Nonparametric Identification of Production Function, Total Factor Productivity, and Markup from Revenue Data

Hiroyuki Kasahara[†] Yoichi Sugita[‡]

November 3, 2020

Abstract

Commonly used methods of production function and markup estimation assume that a firm's output quantity can be observed as data, but typical datasets contain only revenue, not output quantity. We examine the nonparametric identification of production function and markup from revenue data when a firm faces a general nonparametric demand function under imperfect competition. Under standard assumptions, we provide the constructive nonparametric identification of various firm-level objects: gross production function, total factor productivity, price markups over marginal costs, output prices, output quantities, a demand system, and a representative consumer's utility function.

1 Introduction

The estimation of production function and markup is a core tool used in empirical analyses of market outcomes.[1] The residual of an estimated production function, total factor productivity (TFP), is widely used to measure firm-level technological efficiency (see Bartelsman and Doms (2000) and Syverson (2011) for recent surveys) and its contribution to aggregate efficiency (e.g., Olley and Pakes, 1996). Researchers often estimate the elasticity of production functions to analyze technological changes (e.g., Van Biesebroeck, 2003;

Acknowledgement: We are grateful to comments made at seminars and conferences with regard to an earlier version of this work. We thank Zheng Han, Yoko Sakamoto and Makoto Tanaka for excellent research assistance. Sugita acknowledges financial supports from JSPS KAKENHI (grant numbers: 17H00986, 19H01477, and 19H00594).

[†]Department of Economics, University of British Columbia, Canada. (Email: hkasahar@mail.ubc.ca)

[‡]Graduate School of Economics, Hitotsubashi University, Japan. (E-mail: yoichi.sugita@r.hit-u.ac.jp)

[1]Griliches and Mairesse (1999) and Ackerberg, Benkard, Berry, and Pakes (2007) provide excellent surveys on production function estimations.

Doraszelski and Jaumandreu, 2018) and price markups over marginal costs (e.g., Hall, 1988; De Loecker and Warzynski, 2012). The estimation of firm-level markup via production function has been widely applied in various topics and complements markup estimation via demand function (e.g., Berry, Levinsohn, and Pakes, 1995) in economic analysis of firm's market power.

Commonly used methods of production function and markup estimation assume that a firm's output quantity can be observed as data. However, typical firm-level datasets contain only revenue, not output quantity. Therefore, in practice, many applications use revenue deflated by an industry-level price deflator as output.[2] For production function estimation, this practice may be justified under perfect competition where an output price is exogenous and identical across firms. However, ever since Marschak and Andrews (1944)'s pioneering study, several researchers have voiced cautions and suggested that the practice may not be justified under imperfect competition; they show that using revenue as output can significantly bias the identification of production functions (e.g., Klette and Griliches, 1996; De Loecker, 2011) and TFP (e.g., Foster, Haltiwanger, and Syverson, 2008; Katayama, Lu, and Tybout, 2009; De Loecker, 2011). Furthermore, as shown in Bond, Hashemi, Kaplan, and Zoch, 2020, using revenue in place of output quantity may lead to serious biases in estimation of firm's markups. Despite such criticism, the practice of using revenue in place of output quantity persists in many applications given a lack of output quantity data.

In the existing literature, it is not known whether identifying production functions and markups from firm-level revenue data is possible without imposing parametric assumptions. This paper contributes to the literature of production function and markup estimation by establishing nonparametric identification of production function, TFP, and markup from revenue data. The proof is constructive and the required assumption is similar to the standard assumption in the production function literature except that we impose additional assumptions on firm's demand function.

Following Marschak and Andrews (1944), Klette and Griliches (1996) and De Loecker (2011), we explicitly model a demand function that an individual firm faces as a function of its output and observable characteristics that are excluded from the production function.[3] While each of these earlier studies examines a demand function with a constant and identical demand elasticity—something that implies identical markups across firms—we consider a general nonparametric demand function that generates rich heterogeneity in various firm-level outcomes,

[2] A few studies use firm-level datasets that include output quantity (e.g., Foster, Haltiwanger, and Syverson, 2008; De Loecker, Goldberg, Khandelwal, and Pavcnik, 2016; Lu and Yu, 2015; Nishioka and Tanaka, 2019). However, those quantity datasets are available only for a limited number of countries, industries, and years, and they are not easily accessible to all researchers.

[3] De Loecker, Eeckhout, and Unger (2020) study an alternative approach using an exogenous variable to remove output price variation from revenue data.

including markups; for this reason, we can address the bias from markup heterogeneity across firms that the literature has criticized. In other respects, our method requires the standard assumptions and can be implemented using typical data found in empirical applications.

We develop a three-step identification approach that combines the control function approach developed by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015) and the first-order condition approach recently developed by Gandhi, Navarro, and Rivers (2020).[4] Following Levinsohn and Petrin (2003) and Ackerberg et al. (2015), the inverse function of a material demand function serves as a control function for TFP. In the first step, we identify revenue as a function of inputs and observable demand shifters by using the control function; this first step corresponds to that of Ackerberg et al. (2015). Our novel second step identifies the control function for TFP by applying the nonparametric identification of transformation models (e.g., Horowitz, 1996) examined by Ekeland, Heckman, and Nesheim (2004) and Chiappori, Komunjer, and Kristensen (2015). By identifying the control function, TFP is identified (up to normalization) from the dynamics of inputs, without output data. In the third step, we identify a production function, markup, and a demand function, using the first-order condition for the material and the control function identified in the second step.

Our method identifies various objects from the revenue data. In our main setting, markup and output elasticities are identified up to scale; an output price, an output quantity, a gross production function, and TFP are identified up to scale and location. Identification is cross-sectional so that the identified objects can vary over time. With an additional assumption of *local* constant returns to scale, we identify the levels of markup and output elasticities; we may also identify an output price, an output quantity, a production function, and TFP up to location.[5] Finally, if we are willing to assume monopolistic competition (without imposing free entry), we further identify a demand system and a utility function of a representative consumer—specifically, Matsuyama and Ushchev (2017)'s homothetic demand system with a single aggregator (HSA)—that can be used for a counter-factual analysis and a welfare analysis.[6]

The remainder of this paper is organized as follows. Section 2 summarizes previous studies

[4]These approaches assumed quantity data or perfect competition. Gandhi et al. (2020) also examined an imperfect competition with a constant elastic demand as in Klette and Griliches (1996) and De Loecker (2011) where markups must be constant and identical across firms.

[5]Flynn, Gandhi, and Traina (2019) used *global* constant returns to scale to identify a production function. In subsection 3.4.2, we clarify *local* and *global* constant returns to scale.

[6]One frequently sees within the literature an assumption of market structure for the identification of demand and supply side objects. For example, Berry, Levinsohn, and Pakes (1995) identify firm-level marginal costs by specifying oligopolistic competition; meanwhile, Ekeland, Heckman, and Nesheim (2004) and Heckman, Matzkin, and Nesheim (2010) identify various demand and supply side objects of a hedonic model by exploiting the properties of perfect competition.

on how using revenue as output could bias the identification of production function, TFP, and markup; readers familiar with the literature can skip this section and proceed to Section 3. Subsection 3.1 explains our setting, and subsection 3.2 demonstrates our three-step approach by offering a parametric example. Subsection 3.3 presents our nonparametric identification results, and subsection 3.4 discusses additional assumptions for fixing scale and location normalization. Subsection 3.5 examines the identification of a demand system and a representative consumer's utility function. Both subsection 3.6 and the Appendix present identification results in alternative settings, including endogenous labor input, endogenous firm-level observable demand shifters, unobservable demand shifters, and i.i.d. productivity shocks. Section 4 provides concluding remarks.

2 Biases from Using Revenue as Output Quantity

This section summarizes possible biases in the identification of production function, TFP, and markup when revenue is used as an output quantity. We denote the logarithms of the price, output, and revenue of firm i at time t as p_{it}, y_{it}, and $r_{it} := p_{it} + y_{it}$, respectively. Suppose that these variables are related via the inverse demand function $p_{it} = \psi_{it}(y_{it})$ and the revenue function $r_{it} = \varphi_{it}(y_{it}) := y_{it} + \psi_{it}(y_{it})$. Let $y_{it} = f_t(m_{it}, k_{it}, l_{it}) + \omega_{it}$ be firm i's production function where ω_{it} is TFP and $x_{it} := (m_{it}, k_{it}, l_{it})$ is a vector of the logarithms of material, capital, and labor, respectively. To highlight the sources of biases from using revenue as output, assume that TFP is identical across firms within time t, with $\omega_{it} = \omega_t$ for all i. This simplification eliminates an additional and well-known source of bias, correlations between inputs and TFP.

From the first-order condition for profit maximization, $P_{it}\left(1 + \psi'_{it}(y_{it})\right) = MC_{it}$, the elasticity of revenue with respect to output is equal to the inverse of markup:

$$\frac{d\varphi_{it}(y_{it})}{dy} = \frac{MC_{it}}{P_{it}}. \tag{1}$$

Under perfect competition where $P_{it} = MC_{it}$, the variation in revenue across firms coincides with that of output. However, they are generally different when markups vary across firms.

Suppose that, using revenue as output, a researcher identifies a true relationship between revenue and inputs, $\tilde{\varphi}_{it}(x_{it}) := \varphi_{it}(f_t(x_{it}) + \omega_t)$ to use $\tilde{\varphi}_{it}(x_{it})$ as a proxy for $f_t(x_{it})$. Prior studies show that the use of revenue as output could cause biases in three forms. First, Marschak and Andrews (1944) and Klette and Griliches (1996) establish that, from (1), the elasticity of $\tilde{\varphi}_{it}(x_{it})$ relates to the true elasticity of $f_t(x_{it})$ via markup:

$$\frac{\partial \tilde{\varphi}_{it}(x_{it})}{\partial v_{it}} = \frac{MC_{it}}{P_{it}} \frac{\partial f_t(x_{it})}{\partial v_{it}} \text{ for } v_{it} \in \{m_{it}, k_{it}, l_{it}\}. \tag{2}$$

Thus, output elasticities would be underestimated by the extent of markup.

Second, Katayama et al. (2009) and De Loecker (2011) demonstrated a bias in TFP estimates. Let $d\omega_t$ be a TFP change. Suppose that a TFP change for firm i is estimated as a change in revenue with inputs being fixed, $d\tilde{\omega}_{it} = d\tilde{\varphi}_{it}(x_{it})|_{dx_{it}=0}$. From (1), we see that this TFP estimate relates to the true TFP change via markup:

$$d\tilde{\omega}_{it} = \frac{MC_{it}}{P_{it}} d\omega_t. \tag{3}$$

Therefore, TFP would be underestimated by the extent of markup.

Finally, Bond et al. (2020) show that markup estimates using the method of Hall (1988) and De Loecker and Warzynski (2012) are generally biased when revenue elasticity is used in place of output elasticity. Suppose a firm is a price-taker of flexible input v. Hall (1988) and De Loecker and Warzynski (2012) developed the following equation relating to markup and output elasticity with respect to v as:

$$\frac{P_{it}}{MC_{it}} = \frac{\partial f_t(x_{it})/\partial v_{it}}{\alpha_{it}^v} \tag{4}$$

where α_{it}^v is the ratio of expenditure on input v to revenue. If a researcher uses $\partial \tilde{\varphi}_{it}(x_{it})/\partial m_{it}$ instead of $\partial f_t(x_{it})/\partial m_{it}$ in markup equation (4), then from (2), the estimated markup is 1:

$$\frac{\partial \tilde{\varphi}_{it}(x_{it})/\partial v_{it}}{\alpha_{it}^v} = \frac{\frac{MC_{it}}{P_{it}} \frac{\partial f_t(x_{it})}{\partial v_{it}}}{\alpha_{it}^v} = 1. \tag{5}$$

In such a case, the markup would be underestimated.[7]

Klette and Griliches (1996) and De Loecker (2011) developed methods by which to identify production functions from revenue data, by assuming a constant elastic demand function with an identical elasticity.[8] However, with this specific demand function, markups must be constant

[7] Result (5) by Bond et al. (2020) relies on the assumption that a researcher can correctly identify $\tilde{\varphi}_{it}(x_{it})$. In practice, misspecification of $\tilde{\varphi}_{it}(x_{it})$ could derive markup estimates (5) that contain some information on true markups. For instance, De Loecker and Warzynski (2012, Section VI) show that when f is Cobb–Douglas, it is possible to identify the effect of firm-level variables (e.g., export) on markups.

[8] Katayama et al. (2009) also developed a method by which to identify production functions from revenue data. Their method allows for markup heterogeneity but requires the ability to estimate firm's marginal costs from total costs.

and identical across firms. Studies estimating markups from quantity data report substantial heterogeneity in markups across firms (e.g., De Loecker, Goldberg, Khandelwal, and Pavcnik, 2016; Lu and Yu, 2015; Nishioka and Tanaka, 2019). To address the biases arising from firm-level markup heterogeneity, we extend the approach of Klette and Griliches (1996) and De Loecker (2011) by incorporating a general nonparametric demand function that allows for variable and heterogeneous markups.

