Western University

Scholarship@Western

Centre for Human Capital and Productivity. CHCP Working Papers

Economics Working Papers Archive

2020

2020-2 Prices, Profits, Proxies, and Production

Victor H. Aguiar

Roy Allen

Nail Kashaev

Prices, Profits, Proxies, and Production

by

Victor H. Aguiar, Roy Allen, and Nail Kashaev

Working Paper #2020-2 March 2020

Centre for Human Capital and Productivity (CHCP)

Working Paper Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

Prices, Profits, Proxies, and Production*

Victor H. Aguiar† vaguiar@uwo.ca

Roy Allen† rallen46@uwo.ca

Nail Kashaev† nkashaev@uwo.ca

First version: October 10, 2018

This version: March 15, 2020

Abstract This paper studies nonparametric identification and counterfactual bounds for heterogeneous firms that can be ranked in terms of productivity. Our approach works when quantities and prices are latent rendering standard approaches inapplicable. Instead, we require observation of profits or other optimizing-values such as costs or revenues, and either prices or price proxies of flexibly chosen variables. We extend classical duality results for price-taking firms to a setup with discrete heterogeneity, endogeneity, and limited variation in possibly latent prices. Finally, we show that convergence results for nonparametric estimators may be directly converted to convergence results for production sets.

JEL classification: C5, D24.

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

*An earlier version of this paper was circulated as "Prices, Profits, and Production: Identification and Counterfactuals." We are grateful to Paul Grieco, Lance Lochner, Salvador Navarro, David Rivers, Susanne Schennach, and Al Slivinsky for useful comments and encouragement. We also thank the ceminar participants at University of Montreal, McMaster University, and attendants of NASMES 2019, Empirical Microeconomics Workshop at University of Calgary, MEG 2019, CESG 2019, NAWMES 2020.

†Department of Economics, University of Western Ontario.

Introduction

This paper studies nonparametric identification of production sets and counterfactual bounds for firms, allowing multiple inputs and outputs, in an environment where both quantities and prices can be latent. We assume an analyst has data on the values of an optimization problem, such as profits, costs, or revenues, as well as prices or *price proxies*.

Identifying heterogeneous production sets is challenging in situations where the observability of some outputs/inputs or prices is problematic. For instance, in the housing market output quantities and output prices cannot be directly observed because houses provide different services that are hard to measure. However, housing values that can serve as price proxies may be observed (Epple et al., 2010). Other industries, such as health and banking, suffer from similar issues with unobservable inputs or outputs.[1] The latency of quantities makes standard approaches to estimate production functions not directly applicable. In addition, the latency of prices makes classical approaches using duality theory impossible to apply as well. In contrast, we require observability of values and prices or price proxies. While these variables are not always observed, they are available in many existing data sets.[2]

In order to obtain identification of firm-specific production possibility sets we exploit variation in prices or price proxies across markets and variation of optimization values across firms. Our framework extends classical duality theory by allowing (i) rich forms of complementarity and substitutability between outputs and inputs with discrete heterogeneity across firms, (ii) endogeneity between prices and productivity due to simultaneity and market entry decisions, and (iii) omitted prices of flexibly chosen variables. Classical duality theory focuses on either a nonstochastic or *representative agent* framework in which all prices are observed. Important contributions include Shephard (1953), Fuss & McFadden (1978), and Diewert (1982) among many others.

We assume that firms can be ranked in terms of productivity that can take

[1]In the health industry, it is 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).

[2]See Epple et al. (2010), Combes et al. (2017), and Albouy & Ehrlich (2018) in the context of housing; Burke et al. (2019) in the context of agriculture; Nerlove (1963) and Fabrizio et al. (2007) in the context of electricity generation; Roberts & Supina (1996), Foster et al. (2008), and Doraszelski & Jaumandreu (2013) in the context of manufacturing.

finitely many values. This assumption is key to unpack heterogeneity in multiple output/input production sets across firms from data such as prices or price proxies and scalar values of an optimization problem. We formalize this by assuming that a firm with higher productivity has access to all the production possibilities of a less productive firm, and more. Our framework covers Hicks-neutral heterogeneity in productivity as a special case.

Our approach exploits the rich shape constraints in our environment for identification and counterfactual analysis. With price-taking behavior, the structural value function is a convex and homogeneous function of prices. We present a new technique for identification that leverages these properties, together with the assumption that firms can be ranked according to productivity, to identify the structural value function (e.g. profit function). This technique relies on discrete heterogeneity, but allows flexible forms of selection into market. We require a monotone presence assumption, so that if a firm is present in some market with certain observables, then each more productive firm must be present in some market with the same observables. This handles certain monotone selection rules, e.g. only firms that can make nonnegative profits enter, but is much more general.

We next tackle the important possibility that not all prices are observed. Instead, we use *price proxies*, which are unknown functions of the missing prices. As one example, we show that aggregate market-level quantities can serve as price proxies. We leverage homogeneity of the value function to recover these unknown functions. This technique is new and is applicable to other settings with homogeneity of a structural function, and is therefore of independent interest.

Once the structural value function is identified, we turn to recoverability of the production sets. Here we leverage the classic insight that the value function serves as the support function of the production set. This allows us to characterize the most that can be said about heterogeneous production sets, even when price variation is limited. Building on this, we present a general framework for counterfactual questions such as sharp bounds on quantities or profits at a new price. Importantly, these bounds hold for each level of productivity, and thus characterize features of the distribution of firm behavior.

As mentioned previously, relative to classic work on duality we make several contributions by incorporating heterogeneity, endogeneity due to selection, and potential lack of prices.[3] Even when prices are observed but contain limited variation, we con-

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

tribute by providing new results using structural value functions to recover sets and conduct counterfactual analysis. There is little existing work concerning identification with limited 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.[4] Another paper studying limited price variation is Varian (1984), which works with quantities and prices and does not study unobservable heterogeneity.[5] While observation of prices and quantities implies observation of profits, the reverse is not true.

This paper contributes to the recent literature on identification and estimation of multi-output production with unobservable heterogeneity (e.g., Cunha et al., 2010, De Loecker et al., 2016, and Grieco & McDevitt, 2016). We differ since we do not observe quantities and we do not impose separability or parametric restrictions on the shape of production sets. Because we allow production of multiple outputs in flexible ways, use cross sectional variation, and do not observe quantities, we also differ from an important recent literature studying single output production in dynamic panel settings using quantities data, including Griliches & Mairesse (1995), Olley & Pakes (1996), Levinsohn & Petrin (2003), Ackerberg et al. (2015), and Gandhi et al. (2017).[6]

We also contribute to the literature studying recoverability of sets. We build on the tight relationship between the structural value function and the production possibility sets of firms, by providing an equality relating estimation error of value functions and estimation error of production possibility sets. This result allows one to adapt consistency results for any nonparametric estimators of the value function for the purpose of set estimation. The result is related to a classical result in convex analysis linking the distance of support functions with the distance of the corresponding sets, which has been exploited previously in the literature on partial identification.[7] We cannot apply the classical result since it would require seeing negative prices, which

framework of Allen & Rehbeck (2018).

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

[5] 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.

[6] As noted in Ackerberg et al. (2015), some output and input data often come in the form of sales and expenditures that need to be transformed into quantities. We work directly with total values (e.g. profits, total costs, or revenues).

[7] See, for instance, Beresteanu & Molinari (2008), Beresteanu et al. (2011), Kaido & Santos (2014), Kaido (2016), and Kaido et al. (2019).

requires a generalization.

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. Section 2 shows how to identify the structural value function. In Section 3 we extend our methodology to environments where one observes proxies that determine unobservable prices. Our main identification result for production possibility sets is in Section 4. Section 5 provides a general framework to conduct sharp counterfactual analysis in production environments. In Section 6 we show duality between estimation error in value functions and production sets. We conclude in Section 7. All proofs can be found in Appendix A. Appendix B contains extensions and additional results.

1. Setup

This paper studies recoverability of the technology of heterogeneous firms given data on the value function of their maximization problems, as well as data on prices or price proxies that alter the maximization problems.

The technology of heterogeneous firms is described by a correspondence $Y : E \rightrightarrows \mathbb{R}^{d_y}$. Each set $Y(e)$ describes the possible input/output (or "netput") vectors that are feasible for a firm of type e. The variable e captures unobservable heterogeneity in productivity. Negative components of $Y(e)$ correspond to net demands by the firm and positive components correspond to net supply. This formulation allows us to treat single output and multi-output firms in a common framework.[8] We require the following conditions.

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 $Y(e)$;

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

(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 rule out infinite profits and ensure that the maximization problems we consider have a solution.[9]

We study the general *restricted profit maximization problem*

$$\pi_r(y_{-z}, p_z, e) = \max_{y_z : (y_{-z}, y_z) \in Y(e)} p_z' y_z \,,$$

where y_{-z} is a vector of restricted or fixed variables, y_z denotes the variables of choice, and p_z is a vector of prices of y_z. The variable of choice y_z is constrained to belong to the convex set $Y_r(y_{-z}, e)$ defined as

$$Y_r(y_{-z}, e) = \left\{ y_z \in \mathbb{R}^{d_{y_z}} \ : \ (y_{-z}, y_z) \in Y(e) \right\} \,.$$

We refer to $Y_r(y_{-z}, \cdot)$ as the restricted production correspondence.[10]

The behavioral restriction of this model is that given y_{-z}, the firm chooses y_z to maximize restricted profits, taking prices p_z as given. In the special case where y_{-z} is not present, this is the usual profit maximization setup. When y_{-z} consists of inputs, this covers revenue maximization. When y_{-z} consists of outputs, this is cost minimization once we interpret negative y_z as inputs and write

$$\max_{y_z \colon (y_{-z}, y_z) \in Y(e)} p_z' y_z = - \min_{y_z \colon (y_{-z}, y_z) \in Y(e)} p_z'(-y_z).$$

We emphasize that throughout, y_{-z} can be a vector, and so we cover cost minimization with multiple inputs, and revenue maximization with multiple outputs.

