 $\mathbf{e}$. This choice of normalization allows us to interpret e as the ranking of productivity. See Matzkin (2003) for a discussion of alternative normalizations in other settings described by one-dimensional unobservable heterogeneity.

Theorem 2. *Let Assumption 4 and the assumptions of Theorem 1 hold. Moreover, let* $\tilde{Y}(\cdot)$ *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.*

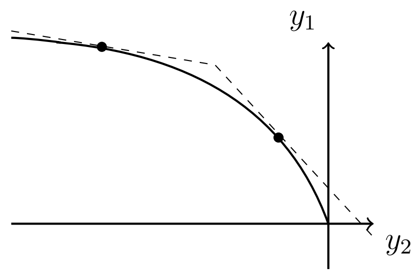

Figure 2 – $\tilde{Y}(e)$ and $Y'(e)$ for $d_y = 2$ and $P = \{p^*, p^{**}\}$. $\tilde{Y}(e)$ is the area under the dashed lines. $Y'(e)$ is the area under the solid curve. Dashed lines correspond to two hyperplanes $p_1^* y_1 + p_2^* y_2 = \pi(p^*, e)$ and $p_1^{**} y_1 + p_2^{**} y_2 = \pi(p^{**},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

(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 $e \in E$.

(iii) If Assumption 3 holds, then $\tilde{Y}(\cdot)$ 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 the production correspondence under our assumptions. These results are related to Varian (1984), Theorem 15.[19] However, Varian (1984) works only with finite datasets, which are comparable to having a finite support of prices in our setting. In addition, Varian (1984) observes prices and quantities while we observe prices and profits. Recall that observing prices and quantities implies observation of profits. Finally, Varian (1984) does not consider unobservable heterogeneity.

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 2-dimensional output/input set, 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 $\{y \in \mathbb{R}^2 : y_1 + y_2 \leq 0\}$.

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

This set induces infinite profits for a price-taking firm whenever $p_1 \neq p_2$. 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.[20]

Theorem 2(iii) is related to Lemma 1, which is the textbook version of recovering production sets from the profit function. In this paper, however, we begin with the distribution of profits and prices. Part (iii) shows that with this distribution, it is possible to identify the distribution of features of $Y(\cdot)$, such as the distribution of possible profit-maximizing quantities. We emphasize that this is true even if quantities are unobservable. An additional manner in which (iii) differs from textbook analysis is that, in econometric settings, it is not always natural to assume that all prices are observed ($P = \mathbb{R}^{d_y}_{++}$). Theorem 2 clarifies the variation in prices sufficient for nonparametric identification of production sets. We note that while Assumption 3 is sufficient for point identification of Y, it is not necessary as illustrated in Appendix B.

The full strength of Assumption 3 may be relaxed if one is only interested in identification of some economically relevant region of the production possibilities frontier. In such cases, it suffices to observe only those prices that are tangential to that region of interest as the following example demonstrates.

Example 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 $0 < \underline{y}_1 \leq \overline{y}_1 < \infty$. Suppose in addition that the unknown production set for some $e \in E$ is given by

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

That is, the production possibilities frontier is $\{y \in \mathbb{R} \times \mathbb{R}_- \ : \ y_1 = \sqrt{-y_2}\}$. Then Theorem 2 implies that it suffices to observe prices only in the set $\{p \in \mathbb{R}^2_{++} \ : \ 2\underline{y}_1 \leq p_1/p_2 \leq 2\overline{y}_1\}$. Note that in this example, even if $Y(\cdot)$ is unknown, it is possible to check whether price variation is rich enough to identify the relevant part of the frontier.

Remark 1. Our identification analysis does not impose any *a priori* restrictions that certain dimensions of $Y(e)$ correspond to inputs, i.e. weakly negative numbers. This additional restriction can be imposed by modifying the set constructed in Theorem 2. Specifically, the set $\tilde{Y}(e)$ constructed in this theorem may be intersected with an appropriate half-space that encodes that certain dimensions (corresponding to inputs) must be nonpositive. We note that an analogous restriction for outputs is not informative because of the assumption of free disposal.

[20]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.

2.2. Supply Function

Building on this identification analysis, we now provide a formula establishing constructive identification of optimal quantities of outputs/inputs from the distribution of profits and prices. To formalize this we introduce the supply function, which exists whenever the profit maximization problem has a unique solution.

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

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

To connect the supply function with the profit function, 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.

Proposition 1. *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*

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

When the assumptions of Theorem 1 hold, we may state this result directly in terms of the distribution of profits and prices. Specifically,

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

One implication of this is that the conditional distribution of quantities given prices can be identified from the conditional distribution of profits given prices.

Note that this formula may be used with continuous prices, since then we may take a derivative of the profit function. This formula may also be used with discrete prices as long as they are sufficiently rich. To see this, consider $P = \{1, 2, \ldots\}^{d_y}$. Using homogeneity, as argued previously one can identify $\pi(\cdot, e)$ over a dense set, and thus it is possible to identify derivatives.

If quantities are observed in addition to prices and profits, Equation 1 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. Sharp Counterfactual Bounds

Theorem 2 makes use of a shape restriction to characterize the identified set of the production correspondence for profit-maximizing, price-taking firms. This shape restriction may be used for a dual purpose of providing sharp counterfactual bounds. In this section we provide several such bounds including bounds on profits or quantities for new prices outside of the support of the data.

Since homogeneity and convexity of the heterogeneous profit function allow us to identify it over $\mathrm{cl}\left(\bigcup_{\lambda>0}\{\lambda p \,:\, p \in P\}\right)$, we can associate the support P with the set where $\pi(\cdot, e)$ is identified. That is why, for notational simplicity and in this section only, we assume that P is a closed subset of the unit sphere $\mathbb{S}^{d_y-1}$ and we consider counterfactual prices with norm normalized to 1.

We first present a result characterizing quantities consistent with profit maximization. Theorem 2(ii) is the basis for the following proposition.

Proposition 2. *Let P be a finite subset of the unit sphere $\mathbb{S}^{d_y-1}$. Given P and $\{\pi(p,\cdot)\}_{p\in P}$, the set of output/input functions $\{y_p(\cdot)\}_{p\in P}$ can generate $\{\pi(p,\cdot)\}_{p\in P}$ if and only if*

$$
\begin{aligned}
&p' y_p(e) = \pi(p, e)\,, \quad \forall p \in P, e \in E\,,\\
&p^{*\prime} y_{p^*}(e) \geq p^{*\prime} y_p(e)\,, \quad \forall p, p^* \in P, e \in E\,.
\end{aligned}
$$

The vector $y_p(e)$ is interpreted as a candidate supply vector given price p and productivity e; it need not be unique and thus may not be equivalent to the supply function. Recall that as discussed in Remark 1, we do not impose *a priori* restrictions that certain components of $Y(e)$ are inputs; this would correspond to imposing additional sign restrictions on the functions $y_p(\cdot)$ described in the proposition.

Proposition 2 essentially states that for each e there must exist output/input vectors such that the weak axiom of profit maximization holds (Varian (1984)). We note, however, that the primitive observables of our paper are the *distribution* of profits and prices. In particular, since Theorem 1 provides a formula for $\pi(p, e)$, this result may be equivalently stated in terms of the joint distribution of profits and prices.

The equality restrictions in Proposition 2 state that the hypothesized output/input vectors should equal the given maximal profits. The inequality restrictions require that there are no strictly more profitable output/input vectors. See Figure 3 for an illustration.

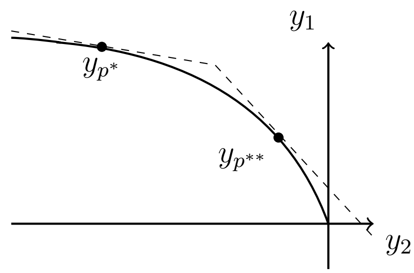

Figure 3 – $P = \{p^*, p^{**}\}$. $\tilde{Y}(e)$ is the area under the dashed lines. $Y'(e)$ is the area under the solid curve. Dashed lines correspond to two hyperplanes $p^{*\prime}y = \pi(p^*, e)$ and $p^{**\prime}y = \pi(p^{**}, e)$. y_{p^*} and $y_{p^{**}}$ can generate $\pi(p, e)$, $p \in P$.

Proposition 2 provides a full characterization of the output/input vectors that are consistent with a given set of prices P and corresponding set of profits $\{\pi(p, e)\}_{p \in P}$; recall that combined with Theorem 1 it may be equivalently stated in terms of the joint distribution of profits and prices. Varian (1982, 1984) has exploited the close connections between empirical content, recoverability of structural functions, and counterfactuals.[21] In our setting, analysis of sharp identification of production sets or output/input vectors consistent with the data facilitates sharp bounds on counterfactual analysis. To illustrate this, suppose we want to check whether a new counterfactual combination of prices and quantities $(p^{\mathrm{c}}, y_{p^{\mathrm{c}}})$ is consistent with profit-maximizing behavior. It is necessary and sufficient to check whether the constraints from Proposition 2 are satisfied with P replaced by $P \cup \{p^{\mathrm{c}}\}$.[22] Thus for a given p^{c} we can find the set of all $y_{p^{\mathrm{c}}}$ that fulfills the constraints from Proposition 2 for the counterfactual-augmented set $P \cup \{p^{\mathrm{c}}\}$. This method provides a sharp characterization of counterfactual quantities consistent with the model.