3 Identification

3.1 Setting

We denote the logarithm of physical output, material, capital, and labor as y_{it}, m_{it}, k_{it}, and l_{it}, respectively, with their respective supports denoted as $\mathscr{Y}$, $\mathscr{M}$, $\mathscr{K}$, and $\mathscr{L}$. We collect the three inputs (material, capital, and labor) into a vector as $x_{it} := (m_{it}, k_{it}, l_{it})' \in \mathscr{X} := \mathscr{M} \times \mathscr{K} \times \mathscr{L}$.

At time t, output y_{it} relates to inputs $x_{it} = (m_{it}, k_{it}, l_{it})'$ via the production function:

$$y_{it} = f_t(x_{it}) + \omega_{it}, \tag{6}$$

where the firm's TFP ω_{it} follows an exogenous first-order stationary Markov process given by

$$\omega_{it} = h(\omega_{it-1}) + \eta_{it}, \tag{7}$$

where we assume that neither $h(\cdot)$ nor the marginal distribution of η_{it} change over time.[9]

The demand function for a firm's product is strictly decreasing in its price, and its inverse demand function is given by

$$p_{it} = \psi_t(y_{it}, z_{it}), \tag{8}$$

where p_{it} is the logarithm of output price and $z_{it} \in \mathscr{Z}$ is an observable firm characteristic that affects firm's demand (e.g., export status in De Loecker and Warzynski (2012)). z_{it} can be either a continuous or discrete vector; in the main text below, z_{it} is assumed to be continuous and exogenous—that is, $z_{it} \perp \eta_{it}$. In subsection 3.6 and the Appendix, we present the identification results when z_{it} is discrete and/or may correlate with η_t.

The inverse demand function (8) generalizes the constant elastic demand function examined by Marschak and Andrews (1944), Klette and Griliches (1996) and De Loecker (2011). Although ψ_t is nonparametric, (8) implicitly makes two assumptions. First, $\psi_t(\cdot)$ is a common function for all firms once the observed characteristics z_{it} are controlled for. This implies

[9] $h(\cdot)$ can include a firm's observable exogenous characteristics.

that unobserved demand shifters must be common for all firms—that is, ψ_t can be written as $\psi_t(y_{it}, z_{it}, A_t)$ where A_t is a vector of unobserved variables and can include an aggregate price/quantity index. In subsection 3.6, we discuss the case where $\psi_t(\cdot)$ includes a firm-level unobservable demand shifter such as quality. Second, $\psi_t(\cdot)$ represents a demand curve that an individual firm takes as given. This is satisfied in the case of monopolistic competition (without free entry) where each firm takes A_t as given.

Let $\bar{r}_{it}$ and $\bar{\mathscr{R}}$ be the logarithm of (true) revenue and its support, respectively. Revenue r_{it} in the data is observed with a measurement error ε_{it}, $r_{it} = \bar{r}_{it} + \varepsilon_{it}$. Then, from (6), the observed revenue relates to output and input as follows:

$$\begin{aligned} r_{it} &= \varphi_t(y_{it}, z_{it}) + \varepsilon_{it} \\ &= \varphi_t(f_t(m_{it}, k_{it}, l_{it}) + \omega_{it}, z_{it}) + \varepsilon_{it} \end{aligned} \tag{9}$$

where $\varphi_t(y_{it}, z_{it}) := \psi_t(y_{it}, z_{it}) + y_{it}$.

We assume that l_{it} and k_{it} are predetermined at the end of the last period $t-1$, while m_{it} is flexibly chosen after observing ω_{it}.[10] Specifically, $m_{it} = \mathbb{M}_t(\omega_{it}, k_{it}, l_{it}, z_{it})$ is chosen at time t by:

$$\mathbb{M}_t(\omega_{it}, k_{it}, l_{it}, z_{it}) \in \arg\max_m \exp(\varphi_t(f_t(m, k_{it}, l_{it}) + \omega_{it}, z_{it})) - \exp(p_t^m + m), \tag{10}$$

where p_t^m denotes the logarithm of the material input price at time t, which is common to all firms. A firm is assumed to be a price-taker for material input.

Equation (9) highlights two identification issues raised by Marschak and Andrews (1944). First, m_{it} correlates with the unobservable ω_{it}. Second, r_{it} relates to $x_{it} = (m_{it}, k_{it}, l_{it})$ via two unknown nonlinear functions $\varphi_t(\cdot, z_{it})$ and $f_t(\cdot)$, and two unobservables ω_{it} and ε_{it}.[11] To address these issues via a control function and a transformation model, we first make the following assumptions.

Assumption 1. *(a) $f_t(\cdot)$ is continuously differentiable with respect to (m, k, l) on $\mathscr{M} \times \mathscr{K} \times \mathscr{L}$ and strictly increasing in m. (b) For every $z \in \mathscr{Z}$, $\varphi_t(\cdot, z)$ is strictly increasing and invertible with its inverse $\varphi_t^{-1}(\bar{r}, z)$, which is continuously differentiable with respect to $(\bar{r}, z)$ on $\bar{\mathscr{R}} \times \mathscr{Z}$. (c) For every $(k, l, z) \in \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$, $\mathbb{M}_t(\cdot, k, l, z)$ is strictly increasing and invertible with its inverse $\mathbb{M}_t^{-1}(m, k, l, z)$, which is continuously differentiable with respect to (m, k, l, z) on $\mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$. (d) ε_t is mean independent of x_t and z_t with $E[\varepsilon_t | x_t, z_t] = 0$.*

[10] In subsection 3.6, we present identification when l_{it} also correlates with ω_{it}.

[11] In subsection 3.6 and the Appendix, we present identification when a firm receives an i.i.d. shock e_{it} to output and then, the firm's revenue includes a non-additive error, $r_{it} = \varphi_t(f_t(x_{it}) + \omega_{it} + e_{it}, z_{it})$.

Assumptions 1 (a) and (b) are standard assumptions about smooth production and demand functions. Assumption 1 (b) $\partial \varphi_t(y,z)/\partial y > 0$ is equivalent to that the elasticity of demand with respect to price, $-(\partial \psi_t(y,z)/\partial y)^{-1}$, is greater than 1; this necessarily holds under profit maximization. Therefore, Assumption 1 (b) is innocuous as long as we analyze the outcomes of profit maximization. Assumption 1 (c) is a standard assumption in the control function approach that uses material as a control function for TFP (Levinsohn and Petrin, 2003; Ackerberg et al., 2015).

The inverse function of the material demand function with respect to TFP

$$\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_t)$$

is used as a control function for ω_{it}. Since $\partial \varphi_t(y_t, z_t)/\partial y_t > 0$, there exists the inverse function $\varphi_t^{-1}(\cdot, z_t)$ so that the revenue function $\bar{r}_{it} = \varphi_t(f_t(x_{itt}) + \omega_{it})$ can be written as:

$$\varphi_t^{-1}(\bar{r}_{it}, z_{it}) = f_t(x_{it}) + \mathbb{M}_t^{-1}(x_{it}, z_{it}). \tag{11}$$

In the following, we identify $\varphi_t^{-1}(\cdot)$, $f_t(\cdot)$, and $\mathbb{M}_t^{-1}(\cdot)$ from the distribution of variables in the data. Let $v_t := (k_t, l_t, z_t, x_{t-1}, z_{t-1})' \in \mathscr{V} := \mathscr{K} \times \mathscr{L} \times \mathscr{Z} \times \mathscr{X} \times \mathscr{Z}$. Data includes a random sample of firms $\{r_{it}, v_{it}\}_{i=1}^N$ from the population. For instance, the variable x_{it} of firm i is considered as a realization of the random variable x_t. Given a sufficiently large N, an econometrician can recover their joint distributions.

Assumption 2. *The following information at time t is known: (a) the conditional distribution $G_{m_t|v_t}(m_t|v_t)$ of m_t given v_t; (b) the conditional expectation $E[r_t|x_t, z_t]$ of r_t given (x_t, z_t); (c) firm's expenditure on material* $\exp(p_t^m + m_{it})$.

Assumption 2 (a) is required for the identification of $\mathbb{M}_t^{-1}(\cdot)$. Assumptions 2 (b) and (c) are additionally required for the identification of $\varphi_t^{-1}(\cdot)$ and $f_t(\cdot)$. Typical production datasets include those variables in Assumption 2.

Let $\{\varphi_t^{*-1}(\cdot), f_t^*(\cdot), \mathbb{M}_t^{*-1}(\cdot)\}$ be the true model structure that satisfies (11). Then, for any $(a_{1t}, a_{2t}, b_t) \in \mathbb{R}^2 \times \mathbb{R}_{++}$,

$$\varphi_t^{-1}(\bar{r}_t, z_t) = (a_{1t} + a_{2t}) + b_t \varphi_t^{*-1}(\bar{r}_t, z_t),\ f_t(x_t) = a_{1t} + b_t f_t^*(x_t),$$
$$\text{and } \mathbb{M}_t^{-1}(x_t, z_t) = a_{2t} + b_t \mathbb{M}_t^{*-1}(x_t, z_t) \tag{12}$$

also satisfy (11), and the true structure $\{\varphi_t^{*-1}(\cdot), f_t^*(\cdot), \mathbb{M}_t^{*-1}(\cdot)\}$ is observationally equivalent to the structure (12). That is, the structure $\{\varphi_t^{-1}(\cdot), f_t(\cdot), \mathbb{M}_t^{-1}(\cdot)\}$ is identified only up to location and scale normalization (a_{1t}, a_{2t}, b_t) from restriction (11).

Therefore, identification requires location and scale normalization. We fix (a_{1t}, a_{2t}, b_t) in (12) by fixing the values of $\{\varphi_t^{-1}(\cdot), f_t(\cdot), \mathbb{M}_t^{-1}(\cdot)\}$ at some points. Specifically, choosing two points $(m_{t1}^*, k_t^*, l_t^*, z_t^*)$ and $(m_{t0}^*, k_t^*, l_t^*, z_t^*)$ on the support $\mathscr{X} \times \mathscr{Z}$ where $m_{t0}^* < m_{t1}^*$, we denote

$$c_{1t} := f_t(m_{t0}^*, k_t^*, l_t^*),\ c_{2t} = \mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, l_t^*, z_t^*),\ \text{and}\ c_{3t} := \mathbb{M}_t^{-1}(m_{t1}^*, k_t^*, l_t^*, z_t^*). \tag{13}$$

Note that $\partial \mathbb{M}_t^{-1} / \partial m_t > 0$ implies that $c_{2t} < c_{3t}$. Then, there exists a unique one-to-one mapping between (c_{1t}, c_{2t}, c_{3t}) in (13) and (a_{1t}, a_{2t}, b_t) in (12) such that $b_t = (c_{3t} - c_{2t}) / \left(\mathbb{M}_t^{*-1}(m_{t1}^*, k_t^*, l_t^*, z_t^*) - \mathbb{M}_t^{*-1}(m_{t0}^*, k_t^*, l_t^*, z_t^*)\right)$, $a_{1t} = c_{1t} - b_{1t} f_t^*(m_{t0}^*, k_t^*, l_t^*)$ and $a_{2t} = c_{2t} - b_{1t}\mathbb{M}_t^{*-1}(m_{t0}^*, k_t^*, l_t^*, z_t^*)$. Thus, we can fix the value of (a_{1t}, a_{2t}, b_t) by choosing arbitrary values $(c_{1t}, c_{2t}, c_{3t}) \in \mathbb{R}^3$ that satisfies $c_{2t} < c_{3t}$. In particular, we impose the following normalization that corresponds to (N2) in Chiappori et al. (2015).

Assumption 3. *(Normalization) The support $\mathscr{X} \times \mathscr{Z}$ includes two points $(m_{t1}^*, k_t^*, l_t^*, z_t^*)$ and $(m_{t0}^*, k_t^*, l_t^*, z_t^*)$ such that $c_{1t} = c_{2t} = 0$ and $c_{3t} = 1$ in (13).*

As Chiappori et al. (2015) demonstrates, this choice of normalization makes the identification proofs transparent.