Overall, we consider firms that are price-taking in the variables of choice y_z, and study a static problem without uncertainty. We note though that in principle the production set $Y(e)$ is general enough to describe paths of production possibilities throughout time, as would arise if there is investment.

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

[10]More formally, it is only a multi-valued mapping because it can be empty for certain combinations of y_{-z} and e.

1.1. Setting and Data

We study identification in settings in which an analyst observes many realizations of certain values of the restricted profit maximization problem as prices vary. In the most general version, we observe noisy measurements of restricted profits, which are the values of the restricted problem. Specifically, we consider the setup

$$\boldsymbol{\pi}_r = \pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e}) + \boldsymbol{\eta} \text{ a.s.},$$

where $\mathbf{y}_{-z}$ is observed,[11] $\boldsymbol{\eta}$ is unobserved measurement error, and $\mathbf{e}$ is unobservable productivity level. For each component of $\mathbf{p}_z$, the analyst either observes the corresponding price, or more generally observes a price proxy $\mathbf{x}_j$ that is linked to the unobserved price by the relationship $\mathbf{p}_{z,j} = g_j(\mathbf{x}_j, \tilde{\mathbf{x}})$, where $\tilde{\mathbf{x}}$ consists of some control variables. We provide further examples and discussion of such proxies in Section 3.

As an example of observables for cost minimization of hospitals (Bilodeau et al., 2000), the analyst observes total cost (possibly measured with error) on variable inputs y_z (labor, supplies, food for patients, drugs, and energy), input prices or input-price proxies, fixed outputs (inpatient car and outpatient visits), and the fixed inputs (number of physicians and capital). We emphasize that we do not need to observe the quantities $\mathbf{y}_z$ of the flexibly chosen variables.[12]

Now we turn to the description of the sources of variation in our setup. Although we do not fully flesh out an equilibrium model incorporating selection we provide an informal discussion of these forces. First, prices vary because of variation across markets. Our results apply when an analyst observes a single firm from each market, and has observations from many markets. Our results also apply when an analyst observes multiple firms in each market. We focus on the former case to simplify presentation, so that we can avoid market-level subscripts.

To further describe why prices can vary, suppose in each market a consumer facing prices p, income m, and with preferences ξ has net demand $y_j^d(p, m, \xi)$ for each of the j flexibly chosen goods. Given a restricted variable y_{-z}, let $y_j(y_{-z}, p_z, e)$ denote the net supply of the flexibly chosen variable j for a firm of type e and facing prices p_z. In each market, market clearing for the j-th good is then written

$$\int y_j^d(p, m, \xi) dF_{\mathbf{m},\boldsymbol{\xi}}(m, \xi) = \int y_j(y_{-z}, p_z, e) dF_{\mathbf{y}_{-z},\mathbf{e}}(y_{-z}, e) + \omega_j,$$

[11] We use bold font for random variables and vectors and regular font for their realizations.

[12] As discussed in the introduction, for additional data sets, see Nerlove (1963), Roberts & Supina (1996), Fabrizio et al. (2007), Foster et al. (2008), Epple et al. (2010), Doraszelski & Jaumandreu (2013), Combes et al. (2017), Albouy & Ehrlich (2018), and Burke et al. (2019).

where $F_{\boldsymbol{m},\boldsymbol{\xi}}$ is the joint distribution over income and preferences of consumers, and $F_{\mathbf{y}_{-z},\mathbf{e}}$ is the joint distributions of restricted variables and productivity. Here, the endowment of good j is denoted ω_j.

In our most general analysis, equilibrium prices can vary across markets due to variation in endowments (ω_j), income and tastes of consumers ($F_{\boldsymbol{m},\boldsymbol{\xi}}$), or variation in productivity or determination of the restricted variables ($F_{\mathbf{y}_{-\mathbf{z}},\mathbf{e}}$). In particular, the determination of restricted variables can vary across markets due to different forms of competition in the restricted variables.

2. Recoverability of Restricted Profit Function

Our ultimate goal is to learn about the production correspondence. We proceed in three steps. In this section, we first identify the restricted profit function (or value function) for heterogeneous firms assuming that the prices are perfectly observed. In Section 3 we show how to apply our analysis to the general case with unobserved prices. In subsequent sections we show how to use information on the restricted profit function to recover features of the production correspondence and describe the most that can be learned concerning counterfactual questions.

Identifying the restricted profit function for heterogeneous firms is challenging. The value function is nonseparable in latent productivity. Both the restricted variables $\mathbf{y}_{-z}$ and prices $\mathbf{p}_z$ may be endogenous. This leads to simultaneity and selection biases. We consider a setting without panel data or instruments. We present a new technique to identify the restricted profit function that addresses these challenges. The key restrictions of the technique are that (i) heterogeneity is one dimensional and allows us to rank firms, and (ii) there are finitely many types of firms.

2.1. Production Monotonicity

It is well-known that the firm problem admits a representative agent, and in principle this observation can be used to recover a *representative agent* restricted profit function. Even a representative agent analysis here is nontrivial because of challenging selection/simultaneity issues discussed previously. Here, we wish to recover not only a representative agent restricted profit function, but also recover the heterogeneous

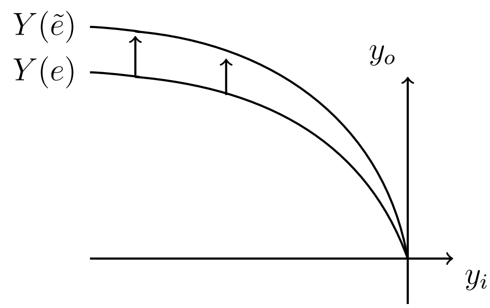

Figure 1 – Nested Production Sets. $\tilde{e} > e$.

structural restricted profit functions. Recovering heterogeneous structural functions allows us to a conduct rich counterfactual analysis concerning how different types of firms are differentially affected by a policy.

To get traction on this problem, we assume firms are rankable in terms of productivity. We think of heterogeneous productivity as an ability to produce more with a given level of inputs (or produce the same output using lower levels of inputs). In other words, the production set of a firm with lower value of e is a subset of the production set with a higher value of e (see Figure 1). Note that $Y_r(y_{-z}, e) \subseteq Y_r(y_{-z}, \tilde{e})$ if and only if $\pi_r(y_{-z}, p_z, e) \leq \pi_r(y_{-z}, p_z, \tilde{e})$ for all p_z. This means that more productive firms have access to a bigger set of production possibilities, and will make more profits or pay lower costs given prices. We formalize this monotonicity by the following ranking assumption on the restricted profit function.

Assumption 1 (Strict Monotonicity)**.** *For every y_{-z}, p_z, e, and $\tilde{e}$ in the support, if $e < \tilde{e}$, then $\pi_r(y_{-z}, p_z, e) < \pi_r(y_{-z}, p_z, \tilde{e})$.*

Assumption 1 is satisfied in many settings. For instance, it is satisfied in a standard single output production function setting with Hicks-neutral productivity. To be more specific, let the single output be y_o and let inputs be l and k, interpreted as labor and capital. Then the set $Y(e)$ is described by tuples $(y_o, -l, -k)$ that satisfy $y_o \leq f(l, k, e)$, where f is the production function. If $f(l, k, e) = A(e)\bar{f}(l, k)$ for some nonnegative, strictly increasing function A, and $\bar{f}$ is always nonnegative strictly convex function, then $f(l, k, e)$ is strictly increasing in e. In this case, $\pi(p, \cdot)$ satisfies Assumption 1.

More generally, the function $f(l, k, e) = A_o(e)\bar{f}(A_l(e)l, A_k(e)k)$ for strictly increasing functions A_o, A_l, and A_k fits into our setup.[13] A more general setup would allow

[13] Li & Sasaki (2017) study a related setup with random coefficients Cobb-Douglas technology,

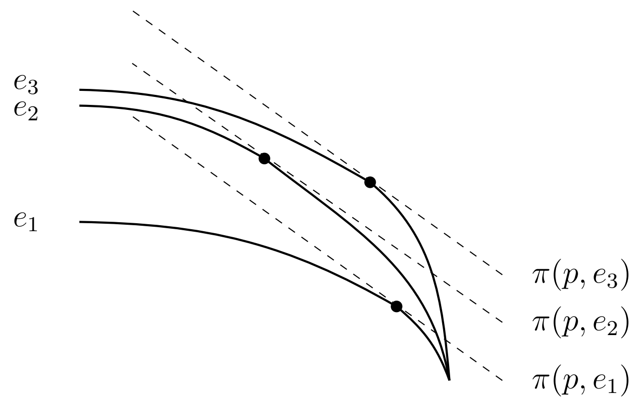

Figure 2 – Nonmonotonic supply.

a different shock to enter A_o, A_l, and A_k (e.g. Doraszelski & Jaumandreu, 2018) and would be outside of our framework. Overall, while Hicks-neutral heterogeneity is a special case of our framework when there is a *single* output, it is considerably more restrictive than needed for the monotonicity assumption to hold.

The assumption that production sets are nested in e is equivalent to the profit function being weakly increasing in e. Thus, value functions are the "right" structural function in which to impose monotonicity if we think of higher productivity as leading to more production possibilities. One may draw the intuition that in general other structural functions are monotone in unobservable heterogeneity. This intuition is false without more structure.