Building on the full characterization of the identified set of the production correspondence, we can construct sharp bounds for any function of counterfactual prices and quantities, potentially with additional restrictions. The upper bound on a functional C given a restriction r and heterogeneity level e is given by

$$\begin{aligned} \overline{C}_r(e) = & \sup_{p^{\mathrm{c}}, y_{p^{\mathrm{c}}}, \{y_p\}_{p \in P}} C(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) \,, \\ & \text{s.t. } r(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) = 0 \,, \\ & \qquad p' y_p = \pi(p, e) \,, \quad \forall p \in P \,, \end{aligned}$$

[21]Recent work in demand analysis building on these connections includes Blundell et al. (2003), Blundell et al. (2017), Allen & Rehbeck (2018), and Aguiar & Kashaev (2018).

[22]Take $\pi(p^{\mathrm{c}}, e)$ to be $p^{\mathrm{c}\prime} y_{p^{\mathrm{c}}}$.

$$p^{*\prime} y_{p^*} \geq p^{*\prime} y_p \,, \qquad \forall p, p^* \in P \cup \{p^{\mathrm{c}}\} \,.$$

The lower bound is given by

$$\begin{aligned} \underline{C}_r(e) = & \inf_{p^{\mathrm{c}}, y_{p^{\mathrm{c}}}, \{y_p\}_{p \in P}} C(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) \,, \\ \text{s.t. } & r(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) = 0 \,, \\ & p' y_p = \pi(p, e) \,, \quad \forall p \in P \,, \\ & p^{*\prime} y_{p^*} \geq p^{*\prime} y_p \,, \qquad \forall p, p^* \in P \cup \{p^{\mathrm{c}}\} \,. \end{aligned}$$

We provide some examples covered by this general setup. Note that these bounds hold for each e, and thus one may also bound the distribution of $\overline{C}_r(\mathbf{e})$ and $\underline{C}_r(\mathbf{e})$. We reiterate that these upper and lower bounds apply to prices on the unit sphere, though they may be adapted for prices off the unit sphere as illustrated in the following examples.

Example 3 (Profit bounds for a counterfactual price)**.** Suppose that we are interested in upper and lower bounds for profits at a given counterfactual price $\overline{p}^{\mathrm{c}}$. When prices p^{c} are on the unit sphere, we may specify $C(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) = p^{\mathrm{c}\prime} y_{p^{\mathrm{c}}}$ and $r(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) = p^{\mathrm{c}} - \overline{p}^{\mathrm{c}}$. Then the problem can be simplified to get

$$\begin{aligned} \overline{C}_r(e) &= \sup_{y \in \tilde{Y}(e)} \overline{p}^{\mathrm{c}\prime} y \,, \\ \underline{C}_r(e) &= \max_{p \in P} \inf_{y \in \tilde{Y}(e) \,:\, p' y = \pi(p,e)} \overline{p}^{\mathrm{c}\prime} y \,, \end{aligned}$$

where $\tilde{Y}(e)$ is the envelope of all production possibility sets consistent with the data defined in Theorem 2. The above bounds are sharp in the following sense: if $\overline{C}_r(e)$ is finite, then it is feasible, i.e. there exists a production set that can generate $\overline{C}_r(e)$. If $\overline{C}_r(e)$ is not finite, then for any finite level K there exists a production set that can generate $C(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) > K$. Analogous statements hold for the lower bounds $\underline{C}_r(e)$.

Recall that we assume the support of prices P is a subset of the unit sphere. This may be imposed in empirical settings by replacing prices with normalized prices $p / \|p\|$. For counterfactual questions involving a price off the unit sphere $\overline{p}^{\mathrm{c}}$, one can bound counterfactual profits at price $\overline{p}^{\mathrm{c}} / \|\overline{p}^{\mathrm{c}}\|$ and then multiply the upper and lower bounds by $\|\overline{p}^{\mathrm{c}}\|$.

Example 4 (Quantity bounds for a counterfactual price)**.** Suppose that we are interested in the upper and lower bounds for $u' y_{p^{\mathrm{c}}}$ for a given counterfactual price $\overline{p}^{\mathrm{c}}$, where u is a unit vector. For example, with $u = (1, 0, \ldots, 0)'$ we are interested in bounds on the first component of y. Then $C(p^c, y_{p^c}) = u' y_{p^c}$ and $r(p^c, y_{p^c}) = p^c - \overline{p}^c$.

Example 5 (Profit bounds for a counterfactual quantity)**.** Suppose a regulator is considering imposing a new regulation that the first component of the output/input vector is fixed at $\overline{y}_1^c$. For example, in analysis of health care (Bilodeau et al. (2000)) a hospital may be required to treat a certain number of patients. To bound profits we may write the objective function as $C(p^c, y_{p^c}) = p^{c\prime} y_{p^c}$. The constraint is given by $r(p^c, y_{p^c}) = y_{1,p^c} - \overline{y}_1^c$.[23] Bounds on profits with this quantity may be useful for a regulator wondering whether a hospital of type e would be profitable with the hypothetical regulation. If the upper bound on profits is negative, the answer is definitively no. If the lower bound on profits is positive, the answer is definitively yes.[24] An additional question a regulator might ask is which types of firms could still be profitable. This can be addressed by studying functions $\overline{C}_r(\cdot)$ and $\underline{C}_r(\cdot)$ as e varies. Note that the constraints r are general, and inequality constraints may be incorporated as well by using indicator functions.

Example 6 (Output bounds for a counterfactual profit level)**.** Suppose that we are interested in the upper and lower bounds for the first component of the output/input vector given a fixed level of profits $\overline{\pi}^c$. Then $C(p^c, y_{p^c}) = (1, 0, \ldots, 0)' y_{p^c}$ and $r(p^c, y_{p^c}) = p^{c\prime} y_{p^c} - \overline{\pi}^c$.

Since P is finite, computing bounds in the first two examples is straightforward since they are the values of linear programs. In the last two examples the problem is quadratic since some constraints are quadratic (e.g. $r(p^c, y_{p^c}) = p^{c\prime} y_{p^c} - \overline{\pi}^c = 0$).

4. 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 possibility set for a firm with productivity level e. The main result in this section relates the estimation error

[23] Note that the problem may not have a solution since the set of parameters that satisfy restrictions may be empty.

[24] This maintains the assumptions of price-taking, profit-maximizing behavior with a technology that is described by a production correspondence.

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 $P = \mathbb{R}^{d_y}_{++}$ (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 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 7. Suppose that $d_y = 2$ and for some $e \in E$,

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

Note that although $\lim_{m \to \infty} (1 - 1/m)\sqrt{-y_2} = \sqrt{-y_2}$ for every finite $y_2 \leq 0$, the Hausdorff distance between these sets is equal to $\sup_{y_2 \in \mathbb{R}_-} \sqrt{-y_2}/m = \infty$ for every finite $m \in \mathbb{N}$.

Example 7 illustrates a technical concern with the Hausdorff distance that arises because of the unboundedness of production possibility sets. However, in empirical applications one may be interested in production possibility sets in regions that correspond to prices that are bounded away from zero. Thus, instead of working with all possible prices we will work only with certain empirically relevant compact convex subsets of $\mathbb{R}^{d_y}_{++}$. We consider the Hausdorff distance between extensions such as

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

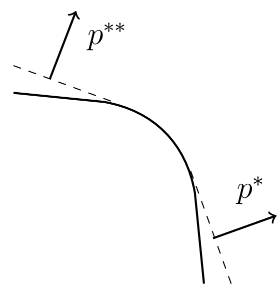

Figure 4 – $Y(e)$ and $Y_{\bar{P}}(e)$ for $d_y = 2$ and $\bar{P} = \{p \in P \,:\, \delta \leq p_2/p_1 \leq 1/\delta,\, \|p\| \leq 1\}$, $0 < \delta < 1$. $Y(e)$ is the area under the solid curve. $Y_{\bar{P}}(e)$ is the area under the dashed lines. Dashed lines correspond to two hyperplanes $p^{*\prime}y = \pi(p^*, e)$ and $p^{**\prime}y = \pi(p^{**}, e)$. They are tangential to the solid curve. p^* is such that $p_2^*/p_1^* = \delta$ and p^{**} is such that $p_2^{**}/p_1^{**} = 1/\delta$.

where $\bar{P} \subseteq P$ is convex and compact. These sets nest the original sets (e.g. $Y(e) \subseteq Y_{\bar{P}}(e)$) because the inequalities hold only for $p \in \bar{P}$, not for every $p \in P$. Moreover, the parts of the production possibility frontiers of the sets $Y(e)$ and $Y_{\bar{P}}(e)$ coincide at points that are tangential to price vectors from $\bar{P}$ (see Figure 4).