3.2 Identification in a Parametric Example

Before presenting the nonparametric identification results, we demonstrate our identification approach by applying it to a simple parametric example. Consider a monopolistically competitive market where each firm i faces the following constant elastic inverse demand function:

$$p_{it} = \alpha_t(z_{it}) + (\rho(z_{it}) - 1)y_{it}, \tag{14}$$

where $\alpha_t(z_{it})$ and $0 < \rho(z_{it}) \le 1$ are unknown parameters.[12] The markup equals $1/\rho(z_{it})$ and depends on the exogenous scalar $z_{it} \in \mathscr{Z} := \{1, 0\}$ such that $z_{it} \perp \eta_{it}$. Firm i has a Cobb–Douglas production function and ω_{it} follows a first-order autoregressive (AR(1)) process:

$$\begin{aligned} y_{it} &= \theta_0 + \theta_m m_{it} + \theta_k k_{it} + \theta_l l_{it} + \omega_{it}, \\ \omega_{it} &= h_0 + h_1 \omega_{it-1} + \eta_{it}, \end{aligned} \tag{15}$$

[12] The demand function (14) can be derived from a constant elasticity of substitution (CES) utility function; $a_t(z_t)$ implicitly includes aggregate expenditure and an aggregate price index.

where $\{\theta_0, \theta_m, \theta_k, \theta_l, h_0, h_1\}$ are unknown parameters. The firm's revenue function is expressed as:

$$r_{it} = \alpha_t(z_{it}) + \rho(z_{it})\theta_0 + \rho(z_{it})\theta_m m_{it} + \rho(z_{it})\theta_k k_{it} + \rho(z_{it})\theta_l l_{it} + \rho(z_{it})\omega_{it} + \varepsilon_{it}. \tag{16}$$

The first-order condition for (10),

$$\rho(z_{it})\theta_m = \frac{\exp(p_t^m + m_{it})}{\exp(r_{it} - \varepsilon_{it})}, \tag{17}$$

determines the control function for ω_{it} as

$$\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it}) = \beta_t(z_{it}) + \beta_m(z_{it})m_{it} + \beta_k k_{it} + \beta_l l_{it} \tag{18}$$

where $\beta_t(z_{it}) := \left(p_t^m - \alpha_t(z_{it}) - \theta_0 - \ln \rho(z_{it})\theta_m\right)/\rho(z_{it})$, $\beta_m(z_{it}) := (1 - \rho(z_{it})\theta_m)/\rho(z_{it}) > 0$, $\beta_k := -\theta_k$ and $\beta_l := -\theta_l$.

For notational brevity, assume that the support $\mathscr{X} \times \mathscr{Z}$ includes two points $(m_{t1}^*, k_t^*, l_t^*, z_t^*) = (0,0,0,0)$ and $(m_{t0}^*, k_t^*, l_t^*, z_t^*) = (1,0,0,0)$. Following Assumption 3, we fix the location and scale of $f_t(\cdot)$ and $\mathbb{M}_t^{-1}(\cdot)$ by imposing the following normalization:

$$\begin{aligned} &0 = f_t(0,0,0) = \theta_0,\ 0 = \mathbb{M}_t^{-1}(0,0,0,0) = \beta_t(0), \\ &1 = \mathbb{M}_t^{-1}(1,0,0,0) = \beta_t(0) + \beta_m(0) \end{aligned} \tag{19}$$

which implies $\theta_0 = 0$, $\beta_t(0) = 0$, and $\beta_m(0) = 1$.

Our identification approach follows three steps.

Step 1: Identification of Measurement Errors The first step removes the measurement error ε_{it} in the spirit of Ackerberg et al. (2015). Substituting (18) into (16) and using $\theta_0 = 0$, we obtain two expressions of r_{it} as follows:

$$\begin{aligned} r_{it} =& (\alpha_t(z_{it}) + \rho(z_{it})\beta_t(z_{it})) + \rho(z_{it})(\theta_m + \beta_m(z_{it}))m_{it} \\ &+ \rho(z_{it})(\theta_k + \beta_k)k_{it} + \rho(z_{it})(\theta_l + \beta_l)l_{it} + \varepsilon_{it} && (20) \\ =& \phi(z_{it}) + m_{it} + \varepsilon_{it}, && (21) \end{aligned}$$

where $\phi(z_{it}) := \alpha_t(z_{it}) + \rho(z_{it})\beta_t(z_{it})$. Applying the conditional moment restriction $E[\varepsilon_{it}|m_t, z_t] = 0$ for the second expression (21), we identify $\phi(z_{it})$, $\bar{r}_{it}$ and ε_{it} by

$$\phi(z_t) = E[r_{it} - m_{it}|m_t, z_t], \bar{r}_{it} = \phi(z_{it}) \text{ and } \varepsilon_{it} = r_{it} - m_{it} - \phi(z_{it}).$$

Step 2: Identification of Control Function and TFP The second step identifies the control function $\mathbb{M}_t^{-1}(\cdot)$. Substituting (18) into the AR(1) process (15) leads to

$$\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it}) = h_0 + h_1 \mathbb{M}_{t-1}^{-1}(m_{it-1}, k_{it-1}, l_{it-1}, z_{it-1}) + \eta_{it}. \tag{22}$$

Since $\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ is linear in m_{it} from (18), we can rearrange (22) as:

$$\begin{aligned} m_{it} = \gamma(z_{it}, z_{t-1}) + \gamma_k(z_{it})k_{it} + \gamma_l(z_{it})l_{it} + \delta_m(z_{it}, z_{it-1})m_{it-1} \\ + \delta_k(z_{it})k_{it-1} + \delta_l(z_{it})l_{it-1} + \tilde{\eta}_{it}, \end{aligned} \tag{23}$$

where

$$\begin{aligned} \gamma(z_{it}, z_{it-1}) &:= \frac{h_0 - \beta_t(z_{it}) + h_1\beta_{t-1}(z_{it-1})}{\beta_m(z_{it})}, \gamma_k(z_{it}) := -\frac{\beta_k}{\beta_m(z_{it})}, \gamma_l(z_{it}) := -\frac{\beta_l}{\beta_m(z_{it})}, \\ \delta_m(z_{it}, z_{it-1}) &:= \frac{h_1\beta_m(z_{it-1})}{\beta_m(z_{it})}, \ \delta_k(z_{it}) := \frac{h_1\beta_k}{\beta_m(z_{it})}, \ \delta_l := \frac{h_1\beta_l}{\beta_m(z_{it})}, \tilde{\eta}_{it} := \frac{\eta_{it}}{\beta_m(z_{it})}. \end{aligned} \tag{24}$$

For a given (z_{it}, z_{it-1}), (23) is a linear model. Since $E[\tilde{\eta}_{it}|v_{it}] = E[\eta_{it}|v_{it}]/\beta_m(z_{it}) = 0$, where $v_{it} := (k_{it}, l_{it}, x_{it-1}, z_{it}, z_{it-1})$, we can identify $\{\gamma(z_{it}, z_{t-1})$, $\gamma_k(z_{it})$, $\gamma_l(z_{it})$, $\delta_m(z_{it}, z_{it-1})$, $\delta_k(z_{it})$, $\delta_l(z_{it})\}$ in (23) from the conditional moment restriction $E[\tilde{\eta}_{it}|v_{it}] = 0$.

From (19) and (24), we identify the parameters of the control function (under the normalization (19)) as:

$$\beta_t(1) = \gamma(0,0) - \gamma(1,0)\frac{\gamma_k(0)}{\gamma_k(1)}, \ \beta_m(1) = \frac{\gamma_k(0)}{\gamma_k(1)}, \ \beta_k = \gamma_k(0) \text{ and } \beta_l = \gamma_l(0).$$

Step 3: Identification of Production Function and Markup The final step identifies the parameters of the demand and production functions. Comparing the two expressions of r_{it} in (20) and (21), we obtain the following relationships:

$$\begin{aligned} \alpha_t(z_{it}) + \rho(z_{it})\beta_t(z_{it}) = \phi(z_{it}), \ \rho(z_{it})(\theta_m + \beta_m(z_{it})) = 1, \\ \theta_k = -\beta_k \text{ and } \theta_l = -\beta_l. \end{aligned} \tag{25}$$

Given that $(\beta_t(z_t), \beta_m(z_t), \beta_k, \beta_l)$ are identified in step 2, the first line in (25) contains four equations (two equations for two values of $z_{it} \in \{0,1\}$) and five parameters $(\alpha_t(0), \alpha_t(1), \rho(0), \rho(1), \theta_m)$. Therefore, to identify these parameters, we need a further restriction.

Following Gandhi et al. (2020), we use as an additional restriction the first-order condition for material (17). The first-order condition (17) implies that the revenue share of material expenditure on the right hand side of (17) is a function of z_{it}. Using ε_{it}, we obtain the revenue share of material expenditure $\exp(p_t^m + m_{it})/\exp(r_{it} - \varepsilon_{it})$ and identify it as a function of z_{it} by taking its expectation conditional on z_{it}:

$$s(z_t) := E\left[\frac{\exp(p_t^m + m_{it})}{\exp(\bar{r}_{it})} \middle| z_t \right].$$

Then, we obtain an additional restriction on the parameters:

$$\rho(z_{it})\theta_m = s(z_{it}). \tag{26}$$

From (25) and (26), we identify the parameters of the demand and production functions as follows

$$\begin{aligned} \rho(0) = 1 - s(0),\ \rho(1) = \frac{1 - s(1)}{\beta_m(1)}, \\ \alpha_t(0) = \phi(0),\ \alpha_t(1) = \phi(1) - \rho(1)\beta_t(1), \\ \theta_0 = 0,\ \theta_m = \frac{s(0)}{1 - s(0)},\ \theta_k = -\beta_k \text{ and } \theta_l = -\beta_l. \end{aligned}$$

Note that the parameters are identified under the scale and location normalization of $f_t(\cdot)$ and $\mathbb{M}_t^{-1}(\cdot)$ in (19). Let θ_i $(i = 0, m, k, l)$ and $\beta_j(z_t)$ $(j = t, m, k, l)$ be those parameters identified above and let θ_j^* and $\beta_i^*(z_t)$ be the true parameters. Then, there exist unknown normalization parameters $(a, b) \in \mathbb{R} \times \mathbb{R}_+$ such that $\theta_0 = a + b\theta_0^*$, $\beta_t = a + b\beta_t^*$, $\theta_i = b\theta_i^*$, $\beta_j(z_t) = b\beta_j^*(z_t)$. We can fix the normalization by imposing further restrictions. For instance, if constant returns to scale $\theta_m^* + \theta_k^* + \theta_l^* = 1$ are imposed, then the scale parameter b can be identified as follows:

$$b = b\left(\theta_m^* + \theta_k^* + \theta_l^*\right) = \theta_m + \theta_k + \theta_l = \frac{s(0)}{1 - s(0)} - \beta_k - \beta_l.$$

We discuss in subsection 3.4 additional assumptions for fixing normalization.

The above identification argument is illustrative, but it relies on the linearity of $\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ in m_{it}, which holds only under restrictive parametric assumptions. Extending the argument, the following subsection establishes nonparametric identification.

3.3 Nonparametric Identification

3.3.1 Step 1: Identification of Measurement Error

The first step removes the measurement error ε_{it}. Substituting the control function $\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$, the revenue function (9) can be written as:

$$\begin{aligned} r_{it} &= \varphi_t\left(f_t(x_{it}) + \mathbb{M}_t^{-1}(x_{it}, z_{it}), z_{it}\right) + \varepsilon_{it} \\ &= \phi_t(x_{it}, z_{it}) + \varepsilon_{it}, \end{aligned}$$

where $\phi_t(x_t, z_t) := \varphi_t\left(f(x_t) + \mathbb{M}_t^{-1}(x_t, z_t), z_t\right)$. From Assumption 1, $\phi_t(\cdot)$ is continuously differentiable. From $E[\varepsilon_{it}|x_t, z_t] = 0$, we can identify $\phi_t(\cdot)$, $\bar{r}_{it}$, and ε_{it} as:

$$\phi_t(x_t, z_t) = E[r_{it}|x_t, z_t], \; \bar{r}_{it} = \phi_t(x_{it}, z_{it}) \text{ and } \varepsilon_{it} = r_{it} - \phi_t(x_{it}, z_{it}). \tag{27}$$

Lemma 1. *Suppose that Assumptions 1–2 hold. Then, we can identify $\phi_t(\cdot)$, $\bar{r}_{it}$, and ε_{it} as in (27).*

Hereafter, $\phi_t(\cdot)$, $\bar{r}_{it}$, and ε_{it} are assumed to be known.[13]

3.3.2 Step 2: Identification of Control Function and TFP

From (7), the control function $\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ satisfies

$$\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it}) = \bar{h}_t(x_{it-1}, z_{it-1}) + \eta_{it}, \tag{28}$$

where $\bar{h}_t(x_{t-1}, z_{t-1}) := h\left(\mathbb{M}_{t-1}^{-1}(m_{t-1}, k_{t-1}, l_{t-1}, z_{t-1})\right)$. As $\partial \mathbb{M}_t^{-1}/\partial m_{it} > 0$, given the values of (k_{it}, l_{it}, z_{it}), the dependent variable in (28) is a monotonic transformation of m_{it}. Therefore, the model (28) belongs to a class of transformation models, the identification of which Chiappori et al. (2015) analyze.