Example 1 (Nonmonotonicity of Inputs/Outputs)**.** Consider the production sets depicted in Figure 2. Each production set is given by $Y(e_i) = \{(y_o, l)' \in \mathbb{R} \times \mathbb{R}_+ : y_o \leq f(l, e_i)\}$, where $f(l, e_1) < f(l, e_2) < f(l, e_3)$ for all $l > 0$. Here, $\pi(p, e_1) < \pi(p, e_2) < \pi(p, e_3)$ for all positive p and Assumption 1 is satisfied. Given the price vector $p = (p_o, p_k)'$ in Figure 2, the optimal levels of inputs and outputs are nonmonotone in productivity since $l^*(p, e_1) < l^*(p, e_3) < l^*(p, e_2)$ and $y_o^*(p, e_1) < y_o^*(p, e_3) < y_o^*(p, e_2)$. For a numerical example see Appendix B.2.

Failures of monotonicity in the optimal choice of input or output have been discussed as well in Pakes (1996, Section 4). Thus, rather than focus on the structural functions describing optimal input/output choices, this paper focuses instead on the restricted profit function, which *is* monotone in a scalar unobservable under the assumption that production sets are nested in e.

imposing that the ratio of random coefficients is a monotone function of a single latent scalar random variable.

2.2. Discrete Heterogeneity and Monotone Selection

With this setup, we consider a new technique to identify the restricted profit function allowing endogeneity. The reason endogeneity is a central concern in such problems is that constraints may be endogenous. For example, in the cost minimization problem, output ($\mathbf{y}_{-z} = \mathbf{y}_o$) is typically a choice variable for the firm. Endogeneity in prices $\mathbf{p}_z$ is also a potential concern if firms with different productivity can choose in which markets to operate (selection into markets). As discussed in Section 1.1 price variation in our setting arises because firms operate in different markets, which have difference endowments or consumer tastes.

The key restriction we impose is that there are finitely many types of firms. We formalize this as follows.

Assumption 2 (Finite Heterogeneity)**.** *$E = \{1, 2, \ldots, d_e\}$ with d_e finite and unknown to the researcher.*

This assumption allows us to identify structural functions without instruments. If instruments are available, continuous heterogeneity can be tackled by existing techniques *provided* there is no measurement error; see for example Appendix B.1. We emphasize that heterogeneity here is in terms of the production types, but due to measurement error in the data we may see continuous distributions of the restricted values, even when we condition on all other observables. In this modeling decision we are close to structural dynamic discrete choice literature that often assumes unobserved discrete heterogeneity that is smoothed out by some continuous idiosyncratic noise (e.g. extreme value distributed preference shock). See, for instance, Arcidiacono & Miller (2011).[14] We are not aware of any identification results that allow for both measurement error and continuous nonseparable structural unobserved heterogeneity in cross sectional data.

We allow rich selection into markets, but impose a monotonicity restriction relating the types of firms that can be present, conditional on certain observables.

Assumption 3 (Monotone Presence)**.**

$$\mathbb{P}\left(\mathbf{e} = e | \mathbf{y}_{-z} = y_{-z}, \mathbf{p}_z = p_z\right) > 0 \implies \mathbb{P}\left(\mathbf{e} = \tilde{e} | \mathbf{y}_{-z} = y_{-z}, \mathbf{p}_z = p_z\right) > 0$$

for all y_{-z}, p_z, e, and $\tilde{e}$ in the support such that $e < \tilde{e}$.

[14]For applications of discrete unobserved heterogeneity in multinomial choice models with random coefficients and panel data estimators see Fox & Gandhi (2016) and Bonhomme & Manresa (2015), respectively.

This means that if we see a firm of type e active in *some* market and producing y_{-z}, then there has to be some market with the same conditioning variables such that any higher type $\tilde{e}$ is present. In principle, this other "market" could be the same market in which e is present. The key restriction is that since we also condition on quantities, we need the higher type to *also* produce the same quantities.

As an example, consider the (unrestricted) profit function, if entry depends on whether a firm obtains nonnegative profits. Specifically,

$$e \text{ enters} \quad \Longleftrightarrow \quad \pi(p, e) \geq 0,$$

where there are no restricted variables. Since we assume monotonicity of π in e, this is a monotone threshold rule, and satisfies Assumption 3.

Assumption 3 is considerably more general than a one-sided selection rule. Importantly, it is only about the *support* of e conditional on some other variables. The reason we require this is that while reasonable selection rules into *markets* may result in a one-sided threshold rule, here we also need to allow selection into the quantities of the restricted variables $\mathbf{y}_{-z}$. For example, as e increases the optimal quantity of the restricted variables may change. Assumption 3 allows this and is satisfied if, for example, there are other unobserved variables that shift the optimal choice of restricted variables y_{-z} (e.g. unobserved prices of the restricted variables).

2.3. Identification

We now turn to identification of the restricted profit function. First, recall that we observe potentially mismeasured restricted profits:

$$\boldsymbol{\pi}_r = \pi_r(\mathbf{y}_{-r}, \mathbf{p}_{-z}, \mathbf{e}) + \boldsymbol{\eta}.$$

Let $\Delta\pi_r(y_{-z}, p_z, e) = \pi_r(y_{-z}, p_z, e) - \pi_r(y_{-z}, p_z, e-1)$ denote the restricted profit difference between firms with adjacent productivity. We impose the following assumption on the measurement error $\boldsymbol{\eta}$.

Assumption 4. *(i) $\boldsymbol{\eta}$ is independent of $\mathbf{y}_{-z}$, $\mathbf{p}_z$, and $\mathbf{e}$, mean zero, and satisfies $\mathbb{P}(|\boldsymbol{\eta}| \leq K/2) = 1$ for some $K < \infty$;*

(ii) (Separatedness) There exists a known (y^*_{-z}, p^*_z, e^*) in their support such that

$$K < \begin{cases} \Delta\pi_r(y^*_{-z}, p^*_z, e^*+1), & \textit{if } e^* = 1, \\ \Delta\pi_r(y^*_{-z}, p^*_z, e^*), & \textit{if } e^* = d_e, \\ \min\left\{\Delta\pi_r(y^*_{-z}, p^*_z, e^*+1),\ \Delta\pi_r(y^*_{-z}, p^*_z, e^*)\right\}, & \textit{otherwise.} \end{cases}$$

We note that multiplicative measurement error can be handled by similar independence and separatedness assumptions.[15]

Assumption 4(i) means that the measurement error is classical. It also imposes a location normalization on the boundedly-supported measurement error. Assumption 4(ii) is more substantial. It assumes that we can find a firm with a particular productivity such that after conditioning on observables the measurement error does not break the ranking imposed by Assumption 1. Note that Assumption 4(ii) has to be imposed on one triplet (y^*_{-z}, p^*_z, e^*) only. Thus, in general the measurement error may completely change the ranking of restricted profits. A simple sufficient condition for Assumption 4(ii) that uses shape restrictions of the restricted profit function is stated in the following result.

Lemma 1 (Rich Support)**.** *If Assumption 1 holds and there exist y^*_{-z} and p^*_z such that $\cup_{\lambda>0}\{\lambda p^*_z\}$ is in the support of $\mathbf{p}_z$ conditional on $\mathbf{y}_{-z} = y^*_{-z}$, then Assumption 4(ii) is satisfied.*

This exploits homogeneity in prices, i.e. $\pi_r(y^*_{-z}, \lambda p^*_z, e) = \lambda\pi_r(y^*_{-z}, p^*_z, e)$ for all e and $\lambda > 0$. The idea behind Lemma 1 is that although the difference between profits evaluated at a particular price may not be big enough to offset the effect of the measurement error (e.g. $\Delta\pi_r(y^*_{-z}, p^*_z, e^*+1) \leq K$) by exploiting homogeneity we always can find λ^* big enough such that

$$\Delta\pi_r(y^*_{-z}, \lambda^* p^*_z, e^*+1) = \lambda^*\Delta\pi_r(y^*_{-z}, p^*_z, e^*+1) > K.$$

The conditions of Lemma 1 guarantee that an extreme price $\lambda^* p^*_z$ can be found in the support for every finite K. Thus, the support of of prices does not have to be unbounded, just sufficiently large relative to the initial difference.

Now we can state our main identification result for the restricted profit function.

Theorem 1. *Suppose Assumptions 1-4 hold. Then π_r is identified from $F_{\boldsymbol{\pi}_r|\mathbf{y}_{-z},\mathbf{p}_z}$ over the joint support of $\mathbf{y}_{-z}$, $\mathbf{p}_z$, and $\mathbf{e}$.*

[15] The bounded support and separatedness conditions in Assumption 4 can be relaxed using results in Schennach (2016) if one has access to repeated cross sections.

Here, we may not be able to identify the structural restricted profit function for certain arguments outside of the support. This is particularly relevant for low types; there many be many combinations of prices and quantities such that low types do not produce either because it is infeasible for them or unprofitable.

Importantly, Theorem 1 only imposes a mild restriction on the stochastic dependence between unobservable heterogeneity $\mathbf{e}$ and observed $\mathbf{y}_{-z}$ and $\mathbf{p}_z$. In particular, in cost minimization settings, the output level and input prices can be related to the distribution of productivity in flexible ways. What is key is the monotonicity restriction on selection into markets described in Assumption 3.

The intuition behind Theorem 1 is that without restricting the dependence structure, monotonicity in the restricted profit function implies that firms always can be ranked. The assumption of the discrete heterogeneity allows us to match firms with the same ranking across different markets, and thereby construct the restricted profit function.