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

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

To state the following result, let $\bar{\mathcal{P}}$ be a collection of all compact, convex, and nonempty subsets of P.

Theorem 3. *Maintain the assumption that $\pi(\cdot, e)$ is homogeneous of degree 1 and convex.*[25] *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 nontrivial extension of a well-known relation between the Hausdorff distance and the support functions of convex *compact* sets to convex, closed, and

[25] Recall that this is equivalent to price-taking, profit-maximizing behavior with technology described by a production correspondence.

unbounded sets.[26] Homogeneity of an estimator can be imposed by rescaling the data by dividing by one of the prices. Unfortunately, convexity can be 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 3. *Maintain the assumption that $\pi(\cdot, e)$ is homogeneous and convex. 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,

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

Convexity of an estimator is difficult to impose in general, in which case Proposition 3 is relevant. It is computationally feasible to impose convexity for certain functional forms of π, which allows one to invoke the stronger Theorem 3. 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

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

where $b_{s,j}(\cdot) = b_{j,s}(\cdot)$ for all s, j. The original class of Diewert (1973) considers a deterministic model or representative agent model, in which each $b_{s,j}(\cdot)$ is constant. We allow unobservable heterogeneity by allowing $b_{s,j}(\cdot)$ to be a function of e. This functional form exhibits several desirable properties: (i) it is linear in the coefficients $b_{s,j}(e)$; (ii) monotonicity of $\pi(p, \cdot)$ can be imposed by assuming that each $b_{s,j}(\cdot)$ is weakly increasing;[27] (iii) convexity can be also imposed using linear inequalities on the

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

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

coefficients;[28] (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 s is described by the formula

$$y_s(p, e) = \sum_{j=1}^{d_y} b_{s,j}(e)(p_j/p_s)^{1/2} .$$

Thus, if quantities are observed in addition to prices and profits, then this equation provides overidentifying information.

5. Unobservable Prices and Attributes

In many empirical applications not all prices are observed. This may cause concern about *omitted price bias* (Zellner et al. (1966), Epple et al. (2010)), which implies a failure of identification of the profit function (or more generally, the value function). This section considers a solution to the omitted price bias that applies when the researcher has access to some observable attributes that are informative about unobservable prices. For example, the rental rate of capital may be linked to market-specific attributes such as short-term and long-term interest rates. Wages may be linked to the unemployment level. De Loecker et al. (2016) uses output price, market shares, product dummies, firm location, and export status as attributes for unobservable input prices. In the housing market, an analyst may use location as a price attribute for a house as in Combes et al. (2017). This section provides identification results if unobservable prices are unknown functions of these attributes.[29] Our technique makes use of the fact that the profit function is homogeneous in prices. We show that we can use Euler's homogeneous function theorem to identify the unknown link functions. This technique may be of independent interest since homogeneity is a common shape restriction.

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

$$p_j = g_j(x_j)$$

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

[29] Hedonic pricing models also exhibit similar structure. However, in that literature it is assumed that both prices and attributes are observed. See, for instance, Ekeland et al. (2004).

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 attribute to simplify the notation. We can also allow for existence of additional attributes that enter every g_j. In this case the analysis below proceeds if we condition on a fixed value of those common attributes.[30]

Note that we are assuming that prices are not a function of e. In our setup prices vary across markets but are constant within a given market. Price-taking behavior together with the assumption that the distribution of productivity is the same across markets imply that prices cannot be a function of e. In fact, prices are determined by market clearing conditions making $g_j(\cdot)$ a function only of market characteristics (x). For an illustration of this statement see Example 10.[31]

We denote $x = (x_j)_{j=1,\dots,d_y} \in X$ and $g(x) = (g_j(x_j))_{j=1,\dots,d_y}$. 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. What remains is to identify g.

We present an informal outline how to identify g 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

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

Replacing prices with price attributes, we obtain

$$\sum_{j=1}^{d_y} \partial_{p_j} \pi(g(x), e) g_j(x_j) = \pi(g(x), e)\,. \tag{2}$$

Define $\tilde{\pi}(x, e) = \pi(g(x), e)$. We thus have

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

Plugging this in to (2) we obtain

$$\sum_{j=1}^{d_y} \partial_{x_j} \tilde{\pi}(x, e) \frac{g_j(x_j)}{\partial_{x_j} g_j(x_j)} = \tilde{\pi}(x, e)\,. \tag{3}$$

[30]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.

[31]We can extend our analysis to allow for market-specific unobservable attributes. However, such an extension would require availability of richer datasets (e.g., grouped or panel datasets).

Assume for now that $\tilde{\pi}(\cdot, e)$ is identified (we establish formal identification in Lemma 3). Thus the only unknowns involve g. By varying x, holding everyone else fixed, Equation 3 can be used to generate a system of equations. We show that when a certain rank condition is satisfied, it is possible to identify the entire function g using an appropriate scale/location normalization. We note that if all prices are observed except, say, $j = 1$, then we may directly apply Equation 3 to learn about g_j.

We now formalize this intuition. The remaining steps are to show that the function $\tilde{\pi}$ can be identified, state our location/scale normalization, and the rank condition that can be applied to the system of equations generated from (3).

First, we establish identification of $\tilde{\pi}(\cdot, \cdot) = \pi(g(\cdot), \cdot)$. We impose an independence restriction that implies Assumption 2, and is implied by Assumption 2 if g is invertible. In Section 6 we discuss how to relax this independence restriction.

Assumption 5. *The unobservable shocks* $\mathbf{e}$ *are independent from attributes* $\mathbf{x}$*. 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 3. *Suppose that Assumptions 4 and 5 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 from* $F_{\pi|\mathbf{x}}$*. In particular, for every* $x \in X$ *and* $e \in E$*,*

$$\tilde{\pi}(x, e) = \inf \left\{ \underline{\pi} \; : \; e \leq F_{\boldsymbol{\pi}|\mathbf{x}}(\underline{\pi}|x) \right\}.$$

Next we set location/scale normalizations and some regularity conditions on g.

Assumption 6. *(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 6(i)-(ii) allow us to identify the scale and the location, respectively, of the multivariate function g. Since we can always relabel both outputs and inputs, Assumption 6(i) is equivalent to assuming that at least one price (not necessary p_{d_y}) is observed.

We now turn to our rank condition. This condition ensures that the system of equations generated from (3) has sufficient variation to recover terms such as $g_j(x_j)/\partial_{x_j} g_j(x_j)$.

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

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

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 attribute 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 3 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) $\mathbb{E}[\boldsymbol{\pi}|\mathbf{x} = \cdot]$ 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 $\mathbb{E}[\boldsymbol{\pi}|\mathbf{x}=\cdot]$. We note that because homogeneity is a shape restriction that is preserved under expectations, identification of g from the conditional mean does not require the assumption that firms can be ranked in terms of productivity.[32] Thus, the technique in this section may be applied to representative agent analysis as well, as formalized in part (ii) of Theorem 4.

To further interpret the rank condition, we study it in parametric examples. We show that the rank condition can be satisfied for the Diewert (1973) profit function presented in Section 4, but can fail for every possible parameter value with Cobb-Douglas technology.

Example 8 (Diewert function, $d_y = 3$)**.** Let

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

Suppose that p_3 is observed, and $p_1 = g_1(x_1)$ and $p_2 = g_2(x_2)$. Assume, moreover, that $\partial_{x_s} g_s(x_s) \neq 0$, for all x_s and $s = 1, 2$. 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 8 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 9 (Cobb-Douglas)**.** For a fixed e, let $y_o \leq k^\alpha l^\beta$ be such that $\alpha+\beta<1$ and $\alpha, \beta > 0$. Then

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

[32] In Appendix C we present an alternative methodology to identify g. While this section uses homogeneity of $\pi(\cdot, e)$ to identify g, that methodology uses convexity to identify g. This methodology also does not require scalar heterogeneity, but requires observing quantities in place of profits.

where $p = (p_o, p_k, p_l)'$. Suppose that only p_o is perfectly observed. Suppose $p_k = g_k(x_k)$ 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 quantity not described in the ratio. Indeed, recall that

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

Thus, if $y_j(g(x), e)/y_s(g(x), e)$ does not depend on observed price p_{d_y}, then the s-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.

6. 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). These results will be applied as well in Section 7 when we consider certain constrained profit maximization problems. The reason endogeneity is a central concern in such problems is that constraints may be endogenous. For example, in the cost minimization problem, the output needed may be a choice variable for the firm. We note that endogeneity is not always a concern. For instance, output quantities may be determined by a regulator (Nerlove (1963)).

In addition to analysis of constrained problems, endogeneity is also a potential concern with the unconstrained profit maximization problem. Recall our benchmark model, with profits, 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 outputs and inputs 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 unobservable productivity. These instruments have to affect prices but must not be related to productivity.[33]

The following example illustrates how prices vary in a cross section of markets, and provides an instance in which price are independent of productivity. At the same time, it also shows how analysis of cost minimization may suffer from endogeneity even when profit maximization does not.