We make the following assumption, which corresponds to Assumptions A1–A3, A5, and A6 in Chiappori et al. (2015).[14]

Assumption 4. *(a) The distribution $G_\eta(\cdot)$ of η is absolutely continuous with a density function $g_\eta(\cdot)$ that is continuous on its support. (b) η_t is independent of $v_t := (k_t, l_t, z_t, x_{t-1}, z_{t-1})' \in \mathscr{V} := \mathscr{K} \times \mathscr{L} \times \mathscr{Z} \times \mathscr{X} \times \mathscr{Z}$ with $E[\eta_t|v_t] = 0$. (c) v_t is continuously distributed on $\mathscr{V}$. (d) Support Ω*

[13] As will be shown, ω_{it} is identified in step 2 independently of step 1. Therefore, one can think of an alternative approach that first identifies ω_{it} and then regresses r_{it} on $(x_{it}, z_{it}, \omega_{it})$ to obtain $E[r_{it}|x_{it}, z_{it}, \omega_{it}]$ instead of $E[r_{it}|x_{it}, z_{it}]$. However, it is not possible to identify $E[r_{it}|x_{it}, z_{it}, \omega_{it}]$ because $\omega_{it} = \mathbb{M}_t^{-1}(x_{it}, z_{it})$ is a deterministic function of (x_{it}, z_{it}). Once (x_{it}, z_{it}) are conditioned, there is no remaining source of variation in ω_{it}.

[14] Assumption 1 (c) corresponds to Assumption A4 of Chiappori et al. (2015).

of ω_t is an interval $[\underline{\omega}, \bar{\omega}] \subset \mathbb{R}$ where $\underline{\omega} < 0$ and $1 < \bar{\omega}$. (e) $h(\cdot)$ is continuously differentiable with respect to ω on Ω. (f) The set $\mathscr{A}_{q_{t-1}} := \{(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z} : \partial G_{m_t|v_t}(m_t|v_t)/\partial q_{t-1} \neq 0$ for all $(m_t, k_t, l_t, z_t) \in \mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}\}$ is nonempty for some $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$.

We can relax Assumption 4(b) by allowing z_t and l_t to correlate with η_t, which we discuss this in subsection 3.6. Assumption 4(d) holds without loss of generality because we can choose any two points on the support of ω_t without changing the essence of our argument. Assumption 4(f) can be interpreted as a generalized rank condition, thus implying that a given exogenous variable q_{t-1} has a causal impact on (m_t, k_t, l_t, z_t). Suppose $g_\eta(\eta) > 0$ for all $\eta \in \mathbb{R}$. Then, as will be shown below (in (30)), Assumption 4(f) holds if and only if

$$\frac{\partial \bar{h}(\tilde{x}_{t-1}, \tilde{z}_{t-1})}{\partial q_{t-1}} = h'\left(\mathbb{M}_{t-1}^{-1}(\tilde{x}_{t-1}, \tilde{z}_{t-1})\right) \frac{\partial \mathbb{M}_{t-1}^{-1}(\tilde{x}_{t-1}, \tilde{z}_{t-1})}{\partial q_{t-1}} \neq 0$$

for some $(\tilde{x}_{t-1}, \tilde{z}_{t-1})$ and some $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$. This condition is equivalent to (1) ω_{t-1} has a causal impact on ω_t ($h'(\omega_{t-1}) \neq 0$) and (2) q_{t-1} has a causal impact on m_{t-1}, ($\partial \mathbb{M}_{t-1}/\partial q_{t-1} \neq 0$). These conditions must be satisfied for at least one exogenous variable q_{t-1} and some point $(\tilde{x}_{t-1}, \tilde{z}_{t-1})$.

Proposition 1 shows that the control function is identified from the distribution of (m_{it}, v_{it}).

Proposition 1. *Suppose that Assumptions 1–4 hold. Then, we can identify $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$ up to scale and location and $G_\eta(\cdot)$ up to the scale normalization of η_t.*

Proof. The proof follows the proof of Theorem 1 in Chiappori et al. (2015). In view of equation (28), the conditional distribution of m_t given v_t satisfies

$$\begin{aligned} G_{m_t|v_t}(m_t|v_t) &= G_{\eta_t|v_t}\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})|v_t\right) \\ &= G_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right), \end{aligned}$$

where the second equality follows from $\eta_t \perp v_t$ in Assumption 4(b). Let $q_t \in \{m_t, k_t, l_t, z_t\}$ and $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$. The derivatives of $G_{m_t|v_t}(m_t|v_t)$ are

$$\frac{\partial G_{m_t|v_t}(m_t|v_t)}{\partial q_t} = \frac{\partial \mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)}{\partial q_t} g_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right), \tag{29}$$

$$\frac{\partial G_{m_t|v_t}(m_t|v_t)}{\partial q_{t-1}} = -\frac{\partial \bar{h}(x_{t-1}, z_{t-1})}{\partial q_{t-1}} g_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right). \tag{30}$$

Using Assumption 4(f), we can choose $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$ and $(\tilde{x}_{t-1}, \tilde{z}_{t-1}) \in \mathscr{A}_{q_{t-1}}$ such that $\partial G_{m_t|v_t}(m_t|k_t, l_t, z_t, \tilde{x}_{t-1}, \tilde{z}_{t-1})/\partial q_{t-1} \neq 0$ for all $(m_t, k_t, l_t, z_t) \in \mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$.

Dividing (29) by (30), we derive

$$\frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial q_t} = -\frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}} \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}. \tag{31}$$

Then, from (31) for $q_t = m_t$ and the normalization in Assumption 3, we obtain

$$\begin{aligned} 1 &= \mathbb{M}_t^{-1}(m_{t1}^*,k_t^*,l_t^*,z_t^*) - \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,z_t^*) \\ &= -\frac{1}{S_{q_{t-1}}} \frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}}, \end{aligned} \tag{32}$$

where

$$S_{q_{t-1}} := \left(\int_{m_{t0}^*}^{m_{t1}^*} \frac{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z_t^*,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial m_t}{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z_t^*,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial q_{t-1}} dm \right)^{-1}.$$

Then, we identify $\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1} = -S_{q_{t-1}}$. Substituting this into (31), $\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)/\partial q_t$ for $q_t \in \{m_t,k_t,l_t,z_t\}$ are identified as follows:

$$\frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial q_t} = S_{q_{t-1}} \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}. \tag{33}$$

Integrating (33) with respective to $q_t \in \{m_t,k_t,l_t,z_t\}$ obtains

$$\begin{aligned} \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t) =& \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t) - \mathbb{M}_t^{-1}(m_{t0}^*,k_t,l_t,z_t) \\ &+ \mathbb{M}_t^{-1}(m_{t0}^*,k_t,l_t,z_t) - \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t,z_t) \\ &+ \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t,z_t) - \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,z_t) \\ &+ \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,z_t) - \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,z_t^*) \\ =& \int_{m_{t0}^*}^{m_t} \frac{\partial \mathbb{M}_t^{-1}(s,k_t,l_t,z_t)}{\partial m_t} ds + \int_{k_t^*}^{k_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*,s,l_t,z_t)}{\partial k_t} ds \\ &+ \int_{l_t^*}^{l_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,s,z_t)}{\partial l_t} ds + \int_{z_t^*}^{z_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,s)}{\partial z_t} ds, \end{aligned} \tag{34}$$

where the first equality follows from $\mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t^*,z_t^*) = 0$ in Assumption 3. Substituting the identified derivatives of $\mathbb{M}_t^{-1}(\cdot)$ in (33) into (34), we can identify $\mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)$ for all (m_t,k_t,l_t,z_t).

Finally, from $\omega_{it} = \mathbb{M}_t^{-1}(m_{it},k_{it},l_{it},z_{it})$, we can identify $\bar{h}_t(x_{t-1},z_{t-1}) = E[\omega_{it}|x_{t-1},z_{t-1}]$ and $\eta_{it} = \omega_{it} - \bar{h}_t(x_{it-1},z_{it-1})$. Thus, we can identify the distribution of η_t, $G_{\eta_t}(\eta)$. □

3.3.3 Step 3: Identification of Production Function and Markup

The final step identifies production function, markup and other remaining objects. From $\bar{r} = \phi_t(x_t, z_t) = \varphi_t(f_t(x_t) + \mathbb{M}_t^{-1}(x_t, z_t), z_t)$ and the monotonicity of φ_t, differentiating $\varphi_t^{-1}(\phi(x_t, z_t), z_t) = f_t(x_t) + \mathbb{M}_t^{-1}(x_t, z_t)$ with respect to $q_t \in \{m_t, k_t, l_t\}$ and z_t gives:

$$\frac{\partial \varphi_t^{-1}(\bar{r}_t, z_t)}{\partial \bar{r}_t} \frac{\partial \phi_t(x_t, z_t)}{\partial q_t} = \frac{\partial f_t(x_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial q_t}, \tag{35}$$

$$\frac{\partial \varphi_t^{-1}(\bar{r}_t, z_t)}{\partial \bar{r}_t} \frac{\partial \phi_t(x_t, z_t)}{\partial z_t} = \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial z_t} - \frac{\partial \varphi_t^{-1}(\bar{r}_t, z_t)}{\partial z_t}. \tag{36}$$

Note that $\partial \varphi_t^{-1}(\bar{r}_t, z_t)/\partial \bar{r}_t = (\partial \varphi_t(y_t, z_t)/\partial y_t)^{-1}$ represents the markup from (1). If the markup $\partial \varphi_t^{-1}(\bar{r}_t, z_t)/\partial \bar{r}_t$ were known, then equations (35) and (36) could identify $\partial f_t(x_t)/\partial q_t$ and $\partial \varphi_t^{-1}(\bar{r}_t, z_t)/\partial z_t$ given that $\mathbb{M}_t^{-1}(x_t, z_t)$ is identified. However, since the markup is unknown, identification requires further restriction. Following Gandhi et al. (2020), we use the first-order condition with respect to the material as an additional restriction.

Assumption 5. *The first-order condition with respect to material for the profit maximization problem (10)*

$$\frac{\partial f_t(x_{it})}{\partial m_{it}} = \frac{\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})}{\partial \bar{r}_t} \frac{\exp(p_t^m + m_{it})}{\exp(\bar{r}_{it})} \tag{37}$$

holds for all firms.

Rearranging the first-order condition, we obtain the Hall-De Loecker-Warzynski markup equation:

$$\frac{\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})}{\partial \bar{r}_t} = \frac{\partial f_t(x_{it})/\partial m_{it}}{\exp(p_t^m + m_{it})/\exp(\bar{r}_{it})}. \tag{38}$$

We establish the following proposition.

Proposition 2. *Suppose that Assumptions 1–5 hold. Then, we can identify $\varphi_t^{-1}(\cdot)$ and $f_t(\cdot)$ up to scale and location and each firm's markup $\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})/\partial \bar{r}_t$ up to scale.*

Proof. From (35) and (37), the markup $\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})/\partial \bar{r}_t$ is identified as

$$\frac{\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})}{\partial \bar{r}_t} = \frac{\partial \mathbb{M}_t^{-1}(x_{it}, z_{it})}{\partial m_t} \left(\frac{\partial \phi_t(x_{it}, z_{it})}{\partial m_t} - \frac{\exp(p_t^m + m_{it})}{\exp(\bar{r}_{it})} \right)^{-1}. \tag{39}$$

From $\bar{r}_t = \phi_t(x_t, z_t)$ and (39), the markup is also identified as a function of (x_t, z_t) as

$$\begin{aligned}\mu_t(x_t, z_t) &:= \frac{\partial \varphi_t^{-1}(\phi_t(x_t, z_t), z_t)}{\partial r_t} \\ &= \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial m_t}\left(\frac{\partial \phi_t(x_t, z_t)}{\partial m_t} - \frac{\exp(p_t^m + m_t)}{\exp(\phi_t(x_t, z_t))}\right)^{-1} \end{aligned} \tag{40}$$

Substituting (40) into (35), we identify $\partial f_t(x_t)/\partial q_t$ for $q_t \in \{m_t, k_t, l_t\}$ as follows:

$$\frac{\partial f_t(x_t)}{\partial q_t} = \mu_t(x_t, z_t)\frac{\partial \phi_t(x_t, z_t)}{\partial q_t} - \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial q_t}. \tag{41}$$

Using $f_t(m_{t0}^*, k_t^*, l_t^*) = 0$ in Assumption 3, we identify $f_t(x_t)$ by integration:

$$\begin{aligned} f_t(m_t, k_t, l_t) = &\int_{m_{t0}^*}^{m_t} \frac{\partial f_t(s, k_t, l_t)}{\partial m_t} ds + \int_{k_t^*}^{k_t} \frac{\partial f_t(m_{t0}^*, s, l_t)}{\partial k_t} ds \\ &+ \int_{l_t^*}^{l_t} \frac{\partial f_t(m_{t0}^*, k_t^*, s)}{\partial l_t} ds. \end{aligned} \tag{42}$$

Let $\bar{\mathscr{R}} := \{\bar{r}_t : \bar{r}_t = \phi_t(x_t, z_t) \text{ for some } (x_t, z_t) \in \mathscr{X} \times \mathscr{Z}\}$ be the support of $\bar{r}_t$. For given $(\bar{r}_t, z_t) \in \bar{\mathscr{R}} \times \mathscr{Z}$, $X_t(\bar{r}_t, z_t) := \{x_t \in \mathscr{X} : \phi_t(x_t, z_t) = \bar{r}_t\}$ is non-empty by the construction of $\bar{\mathscr{R}}$. Then, because $f_t(x_t)$ and $\mathbb{M}_t^{-1}(x_t, z_t)$ are identified, the output quantity $\varphi_t^{-1}(\bar{r}_t, z_t)$ for any $(\bar{r}_t, z_t) \in \bar{\mathscr{R}} \times \mathscr{Z}$ is identified by

$$\varphi_t^{-1}(\bar{r}_t, z_t) = f_t(x_t) + \mathbb{M}_t^{-1}(x_t, z_t) \text{ for } x_t \in X_t(\bar{r}_t, z_t).$$

□

The output price for individual firms is identified as

$$p_{it} := \bar{r}_{it} - \varphi_t^{-1}(\bar{r}_{it}, z_{it}).$$

Corollary 1. *Suppose that Assumptions 1–5 hold. Then, the production function, output quantities, output prices, and TFP are identified up to scale and location; markups and output elasticities are identified up to scale.*

Remark 1. Examination of the proofs reveals that we have over-identifying restrictions. In particular, the proof of Proposition 1 goes through with any choice of $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$ in (33). Furthermore, the proof of Proposition 2 does not rely on the restriction in (36) for identifying $\varphi_t^{-1}(\cdot)$. These over-identifying restrictions can be useful in developing a specification test for the model as well as for efficiently estimating the model.