Theorem 1 can be used to weaken assumptions usually made in analysis of restricted profit maximizing behavior. For instance, with cost minimization, Bilodeau et al. (2000) focuses on a parametric setup with additively separable heterogeneity and assumes that fixed variables are exogenous. While working with the same observables, our methodology does not require parametric restrictions, and does not assume exogeneity.

3. Unobservable Prices and Proxies

In Section 2 we showed how to identify the restricted profit function when the entire vector of prices of flexibly chosen variables, p_z, is observed. 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). However, the researcher may have access to some observable proxies that are informative about unobservable prices. For example, the rental rate of capital may be linked to market-specific characteristics such as short-term and long-term interest rates. Wages may be linked to the unemployment level or aggregate labor supply. De Loecker et al. (2016) uses output price, market shares, product dummies, firm location, and export status as proxies for unobservable input prices. In the housing market, an analyst may use location as a price proxy for a house as in Combes et al. (2017).[16]

This section studies how to identify the function linking prices proxies to unobserved prices through

$$p_j = g_j(x_j, \tilde{x}),$$

where g_j is an unknown function and p_j is a component of p_z. We assume that every price has its own *excluded* proxy x_j, which is a proxy that affects its own price and does not affect any other prices. The vector of common proxies $\tilde{x}$ may include common market characteristics such as size of the market or other macroeconomic characteristics. Importantly, since g_j is fully nonparametric, $\tilde{x}$ can include categorical variables such as location (e.g. country or state) and time (e.g. month or year) identifiers. The above formulation covers the case when price is observed. In that case $g_j(x_j, \tilde{x}) = x_j$, where x_j is the price of y_j. To simplify the exposition we drop $\tilde{x}$ from the notation, and analysis may be interpreted conditional on $\tilde{x}$. For instance, we write $g_j(x_j)$ instead of $g_j(x_j, \tilde{x})$.

Note that we assume prices are not a function of e or any other unobservables. In our setup prices vary across markets but are constant within a given market. Price-taking behavior implies that prices can be a function of the distribution of $\mathbf{e}$ in a market, but not the firm-specific productivity e. More generally, prices are determined by market clearing conditions, where preferences and productivity are integrated out, making g_j a function of market characteristics (x). For specific examples of a structure with a function g mapping market characteristics to prices see Sections 3.2 and 3.3.

We first present an informal outline how to identify g when one observes profits. We denote $x = (x_j)_{j=1,\ldots,d_{yz}} \in X$ and $g(x) = (g_j(x_j))_{j=1,\ldots,d_{yz}}$. Profits are given by $\pi(g(x), e)$. If the function g were known, we could identify π directly by previous arguments. What remains is to identify g. 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

$$\sum_{j=1}^{d_y} \partial_{p_j} \pi(p, e) p_j = \pi(p, e)\,.$$ [17]

Replacing prices with price proxies, we obtain

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

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

[17] Recall that we work with the unrestricted profit function for notational simplicity, but the restricted profit function is also homogeneous of degree 1 in prices.

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 (1) we obtain

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

Assume for now that $\tilde{\pi}(\cdot,e)$ is identified. Thus the only unknowns involve g. By varying x, holding everything else fixed, Equation 2 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 one, then we may directly apply Equation 2 to learn about g_j.

To formalize this, we impose location/scale conditions and some regularity conditions on g.

Assumption 5. *(i) $g_{d_{y_z}}(x_{d_{y_z}}) = x_{d_{y_z}}$ for all $x_{d_{y_z}}$, i.e. the price of the d_{y_z} flexibly chosen variable is observed;*

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

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

(iv) $g_j(\cdot)$ is differentiable on the interior of X_j, and the set

$$\left\{x_j \in X_j \ : \ \partial_{x_j}g(x_j) = 0\right\}$$

has Lebesgue measure zero for every j.

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

We now turn to our rank condition. This condition ensures that the system of equations generated from (2) has sufficient variation to recover terms such as $g_j(x_j)/\partial_{x_j}g_j(x_j)$.

Definition 2. We say that $h : \prod_{j=1}^{d_{yz}} X_j \to \mathbb{R}$ satisfies the rank condition at a point $x_{-d_{yz}} \in \mathbb{R}^{d_{yz}-1}$ if there exists a collection of $\{x_{d_{yz},l}\}_{l=1}^{d_{yz}-1}$ such that

(i) $x_l^* = (x'_{-d_{yz}}, x_{d_{yz},l})' \in \prod_{j=1}^{d_{yz}} X_j$;

(ii) The square matrix

$$\begin{bmatrix} \partial_{x_1} h(x_1^*) & \dots & \partial_{x_{d_{yz}-1}} h(x_1^*) \\ \partial_{x_1} h(x_2^*) & \dots & \partial_{x_{d_{yz}-1}} h(x_2^*) \\ \dots & \dots & \dots \\ \partial_{x_1} h(x_{d_{yz}-1}^*) & \dots & \partial_{x_{d_{yz}-1}} h(x_{d_{yz}-1}^*) \end{bmatrix}$$

is nonsingular.

We will apply this rank condition to $\tilde{\pi}$ in place of h. 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 $\tilde{\pi}$ may equivalently be interpreted as a rank condition involving the supply function for the goods as well as certain derivatives of g (i.e., variation in observed prices should induce enough variation in supply of goods with unobserved prices).

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

Theorem 2. *Suppose Assumption 5 holds. Then g is identified over the support of $\mathbf{x}$ if for some y_{-z}^* one of the following conditions holds:*

(i) $\tilde{\pi}(x, e) = \pi_r(y_{-z}^*, g(x), e)$ is identified for each x and e. In addition, for every $x_{-d_{yz}}$, there exists $e^* \in [0, 1]$ such that $\tilde{\pi}(\cdot, e^*)$ satisfies the rank condition at x_{-d_y};

(ii) $\boldsymbol{\pi}_r^* = \pi_r(\mathbf{y}_{-z}, g(\mathbf{x}), \mathbf{e}) + \boldsymbol{\eta}$ a.s., where $\boldsymbol{\pi}_r^*$ is observed, π_r is homogeneous of degree 1 in the second argument; $F_{\mathbf{e}|\mathbf{p}_z, \mathbf{y}_{-z}}(e|p_z, y_{-z}^*)$ is homogeneous of degree 0 in p_z; $\boldsymbol{\eta}$ satisfies Assumption 4(i); and $\mathbb{E}\left[\boldsymbol{\pi}_r^* | \mathbf{x} = \cdot, \mathbf{y}_{-z} = y_{-z}^*\right]$ satisfies the rank condition at every x_{-d_y}.

To interpret (i), recall that Theorem 1 provides conditions under which $\tilde{\pi}$ is identified from the conditional distribution of $\pi_r(\mathbf{y}_{-z}, g(\mathbf{x}), \mathbf{e})$ conditional $\mathbf{x}$ and $\mathbf{y}_{-z}$. To apply those results one just needs to replace $\mathbf{p}_z$ by $\mathbf{x}$. Here we clarify that given *some* way to identify a structural function of the form of $\tilde{\pi}$, we can identify g.

Part (ii) requires different structure. We state it because our technique is new and this result may be of independent interest. In particular, it does *not* require discreteness of $\mathbf{e}$ and monotonicity of the restricted profit function in the unobservable e, and thus applies to more general forms of heterogeneity than we consider. However, such generality comes with the cost of assuming homogeneity of degree 0 in prices of the conditional distribution of productivity conditional on prices and quantities, which was not required by part (i). Homogeneity of the distribution function in prices means that the distribution of productivity in the market depends only on relative prices. This trivially happens if productivity is independent from flexible prices conditional on fixed quantities. It also may naturally happen in profit maximizing environments where entry decisions are driven by the threshold rule where firms with nonnegative profits enter. Since the profit function is homogeneous of degree 1 in prices, we can deduce that the entry decision is only determined by the direction of the price vector, not by its norm: $\pi(p, e) \geq 0$ if and only if $\pi(p/\|p\|, e) \geq 0$.

To further interpret the rank condition, we study it in two parametric examples in Appendix B.3. There we show that the rank condition can be satisfied for the Diewert (1973) profit function, but can fail for every possible parameter value with Cobb-Douglas technology.

We conclude this section by noting it is straightforward to generalize our technique to a homogeneous function of any degree $\alpha \geq 0$ (e.g. the supply function) since the main identifying equation (2) can be rewritten as

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

3.1. Other Observables

Theorem 1 applies when (only) values of the restricted profit function, restricted variables, and price proxies are observed. When other variables are observed, it can be adapted to handle other settings.

To illustrate this suppose that p_1 is not observed, does not have a proxy, and does not vary across markets. Suppose further that for each good $j \geq 3$, x_j is a price proxy conditional on p_2. That is, $p_j = g_j(x_j, p_2)$ for $j \geq 3$. The fact that p_2 is in this function means it violates our previous exclusion restriction and so Theorem 1 cannot directly be applied. Nonetheless, we can adapt the technique to cover this case.