Example 10. Consider a collection of competitive markets. Each market is characterized by a mass of consumers $\eta > 1$. Preferences over a consumption good y_o and a numeraire m are given by $\mathbf{u}(y_o, m) = 2\boldsymbol{\alpha}^{1/2} y_o^{1/2} + m$, where $\boldsymbol{\alpha}$ is uniformly distributed on $[0, 1]$. There is a unit mass of firms in each market with cost function $c(y_o, p_i, e) = -y_o^2 p_i / 2e$,[34] where p_i is the price of labor. The productivity term $\mathbf{e}$ is uniformly distributed on $[0, 1]$. Assume that (i) p_i is exogenously determined (e.g., minimum wage), and (ii) the endowment of the consumption good in each economy is zero. Then the market clearing condition is

$$\int_0^1 y_0((p_o, p_i)', e) de = \eta \int_0^1 x(p_o, \alpha) d\alpha \,,$$

where the individual supply is $y_o((p_o, p_i)', e) = e p_o / p_i$, and the individual demand is $x(p_o, \alpha) = \alpha p_o^{-2}$.[35] The unique equilibrium of each market satisfies

$$p_o = (\eta p_i)^{1/3} \,.$$

In this example, objects of interest are the cost function $c(y_o, p_i, e) = -y_o^2 p_i / 2e$, and the profit function $\pi(p, e) = e p_o^2 / (2 p_i)$. Recoverability of production sets from these functions is possible from our previous results. Note that in this example the equilibrium prices in each market $p = (p_o, p_i)'$ do not depend on productivity of firms. Thus we may use Theorem 1 to identify the profit function. However, the quantity

[33] We note endogenous market entry/exit is less of a concern as a source of endogeneity in a static setting.

[34] The cost function in our setup is negative because inputs are negative quantities. We formally define the cost function and provide additional details in Section 7.

[35] The average demand is multiplied by the mass of consumers in the given market to obtain the aggregate demand.

produced by firms, $y_o = y_o(p, e)$, is a function of productivity (i.e. endogenous because of transmission bias). Hence, addressing endogeneity is important to identify the cost function. Output prices may be used as an instrument for y_o following the methodology we will outline shortly.

We note that endogeneity in prices is different from omitted price bias. Even when prices are exogenous, a price might be unobservable. In this case, one can follow an approach similar to Section 5 to address omitted price bias. Specifically, the vector of attributes is $x = (\eta, p_i)$ and the link function is $p_o = g_o(\eta, p_i) = (\eta p_i)^{1/3}$. Note that p_o is not a function of e.

We now return to our general setup. 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 $(\boldsymbol{\pi}, \mathbf{p}', \mathbf{w}')'$, 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 unobservable heterogeneity $\mathbf{e}$.

Assumption 7. *$F_{\mathbf{e}|\mathbf{w}}(\cdot|w) = F_{\mathbf{e}}(\cdot)$ for all $w \in W$.*

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

$$\mathbb{P}\left(\boldsymbol{\pi} \leq \pi(\mathbf{p}, e) | \mathbf{w} = w\right) = e \tag{4}$$

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

This lemma says that in the presence of endogeneity, we can still rank firms conditional on the instrumental variable. While our previous analysis uses weak monotonicity of $\pi(p, \cdot)$ in e, we now impose strict monotonicity. Note that Equation 4 is an integral equation that connects the unknown profit function, the distribution of observables, and productivity e. Indeed, Equation 4 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 $\mathbf{p}$ conditional on $\mathbf{w} = w$ and we assume the conditional p.d.f. of $\mathbf{p}$ conditional $\mathbf{w} = w$ exists for all w. 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.[36] In Appendix D we establish identification of $\pi(\cdot)$ based on the results of Chernozhukov & Hansen (2005).

Note that if the heterogeneous 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. In addition, the counterfactual bounds of Section 3 can be applied. Finally, one can apply the same argument to endogenous price attributes in order to identify the composite profit function $\tilde{\pi}$ (Lemma 3) and then obtain the results from Section 5 without imposing Assumption 5.

7. Constrained Profit Maximization, Cost Minimization, and Revenue Maximization

In the preceding sections we have studied identification of the production correspondence given profits and prices or attributes. We now turn to constrained problems such as cost minimization; we call this a constrained problem because it involves optimization fixing a level of output. Such problems are closely related to profit maximization provided the firm is a price taker regarding the choice variables in the constrained problem. Our previous analysis can be adapted to such settings.[37]

One difference between analysis of the unconstrained profit maximization problem

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

[37] Constrained profit maximization problems, including the cost minimization and revenue maximization problems, are well-known variants of the profit maximization problem. For a textbook treatment see Fuss & McFadden (1978).

and constrained problems is that variables describing the constraints may be choice variables, and hence endogenous. For example, with cost minimization the given quantity needed may come from a profit maximization problem. Thus, Theorem 1 cannot directly be applied. To address endogeneity we may apply the results in Section 6, adapted to a constrained setting. To fix ideas, we now describe the cost minimization and revenue maximization problems in detail.

Cost Minimization Assume that firms are minimizing the cost of production for a given vector of outputs y_o. This is compatible with firms having market power in output markets, but we still require that firms be input price takers. In our terminology, the objective of the firm (inputs y_i are assumed to be nonpositive) is

$$c(y_o, p_i, e) = \max_{y_i \in Y_i(e, y_o)} p_i' y_i \,,$$

where $Y_i(e, y_o) = \{y_i \,:\, (y_o', y_i')' \in Y(e)\}$ is the set of input quantities that make output vector y_o available for production, and p_i is a vector of input prices. In this formulation, given that we treat inputs as nonpositive quantities and prices as positive, $c(y_o, p_i, e)$ is nonpositive by construction. Thus, in the classical sense, $c(y_o, p_i, e)$ represents negative costs. Note that if the correspondence $Y_i(\cdot, y_o)$ is not empty, then it is a production correspondence (Definition 1). The function c is well-defined as long as $Y_i(e, y_i)$ is nonempty.[38]

In this setting, we need to observe total cost, input prices, and output quantities. Note that these observables are different from our benchmark analysis of unconstrained profit-maximization, but are closely related. Cost now replaces profits as the value of an optimization problem. Our previous results go through whenever the cost function is identified. In some settings (e.g. Bilodeau et al. (2000)), outputs are chosen exogenously. In such cases, if an analyst assumes that conditional on y_o, prices and heterogeneity are independent, Theorem 1 may be applied to identify the cost function by conditioning on y_o. Once the cost function is identified, recoverability of the cost function allows one to recover the input requirement set $Y_i(e, y_o)$.

Revenue Maximization Assume that firms maximize revenue among all possible output combinations, fixing a given level of input y_i. This is the mirror image of cost minimization and only requires price-taking in the output prices. The objective of

[38]If $Y_i(e, y_o)$ is empty, then we can set $c(y_o, p_i, e) = -\infty$.

the firm is

$$r(y_i, p_o, e) = \max_{y_o \in Y_o(e,y_i)} p_o' y_o \,,$$

where $Y_o(e, y_i) = \{y_o : (y_o', y_i')' \in Y(e)\}$ is the set of output quantities producible with the input vector y_i, and p_o is a vector of output prices. Note that if the correspondence $Y_o(\cdot, y_i)$ is not empty, then it is a production correspondence (Definition 1).[39] The revenue function can be seen as a multi-output generalization of the classical production function approach. In fact, for the single-output case, the revenue function is equivalent to the production function multiplied by the price of the output. The main advantage of the revenue maximization approach is that one can cover multi-output production; a disadvantage is one must assume price-taking behavior. The key feature in analyzing revenue maximization is that the output vector y_o is reduced to a scalar object (the firm's revenue, $p_o' y_o$). We show that this reduction of dimensionality (i.e. we may only observe revenue not its parts) does not prevent us from recovering heterogeneous production sets. Nor does it prevent us from providing counterfactual bounds, such as bounds on counterfactual revenue at new prices. Our analysis builds on and extends classical duality techniques (Fuss & McFadden (1978)) to a setting with unobservable heterogeneity and limited variation in prices.

7.1. Examples

We now provide several examples of constrained and unconstrained problems examined in existing work. The unifying structure of these disperse examples is that with data on the value function of a problem, it is possible to recover a production set nonparametrically in the presence of nonseparable heterogeneity. These existing papers either study a representative agent problem (without unobservable heterogeneity) or heterogeneity that is additively separable, and typically impose parametric restrictions.