3.3.4 Comparison to Existing Identification Approaches

Our approach follows the spirits of existing identification approaches, but it does differ from them in terms of implementations. First, step 2 distinguishes our approach from the standard control function approach (e.g., Ackerberg et al., 2015). In step 2, we identify the control function from the dynamics of the inputs, and without using any output measure. To clarify why this approach is necessary, consider an alternative approach that uses an output measure. Specifically, in the second step, we substitute $\omega_{it} = \varphi_t^{-1}(\bar{r}_{it}, z_{it}) - f_t(x_{it})$ into (28) and obtain the alternative transformation model:

$$\varphi_t^{-1}(\bar{r}_{it}, z_{it}) = f_t(x_{it}) + \tilde{h}_t(\bar{r}_{it-1}, x_{it-1}, z_{t-1}) + \eta_{it}$$

where $\tilde{h}_t(\bar{r}_{t-1}, x_{t-1}, z_{t-1}) := h(\varphi_{t-1}^{-1}(\bar{r}_{t-1}, z_{t-1}) - f_t(x_{t-1}))$. Since this model also belongs to the class of transformation models examined by Chiappori et al. (2015), one might think that we could have identified $\varphi_t(\cdot)$ and $f_t(\cdot)$ from the conditional distribution function $G_{\bar{r}_t|w_t}(\bar{r}_t|w_t)$ of $\bar{r}_t$ given $w_t := (x_t, z_t, \bar{r}_{t-1}, x_{t-1}, z_{t-1})$. This is not possible, however, because once (x_t, z_t) is conditioned on, $\bar{r}_t = \phi_t(x_t, z_t)$ loses all variations. Therefore, the derivatives of $G_{\bar{r}_t|w_t}$ with respect to past variables $(\bar{r}_{t-1}, x_{t-1}, z_{t-1})$ are always 0, which violates the condition corresponding to Assumption 4 (f).

Second, Ackerberg et al. (2015) identify a structural value-added function, $y_{it} = v_t(k_{it}, l_{it}) + \omega_{it}$, which under perfect competition derives from a Leontief production function $y_{it} = \min\{v_t(k_{it}, l_{it}) + \omega_{it}, a + m_{it}\}$. However, the structural value-added function is difficult to employ under imperfect competition because $y_{it} < v_t(k_{it}, l_{it}) + \omega_{it}$ can occur. Note that the maximum output capacity $y_{it}^* := v_t(k_{it}, l_{it}) + \omega_{it}$ is determined before a firm chooses m_{it} and y_{it}. Therefore, if y_{it}^* is large—due, for example, to a large shock on ω_{it}—then the profit maximizing output y_{it} can be lower than y_{it}^*.[15] Intuitively speaking, when increases in TFP double, a firm can preclude a price drop by increasing its output by less than double.

Third, our approach uses the first-order condition for material in a way different from that seen in Gandhi et al. (2020), whose step identifies the material elasticity $\partial f_t(x_t)/\partial m_t$ from

[15] As Ackerberg et al. (2015) explains, under perfect competition, if $y_{it} < y_{it}^*$, then the optimal output is 0 since the output becomes linear in material. Since firms in a dataset have positive outputs, $y_{it} = y_{it}^*$ holds for firms observed in a dataset. However, under imperfect competition, it is possible to have $y_{it} < y_{it}^*$ and the optimal output is strictly positive.

the first-order condition (37):

$$\ln \frac{\exp(p_t^m + m_{it})}{\exp(r_{it})} = \ln \frac{\partial f_t(x_{it})}{\partial m_{it}} - \ln \frac{\partial \varphi_t^{-1}(r_{it} - \varepsilon_{it}, z_{it})}{\partial r_t} - \varepsilon_{it}$$

under the assumption of perfect competition where $\ln \partial \varphi_t^{-1}(r_{it} - \varepsilon_{it}, z_{it})/\partial r_{it} = 0$ for all i. Under imperfect competition, when the markup depends on revenue $r_{it} - \varepsilon_{it}$, $\partial f_t(x_t)/\partial m_t$ cannot be identified solely from the first-order condition.

3.4 Fixing Normalization across Periods

Let $(\varphi_t^{-1}(\cdot), f_t(\cdot), \mathbb{M}_t^{-1}(\cdot))$ be a model structure for period t identified by using Propositions 1 and 2 under the normalization in Assumption 3. Let $(\varphi_t^{*-1}(\cdot), f_t^*(\cdot), \mathbb{M}_t^{*-1}(\cdot))$ denote the true model structure. Since the structure is identified up to scale and location normalization, there exist period-specific location and scale parameters $(a_{1t}, a_{2t}, b_t) \in \mathbb{R}^2 \times \mathbb{R}_+$ such as

$$\begin{aligned} &\varphi_t^{-1}(r_t, z_t) = a_{1t} + a_{2t} + b_t \varphi_t^{*-1}(r_t, z_t),\ f_t(x) = a_{1t} + b_t f_t^*(x_t), \\ &\mathbb{M}_t^{-1}(x_t, z_t) = a_{2t} + b_t \mathbb{M}_t^{*-1}(x_t, z_t). \end{aligned} \tag{43}$$

Generally speaking, the location and scale normalization differ across periods—that is, $(a_{1t}, a_{2t}, b_t) \neq (a_{1t+1}, a_{2t+1}, b_{t+1})$. For the identified objects to be comparable across periods, we need to fix normalization across periods by assuming that some object in the model is time-invariant. The subsection discusses these additional assumptions.[16]

3.4.1 Scale Normalization

From (43), the ratio of identified markups across two periods relates to the ratio of true markups as

$$\frac{\partial \varphi_{t+1}^{-1}(r,z)/\partial r}{\partial \varphi_t^{-1}(r,z)/\partial r} = \frac{b_{t+1}}{b_t} \frac{\partial \varphi_{t+1}^{*-1}(r,z)/\partial r}{\partial \varphi_t^{*-1}(r,z)/\partial r}.$$

Therefore, the ability to identify how true markups change over two periods requires identification of the ratio of scale parameters, b_{t+1}/b_t. Similarly, the ratio of identified output elasticities across periods and that of identified TFP deviation from the mean are related to

[16]Klette and Griliches (1996) and De Loecker (2011) identify the levels of markups and output elasticities from revenue data by using a functional form property of a demand function. They consider a constant elastic demand function leading to $\varphi_t(y_{it}, z_{it}) = \alpha y_{it} - (\alpha - 1) z_{it}$ where z_{it} is an aggregate demand shifter, which is an weighted average of revenue across firms, and α is an unknown parameter. This formulation implies $\varphi_t^{-1}(r_{it}, z_{it}) = (1/\alpha) r_{it} + (1 - 1/\alpha) z_{it}$ and imposes a linear restriction $\partial \varphi_t^{-1}(r_{it}, z_{it})/\partial r_{it} + \partial \varphi_t^{-1}(r_{it}, z_{it})/\partial z_{it} = 1$, which fixes the scale parameter b_t.

their true values via the ratio of scale parameters:

$$\frac{\partial f_{t+1}(x)/\partial q}{\partial f_t(x)/\partial q} = \frac{b_{t+1}}{b_t}\frac{\partial f^*_{t+1}(x)/\partial q}{\partial f^*_t(x)/\partial q} \text{ and } \frac{\omega_{it+1} - E\left[\omega_{it+1}\right]}{\omega_{it} - E\left[\omega_{it}\right]} = \frac{b_{t+1}}{b_t}\left(\frac{\omega^*_{it+1} - E\left[\omega^*_{it+1}\right]}{\omega^*_{it} - E\left[\omega^*_{it}\right]}\right)$$

for $q \in \{m, k, l\}$.

To identify b_{t+1}/b_t, we consider the following assumptions.

Assumption 6. *At least one of the following conditions (a)–(c) holds. (a) The unconditional variance of η_{it} does not change over time. (b) For some known interval $\mathscr{B}$ of $\mathscr{X}$, the output elasticity of one of the inputs does not change over time for all $x \in \mathscr{B}$. (c) For some known interval $\mathscr{B}$ of $\mathscr{X}$, the sum of output elasticities of the three inputs does not change over time for all $x \in \mathscr{B}$.*

Assumption 6 (a) holds, for example, if the productivity shock ω_{it} follows a stationary process because stationarity requires that the distribution of η_{it} does not change over time. Assumption 6 (b) assumes that the elasticity of output with respect to one input does not change over time for some known interval; meanwhile, under Assumption 6 (c), returns to scale in production technology does not change for some known interval of inputs.

Proposition 3. *Suppose that Assumptions 1–6 hold for time t and $t+1$. Then, we can identify the ratio of markups between two periods t and $t+1$, the ratio of output elasticities between t and $t+1$, and the ratio of TFP deviation from the mean between t and $t+1$.*

Proof. Suppose that Assumption 6(a) holds. Let $\mathrm{var}(\eta_t)$ and $\mathrm{var}(\eta_{t+1})$ be the variance of η_t and η_{t+1} identified under the period-specific normalization in Assumption 3 for t and $t+1$, respectively. From (28) and (43), $\mathrm{var}(\eta_t) = b_t^2 \mathrm{var}(\eta^*_t)$ and $\mathrm{var}(\eta_{t+1}) = b^2_{t+1}\mathrm{var}(\eta^*_{t+1})$. From $\mathrm{var}(\eta^*_t) = \mathrm{var}(\eta^*_{t+1})$, b_{t+1}/b_t is identified as $b_{t+1}/b_t = \sqrt{\mathrm{var}(\eta_t)/\mathrm{var}(\eta_{t+1})}$.

Let $\partial f_t(x)/\partial q$ and $\partial f_{t+1}(x)/\partial q$ be those elasticities identified under the period-specific normalization in Assumption 3 for t and $t+1$, respectively, and $\partial f^*_t(x)/\partial q$ and $\partial f^*_{t+1}(x)/\partial q$ be the true elasticities. From (43), $\partial f_t(x)/\partial q = b_t \partial f^*_t(x)/\partial q$ and $\partial f_{t+1}(x)/\partial q = b_{t+1}\partial f^*_{t+1}(x)/\partial q$ hold.

Suppose that Assumption 6(b) holds. Then, $\partial f^*_t(x)/\partial q = \partial f^*_{t+1}(x)/\partial q$ for some input $q \in \{m, k, l\}$ and $x \in \mathscr{B}$. Then, b_{t+1}/b_t is identified as $b_{t+1}/b_t = (\partial f_{t+1}(x)/\partial q)/(\partial f_t(x)/\partial m)$ for $x \in \mathscr{B}$.

Suppose that Assumption 6(c) holds, implying

$$1 = \frac{\partial f^*_{t+1}(x)/\partial m + \partial f^*_{t+1}(x)/\partial k + \partial f^*_{t+1}(x)/\partial l}{\partial f^*_t(x)/\partial m + \partial f^*_t(x)/\partial k + \partial f^*_t(x)/\partial l} \text{ for } x \in \mathscr{B}.$$

Then, b_{t+1}/b_t is identified as

$$\frac{b_{t+1}}{b_t} = \frac{\partial f_{t+1}(x)/\partial m + \partial f_{t+1}(x)/\partial k + \partial f_{t+1}(x)/\partial l}{\partial f_t(x)/\partial m + \partial f_t(x)/\partial k + \partial f_t(x)/\partial l} \quad \text{for } x \in \mathscr{B}.$$

□

3.4.2 Local Constant Returns to Scale

We consider the following local constant returns to scale that strengthens Assumption 6 (c).