To see this, recall Euler's homogeneous function theorem states

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

while Hotelling's lemma reads

$$y_j(p,e) = \partial_{p_j} \pi(p,e).$$

These imply

$$\sum_{j=3}^{d_y} \partial_{p_j} \pi(p,e) p_j = \pi(p,e) - p_1 y_1(p,e) - p_2 y_2(p,e)\,.$$

Moreover, for $j \neq 2$

$$\partial_{p_j} \pi(g(x,p_2),e) \partial_{x_j} g_j(x_j,p_2) = \partial_{x_j} \tilde{\pi}(x,p_2,e)\,.$$

Hence we obtain

$$\sum_{j=3}^{d_y} \partial_{x_j} \tilde{\pi}(x,p_2,e) \frac{g_j(x_j,p_2)}{\partial_{x_j} g_j(x_j,p_2)} = \tilde{\pi}(x,p_2,e) - \tilde{r}(x,p_2,e)\,, \tag{3}$$

where

$$\tilde{r}(x,p_2,e) = p_1 y_1(p_1,p_2,g(x,p_2),e) + p_2 y_2(p_1,p_2,g(x,p_2),e)$$

is the contribution of goods 1 and 2 to profits. The difference from Equation 2 is that in order to build a system of ordinary differential equations that identifies g we need to identify

$$\tilde{\pi}(x,p_2,e) - \tilde{r}(x,p_2,e)$$

as well. If $\tilde{\pi}(x,p_2,e)$ and $\tilde{r}(x,p_2,e)$ can be identified, we are done. More generally, it is not necessary to identify the heterogeneous structural functions separately. Instead, it is enough to identify their aggregate versions, because homogeneity aggregates (recall Theorem 1(ii) and the subsequent discussion). In sum, it is possible to identify prices that vary across markets even if there are prices that are unobserved but are fixed across markets (like p_1) and there are observed prices that do not satisfy the exclusion restriction (like p_2). We will use this insight about prices that are unobserved but fixed across markets in Section 3.3 to show how our approach can be used to generalize Epple et al. (2010).

3.2. Aggregate Quantities as Proxies

This section provides a foundation for use of aggregate quantities in a market as price proxies. This applies even if the firm-level quantities are not observed. We show this for an equilibrium model in which variation in aggregate quantities or (possibly unobserved) prices occurs due to variation in endowments.

Consumers have preferences over quantities y described by the utility function $u(y, \xi)$, where ξ represents unobservable heterogeneity. The key assumption we make is that preferences are separable in the goods for which we require price proxies.

Assumption 6. *Preferences are quasilinear and additively separable, i.e.*

$$u(y, \xi) = \sum_{j=1}^{d_y} u_j(y_j, \xi) + y_{d_y+1}.$$

The budget constraint takes the form

$$\sum_{j=1}^{d_y} p_j y_j + y_{d_y+1} \leq m,$$

where m is income and the choice of y_{d_y+1} can be negative.

Allowing y_{d_y+1} to be negative (or assuming m is high enough) is standard so that the model does not have income effects.

Suppose further that the quantity demanded is unique (for almost every ξ given the distribution of heterogeneity ξ), so that we can write

$$y^d(p, \xi) = \arg\max_y \sum_{j=1}^{d_y} u_j(y, \xi) - \sum_{j=1}^{d_y} p_j y_j.$$

Because of the separable preferences, the quantity demanded of good j can be written as $y_j^d(p_j, \xi)$. This is weakly decreasing in prices by standard arguments, and weak monotonicity is preserved under expectations.

Building on this, we have the following result.

Lemma 2. *Let Assumption 6 hold. In addition, assume the distribution of preferences $F_{\boldsymbol{\xi}}$ is the same across markets, and define the aggregate consumer demand*

$$\overline{x}_j(p_j) = \int y_j^d(p_j, \xi) dF_{\boldsymbol{\xi}}.$$

If $\overline{x}_j$ is strictly increasing, then $\overline{x}_j$ is a price proxy for p_j. That is,

$$p_j = g_j(\overline{x}_j)$$

for some function g_j.

In particular, g_j is the inverse demand for good j. Importantly, this argument states that market-level aggregate consumer demand of good j is a valid price proxy for good j, *provided* the aggregate quantity is constructed using the same measure across markets. Thus, the distribution of unobservable demand heterogeneity has to be the same across markets in order to apply this type of proxy. Importantly, this approach does not rule out selection of firms into markets.

3.3. Value as Proxy

While the previous subsection showed that restrictions on the *demand side* allow us to use aggregate consumer demand as price proxies, we now show an example building on Epple et al. (2010) in which restrictions on the *supply side* allow us to use certain values as price proxies.

Epple et al. (2010) consider the production of housing in which the analyst summarizes all goods and services provided by a house per-acre (i.e., per unit of land) as a single output y_o. 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. Instead, the analyst observes total revenue of selling a house $p_o y_o$, where p_o is the price of housing, and the price of land p_l. Variation in these observables is driven by market variation.[18]

In contrast to Epple et al. (2010), who worked with a representative firm, we study identification in the presence of heterogeneity. Assume that for each $e \in E$, production of housing satisfies constant returns to scale in inputs, so that we can write

$$y_o = f(m, e),$$

where f is the production function per-acre and m are materials per-acre used in construction. The production set associated with this production function is $Y(e) =$

[18] Epple et al. (2010) observes the exact value of the house and the land quantity. Since they assume constant returns to scale, the problem can be reformulated relative to units of land.

$\{(y_o, -m) : y_o \le f(m, e)\}$. Then the profit function per-acre is

$$\pi(p_o, p_m, p_l, e) = \max_{(y_o, -m) \in Y(e)} p_o y_o - p_m m - p_l.$$

Epple et al. (2010) assume that p_m is the same across markets and equals 1 since p_m is unobserved. We will make the same assumption and drop p_m from the notation. As a result there is no variation in p_m and we cannot use homogeneity of the profit function in all prices. Nonetheless, as we have explained in Section 3.1, using Hotelling's lemma we can still leverage the intuition behind Theorem 2.[19]

Since we consider per-acre production, the optimal $y_o(p_o, e)$ and $m(p_o, e)$ will not depend on p_l. Hence, the value of housing $v(p_o, e) = p_o y_o(p_o, e)$ and the average value of housing in a market $\overline{v}(p_o) = \int v(p_o, e) dF_{\mathbf{e}}(e)$ do not depend on price of land p_l. Since $y_o(p_o, e)$ is monotone in p_o, then $\overline{v}(p_o)$ is also monotone in p_o. Importantly, $\overline{v}$ is identified when we observe total revenue $p_o y_o$.

Lemma 3. *Suppose the distribution of firm productivity $F_{\mathbf{e}}$ is the same across markets and the other assumptions of this section hold. If $\overline{v}(p_o)$ is strictly increasing in p_o, then average value of housing per market $\overline{v}$ is a price proxy, i.e. there exists a function $\tilde{g}$ such that*

$$p_o = g(\overline{v}).$$

This equation is analogous to Equation 6 in Epple et al. (2010) if we interpret their results as a representative agent analysis. Moreover, p_l is an observable price that does not affect p_o.

We will use a generalization of the zero-profit assumption from Epple et al. (2010). While they assume a single type of firm, which attains zero profits, we assume that profits are zero *on average* in a given market[20]:

$$\int \pi(p_o, p_l, e) dF_{\mathbf{e}}(e) = p_o \overline{y}_o(p_o) - \overline{m}(p_o) - p_l = 0,$$

where $\overline{y}_o$ and $\overline{m}$ are the aggregate output per-acre and the aggregate demand for materials per-acre in a given market. Since p_l and $\overline{v}$ are observed, the equilibrium assumption nonparametrically recovers a reduced form revenue function from produc-

[19]In the notation of Section 3.1 the price of land p_m is p_1.

[20]Melitz & Redding (2014) show that free-entry and constant returns of scale imply that ex-ante expected profits are zero, net of entry cost. Here we can assume entry cost is zero. In equilibrium, firms will have zero-profits on average *just before* firms with negative profits leave the market.

tion minus materials cost (recall that $p_m = 1$)

$$p_l = \tilde{\pi}(\overline{v}) := g(\overline{v})\overline{y}_o(g(\overline{v})) - \overline{m}(g(\overline{v})).$$

Moreover, since $g(\overline{v})\overline{y}_o(g(\overline{v})) = \overline{v}$ by definition, we also identify material costs

$$\tilde{r}(\overline{v}) = -\overline{m}(g(\overline{v})).$$

Similar to Equation 3 we identify function g since we identify $\tilde{\pi}(\overline{v})$ and $\tilde{\pi}(\overline{v}) - \tilde{r}(\overline{v})$. In particular, g will solve the following differential equation, which is implied by Equation 3:

$$\frac{\partial_{\overline{v}} g(\overline{v})}{g(\overline{v})} = \frac{\partial_{\overline{v}} \tilde{\pi}(\overline{v})}{\tilde{\pi}(\overline{v}) - \tilde{r}(\overline{v})} = \frac{\partial_{\overline{v}} \tilde{\pi}(\overline{v})}{\overline{v}}.$$

Knowing g we can identify $y_o(p_o, e)$ for different levels of heterogeneity since the observed v is equal to $g(\overline{v})y_o(g(\overline{v}), e)$. Thus, our approach generalizes Epple et al. (2010) to allow for unobserved heterogeneity in productivity.

4. Identification of the Production Correspondence

In Section 2 we showed how to identify the restricted profit function allowing endogenous entry and correlation between fixed quantities and productivity, without requiring instruments. Section 3 extends this result to settings when some prices are not observed but the analyst has price proxies, and provides examples of such proxies.

We now focus on how any of these identification results (or results not presented here) for the restricted profit function can be used to identify the primitive object of interest: the production correspondence. For the sake of notational simplicity from now on, we focus on the profit function though the results can be adapted to the restricted profit function by conditioning.

Recall that we can identify the profit function $\pi(p, \cdot)$ only over the support of prices (or more generally over the support of $g(x)$, where x is the vector of price proxies). The support of prices may consist of all nonnegative numbers, or may be much smaller, i.e. finite. We present a sharp identification result for the production correspondence that covers both cases.