Example 11 (Profit, Berger et al. (1993))**.** Consider analysis of production of loans by a commercial bank. The vector of outputs y_o is composed of business loans and consumer loans. The vector of flexible inputs y_i includes labor and purchased funds; recall y_i itself is weakly negative because the full output/input vector y is a net output vector. The vector of fixed inputs y_f includes core deposits and physical capital. The analyst observes profits, the prices of the outputs p_o, the prices of the flexible inputs p_i, and the quantities of all outputs and inputs including the

[39] Similar to the cost minimization problem, we can set $r(y_i, p_o, e) = -\infty$ if $Y_o(e, y_i) = \emptyset$.

fixed inputs y_f.[40] Banks are assumed to be profit maximizers and price takers.[41] Our analysis shows that with the cross-section observed by Berger et al. (1993) it is possible to nonparametricaly identify the production sets of banks, indexed by the level of productivity for a given level of fixed inputs. In contrast, Berger et al. (1993) considers a parametric framework with additively separable error.

Example 12 (Cost, Bilodeau et al. (2000))**.** Consider a hospital that minimizes cost for a vector of outputs y_o given fixed inputs y_f (including the number of physicians and capital). Flexible inputs are denoted by y_i and the prices of the flexible inputs are denoted by p_i.[42] The hospital associated with productivity level $e \in [0, 1]$ minimizes cost:

$$c(y_o, y_f, p_i, e) = \max_{y_i \in Y(e, y_o, y_f)} p_i' y_i \,,$$

where $Y(e, y_o, y_f)$ denotes the set of flexible input quantities that make y_o available given y_f. For simplicity, we stack quantities in the vector $h = (y_o', y_f')'$ so we can write $c(h, p_i, e)$. Bilodeau et al. (2000) observes total costs, quantities of fixed outputs and inputs, as well as prices for all inputs. We show that using our results we achieve identification of a fully nonparametric cost function with nonseparable heterogeneity, in contrast with Bilodeau et al. (2000), which focuses on a parametric setup with additively separable heterogeneity. Bilodeau et al. (2000) studies hospitals run by a regulator which means that outputs and fixed inputs can be thought as exogenous, in the sense that they are independent from productivity.[43] However, prices of fixed inputs can be used as instruments for y_f, and prices of outputs can be used as instruments for y_o, in cases where markets influence the choice of outputs and fixed inputs.

Example 13 (Zero Profits and Revenue Observed, Combes et al. (2017))**.** Consider the production of housing, in which the analyst summarizes all goods and services provided by a house as a single output y_o. The production function satisfies $y_o = f(-y_i, e) = f((k, l)', e)$, where k is capital used and l is land, and inputs are collected

[40] They observe flows and balances that they divide to obtain prices. Profits are computed as total revenues minus total cost.

[41] Berger et al. (1993) assumes that the fixed inputs do not depend on the contemporaneous productivity, treating them as exogenous. If this assumption is relaxed, then prices of this fixed inputs can be used as instruments.

[42] Outputs include all services provided by the hospitals of interest (e.g., inpatient car, outpatient visits), variable inputs include labor, supplies, food for patients, drugs, energy. They observe all inputs and outputs.

[43] Nerlove (1963) takes a similar approach for the electricity industry, where the output quantity is treated as exogenous since it is fixed by a regulator.

as $y_i = -(k, l)'$. In Combes et al. (2017), the analyst does not observe housing goods and services y_o, which is recognized as an important problem for the estimation of a production function for housing (e.g. Epple et al. (2010)). Assume that firms maximize profits and are in a long-term equilibrium with zero profits. A necessary condition for profit maximization is that for fixed level of l and k, the firms must maximize revenue per unit of land

$$r(y_i, p_o, e) = \max_{y_o \in Y_i(e, y_i)} p_o y_o \,.$$

Then note that by the assumption of profit maximization,

$$\pi((p_o, p_k, p_l)', e) = \max_{y_i} r(y_i, p_o, e) - p_k k - p_l l.$$

The zero profit condition implies that in equilibrium,

$$p_k k + p_l l = r(-(k, l)', p_o, e)\,.$$

Thus, observed input prices $(p_k, p_l)'$ and quantities y_i allow us to compute the value of the revenue function. The second issue is that p_o may not be observed either (Epple et al. (2010)). Combes et al. (2017) address this by assuming that the price of the output is a deterministic function of location, such that $p_o = g(x)$, where x is location. We can then write the revenue equation of interest as

$$p_k k + p_l l = r(y_i, g(x), e)\,.$$

Our results covering endogeneity (Section 6) may be used to identify the structural function $\tilde{r}(y_i, x, e) = r(y_i, g(x), e)$. One can use local average measures of quality of life as an instrument, as proposed by Albouy & Ehrlich (2018). Recall that using homogeneity of $r(y_i, \cdot, e)$ in prices, we can identify g and r as in Section 5 from the equation

$$\partial_x \tilde{r}(y_i, x, e) \frac{g(x)}{\partial_x g(x)} = \tilde{r}(y_i, x, e)\,.$$

Using this equation, g can be identified up to a location/scale normalization. Recall that

$$r(y_i, g(x), e) = g(x) f(-y_i, e)\,,$$

where f is the production function. Thus, identification of r (as y_i and x vary) identifies f as y_i varies. We note that our identification analysis applies as well when there are multiple outputs with unobservable prices of these outputs. In this case one would need an output-specific price attribute for each output with a missing price.

In this example, our framework provides a new identification result for the housing production function and the price link function g with the same observables as in Combes et al. (2017), namely $(p_l, l, k, x, w')'$, where w is a vector of instruments.[44] Our identification results allow rich nonseparable heterogeneity. This also provides an alternative methodology to the identification results of Epple et al. (2010) when capital and land are observed.[45] The revenue function can be used to recover the production set and thus the production function of housing. Also, we recover the pricing function g that maps locations to prices of outputs. Recal that our general analysis treats multiple outputs and single outputs in a common setup. Thus, the analysis in this example may be adapted to handle multiple outputs.

Example 14. (Zero Profits and Cost Observed, Albouy & Ehrlich (2018)) Consider the production of one house using a technology that uses as inputs land and materials.[46] The analyst is interested in identifying the substitution/complementarity patterns of land and materials for the production of a house. As in Albouy & Ehrlich (2018), one may have access to a dataset with price data and no quantities. Following Albouy & Ehrlich (2018) we make two economic assumptions to facilitate analysis. First, we assume average cost is equal to marginal cost (under constant returns to scale), and second, we assume zero profits. The unit cost function is given by $c(p_i, e) = \max_{y_i \in Y_i(e,1)} p_i' y_i$, where p_i is the input price vector that consists of the price of land and the price of materials. Under the assumption of zero-profits, equilibrium conditions imply that $p_o = -c(p_i, e)$. Note also that this equilibrium condition implies that prices can be used to recover costs. Recall that here, a firm either produces one or zero units of housing, and this cost function is evaluated at $y_o = 1$ unit of housing. Our results show that with the same economic assumptions and observables as in Albouy & Ehrlich (2018), namely price data $(p_o, p_i')'$, the unit cost function is identified in an environment with rich heterogeneity. The unit cost function can then be

[44] Combes et al. (2017) observes for each transaction the size and value of the land parcel $((p_l l, l)')$. They also observe the cost of construction of the house, here we use it as $p_k k$, where k is broadly understood as a composite input representing capital and materials. In Combes et al. (2017), following Epple et al. (2010), a fixed price of capital p_k is assumed to be known, and then they use the total cost of construction to get k. Combes et al. (2017) uses as instruments for l and k, (i) the distance to the center of each city or location and urban fixed effects, (ii) mean income and the dispersion of income in each location, (iii) geophysical variables of the terrain, and (iv) the share of the population with a university degree.

[45] We note their setup does not study identification with unobservable heterogeneity.

[46] The technology is encoded in the constraint set $Y_i(e, y_o)$ where the output quantity is $y_o = 1$.

used to recover the technology of production of a house, or to provide counterfactual bounds if the cost function is identified only at a limited set of input prices.

7.2. General Formulation

We now describe how our results may be adapted to a general setup that includes cost minimization and profit maximization as a special case. We consider a general constrained profit maximization problem,

$$\pi_c(h, p_z, e) = \max_{z \in Y(e,h)} p_z' z \,,$$

where h is a vector of constrained or fixed variables, z is the variable of choice, and p_z is a vector of prices of z, supported on P_z. The function π_c is the restricted profit function (Fuss & McFadden (1978)) for a firm with productivity level e. We note that the value of this problem involves p_z and z and not any revenues or costs arising from fixed variables. The vector of outputs and inputs $y \in Y(e)$ can be rearranged to have fixed components first, and variable components second, i.e. $y = (h', z')'$. The variable of choice z is constrained to belong the convex set $Y(e, h)$ defined as

$$Y(e, h) = \Big\{ z \in \mathbb{R}^{d_z} \ : \ (h', z')' \in Y(e) \Big\} \,.$$

We refer to $Y(\cdot, h)$ as the constrained production correspondence. Note that all properties (e.g., convexity, closedness, free disposal, the recession cone property, monotonicity, hemicontinuity) of the production correspondence are inherited when we consider $Y(\cdot, h)$. Note that in some settings, this set may be empty for certain values of e. For example with cost minimization, if h is a given level of outputs, it may be that for firms with sufficiently low productivity, h is not attainable for *any* level of inputs.[47] This concern is ruled out by typical parametric families, but is still present in nonparametric settings.