Assumption 7. *(Local Constant Returns to Scale) For some known interval $\mathscr{B}$ of $\mathscr{X}$, the sum of the output elasticities of the three inputs equals to 1 for all $x \in \mathscr{B}$.*

Assumption 7 is stronger than Assumption 6(c), but it is weaker than the assumptions used in some other studies on markups. Markup is sometimes estimated as the ratio of revenue $\exp(r_{it})$ to total costs TC_{it} under the assumption that a cost function is linear in output $TC_{it} = MC_{it} y_{it}$ with constant marginal costs MC_{it}. The linear cost function requires the following assumptions that are stronger than Assumption 7: (1) constant returns to scale *globally* holds for all $x \in \mathscr{B}$; (2) all three inputs are flexible and (3) a firm is a price taker of all three inputs. Under Assumption 7, marginal costs may increase in output, especially in the short run, when dynamic inputs such as capital require adjustment costs.

With Assumption 7, the scale normalization parameter b_t can be identified for all periods as follows. Let $f_t(x)$ be the identified production function and $f_t^*(x)$ be the true one where $f_t(x_t) = a_t + b_t f_t^*(x_t)$ from (43). For $x \in \mathscr{B}$, we have

$$b_t = b_t \left(\frac{\partial f_t^*(x)}{\partial m} + \frac{\partial f_t^*(x)}{\partial k} + \frac{\partial f_t^*(x)}{\partial l} \right) = \frac{\partial f_t(x)}{\partial m} + \frac{\partial f_t(x)}{\partial k} + \frac{\partial f_t(x)}{\partial l}.$$

Given that we have identified the scale parameter b_t in (43), we have established the following proposition.

Proposition 4. *Suppose that Assumptions 1–5 and 7 hold. Then, $\varphi_t(\cdot)$ and $f_t(\cdot)$ can be identified up to location. The levels of markup and output elasticities can be identified. Output quantity, output price, and TFP can be identified up to location.*

3.4.3 Location Normalization

Suppose that scale normalization b_t is already identified—for example, from Proposition 4. Define

$$\tilde{\varphi}_t^{-1}(r_t, z_t) := \varphi_t^{-1}(r_t, z_t)/b_t, \ \tilde{f}_t(x) := f_t(x)/b_t, \ \tilde{\omega}_t := \omega_t/b_t,$$
$$\tilde{a}_{1t} := a_{1t}/b_t, \text{ and } \tilde{a}_{2t} := a_{2t}/b_t. \tag{44}$$

Then, (43) is written as

$$\tilde{\varphi}_t^{-1}(r_t, z_t) = \tilde{a}_{1t} + \tilde{a}_{2t} + \varphi_t^{*-1}(r_t, z_t), \ \tilde{f}_t(x) = \tilde{a}_{1t} + f_t^*(x_t), \ \tilde{\omega}_t = \tilde{a}_{2t} + \omega_t^*. \tag{45}$$

From (43), the growth rates (log differences) of the identified output and TFP between t and $t+1$ are related to their true values as follows:

$$\tilde{\varphi}_{t+1}^{-1}(\bar{r}_{it+1}, z_{it+1}) - \tilde{\varphi}_t^{-1}(\bar{r}_{it}, z_{it}) = \tilde{a}_{1t+1} + \tilde{a}_{2t+1} - \tilde{a}_{1t} - \tilde{a}_{2t} + \varphi_{t+1}^{*-1}(\bar{r}_{it+1}, z_{it+1}) - \varphi_t^{*-1}(\bar{r}_{it}, z_{it}),$$
$$\tilde{f}_{t+1}(x_{t+1}) - \tilde{f}_t(x_t) = \tilde{a}_{1t+1} - \tilde{a}_{1t} + f^*(x_{t+1}) - f_t^*(x_t),$$
$$\tilde{\omega}_{it+1} - \tilde{\omega}_{it} = \tilde{a}_{2t+1} - \tilde{a}_{2t} + \omega_{it+1}^* - \omega_{it}^*. \tag{46}$$

Therefore, to identify the growth rates of output and TFP, we need to identify the changes in the location parameters. To do so, we can use an industry-level producer price index P_t^*, which is often available as data, to identify the change in the location parameters. Suppose that P_t^* is a Laspeyres index

$$P_t^* := \frac{\sum_{i \in \tilde{N}} \exp(p_{it}^* + y_{i0}^*)}{\sum_{i \in \tilde{N}} \exp(p_{i0}^* + y_{i0}^*)}, \tag{47}$$

where $\tilde{N}$ is a known set (or a random sample) of products. p_{i0}^* and y_{i0}^* are firm i's log true price and log true output at the base period, respectively. The following argument holds for forms of a price index (other than Laspeyres) as long as the price index is a known function of prices that is homogenous of degree 1; this condition is usually satisfied.

Assumption 8. *(a) The industry-level producer price index P_t^* is known as data. (b) For some known point $\bar{x} \in \mathcal{X}$, the true production functions of t and $t+1$, $f_t^*(\cdot)$ and $f_{t+1}^*(\cdot)$, satisfy $f_t^*(\bar{x}) = f_{t+1}^*(\bar{x})$.*

Assumption 8(b) is innocuous, implying that any output change between t and $t+1$ when inputs are fixed at $\bar{x}$ is attributed to a TFP change.

Using the aggregate price index, we can identify the change in the location parameters and identify the growth of TFP and output.

Proposition 5. *Suppose Assumptions 1–5, 7, and 8 hold. Then, the true growth rate of output* $\varphi_{t+1}^{*-1}(\bar{r}_{it+1}, z_{it+1}) - \varphi_t^{*-1}(\bar{r}_{it}, z_{it})$ *and that of TFP* $\omega_{it+1}^* - \omega_{it}^*$ *can be identified for each firm.*

Proof. Let $\tilde{p}_{it} := \bar{r}_{it} - \tilde{\varphi}_t^{-1}(\bar{r}_{it}, z_{it})$ and $\tilde{y}_{it} := \tilde{\varphi}_t^{-1}(\bar{r}_{it}, z_{it})$ be an output price and an output quantity identified under the normalization in (44) and Assumption 3, respectively. Using these, we calculate an industry-level producer price index with them:

$$P_t := \frac{\sum_{i \in \tilde{N}} \exp(\tilde{p}_{it} + \tilde{y}_{i0})}{\sum_{i \in \tilde{N}} \exp(\tilde{p}_{i0} + \tilde{y}_{i0})}.$$

From (45) and (47), P_t is written as

$$\begin{aligned} P_t &= \frac{\sum_{i \in \tilde{N}} \exp(-(\tilde{a}_{1t} + \tilde{a}_{2t}) + p_{it}^* + \tilde{a}_{1,0} + \tilde{a}_{2,0} + y_{i0}^*)}{\sum_{i \in \tilde{N}} \exp(p_{i0}^* + y_{i0}^*)} \\ &= \exp(\tilde{a}_{1,0} + \tilde{a}_{2,0} - (\tilde{a}_{1t} + \tilde{a}_{2t}))P_t^*. \end{aligned}$$

Therefore, $\tilde{a}_{1t+1} + \tilde{a}_{2t+1} - \tilde{a}_{1t} - \tilde{a}_{2t}$ is identified as:

$$\tilde{a}_{1t+1} + \tilde{a}_{2t+1} - \tilde{a}_{1t} - \tilde{a}_{2t} = \ln P_{t+1}^* - \ln P_{t+1} - \left(\ln P_t^* - \ln P_t\right) \tag{48}$$

From (46), we identify the output growth rate $\varphi_{t+1}^{*-1}(\bar{r}_{it+1}, z_{it+1}) - \varphi_t^{*-1}(\bar{r}_{it}, z_{it})$.

Evaluating the second equation in (46) at $x_{t+1} = x_t = \bar{x}$ in Assumption 8(b), we identify $\tilde{a}_{1t+1} - \tilde{a}_{1t}$ as:

$$\begin{aligned} \tilde{a}_{1t+1} - \tilde{a}_{1t} &= \tilde{a}_{1,t+1} + f_{t+1}^*(\bar{x}) - \left(\tilde{a}_{1,t} + f_t^*(\bar{x})\right) \\ &= \tilde{f}_{t+1}(\bar{x}) - \tilde{f}_t(\bar{x}). \end{aligned}$$

From (48), $\tilde{a}_{2t+1} - \tilde{a}_{2t}$ is also identified as

$$\tilde{a}_{2t+1} - \tilde{a}_{2t} = \ln P_{t+1}^* - \ln P_{t+1} - \left(\ln P_t^* - \ln P_t\right) - \left(\tilde{f}_{t+1}(\bar{x}) - \tilde{f}_t(\bar{x})\right).$$

Therefore, from (46), the true TFP growth rate $\omega_{it+1}^* - \omega_{it}^*$ is also identified. □

3.5 Identification of Demand System and Utility Function

Given that we have identified each firm's output price and quantity, it is possible to identify with additional assumptions a system of demand functions and a homothetic utility function of a representative consumer. The identified demand system and the identified utility function can be used to undertake counterfactual analysis and welfare analysis.

We consider an HSA system (Matsuyama and Ushchev, 2017), which can be expressed as a system of direct demand functions or of inverse demand functions. The two systems are self-dual in the sense that either can be derived from the other. We consider a system of inverse demand functions. Let $P_{it} := \exp(p_{it})$ and $Y_{it} := \exp(y_{it})$ be the levels of price and quantity of firm i's output at time t, respectively. Let N_t be the set of firms in the industry and $\Phi_t := \sum_{i\in N_t} P_{it}Y_{it}$ be the industry expenditure. The inverse demand function for product i is given by

$$P_{it} = \frac{\Phi_t}{Y_{it}} S_t\left(\frac{Y_{it}}{A_t(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it}\right).$$

where $S_t(\cdot, z_{it})$ provides the budget share of product i, $\mathbf{Y}_t := (Y_{1t}, ..., Y_{Nt}) \in \bar{\mathscr{Y}} := \exp(\mathscr{Y})^N$ is a vector of consumption, $\mathbf{z}_t := (z_{1t}, ..., z_{Nt})$ is a vector of observable demand shifters and $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ is the aggregate quantity index summarizing interactions across products.[17] Since $S_t(\cdot)$ is nonparametric, the HSA system can nest various demand functions used in the literature such as the constant elastic demand from the CES utility, the symmetric translog demand (Feenstra, 2003; Feenstra and Weinstein, 2017), or the constant response demand (MrÃizovÃi and Neary, 2017; 2019).[18]

For identification of a demand system, we make assumptions regarding the market structure.

Assumption 9. *The good market is monopolistically competitive (without free entry)—that is, each firm takes the quantity index $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ as given.*

The assumption of monopolistic competition follows Klette and Griliches (1996) and De Loecker (2011), with the inverse demand function becoming a symmetric function of the firm's own output, as in (8).

The demand elasticity equals $(\mu - 1)/\mu$ when μ is markup. If the markup is identified up to scale, then the demand elasticity is not uniquely identified. Therefore, we need to fix the scale normalization to identify the demand function.

Assumption 10. *$\varphi_t^{-1}(\bar{r}_t, z_t)$ is identified up to location.*

Assumption 10 is satisfied when Proposition 4 holds.

[17]If the utility function is CES $U_t(\mathbf{Y}_t, \mathbf{z}_t) = \left[\sum_{i=1}^N Y_{it}^{\rho(z_{it})}\right]^{1/\rho(z_{it})}$, then the inverse demand function is given by $P_{it} = \frac{\Phi_t}{Y_{it}}\left(\frac{Y_{it}}{U_t(\mathbf{Y}_t, \mathbf{z}_t)}\right)^{\rho(z_{it})}$. In this case, the quantity index is the same as the utility function, but they are generally different.

[18]A HSA version of the constant response demand (MrÃizovÃi and Neary, 2017; 2019) can be formulated as for example, $P_{it} = \frac{\beta\Phi_t}{Y_{it}}\left[\left(\frac{Y_{it}}{A_t(\mathbf{Y}_t, \mathbf{z}_t)}\right)^{\alpha} + \gamma z_{it}\right]^{\delta}$ where firm i's markup is given by $\mu_{it} = \frac{1}{\alpha\beta} + \frac{\gamma z_{it}}{\alpha\beta(Y_{it})^{\alpha}}$. See Matsuyama and Ushchev (2017) regarding how the HSA nests the translog demand.