First, we note that $\pi(\cdot, e)$ is homogeneous of degree 1 in prices. It is also convex

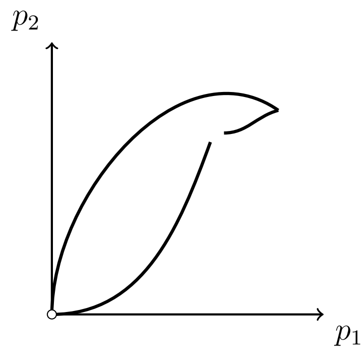

Figure 3 – The set $P(e)$ (depicted by black curve) satisfies Assumption 7 and has an empty interior. Dots represent "holes" in the support. Thus, $P(e)$ is not a connected set.

in prices, hence continuous. These features lead to consideration of the following richness assumption, which ensures $Y(\cdot)$ may be recovered uniquely. Let $P(e)$ denote the conditional support of $\mathbf{p}$ conditional on $\mathbf{e} = e$ (if $\mathbf{p}$ and $\mathbf{e}$ are independent, then $P(e)$ does not vary with e).

Assumption 7.

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

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

The set

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

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.[21] Assumption 7 means that after we consider the implications of homogeneity and continuity, it is as if we have full variation in prices. Figure 3 is an example of a set satisfying this assumption. Another example is the Cartesian product of all natural numbers, $P(e) = \{1, 2, \ldots\}^{d_y}$. Thus, Assumption 7 does not impose that the support of $\mathbf{p}$ contains an open ball.

[21] 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.4.

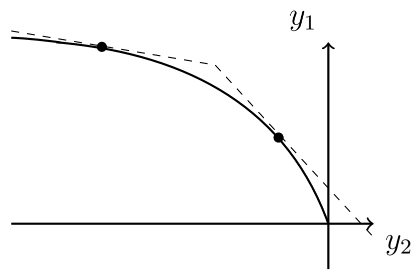

Figure 4 – $\tilde{Y}(e)$ and $Y'(e)$ for $d_y = 2$ and $P(e) = \{p^*, p^{**}\}$. $\tilde{Y}(e)$ is the area under the dashed lines. $Y'(e)$ is the area under the solid curve. Dashed lines correspond to two hyperplanes $p^{*\prime}y = \pi(p^*, e)$ and $p^{**\prime}y = \pi(p^{**}, e)$. They are tangential to the solid curve.

Theorem 3. *Let $\pi(p, e)$ be identified by some previous argument, over the set $p \in P(e)$ for all e. Moreover, let $\tilde{Y}(\cdot)$ be defined via*

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

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.[22]

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

(iii) If Assumption 7 holds, then $\tilde{Y}(\cdot)$ is the only production correspondence that can generate the data.

Parts (i) and (ii) of Theorem 3 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.[23] However, Varian (1984) works

[22]By generate the data we mean that the profit function induced by $\tilde{Y}$ agrees with the identified profit function $\pi(p, e)$ for all $e \in E$ and $p \in P(e)$.

[23]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).

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 3(ii) establishes that $\tilde{Y}(\cdot)$ is the envelope of all production correspondences that can generate the data (see Figure 4). 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)).[24]

Theorem 3(iii) is related to classic work on the identification of a deterministic production set from a deterministic profit function.[25] 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 ($P(e) = \mathbb{R}^{d_y}_{++}$). Theorem 3 clarifies the variation in prices sufficient for nonparametric identification of production sets. We note that while Assumption 7 is sufficient for point identification of Y, it is not necessary as illustrated in Appendix B.4.

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 3. 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.

[24]To see this, suppose that a firm of type $e \in E$ has 2-dimensional output/input set, prices are a constant vector $P(e) = \{(1,1)'\}$, and profits at that price are given by $\pi((1,1)', e) = 0$. Then the set $\tilde{Y}(e)$ is $\left\{y \in \mathbb{R}^2 : y_1 + y_2 \leq 0\right\}$. 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, e.g. Kreps (2012), Proposition 9.7. Note from part (iii), when Assumption 7 holds it follows that $\tilde{Y}$ is a production correspondence, and thus satisfies the recession cone property.

[25]See e.g. Kreps (2012), Corollary 9.18 for a textbook result.

5. Sharp Counterfactual Bounds

Theorem 3 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 $\mathrm{cl}\left(\bigcup_{\lambda>0}\{\lambda p : p \in P(e)\}\right)$, we can associate the conditional support $P(e)$ (of prices condition on e) with the set where $\pi(\cdot, e)$ is identified. That is why, for notational simplicity and in this section only, we assume that $P(e)$ is a closed subset of the unit sphere $\mathbb{S}^{d_y-1}$ for all e, and we consider counterfactual prices with norm normalized to 1.

We first present a result characterizing quantities consistent with profit maximization. Theorem 3(ii) is the basis for the following proposition.

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

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

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 1 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.

We can adapt Proposition 1 to answer counterfactual questions by considering a hypothetical tuple (p^c, y_{p^c}) of prices and quantities. If Proposition 1 applies with these additional counterfactual values, then they are feasible given the theory.[26] In more

[26] Varian (1982, 1984) has exploited the close connections between empirical content, recoverability

detail, we present bounds on counterfactual objects, potentially with additional restrictions. The upper bound on a functional C given a restriction s and heterogeneity level e is given by

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

The lower bound is given by

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

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 $\overline{C}_s(\mathbf{e})$ and $\underline{C}_s(\mathbf{e})$. 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 2 (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 $s(p^{\mathrm{c}}, y_{p^{\mathrm{c}}}) = p^{\mathrm{c}} - \overline{p}^{\mathrm{c}}$. Then the problem can be simplified to get

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

where $\tilde{Y}(e)$ is the envelope of all production possibility sets consistent with the data defined in Theorem 3. The above bounds are sharp in the following sense: if $\overline{C}_s(e)$

of structural functions, and counterfactuals. 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).

is finite, then it is feasible, i.e. there exists a production set that can generate $\overline{C}_s(e)$. If $\overline{C}_s(e)$ is not finite, then for any finite level K there exists a production set that can generate $C(p^c, y_{p^c}) > K$. Analogous statements hold for the lower bounds $\underline{C}_s(e)$. Recall that we assume the support of prices $P(e)$ 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 $\overline{p}^c$, one can bound counterfactual profits at price $\overline{p}^c/\|\overline{p}^c\|$ and then multiply the upper and lower bounds by $\|\overline{p}^c\|$.

Example 3 (Quantity bounds for a counterfactual price)**.** Suppose that we are interested in the upper and lower bounds for $u'y_{p^c}$ for a given counterfactual price $\overline{p}^c$, where u is a 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 4 (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 $s(p^c, y_{p^c}) = y_{1,p^c} - \overline{y}_1^c$.[27] 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.[28] An additional question a regulator might ask is which types of firms could still be profitable. This can be addressed by studying functions $\overline{C}_s(\cdot)$ and $\underline{C}_s(\cdot)$ as e varies. Note that the constraints s are general, and inequality constraints may be incorporated as well by using indicator functions.

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

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

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

6. Estimation and Consistency

While the previous identification results describe how to identify the profit or restricted profit function, this paper does not study estimation of the restricted profit function. Instead, we present a result that links *any* estimator of the restricted profit function to an induced estimator of the corresponding set. As in previous section, for notational convenience we work with the profit function, though the analysis applies to the restricted profit function by conditioning.

We now 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 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 $P(e) = P = \mathbb{R}^{d_y}_{++}$ (we assume Assumption 7). 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 3. 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 5. Suppose that $d_y = 2$ and for some $e \in E$,

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

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 infinite for every finite $m \in \mathbb{N}$.

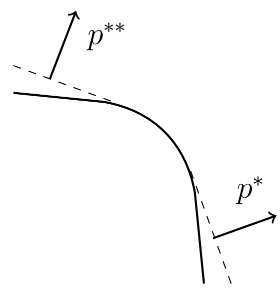

Figure 5 – $Y(e)$ and $Y_{\bar{P}}(e)$ for $d_y = 2$ and $\bar{P} = \{p \in P \ : \ \delta \leq p_2/p_1 \leq 1/\delta, \|p\| \leq 1\}$, $0 < \delta < 1$. $Y(e)$ is the area under the solid curve. $Y_{\bar{P}}(e)$ is the area under the dashed lines. Dashed lines correspond to two hyperplanes $p^{*\prime}y = \pi(p^*, e)$ 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 $p_2^{**}/p_1^{**} = 1/\delta$.

Example 5 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

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

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 5).

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 4. *Maintain the assumption that $\pi(\cdot, e)$ is homogeneous of degree 1 and* *convex.*[29] *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 4 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 *unbounded* sets.[30] 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 2. *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 2 is relevant. It is computationally feasible to impose convexity for certain functional forms of π, which allows one to invoke the stronger Theorem 4. See Appendix B.5 for an example with the flexible functional form of Diewert (1973).

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

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

7. Conclusion

In this paper we have provided an update to classical duality theory in order to identify heterogeneous production sets in the presence of endogeneity, measurement error, omitted prices, and unobservable quantities. Our framework's main strength is to unpack rich heterogeneity as well as rich substitution/complementarity patterns with market level variation, using values of optimization problems. We achieve this by exploiting all shape constraints imposed by the economic environment we consider. This includes a key restriction that firms can be ranked in terms of productivity, and there are finitely many types of firms. Our identification results are constructive and can be applied in many available data sets.

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*.

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*.

Arcidiacono, P. & Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6), 1823–1867.