When P_z consists of all weakly positive prices, the value function $\pi_c(h, p_z, e)$ is the support function of $Y(e, h)$ for fixed h. Namely,

$$Y(e, h) = \{ z \in \mathbb{R}^{d_z} \ : \ p_z' z \le \pi_c(h, p_z, e), \, \forall p_z \in P_z \} \,.$$

In general, however, P_z may be finite, and it may only be possible to learn certain features of $Y(e, h)$. For example, once we identify π_c for values of prices in P_z, our

[47]This has a statistical analogue in that the support of $\mathbf{e}$ may vary when one conditions on $\mathbf{h}$.

previous results may be used to bound the production correspondence $Y(e, h)$ as in Sections 2 or 3. If we further identify the value function as h varies, we may also identify the unconstrained production correspondence $Y(\cdot)$. We do not formally study this case since in different settings there are different restrictions involving the sets $Y(e, h)$ as h varies. For example, with cost minimization suppose $\tilde{h} \geq h$ in the usual partial order, i.e. $\tilde{h}$ is weakly higher along every dimension. With the setup of cost minimization it follows that $Y(e, \tilde{h}) \subseteq Y(e, h)$, reflecting that there are fewer ways to produce $\tilde{h}$ than h due the assumption of free disposal. This imposes additional restrictions that may be used for identification and counterfactual bounds (Varian (1984)).

In order to identify the constrained profit function $\pi_c(\cdot)$, assume that we observe constraints h, either prices p_z or price attributes x (such that $p_{z,s} = g_s(x_s)$), and values π_c such as profits, revenues, or costs. Note that the results of Section 6 do not require any special structure of the profit function such as convexity or homogeneity. Hence, even if constraints h or attributes x are endogenous, one can apply the results of Section 6. Thus we can identify the composite function $\tilde{\pi}_c(\cdot, \cdot, \cdot) = \pi_c(\cdot, g(\cdot), \cdot)$, where $g(\cdot)$ is the unknown pricing function. Note that $\tilde{\pi}_c$ is a generalization of $\tilde{\pi}$ from Section 5. Hence, the results in Section 5 can be used to identify $g(\cdot)$ by analogous arguments, and thus one can identify function π_c. In summary, the previous results can be adapted to analysis of constrained problems, which covers cost minimization and revenue maximization as special cases. We note that while it is key for our analysis that the objective function is linear, the analysis applies outside of the firm problem as well.[48]

8. 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, and with potentially limited variation in prices. Our key assumption on heterogeneity is that firms can be ranked in terms of productivity. This is equivalent to weak monotonicity of the heterogeneous profit function, which we leverage to identify the heterogeneous profit function by generalizing the identification approach

[48] See Cunha et al. (2010) for an example of production function analysis outside of the firm problem.

of Matzkin (2003). Once the heterogeneous profit function is identified, we show how to identify firm production possibility sets under rich variation in prices, and also describe the most that can be learned about such sets with limited price variation. Building on this result, we provide sharp bounds on counterfactual profits at new price as well as bounds on optimal output/input vectors at a new price.

The assumption that firms can be ranked in terms of productivity allows us to present constructive identification results for profit functions, production sets, and sharp counterfactual bounds. We note, however, that the identification results for production sets as well as counterfactual bounds make use of the fact that a structural profit function has *somehow* been identified. Thus, the identification results for sets and counterfactuals apply beyond the setting of scalar heterogeneity *provided* one can identify the structural profit function $\pi(p, e)$. In addition, the identification results for sets and counterfactual bounds apply to a representative agent analysis if we replace $\pi(p, e)$ with $\mathbb{E}[\pi(p, \mathbf{e})]$, where the expectation is taken with respect to $\mathbf{e}$.

In order to extend the applicability of our core analysis, we provide several additional results that further lay a foundation for empirical work. We present a general 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 strictly positive prices. We also provide a constructive identification result showing how to work with price attributes instead of prices. This technique uses Euler's homogeneous function theorem to identify unknown index functions, and may be of indepenent interest. We then describe how the independence conditions in our main analysis may be relaxed in the presence of endogeneity. Finally, we describe how our baseline analysis of profit maximization applies to other constrained maximization problems in which the objective function is linear, such as cost minimization or revenue maximization.

We leave dynamic considerations for future work. Here we focus on a cross section and we do not model the dynamic firm problem. In some cases, however, analysis of a dynamic setting is possible by reducing certain features of the problem to static ones. For example, Gandhi et al. (2017) study a dynamic setting in which certain flexible inputs are chosen similarly to how outputs and inputs are chosen in a static problem.

References

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

Aguiar, V. & Kashaev, N. (2018). Stochastic revealed preferences with measurement error. *arXiv preprint arXiv:1810.05287*.

Albouy, D. & Ehrlich, G. (2018). Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics*.

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

Allen, R. & Rehbeck, J. (2018). Identification with additively separable heterogeneity. *Working Paper*.

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

Berger, A. N., Hancock, D., & Humphrey, D. B. (1993). Bank efficiency derived from the profit function. *Journal of Banking & Finance*, 17(2-3), 317–347.

Bilodeau, D., Cremieux, P.-Y., & Ouellette, P. (2000). Hospital cost function in a non-market health care system. *Review of Economics and Statistics*, 82(3), 489–498.

Blundell, R. W., Browning, M., & Crawford, I. A. (2003). Nonparametric engel curves and revealed preference. *Econometrica*, 71(1), 205–240.

Blundell, R. W., Kristensen, D., & Matzkin, R. L. (2017). *Individual counterfactuals with multidimensional unobserved heterogeneity*. Technical report, cemmap working paper.

Brown, D. J. & Matzkin, R. L. (1996). Testable restrictions on the equilibrium manifold. *Econometrica: Journal of the Econometric Society*, (pp. 1249–1262).

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.

Cherchye, L., Rock, B. D., & Walheer, B. (2016). Multi-output profit efficiency and directional distance functions. *Omega*, 61, 100 – 109.

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).

Combes, P.-P., Duranton, G., & Gobillon, L. (2017). The production function for housing: Evidence from france.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883–931.

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.

Doraszelski, U. & Jaumandreu, J. (2013). R&d and productivity: Estimating endogenous productivity. *Review of Economic Studies*, 80(4), 1338–1383.

Ekeland, I., Heckman, J. J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of political economy*, 112(S1), S60–S109.

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.

Hanoch, G. & Rothschild, M. (1972). Testing the assumptions of production theory: a nonparametric approach. *Journal of Political Economy*, 80(2), 256–275.

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*.

Marschak, J. & Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica, Journal of the Econometric Society*, (pp. 143–205).

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.

Nerlove, M. (1963). Returns to scale in electricity supply. *Measurement in Economics*.

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. (1982). The nonparametric approach to demand analysis. *Econometrica: Journal of the Econometric Society*, (pp. 945–973).

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

Zellner, A., Kmenta, J., & Dreze, J. (1966). Specification and estimation of cobb-douglas production function models. *Econometrica: Journal of the Econometric Society*, (pp. 784–795).

A. Proofs of Main Results

A.1. Proof of Lemma 2

See Rockafellar (1970), Corollary 13.1.1.

A.2. Proof of Theorem 1

Fix some p. For every $\underline{\pi}$ define

$$E(\underline{\pi}) = \{e \in E \, : \, \pi(p, e) \leq \underline{\pi}\} \, .$$

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

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

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

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

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

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

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

$$\pi(p, e) = \inf\left\{\underline{\pi} \, : \, e \leq e^*(\underline{\pi})\right\}$$

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

A.3. Proof of Theorem 2

It is immediate that $\tilde{Y}(e)$ is closed, convex, and satisfies free disposal for every $e \in E$. Moreover, $\max_{y \in \tilde{Y}(e)} p'y = \pi(p, e)$ 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, any given production set $Y'(e)$ can generate the data if and only if $\max_{y \in Y'(e)} p'y = \pi(p, e)$ for every $p \in P$. The set $\tilde{Y}(e)$ is constructed as the largest set (not necessary production set) consistent with profit maximization. This set is closed, convex, and satisfies free disposal. Since a production correspondence *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

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

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

A.4. Proof of Proposition 2

Fix some $e \in E$. To simplify notation we drop e from the objects below (e.g. $\pi(p, e) = \pi(p)$ and $y_p(e) = y_p$). Suppose $\{y_p\}_{p \in P}$ can generate $\{\pi(p)\}_{p \in P}$. Since $\{y_p\}_{p \in P}$ are profit-maximizing output/input vectors we must have $p'y_p = \pi(p)$. To prove that $p^{*\prime} y_{p^{*\prime}} \geq p^{*\prime} y_p$ for all $p, p^* \in P$, assume the contrary. But then y_{p_*} is not maximizing profits at p^* since y_p is available. The contradiction proves necessity.