An HSA demand system can be constructed as follows. Suppose $\varphi_t^{-1}(\bar{r}_t, z_t)$ is identified from Proposition 4; taking its inverse function obtains the revenue function $\varphi_t(y_t, z_t)$. Fixing a realized data point of $\mathbf{Y}_t^0 := (Y_{1t}^0, ..., Y_{Nt}^0) \in \bar{\mathscr{Y}}$ and $\mathbf{z}_t^0 := (z_{1t}^0, ..., z_{Nt}^0) \in \mathscr{Z}^N$, we let $\Phi_t := \sum_{i \in N_t} \exp\left(\varphi_t\left(\ln Y_{it}^0, z_{it}^0\right)\right)$ be the consumer's budget, which is taken as given. For given $(\mathbf{Y}_t, \mathbf{z}_t) \in \bar{\mathscr{Y}} \times \mathscr{Z}^N$, we define a vector of market shares $S_t(\mathbf{Y}_t, \mathbf{z}_t) := (S_t(Y_{1t}, z_{1t}), ..., S_t(Y_{Nt}, z_{Nt}))$ such that

$$S_t(Y_{it}, z_{it}) := \frac{\exp(\varphi_t(\ln Y_{it}, z_{it}))}{\Phi_t}.$$

The quantity index $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ is identified as follows. First, since $\sum_{i \in N_t} S_t\left(Y_{it}^0, z_{it}^0\right) = 1$, by construction, $A_t(\mathbf{Y}_t^0, \mathbf{z}_t^0) = 1$ holds for the data point $(\mathbf{Y}_t^0, \mathbf{z}_t^0)$. For other values $(\mathbf{Y}_t, \mathbf{z}_t) \in \bar{\mathscr{Y}} \times \mathscr{Z}^N$, we can obtain $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ by solving

$$\sum_{i \in N_t} S_t\left(\frac{Y_{it}}{A_t(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it}\right) = 1.$$

Since $S_t(\cdot, z_{it})$ is continuous and strictly increasing, $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ is uniquely determined. Then, we obtain the inverse demand function for all $(\mathbf{Y}_t, \mathbf{z}_t) \in \bar{\mathscr{Y}} \times \mathscr{Z}^N$:

$$P_{it} = \frac{\Phi_t}{Y_{it}} S_t\left(\frac{Y_{it}}{A_t(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it}\right). \tag{49}$$

Applying the result of Matsuyama and Ushchev (2017, Proposition 1 and Remark 3), the following proposition establishes that the HSA demand system (49) constructed above can be derived from a unique consumer preference, and that it is possible to identify an associated utility function. Appendix A.1 supplies the proof.

Proposition 6. *Suppose Assumption 10 holds. (a) There exists a unique monotone, convex, and homothetic rational preference $\succsim$ over $\bar{\mathscr{Y}}$ that generates an HSA demand system (49). (b) This preference $\succsim$ is represented by a homothetic utility function defined by*

$$\ln U_t(\mathbf{Y}_t, \mathbf{z}_t) = \ln A_t(\mathbf{Y}_t, \mathbf{z}_t) + \sum_{i=1}^{N} \int_{c_i(\mathbf{z}_t)}^{Y_{it}/A_t(\mathbf{Y}_t, \mathbf{z}_t)} \frac{S_t(\xi, z_{it})}{\xi} d\xi,$$

where $c(\mathbf{z}_t) := (c_1(\mathbf{z}_t), ..., c_N(\mathbf{z}_t))$ is defined by $U_t(c(\mathbf{z}_t), \mathbf{z}_t) = 1$. (c) The identified demand system $S_t(\cdot)$ and preference $\succsim$ do not depend on the location normalization of $\varphi_t^{-1}(\bar{r}_t, z_t)$.

3.6 Identification in Alternative Settings

3.6.1 Endogenous Labor Input

Identification is possible when l_t correlates with η_t. In the spirits of Ackerberg et al. (2015) and the dynamic generalized method of moment approach (e.g., Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998, 2000), we provide identification using lagged labor l_{t-1} as an instrument for l_t. Specifically, we follow the approach of Ackerberg et al. (2015), which assumes (1) l_t correlates with l_{t-1} and (2) the firm's profit maximization problem regarding m_t conditional on l_t is expressed by (10), which allows the material demand to be written as $m_{it} = \mathbb{M}_t(\omega_{it}, k_{it}, l_{it}, z_{it})$. This approach has the advantage of being consistent with various data generating processes regarding the choice of l_t.[19]

Identifying $\mathbb{M}_t^{-1}(x_t, z_t)$ using l_{t-1} as an instrument for l_t is nontrivial because the model (28) includes l_{t-1} in $\bar{h}_t(x_{t-1}, z_{t-1})$. It is not possible to use the variation of l_{t-1} simultaneously for two purposes (i.e., identifying $\bar{h}_t(x_{t-1}, z_{t-1})$ and instrumenting l_t). Therefore, we proceed to identification in two steps. We first identify $\bar{h}_t(x_{t-1}, z_{t-1})$ (up to location) and then use l_{t-1} to identify $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$.

Identification of $\bar{h}_t(x_{t-1}, z_{t-1})$.

Assumption 11. *(i) Assumptions 4 (a), (d), (e), and (f) hold. (ii) η_t is independent of $\tilde{w}_t := (k_t, z_t, x_{t-1}, z_{t-1})' \in \tilde{\mathscr{W}} := \mathscr{K} \times \mathscr{Z} \times \mathscr{X} \times \mathscr{Z}$ with $E[\eta_t|\tilde{w}_t] = 0$. $\tilde{w}_t$ is continuously distributed on $\tilde{\mathscr{W}}$. (iii) For each $(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z}$, $\mathscr{A}_{m_t}(x_{t-1}, z_{t-1}) = \{(\tilde{x}_t, \tilde{z}_t) \in \mathscr{X} \times \mathscr{Z} | \partial G_{m_t|v_t}(\tilde{m}_t|\tilde{k}_t, \tilde{l}_t, \tilde{z}_t, x_{t-1}, z_{t-1})/\partial m_t > 0\}$ is non-empty.*

Assumptions 11 (i) and (ii) simply modify Assumption 4 such that l_t may correlate with η_t. Assumption 11 (iii) is innocuous because it is satisfied if the firm's survival probability at time t conditional on (x_{t-1}, z_{t-1}) is not 0.

The conditional distribution of m_t given v_t satisfies

$$G_{m_t|v_t}(m_t|v_t) = G_{\eta_t|l_t}\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\,|l_t\right).$$

Taking the derivatives of both sides with respect to $q_t \in \{m_t, k_t, z_t\}$ and $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$ and their ratios, we identify $\partial \mathbb{M}_t^{-1}(m, k_t, l_t, z_t)/\partial q_t$ and $\partial \bar{h}(x_{t-1}, z_{t-1})/\partial q_t$

[19] See Ackerberg et al. (2015) for examples of such data-generating processes. For example, l_t can be chosen at time t with adjustment costs; a firm can face an auto-correlated firm-specific wage; or l_t can be chosen at time $t-1$ or at an intermediate time between t and $t-1$.

as follows:

$$\frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial q_t} = -\frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}} \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}, \tag{50}$$

$$\frac{\partial \bar{h}(x_{t-1},z_{t-1})}{\partial q_{t-1}} = -\frac{\partial \mathbb{M}_t^{-1}(\tilde{m}_t,\tilde{k}_t,\tilde{l}_t,\tilde{z}_t)}{\partial m_t} \frac{\partial G_{m_t|v_t}\left(\tilde{m}_t|\tilde{k}_t,\tilde{l}_t,\tilde{z}_t,x_{t-1},z_{t-1}\right)/\partial q_{t-1}}{\partial G_{m_t|v_t}\left(\tilde{m}_t|\tilde{k}_t,\tilde{l}_t,\tilde{z}_t,x_{t-1},z_{t-1}\right)/\partial m_t}, \tag{51}$$

where $(\tilde{x}_{t-1},\tilde{z}_{t-1}) \in \mathscr{A}_{q_{t-1}}$ and $(\tilde{x}_t,\tilde{z}_t) \in \mathscr{A}_{m_t}(x_{t-1},z_{t-1})$. Note that (50) is the same as in (31). Thus, following the same steps as those in the proof for Proposition 1, we identify $\partial \mathbb{M}_t^{-1}(m,k_t,l_t,z_t)/\partial q_t$ up to scale, and then $\partial \bar{h}(x_{t-1},z_{t-1})/\partial q_{t-1}$ up to scale from (51).

Define $d_l(l_t) := \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t,z_t^*)$ for (m_{t0}^*,k_t^*,z_t^*) in (13) and $d := \bar{h}_t\left(x_{t-1}^*,z_{t-1}^*\right)$ for some point $(x_{t-1}^*,z_{t-1}^*) \in \mathscr{X} \times \mathscr{Z}$. Integrating the identified elasticities in (50) and (51), we obtain

$$\mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t) = d_l(l_t) + \Lambda_{lt}(x_t,z_t), \tag{52}$$

$$\bar{h}(x_{t-1},z_{t-1}) = d + \Lambda_{ht}(x_{t-1},z_{t-1}), \tag{53}$$

where function $d_l(l_t)$ and constant d are unknown objects to be identified; $\Lambda_{lt}(x_t,z_t)$ and $\Lambda_{ht}(x_{t-1},z_{t-1})$ are identified and thus treated as known functions.[20]

Identification of $\mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)$. Defining $H_{it} := \Lambda_{lt}(x_{it},z_{it}) - \Lambda_{ht}(x_{it-1},z_{it-1})$ as a known variable, we rewrite model (28) as

$$H_{it} = d - d_l(l_{it}) + \eta_{it}.$$

From $l_{t-1} \perp \eta_t$, we obtain the following moment condition for nonparametric instrument variable (IV) identification:

$$E\left[H_{it} - d + d_l(l_{it})|l_{it-1}\right] = 0. \tag{54}$$

[20]Specifically, $\Lambda_{lt}(x_t,z_t)$ and $\Lambda_{ht}(x_{t-1},z_{t-1})$ are given by

$$\Lambda_{lt}(x_t,z_t) := \int_{m_{t0}^*}^{m_t} \frac{\partial \mathbb{M}_t^{-1}(s,k_t,l_t,z_t)}{\partial m_t} ds + \int_{k_t^*}^{k_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*,s,l_t,z_t)}{\partial k_t} ds + \int_{z_t^*}^{z_t} \frac{\partial \mathbb{M}_t^{-1}(m_{t0}^*,k_t^*,l_t,s)}{\partial z_t} ds$$

$$\begin{aligned}\Lambda_{ht}(x_{t-1},z_{t-1}) := &\int_{m_{t-1}^*}^{m_{t-1}} \frac{\partial \bar{h}_t(s,k_{t-1},l_{t-1},z_{t-1})}{\partial m_{t-1}} ds + \int_{k_{t-1}^*}^{k_{t-1}} \frac{\partial \bar{h}_t(m_{t-1}^*,s,l_{t-1},z_{t-1})}{\partial k_{t-1}} ds \\ &+ \int_{l_{t-1}^*}^{l_{t-1}} \frac{\partial \bar{h}_t(m_{t-1}^*,k_{t-1}^*,s,z_{t-1})}{\partial l_{t-1}} ds + \int_{z_{t-1}^*}^{z_{t-1}} \frac{\partial \bar{h}_t(m_{t-1}^*,k_{t-1}^*,l_{t-1}^*,s)}{\partial z_{t-1}} ds\end{aligned}$$

For instance, if f_t is Cobb-Douglas as in (15), then $d_l(l_t) = -\theta_l\left(l_t - l_t^*\right)$ from (18), and the moment condition (54) becomes that for linear IV regression:

$$E\left[H_{it} - d - \theta_l(l_{it} - l_{it}^*)|l_{it-1}\right] = 0.$$

A standard procedure of linear IV regression identifies (d, θ_l) if l_{it} sufficiently correlates with l_{it-1}.

Following the literature on nonparametric IV (e.g, Newey and Powell, 2003), we assume that l_{t-1} satisfies the following completeness condition.

Assumption 12. *For all functions $\delta(l_t) : \mathscr{L} \to \mathbb{R}$ such that $E[\delta(l_t)|l_{t-1}] < \infty$, $E[\delta(l_t)|l_{t-1}] = 0$ a.s. implies $\delta(l_t) = 0$ a.s..*

With Assumption 12, the moment condition (54) uniquely identifies $\{d, d_l(l_t)\}$.[21] Since $E[\varepsilon_t|x_t, z_t] = 0$, step 1 continues to identify $\phi_t(\cdot)$. Therefore, once $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$ is identified, step 3 identifies all the same objects as before.

Proposition 7. *Suppose that l_t may correlate with η_t and that Assumptions 1–3, 5, 11 and 12 hold. Then, the production function, output quantities, output prices and TFP are identified up to scale and location; markups and output elasticities are identified up to scale.*

3.6.2 Endogenous Firm Characteristics

Firm characteristics z_t may correlate with η_t. For simplicity, we again assume that l_t is exogenous. We show that even in the absence of any IV for z_t, we can identify the markup and the production function. If valid IVs for z_t are available, all the same objects can be identified as before.