Beresteanu, A., Molchanov, I., & Molinari, F. (2011). Sharp identification regions in models with convex moment predictions. *Econometrica*, 79(6), 1785–1821.

Beresteanu, A. & Molinari, F. (2008). Asymptotic properties for a class of partially identified models. *Econometrica*, 76(4), 763–814.

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.

Bonhomme, S. & Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3), 1147–1184.

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

Burke, M., Bergquist, L. F., & Miguel, E. (2019). Sell low and buy high: arbitrage and local price effects in kenyan markets. *The Quarterly Journal of Economics*, 134(2), 785–842.

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.

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.

Doraszelski, U. & Jaumandreu, J. (2018). Measuring the bias of technological change. *Journal of Political Economy*, 126(3), 1027–1084.

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.

Fabrizio, K. R., Rose, N. L., & Wolfram, C. D. (2007). Do markets reduce costs? assessing the impact of regulatory restructuring on us electric generation efficiency. *American Economic Review*, 97(4), 1250–1277.

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: selection on productivity or profitability? *American Economic Review*, 98(1), 394–425.

Fox, J. T. & Gandhi, A. (2016). Nonparametric identification and estimation of random coefficients in multinomial choice models. *The RAND Journal of Economics*, 47(1), 118–139.

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. (2016). A dual approach to inference for partially identified econometric models. *Journal of econometrics*, 192(1), 269–290.

Kaido, H., Molinari, F., & Stoye, J. (2019). Confidence intervals for projections of partially identified parameters. *Econometrica*, 87(4), 1397–1432.

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.

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

Li, T. & Sasaki, Y. (2017). Constructive identification of heterogeneous elasticities in the cobb-douglas production function. *arXiv preprint arXiv:1711.10031*.

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

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

Melitz, M. J. & Redding, S. J. (2014). Heterogeneous firms and trade. In *Handbook of international economics*, volume 4 (pp. 1–54). Elsevier.

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.

Pakes, A. (1996). Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interaction. In *Advances in Econometrics, Sixth World Congress*, volume 2 (pp. 171–259).: by C. Sims.

Roberts, M. J. & Supina, D. (1996). Output price, markups, and producer size. *European Economic Review*, 40(3-5), 909–921.

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

Schennach, S. M. (2016). Recent advances in the measurement error literature. *Annual Review of Economics*, 8, 341–377.

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 1

Fix y^*_{-z} and p^*_z. By homogeneity of degree 1 of the restricted profit function in prices and Assumption 1,

$$\Delta\pi_r(y^*_{-z}, \lambda p^*_z, e) = \lambda\Delta\pi_r(y^*_{-z}, p^*_z, e) > 0$$

for every e and $\lambda > 0$. Since $\cup_{\lambda>0}\{\lambda p^*_z\}$ in the conditional support, we always can find λ large enough and e^* such that Assumption 4(ii) is satisfied.

A.2. Proof of Theorem 1

Under Assumption 4 we can find an interval $[a, b]$ in the support of $\boldsymbol{\pi}_r$ conditional on $\mathbf{y}_{-z} = y^*_{-z}$, $\mathbf{e} = e^*$, and $\mathbf{p}_z = p^*_z$ such that

$$\mathbb{P}\left(a \leq \pi_r(y^*_{-z}, p^*_z, e^*) + \boldsymbol{\eta} \leq b\right) = 1$$

and

$$\mathbb{P}\left(a \leq \pi_r(y^*_{-z}, p^*_z, e) + \boldsymbol{\eta} \leq b\right) = 0$$

for any $e \neq e^*$. Hence, we identify

$$\pi_r(y^*_{-z}, p^*_z, e^*) = \mathbb{E}\left[\boldsymbol{\pi}_r | a \leq \boldsymbol{\pi}_r \leq b, \mathbf{y}_{-z} = y^*_{-z}, \mathbf{e} = e^*, \mathbf{p}_z = p^*_z\right],$$

where we leverage that $\boldsymbol{\eta}$ has mean zero even after conditioning.

Thus, we can also recover the distribution of $\boldsymbol{\eta}$ by subtracting the identified $\pi_r(y^*_{-z}, p^*_z, e^*)$ from the known distribution of $\boldsymbol{\pi}_r | a \leq \boldsymbol{\pi}_r \leq b, \mathbf{y}_{-z} = y^*_{-z}, \mathbf{e} = e^*, \mathbf{p}_z = p^*_z$. Since $\boldsymbol{\eta}$ and $\pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e})$ have bounded support and are independent conditional on $\mathbf{y}_{-z} = y_{-z}$ and $\mathbf{p}_z = p_z$, we can constructively identify the moment generating function of $\pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e})$ conditional on $\mathbf{y}_{-z} = y_{-z}$ and $\mathbf{p}_z = p_z$ as the ratio of the moment generating functions of $\boldsymbol{\pi}_r$ conditional on $\mathbf{y}_{-z} = y_{-z}$ and $\mathbf{p}_z = p_z$ and $\boldsymbol{\eta}$. Since the distribution of $\pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e})$ conditional on $\mathbf{y}_{-z} = y_{-z}$ and $\mathbf{p}_z = p_z$ is discrete, its moment generating function is sufficient for its identification. Note that the moment generating function of $\boldsymbol{\eta}$ is well-defined and is never equal to zero since $\boldsymbol{\eta}$ is a bounded random variable.

Assumption 3 implies that whenever a type e occurs with positive probability conditional on y_{-z} and p_z, then higher types also occur with positive probability. Assumption 1 then implies that the ranking over restricted profits is equivalent to the ranking over productivity e. As a result, if some firm type e does not operate given y_{-z} and p_z, then it has to a low type. Let $\Pi_r(y_{-z}, p_z)$ be the support of $\pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e})$ conditional on $\mathbf{y}_{-z} = y_{-z}$ and $\mathbf{p}_z = p_z$. Fix some y_{-z} and p_z. Since the support of $\mathbf{e}$ is finite, the set $\Pi_r(y_{-z}, p_z)$ will also be finite. As a result, Assumption 1 implies that

$$\pi_r(y_{-z}, p_z, d_e) = \max\left[\Pi_r(y_{-z}, p_z)\right].$$

That is, the most productive firm will make more profits than any other firm. Note that the firm with productivity $e = d_e - 1$, if it is present in the market, will be the second one in terms of restricted profits :

$$\pi_r(y_{-z}, p_z, d_e - 1) = \max\left[\Pi_r(y_{-z}, p_z, s) \setminus \{\pi_r(y_{-z}, p_z, d_e)\}\right].$$

In general, given y_{-z} and p_z, if the firm with productivity e operates ($|\Pi_r(y_{-z}, p_z)| > d_e - e$), then

$$\pi_r(y_{-z}, p_z, e) = \max\left[\Pi_r(y_{-z}, p_z) \setminus \bigcup_{e' > e} \{\pi_r(y_{-z}, p_z, e')\}\right].$$

Note that we may not be able to identify the structural restricted profit function for arguments in which e is too low.

A.3. Proof of Theorem 2

To prove sufficiency of (i), fix some x_{-d_y}, and take y_z^* from the statement of the theorem and $e^* \in E$ from condition (i). We abuse notation and drop e^* and y_{-z}^*. By homogeneity of degree 1 of $\pi_r(\cdot)$ we have that for every x

$$\sum_{j=1}^{d_y} \partial_{g_j} \pi_r(g(x)) g_j(x_j) = \pi_r(g(x))\,. \tag{4}$$

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

$$\partial_{g_j} \pi_r(g(x)) \partial_{x_j} g_j(x_j) = \partial_{x_j} \tilde{\pi}(x)\,, \tag{5}$$

for every $j = 1, \ldots, d_y$. Combining (4) and (5) we get that

$$\sum_{j=1}^{d_y} \partial_{x_j} \tilde{\pi}(x) \frac{1}{\partial_{x_j}(\log(g_j(x_j)))} = \tilde{\pi}(x) \tag{6}$$

as long as $0 < \left| \dfrac{\partial_{x_j} g_j(x_j)}{g_j(x_j)} \right| < \infty$ for every $j = 1, \ldots, d_y$. This latter condition is satisfied for almost every x_j with respect to Lebesgue measure by Assumption 5(v).

Let $t = \left(\dfrac{1}{\partial_{x_j}(\log(g_j(x_j)))} \right)_{j=1,\ldots,d_y-1}$. 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 (6) can be rewritten as

$$At = b\,, \tag{7}$$

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 5(ii)) we identify $g_j(\cdot)$ for every $j = 1, \ldots, d_y - 1$.

To prove sufficiency of (ii), recall that we assume that for all e and p_z in the support, and $\lambda > 0$

$$F_{\mathbf{e}|\mathbf{p},\mathbf{y}_{-z}}(e|\lambda p_z, y_{-z}^*) = F_{\mathbf{e}|\mathbf{p},\mathbf{y}_{-z}}(e|p_z, y_{-z}^*).$$

Hence, under Assumption 4(i), the function

$$\mathbb{E}\left[\pi_r(\mathbf{y}_{-z}, \mathbf{p}_z, \mathbf{e}) | \mathbf{p}_z = \cdot, \mathbf{y}_{-z} = y_{-z}^*\right]$$

is homogeneous of degree 1 in p_{-z} and

$$\mathbb{E}\left[\boldsymbol{\pi}_r^*|\mathbf{x}=\cdot,\mathbf{y}_{-z}=y_{-z}^*\right]=\mathbb{E}\left[\pi_r(\mathbf{y}_{-z},\mathbf{p}_z,\mathbf{e})|\mathbf{p}_z=g(\cdot),\mathbf{y}_{-z}=y_{-z}^*\right].$$

Thus, the result follows from applying the same arguments as in the proof of sufficiency of (i) to the function $\mathbb{E}\left[\boldsymbol{\pi}_r^*|\mathbf{x}=\cdot,\mathbf{y}_{-z}=y_{-z}^*\right]$ instead of $\tilde{\pi}(\cdot)$.