To prove sufficiency consider

$$Y^* = \text{co}(\{y_p\}_{p \in P}) + \mathbb{R}_-^{d_y} \, ,$$

where $\text{co}(A)$ denotes the convex hull of a set A, i.e. the smallest convex set containing A. The summation is the Minkowski sum. Y^* is sometimes referred to as the free-disposal convex hull of $\{y_p\}_{p \in P}$. In particular, note that Y^* is convex, closed, and satisfies free disposal.

We obtain that for every $p \in \mathbb{R}^{d_y}_{++} \cap \mathbb{S}^{d_y-1}$,

$$\sup_{y \in Y^*} p'y = \sup_{y \in \text{co}(\{y_p\}_{p \in P})} p'y + \sup_{y \in \mathbb{R}^{d_y}_{-}} p'y = \sup_{y \in \text{co}(\{y_p\}_{p \in P})} p'y\,.$$

Because P is finite, $\{y_p\}_{p \in P}$ is bounded. Thus, its convex hull $\text{co}(\{y_p\}_{p \in P})$ is also bounded. This implies that $\sup_{y \in Y'} p'y$ is finite for every $p \in \mathbb{R}^{d_y}_{++} \cap \mathbb{S}^{d_y-1}$, hence the recession cone property is satisfied for the set Y^*.[49]

It is left to show that

$$\pi(p, e) = p'y_p = \sup_{y \in Y^*} p'y$$

for every $p \in P \cap \mathbb{S}^{d_y-1}$. The first equality is assumed. Suppose the second equality is not true for some p^*. Then there exists $\tilde{y} \in Y^*$ such that $p^{*\prime} y_{p^*} < p^{*\prime}\tilde{y}$. Since $\tilde{y} \in Y^*$ it can be represented as a finite convex combination of points from $\{y_p\}_{p \in P}$. But since

$$p^{*\prime} y_{p^*} \geq p^{*\prime} y_p\,,$$

for all $p, p^* \in P$ it has to be the case that

$$p^{*\prime} y_{p^*} \geq p^{*\prime} \tilde{y}.$$

The contradiction completes the proof. Since the choice of e was arbitrary the result holds for all $e \in E$.

A.5. Proof of Theorem 3 and Proposition 3

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

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

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

[49] We note that Varian (1984) studies a result related to this proposition, taking as primitives a deterministic dataset of prices and quantities. He does not verify the recession cone property.

is unbounded in direction u, then $h_A(u) = \infty$.

As preparation, we need a technical lemma. This lemma involves a polar cone, which for a set C is defined by

$$\text{PolCon}(C) = \{u \in \mathbb{R}^{d_y} \ : \ u'p \leq 0, \forall p \in C\}.$$

Lemma 5. *Let $\bar{P} \subseteq \mathbb{S}^{d_y-1}$ be a closed set such that $\cup_{\lambda>0}\{\lambda p, p \in \bar{P}\}$ is a closed, convex cone, and let $a : \mathbb{R}^{d_y} \to \mathbb{R}$ be a convex, homogeneous of degree 1 function. Define*

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

If $\text{PolCon}(\bar{P})$ *is nonempty, then for any $u \in \mathbb{S}^{d_y-1}$,*

$$h_A(u) = \begin{cases} a(u), & \textit{if } u \in \bar{P}, \\ +\infty, & \textit{otherwise.} \end{cases}$$

Proof. Case 1. Take $u \in \bar{P}$. Since $a(\cdot)$ is convex and homogeneous of degree 1 $h_A(u) = a(u)$.

Case 2. Take $u \in \mathbb{S}^{d_y-1} \setminus \bar{P}$. First, we establish that there always exists $u^* \in \text{PolCon}(\bar{P})$ such that $u'u^* > 0$. To prove this suppose to the contrary that for every $u^* \in \text{PolCon}(\bar{P})$, $u'u^* \leq 0$, it follows that $u \in \text{PolCon}(\text{PolCon}(\bar{P}))$. The latter is not possible, since $\text{PolCon}(\text{PolCon}(\bar{P}))$ is the smallest closed convex cone containing $\bar{P}$ (Rockafellar (1970), Theorem 14.1), and $u \notin \bar{P}$ by assumption.

For some u^* that satisfies $u'u^* > 0$, consider $y^m = y^0 + mu^*$, $m = 1, 2, \ldots$, where y^0 is an arbitrary point from A. Since $u^* \in \text{PolCon}(\bar{P})$, by construction $u^{*\prime}p \leq 0$ for all $p \in \bar{P}$. Using this fact, note that $y^m \in A$ for all $m = 1, 2, \ldots$ since

$$p'y^m = p'y^0 + mu^{*\prime}p \leq a(p) + 0$$

for all $p \in \bar{P}$. Finally,

$$h_A(u) \geq u'y^m = u'y^0 + mu'u^*$$

diverges to $+\infty$, since $u'u^* > 0$. ∎

We now provide a key lemma. This result generalizes a classical result that holds for $\bar{P} = \mathbb{S}^{d_y-1}$. To our knowledge this result is new, and it may be of independent interest.

Lemma 6. *Let $d_y \geq 2$ and let the functions $a, b : \mathbb{R}^{d_y}_{++} \to \mathbb{R}$ be convex and homoge-*

neous 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 $\bar{P} \subseteq \mathbb{R}^{d_y}_{++}$ *is convex and compact. Then*

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

Proof. For closed 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,

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

Because $\bar{P}$ is a subset of $\mathbb{R}^{d_y}_{++}$, its polar cone $\text{PolCon}(\overline{P})$ is nonempty; in particular the polar cone contains the negative unit vector $(-1, \ldots, -1)'$. The set $\bar{P}$ satisfies the conditions of Lemma 5, and so we obtain that $h_A(u) = h_{B+\rho\mathbb{B}^{d_y-1}}(u) = h_B(u) = h_{A+\rho\mathbb{B}^{d_y-1}}(u) = \infty$ for all $u \in \mathbb{S}^{d_y-1} \setminus \{p/\|p\| \ , p \in \bar{P}\}$. Hence,

$$\begin{aligned}
&\{\rho \in \mathbb{R}_+ \ : \ A \subseteq B + \rho \mathbb{B}^{d_y-1}, B \subseteq A + \rho \mathbb{B}^{d_y-1}\} \\
&\quad = \{\rho \in \mathbb{R}_+ \ : \ \ h_A(u) \leq h_{B+\rho\mathbb{B}^{d_y-1}}(u), \\
&\qquad\qquad h_B(u) \leq h_{A+\rho\mathbb{B}^{d_y-1}}(u), \forall u \in \{p/\|p\| \ : p \in \bar{P}\}\} \\
&\quad = \{\rho \in \mathbb{R}_+ \ : \ \ h_A(u) \leq h_B(u) + h_{\rho\mathbb{B}^{d_y-1}}(u), \\
&\qquad\qquad h_B(u) \leq h_A(u) + h_{\rho\mathbb{B}^{d_y-1}}(u), \forall u \in \{p/\|p\| \ : \ p \in \bar{P}\}\} \\
&\quad = \{\rho \in \mathbb{R}_+ \ : \ h_A(u) \leq h_B(u) + \rho, h_B(u) \leq h_A(u) + \rho, \forall u \in \{p/\|p\| \ : p \in \bar{P}\}\} \\
&\quad = \{\rho \in \mathbb{R}_+ \ : \ \sup_{u \in \{p/\|p\| \, : \, p \in \bar{P}\}} \|h_A(u) - h_B(u)\| \leq \rho\} \, .
\end{aligned}$$

Now 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. Thus,

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

■

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 3. We first present two lemmas, which are modifications of Lemmas 6 and 7 in Brunel (2016).

Lemma 7. *Assume that $\bar{P} \subseteq \mathbb{S}^{d_y-1} \bigcap P$ is compact and $\cup_{\lambda>0}\{\lambda p : p \in \bar{P}\}$ is convex. Let $a : \bar{P} \to \mathbb{R}$ be a continuous function. Let $A = \{y \in \mathbb{R}^{d_y} : p'y \leq a(p),\, p \in \bar{P}\}$ be nonempty. It follows that 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 $p^* \in \bar{P}$. Note that $h_A(p^*) \leq a(p^*) < \infty$. Since A is closed, by the supporting hyperplane theorem $h_A(p^*) = p^{*\prime}y^*$ for some $y^* \in A$.

The rest of the lemma follows from Theorem 2(b) in López & Still (2007) if we show that $P' = \{p \in \bar{P} : p'y^* = a(p)\}$ is nonempty. By way of contradiction assume that P' is empty. Hence, $p'y^* < a(p)$ for all $p \in \bar{P}$. 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 $p \in \bar{P}$,

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

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. Thus, P' is nonempty. ■

Lemma 8. *Assume that $\bar{P} \subseteq \mathbb{S}^{d_y-1} \bigcap P$ is compact and $\cup_{\lambda>0}\{\lambda p : p \in \bar{P}\}$ is 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*

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

are nonempty 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 7 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

$$\begin{aligned} a(p^*) = h_A(p^*) = h_A\left(\sum_{p\in P_n^*} \lambda_{p,n} p\right) \leq \sum_{p\in P_n^*} \lambda_{p,n} h_A(p) = \sum_{p\in P_n^*} \lambda_{p,n} a(p) \leq \sum_{p\in P_n^*} \lambda_{p,n}(b_n(p) + \eta_n) \\ = \sum_{p\in P_n^*} \lambda_{p,n} p' y_n^* + \eta_n \sum_{p\in P_n^*} \lambda_{p,n} = p^{*\prime} y_n^* + \eta_n \sum_{p\in P_n^*} \lambda_{p,n} = h_{B_n}(p^*) + \eta_n \sum_{p\in P_n^*} \lambda_{p,n}\,. \end{aligned} \tag{5}$$

Moreover,

$$h_{B_n}(p^*) \leq b_n(p^*) \leq a(p^*) + \eta_n\,. \tag{6}$$

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 (6) 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,

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

■

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

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

To prove Proposition 3, note that by Lemma 6, with probability 1,

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

The conclusion then follows by applying Lemma 8 to the right hand side of the equality above.

A.6. Proof of Lemma 3

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

A.7. Proof of Theorem 4

To prove sufficiency of (i), note that $\tilde{\pi}(x, \cdot)$ is identified for every $x \in X$ by Lemma 3.

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 $x \in X$

$$\sum_{j=1}^{d_y} \partial_{g_j}\pi(g(x))g_j(x_j) = \pi(g(x))\,. \tag{7}$$

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

$$\partial_{g_j}\pi(g(x))\partial_{x_j}g_j(x_j) = \partial_{x_j}\tilde{\pi}(x)\,, \tag{8}$$

for every $j = 1, \ldots, d_y$. Combining (7) and (8) we get that

$$\sum_{j=1}^{d_y} \partial_{x_j}\tilde{\pi}(x)\frac{1}{\partial_{x_j}(\log(g_j(x_j)))} = \tilde{\pi}(x) \tag{9}$$

as long as $0 < \left\| \frac{\partial_{x_j} g_j(x_j)}{g_j(x_j)} \right\| < \infty$ for every $j = 1, \ldots, d_y$.

Let $t = \left(\frac{1}{\partial_{x_j}(\log(g_j(x_j)))} \right)_{j=1,\ldots,d_y-1}$. Note that t does not depend on x_{d_y}. Since $\tilde{\pi}$ satisfies the rank condition there exists nonsingular $A(\tilde{\pi}(x^*))$ such that equation (9) can be rewritten as

$$At = b\,, \tag{10}$$

where $b = (b_j)_{j=1,\ldots,d_y-1}$ and $b_j = \tilde{\pi}(x_j^*) - \partial_{x_{d_y}} \tilde{\pi}(x_j^*) x_{d_y,j}$. Since $A(\tilde{\pi}(x^*))$ is of full rank and is identified, t is identified. Since the choice of x_{-d_y} was arbitrary and we know the location (Assumption 6(ii)) we identify $g_j(\cdot)$ for every $j = 1, \ldots, d_y - 1$.

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

$$\mathbb{E}[\boldsymbol{\pi}|\mathbf{x} = x] = \mathbb{E}[\tilde{\pi}(\mathbf{x}, \mathbf{e}) \mid \mathbf{x} = x]\,,$$

and homogeneity is clearly preserved under expectations.

A.8. Proof of Lemma 4

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

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

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

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

The law of iterated expectations together with Assumptions 4 and 7 then imply that

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

B. 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 $\tilde{\pi}$ 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

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

that satisfies the properties described above. 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 15 (Unique Extension without Assumption 3)**.** Consider $\pi(p, e) = e \sum_{j=1}^{d_y} |p_j|$ with $E = [0, M]$, $0 < M < \infty$. This functions is homogeneous of degree 1 and convex in p, and hence the profit function for price-taking firms, indexed by e (Kreps (2012), Proposition 9.14). Let $\Delta^{d_y-1} = \{p \in \mathbb{R}^{d_y}_{++} \;:\; \sum_{j=1}^{d_y} p_j = 1\}$ denote the relative interior of the probability simplex, and let $S = \{p \in \Delta^{d_y-1} \;:\; |y_j - 1/d_y| \leq 1/d_y \text{ for each } j\}$ denote a convex set centered at the midpoint of the simplex. Let P be the probability simplex with the region S removed, i.e. $P = \Delta^{d_y-1} \setminus S$. Note that P is a subset of the affine space $\{p \in \mathbb{R}^{d_y} \;:\; \sum_{j=1}^{d_y} y_j = 1\}$, and $\pi_P(\cdot, e)$ is equal to e over P. Any convex extension of $\pi_P(\cdot, e)$ to the convex hull of P, Δ^{d_y-1}, must also be equal to e. In more detail, 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 $d_y - 1$). Because there is a unique convex extension of $\pi_P(\cdot, e)$ to all of Δ^{d_y-1}, there is a unique convex and homogeneous extension to all of $\mathbb{R}^{d_y}_{++}$. By Lemma 1 the production correspondence is identified even though Assumption 3 fails to hold.

For additional geometric intuition behind this example, consider a line segment from $(0, 0$ to $(1, 0)$ in $\mathbb{R}^2$. If one deletes a chunk out of the middle of this line segment, but maintains each endpoint, then the convex hull of this modified set is actually the original set.

This example also shows that it is possible to uniquely determine $\pi(p, e)$ at values

p that are not in the set $\text{int}\left(\text{cl}\left(\bigcup_{\lambda>0}\{\lambda p\,:\,p\in P\}\right)\right)$. 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.

C. Additional Results for Unobservable Prices

In this section we show that if prices and profits are not observed, but price attributes and the output/input vector are, then we may recover the distribution of what we term *pseudo-profits*. This distribution, conditional on attributes, may be thought of as the 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 6(ii)), we recover the location and scale of profits. 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 possibility sets by our previous arguments.

We make use of a representative-firm assumption, as formalized below.

Assumption 8. *(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 possibility 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 (2018) 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.[50]

By exploiting a symmetry feature that arises due to optimization (cf. Allen & Rehbeck (2018)), 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 6, 5, and 8 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,*

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

Proof. This follows by adapting arguments in Allen & Rehbeck (2018). The envelope theorem applied to the representative firm problem yields Hotelling's lemma,

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

Differentiating, we obtain

$$\partial_{x_k}\mathbb{E}\left[y(g(x), \mathbf{e})\right] = \nabla_{j,k}\overline{\pi}(g(x))\partial_{x_k}g_k(x_k). \tag{11}$$

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 (11) and its counterpart with j, k interchanged to obtain,

$$\frac{\partial_{x_j}\mathbb{E}\left[y_k(g(x), \mathbf{e})\right]}{\partial_{x_k}\mathbb{E}\left[y_j(g(x), \mathbf{e})\right]} = \frac{\partial_{x_j}g_j(x_j)}{\partial_{x_k}g_k(x_k)}. \tag{12}$$

Now set $k = d_y$. Then (12) is valid because we have assumed the global restriction

[50] See Allen & Rehbeck (2018) for a rigorous statement.

$\nabla_{j,d_y}\overline{\pi} \neq 0$ 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 6(i), we identify differences in $\partial_{x_j} g_j(\cdot)$ for all j by integrating (12). By Assumption 6(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. ■

Recall that by Hotelling's lemma, twice differentiability of the aggregate profit function $\overline{\pi}(\cdot)$ amounts to differentiability of the aggregate supply function $\mathbb{E}[y(\cdot,\mathbf{e})]$. 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 possibility set $\overline{Y}$ can be written as a Cartesian product of two nonempty sets, e.g. $\overline{Y} = \overline{Y}^1 \times \overline{Y}^2$.[51]

Once g is identified, profits are identified from the relation $\boldsymbol{\pi} = g(\mathbf{x})'\mathbf{y}$ whenever we observe price attributes and the intput/output vector $\mathbf{y}$. Thus, we may identify the conditional distribution of profits given prices from the conditional distribution of inputs/outputs given price attributes. 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 (2018). Such an approach can be used to identify the distribution of profits given prices up to scale.

D. Endogeneity

In this section, we provide an identification result based on Chernozhukov & Hansen (2005). Note that Equation 4 is equivalent to the IV model of quantile treatment effects of Chernozhukov & Hansen (2005). Thus we can directly invoke their identification result. For some fixed $\delta, \underline{f} > 0$, define the relevant parameter

[51]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 $\overline{Y}^2$.

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 $\mathbf{w}$. The following theorem follows from Theorem 4 in Chernozhukov & Hansen (2005).

Theorem 6. *Suppose that*

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

(ii) Assumptions 4 and 7 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 $e \in E$*, if* $\pi', \pi^* \in \mathcal{P}$ *and* $\mathbb{E}\left[(\pi'(\mathbf{p}, e) - \pi^*(\mathbf{p}, e))\omega(\mathbf{p}, \mathbf{w}; e)|\mathbf{w}\right] = 0$ a.s.*, then* $\pi'(\mathbf{p}, e) = \pi^*(\mathbf{p}, e)$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 $\pi'(\cdot, e) \in \mathcal{P}$ *such that*

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

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