A.4. Proof of Theorem 3

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 $p\in P(e)$ and $e\in E$. Thus, conclusion (i) follows from the fact that $\pi(p,e)$ is identified for each $p\in P(e)$ and $e\in E$ by Theorem 1.

To establish conclusion (ii), recall that under the assumptions of Theorem 1, 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 $p\in P(e)$. 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

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

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

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

When Assumption 7 holds, identification of $Y(\cdot)$ follows from Corollary 9.18 in Kreps (2012).

A.5. Proof of Proposition 1

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

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

where $\mathrm{co}(A)$ 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 $p \in \mathbb{R}_{++}^{d_y} \cap \mathbb{S}^{d_y - 1}$,

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

Because P is finite, $\{y_p\}_{p \in P}$ is bounded. Thus, its convex hull $\mathrm{co}(\{y_p\}_{p \in P})$ 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^*.[31]

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 $e \in E$.

[31] 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.

A.6. Proof of Theorem 4 and Proposition 2

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 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

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

Lemma 4. *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 $\mathrm{PolCon}(\bar{P})$ is nonempty, then for any $u \in \mathbb{S}^{d_y-1}$,

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

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 $u^* \in \mathrm{PolCon}(\bar{P})$ such that $u'u^* > 0$. To prove this suppose to the contrary that for every $u^* \in \mathrm{PolCon}(\bar{P})$, $u'u^* \leq 0$, it follows that $u \in \mathrm{PolCon}(\mathrm{PolCon}(\bar{P}))$. The latter is not possible, since $\mathrm{PolCon}(\mathrm{PolCon}(\bar{P}))$ 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 $u^* \in \mathrm{PolCon}(\bar{P})$, 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 5. *Let $d_y \geq 2$ and let the functions $a, b : \mathbb{R}^{d_y}_{++} \to \mathbb{R}$ be convex and homogeneous of degree 1. Define*

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

where $\bar{P} \subseteq \mathbb{R}^{d_y}_{++}$ 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 $\mathrm{PolCon}(\overline{P})$ is nonempty; in particular the polar cone contains the negative unit vector $(-1, \ldots, -1)'$. The set $\bar{P}$ satisfies the conditions of Lemma 4, 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,

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

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

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

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

for all $p \in \bar{P}$ and $e \in E$. Thus, Theorem 4 is obtained as corollary.

We now turn to the proof of Proposition 2. We first present two lemmas, which are modifications of Lemmas 6 and 7 in Brunel (2016).

Lemma 6. *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 $A = \{y \in \mathbb{R}^{d_y} \,:\, p'y \le a(p),\, p \in \bar{P}\}$ 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 $h_A(p^*) \le a(p^*) < \infty$. 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 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 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 6 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

$$\begin{aligned} a(p^*) = h_A(p^*) = h_A\left(\sum_{p\in P^*_n} \lambda_{p,n} p\right) &\leq \sum_{p\in P^*_n} \lambda_{p,n} h_A(p) = \sum_{p\in P^*_n} \lambda_{p,n} a(p) \leq \sum_{p\in P^*_n} \lambda_{p,n}(b_n(p) + \eta_n) \\ &= \sum_{p\in P^*_n} \lambda_{p,n} p'y^*_n + \eta_n \sum_{p\in P^*_n} \lambda_{p,n} = p^{*\prime}y^*_n + \eta_n \sum_{p\in P^*_n} \lambda_{p,n} = h_{B_n}(p^*) + \eta_n \sum_{p\in P^*_n} \lambda_{p,n}\,. \end{aligned} \tag{8}$$

Moreover,

$$h_{B_n}(p^*) \leq b_n(p^*) \leq a(p^*) + \eta_n\,. \tag{9}$$

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 (9) 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 4 note that since $\pi(\cdot, e)$ and $\hat{\pi}(\cdot, e)$ are homogeneous of degree 1, we have

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

To prove Proposition 2, note that by Lemma 5, with probability 1,

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

The conclusion then follows by applying Lemma 7 to the right hand side of the equality above.

B. Supplemental Results

B.1. Continuous Heterogeneity

In this section we consider the possibility of continuous heterogeneity. Unfortunately, in contrast to our main result in Theorem 1, when facing continuous heterogeneity we will need more observables (instruments) and more assumptions. Moreover, we will need to assume that there is no measurement error η.

For simplicity we state the results for the profit function. The analysis of the general restricted profit function is similar. Assume that the analyst observes $(\boldsymbol{\pi}, \mathbf{p}', \mathbf{w}')'$, where the instrumental variable $\mathbf{w}$ is supported on W and $\boldsymbol{\pi} = \pi(\mathbf{p}, \mathbf{e})$ are perfectly measured profits. We normalize $\mathbf{e}$ to be uniformly distributed.

Assumption 8. *The distribution of* $\mathbf{e}$ *is uniform over* $[0, 1]$.

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

Assumption 9. $F_{\mathbf{e}|\mathbf{w}}(\cdot|w) = F_{\mathbf{e}}(\cdot)$ *for all* $w \in W$.

Assumption 9 together with the requirement that the profit function $\pi(p, \cdot)$ is monotone (Assumption 1) imply the following integral equation familiar from the literature on nonparametric quantile instrumental variable models.

Lemma 8. *If Assumptions 1, 8 and 9 are satisfied, then the following holds:*

$$\mathbb{P}\left(\boldsymbol{\pi} \leq \pi(\mathbf{p}, e)|\mathbf{w} = w\right) = e \tag{10}$$

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

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

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

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

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

The law of iterated expectations together with Assumptions 8 and 9 then imply that

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

$\blacksquare$

This lemma says that in the presence of endogeneity, we can still rank firms conditional on the instrumental variable. Note that Equation 10 is an integral equation that connects the unknown profit function, the distribution of observables, and productivity e. Indeed, Equation 10 can be rewritten as

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

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

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

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.[32] Next we establish identification of $\pi(\cdot)$ based on the results of Chernozhukov & Hansen (2005).

Note that Equation 10 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 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 $\mathbf{p}$

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

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

Theorem 5. *Suppose that*

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

(ii) Assumptions 8 and 9 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*

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

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

B.2. Nonmonotonicity of Supply

Consider the following production sets that correspond to three different levels of productivity. $Y(e_i) = \{(y_o,l)' \in \mathbb{R}\times\mathbb{R}_+ \,:\, y_o \le f_i(l)\}$, where

$$f_1(l) = l^{0.4}, \qquad f_2(l) = 2\cdot l^{0.4}$$

and

$$f_3(l) = \begin{cases} l^{0.2} & 0.01 \ge l \ge 0 \\ 7\cdot(l-0.01) + 0.01^{0.2} & 0.03 \ge l \ge 0.01 \\ 2\cdot l^{0.4} + 7\cdot 0.02 + 0.01^{0.2} - 2\cdot 0.03^{0.4} & 0.03 \le l. \end{cases}$$

Note that by construction $f_1(l) < f_2(l) < f_3(l)$ for all $l > 0$. Hence, $Y(e_1) \subseteq Y(e_2) \subseteq Y(e_3)$ and $\pi(p,e_1) < \pi(p,e_2) < \pi(p,e_3)$ for all positive p. If one takes $p = (p_o,p_l)'$

such that $p_o/p_l = 0.12$, then the optimal levels of inputs and outputs are

$$
\begin{aligned}
0.007 > l_1^* = 0.048^{5/3} > 0.006, \quad 0.2 > y_{o,1}^* = 0.048^{2/3} > 0.1 \\
0.03 > l_2^* = 0.096^{5/3} > 0.02, \quad 0.5 > y_{o,2}^* = 2 \cdot 0.096^{2/3} > 0.41 \\
0.01 > l_3^* = 0.024^{5/4} > 0.009, \quad 0.40 > y_{o,3}^* = 0.024^{1/4} > 0.39.
\end{aligned}
$$

Thus, for this price vector neither the optimal level of the input nor the optimal level of the output are monotone in productivity since $l_1^* < l_3^* < l_2^*$ and $y_{o,1}^* < y_{o,3}^* < y_{o,2}^*$.

B.3. Parametric Examples and Price Proxies

Section 3 shows that if prices are not observed but price proxies are, then it is possible to reproduce price variation from such proxies. The technique requires a high level rank condition. We present two examples to better understand this rank condition.

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

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

$\alpha, \beta > 0$. Then

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

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

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

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.

B.4. Point Identification and Assumption 7

It is natural to wonder when Assumption 7 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 $(p',e)' \in P \times E$.

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

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

that satisfies the 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 8 (Unique Extension without Assumption 7)**.** 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 Corollary 9.18 in Kreps (2012) the production correspondence is identified even though Assumption 7 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 $\operatorname{int}\left(\operatorname{cl}\left(\bigcup_{\lambda>0}\{\lambda p \,:\, p \in P\}\right)\right)$. We are only able to construct "knife edge" examples in which the support restriction of Assumption 7 is *not* equivalent to point identification of $Y(\cdot)$. We note that strict convexity of $\pi(\cdot, e)$ rules out this sort of example.

B.5. Parametric Estimation of the Diewert (1973) Functional Form

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 increasing; (iii) convexity can be also imposed using linear inequalities on the coefficients;[33] (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.

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