We modify Assumption 4 so that z_t may correlate with η_t.

Assumption 13. *(i) Assumptions 4 (a), (d), (e), and (f) hold. (ii) η_t is independent of $\bar{w}_t := (k_t, l_t, x_{t-1}, z_{t-1})' \in \bar{\mathscr{W}} := \mathscr{K} \times \mathscr{L} \times \mathscr{X} \times \mathscr{Z}$. $\bar{w}$ is continuously distributed on $\bar{\mathscr{W}}$. (iii) For each $(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z}$, $\mathscr{A}_{m_t}(x_{t-1}, z_{t-1}) = \{(\tilde{x}_t, \tilde{z}_t) \in \mathscr{X} \times \mathscr{Z} | \partial G_{m_t|v_t}(\tilde{m}_t|\tilde{k}_t, \tilde{l}_t, \tilde{z}_t, x_{t-1}, z_{t-1})/\partial m_t > 0\}$ is non-empty.*

Identification without Instrument Variables. The conditional distribution of m_t given v_t satisfies

[21] The proof is as follows. Suppose $\{\tilde{d}, \tilde{d}_l(l_{it})\}$ also satisfies the moment condition (54). Then, it holds that $E\left[\tilde{d} - d + \tilde{d}_l(l_{it}) - d_l(l_{it})|l_{it-1}\right] = 0$ a.s. The completeness condition implies $\tilde{d} - d + \tilde{d}_l(l_{it}) - d_l(l_{it}) = 0$ a.s. Since $\tilde{d}_l(l_t^*) = d_t(l_t^*) = 0$ from Assumption 3, $\tilde{d} = d$ holds so that $\tilde{d}_0(l_{it}) = d_0(l_{it})$.

$$G_{m_t|v_t}(m|v_t) = G_{\eta_t|z_t}\left(\mathbb{M}_t^{-1}(m, k_t, l_t, z_t) - \bar{h}_t\left(x_{t-1}, z_{t-1}\right)|z_t\right).$$

Taking the derivatives of both sides with respect to m, $q_t \in \{m_t, k_t, l_t\}$ and $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$, we obtain (50) and (51). Following the same steps as in subsection 3.6.1, we identify $\partial\mathbb{M}_t^{-1}(m, k_t, l_t, z_t)/\partial q_t$ and $\partial\bar{h}\left(x_{t-1}, z_{t-1}\right)/\partial q_{t-1}$ up to scale.

Since $E[\varepsilon_t|x_t, z_t] = 0$, Lemma 1 continues to hold and $\phi_t(\cdot)$ is identified. Therefore, using (35) and the first-order condition (37) with the identified derivatives of $\mathbb{M}_t^{-1}(\cdot)$, it is possible to identify markup (39) and output elasticities (41) up to scale. Integrating the output elasticities, we can identify the production function, following (42).

Proposition 8. *Suppose that z_t may correlate with η_t and that Assumptions 2, 3, 5, and 13 hold. Then, we can identify the markup $\partial\varphi_t^{-1}(\bar{r}_{it}, z_{it})/\partial r_t$ of each firm up to scale and the production function $f_t(\cdot)$ up to scale and location.*

Applying Propositions 3 and Proposition 4, it is possible to identify the changes in markup and output elasticities overtime and the levels of markup and elasticities, respectively.

Identification with Instrument Variables. To identify $\varphi_t^{-1}(\cdot)$ and $\mathbb{M}_t^{-1}(\cdot)$, we need a set of IVs ζ_t for z_t. A candidate for ζ_t is z_{t-1} if z_{t-1} correlates with z_t.

Assumption 14. *(a) There exits a set of instruments ζ_t such that $E[\eta_t|\zeta_t] = 0$ a.s. (b) For all functions $\delta(z_t) : \mathscr{Z} \to \mathbb{R}$ such that $E[\delta(z_t)|\zeta_t] < \infty$, $E[\delta(z_t)|\zeta_t] = 0$ a.s. implies $\delta(z_t) = 0$ a.s.*

Following similar steps by which to derive (53), we obtain

$$\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) = d_z(z_t) + \Lambda_{zt}\left(x_t, z_t\right)$$

and (53), where $d_z(z_t) := \mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, l_t^*, z_t)$ is an unknown function to be identified; $\Lambda_{zt}\left(x_t, z_t\right)$ is identified and treated as a known function.[22] Defining $H_{it}^{zh} := \Lambda_{zt}\left(x_{it}, z_{it}\right) - \Lambda_{ht}\left(x_{it-1}, z_{it-1}\right)$ as a known variable, we rewrite model (28) as

$$H_{it}^{zh} = d - d_z(z_{it}) + \eta_{it}.$$

From Assumption 14, the moment condition, $E[\eta_{it}|\zeta_{it}] = E\left[H_{it}^{zh} - d + d_z(z_{it})|\zeta_{it}\right] = 0$, identifies $\{d, d_z(z_t)\}$.

[22]Specifically, $\Lambda_{zt}\left(x_t, z_t\right)$ is given by

$$\Lambda_{zt}\left(x_t, z_t\right) := \int_{m_{t0}^*}^{m_t} \frac{\partial\mathbb{M}_t^{-1}(s, k_t, l_t, z_t)}{\partial m_t} ds + \int_{k_t^*}^{k_t} \frac{\partial\mathbb{M}_t^{-1}(m_{t0}^*, s, l_t, z_t)}{\partial k_t} ds + \int_{l_t^*}^{l_t} \frac{\partial\mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, s, z_t)}{\partial l_t} ds.$$

Proposition 9. *Suppose that z_t may correlate with η_t and a set of IVs ζ_t satisfies Assumption 14. Suppose Assumptions 2, 3, 5, and 13 hold. Then, we can identify $\varphi_t^{-1}(\cdot)$ and $\mathbb{M}_t^{-1}(\cdot)$ up to scale and location and identify $G_\eta(\cdot)$ up to scale. That is, output quantities, output prices and TFP are identified up to scale and location.*

3.6.3 Alternative Settings

The Appendix presents the identification results in three alternative settings. The identification argument remains the same but requires some additional steps.

Discrete Firm Characteristics. Observable firm characteristics z_t may constitute a discrete variable. Appendix A.2 provides a proof.

Unobservable Firm-Level Demand-Shifter. The identification can incorporate an unobserved demand shifter ξ_{it}, which can be called quality. Let $y_{it}^{\dagger} := y_{it} + \xi_{it}$ and $p_{it}^{\dagger} := y_{it} - \xi_{it}$ be the quality-adjusted output and the quality-adjusted price, respectively. We consider the following inverse function and revenue function:

$$\begin{aligned} p_{it}^{\dagger} &= \psi_t\left(y_{it}^{\dagger}, z_{it}\right), \\ \bar{r}_{it} &= \varphi_t\left(y_{it}^{\dagger}, z_{it}\right) = \varphi_t\left(f_t(x_t) + \omega_{it}^{\dagger}, z_{it}\right) \end{aligned} \tag{55}$$

where $\omega_{it}^{\dagger} \equiv \omega_{it} + \xi_{it}$ is a composite of TFP and quality. In Appendix A.3, we show that (55) derives from a representative consumer's maximization problem where $\exp(\xi_{it})$ enters the utility function in a multiplicative manner with quantity. In (55), higher quality allows a firm to earn more revenue for a given output. We assume that $\tilde{\omega}_{it}$ follows a first-order Markov process $\omega_{it}^{\dagger} = h\left(\omega_{it-1}^{\dagger}\right) + \eta_{it}$.

Under the current setting, the model structure becomes identical to the main model where $(p_{it}, y_{it}, \omega_{it})$ are replaced with $(p_{it}^{\dagger}, y_{it}^{\dagger}, \omega_{it}^{\dagger})$. Therefore, applying precisely the same steps, we can identify all functions identified in Section 3 and the quality-adjusted variables $(p_{it}^{\dagger}, y_{it}^{\dagger}, \omega_{it}^{\dagger})$.

IID Productivity Shock. As an alternative error structure, we consider an i.i.d. production shock e_{it} to output instead of a measurement error ε_{it}. Then, the firm's observed revenue r_{it} and inputs x_{it} are related as follows:

$$r_{it} = \varphi_t(f_t(x_{it}) + \omega_{it} + e_{it}, z_{it}). \tag{56}$$

A firm chooses m_{it} at time t by maximizing the expected profit:

$$\begin{aligned} m_{it} &= \mathbb{M}_t(\omega_{it}, k_{it}, l_{it}, z_{it}) \\ &:= \underset{m \in \mathscr{M}}{\operatorname{argmax}}\ E\left[\exp\left(\varphi_t(f_t(m, k_{it}, l_{it}) + \omega_{it} + e_{it}, z_{it})\right) | \mathscr{I}_{it}\right] - \exp(p_t^m + m). \end{aligned}$$

where $\mathscr{I}_{it}$ is the set of information for the firm that includes all past variables and all time$-t$ variables except e_{it}. The identification of the control function $\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ remains the same because $\mathbb{M}_t^{-1}(\cdot)$ continues to be a function of the same variables.

In the second step, the revenue function (56) is written as:

$$\varphi_t^{-1}(r_{it}, z_{it}) = f_t(x_{it}) + \mathbb{M}_t^{-1}(x_{it}, z_{it}) + e_{it}. \tag{57}$$

Model (57) also belongs to the class of transformation models studied by Chiappori et al. (2015). Therefore, by applying the nonparametric identification of a transformation model and using the first-order condition for the material, we can identify $\varphi_t(\cdot)$ and $f_t(\cdot)$ up to scale and location from the conditional distribution of r_{it} given (x_{it}, z_{it}) under the assumptions similar to those for Proposition 2. As an additional complication, the first-order condition includes expectation with respect to e_t. Therefore, we first identify the distribution of e_t to derive the first-order condition. Appendix A.4 provides a proof.

Because of the i.i.d. shock e_{it}, the realized value of $\partial \varphi_t^{-1}(r_{it}, z_{it})/\partial r_t$ no longer equals the markup. We identify the markup from the cost minimization, following Hall (1988) and De Loecker and Warzynski (2012). As shown in Appendix A.4, the equation for the markup μ_{it} becomes

$$\mu_{it} = \frac{\partial f_t(x_{it})/\partial m_{it}}{\exp(p_t^m + m_{it})/\exp(r_{it} - e_{it})}.$$

The difference from the original Hall-De Loecker-Warzynski markup equation (38) is $\exp(r_{it} - e_{it})$ instead of $\exp(\bar{r}_{it}) = \exp(r_{it} - \varepsilon_{it})$. While $\bar{r}_t = \phi_t(x_t, z_t)$ in (38) is a deterministic function of (x_t, z_t), $r_t - e_t$ is generally not. Therefore, the markups are different across firms even after being conditioned on (x_t, z_t).

4 Concluding Remarks

The current study developes constructive nonparametric identification of production function and markup from revenue data. Our method simultaneously addresses two fundamental identification issues raised in the literature of production function estimation since Marschak and Andrews (1944)—namely, correlations between inputs and TFP, and biases from markup heterogeneity when revenue is used as output. Under standard assumptions, when revenue is modeled as a function of output (rather than a mere proxy for output) and firm's observed characteristics, various economic objects of interest can be identified from revenue data. In an ongoing follow-up research, we provide an estimation procedure and plan to estimate these objects from an actual dataset.

References

Ackerberg, D., Benkard, C. L., Berry, S., and Pakes, A. (2007), "Chapter 63 Econometric Tools for Analyzing Market Outcomes," Elsevier, vol. 6 of *Handbook of Econometrics*, pp. 4171 – 4276. 1

Ackerberg, D. A., Caves, K., and Frazer, G. (2015), "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411–2451. 1, 3.1, 3.2, 3.3.4, 15, 3.6.1, 19

Arellano, M. and Bond, S. (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58, 277–297. 3.6.1

Arellano, M. and Bover, O. (1995), "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68, 29–51. 3.6.1

Bartelsman, E. J. and Doms, M. (2000), "Understanding productivity: Lessons from longitudinal microdata," *Journal of Economic literature*, 38, 569–594. 1

Berry, S., Levinsohn, J., and Pakes, A. (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890. 1, 6

Blundell, R. and Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115–143. 3.6.1

— (2000), "GMM estimation with persistent panel data: an application to production functions," *Econometric Reviews*, 19, 321–340. 3.6.1

Bond, S., Hashemi, A., Kaplan, G., and Zoch, P. (2020), "Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data," NBER Working Paper w27002. 1, 2, 7

Chiappori, P.-A., Komunjer, I., and Kristensen, D. (2015), "Nonparametric Identification and Estimation of Transformation Models," *Journal of Econometrics*, 188, 22–39. 1, 3.1, 3.1, 3.3.2, 14, 3.3.2, 3.3.4, 3.6.3, A.4.2

De Loecker, J. (2011), "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79, 1407–1451. 1, 4, 2, 2, 3.1, 16, 3.5

De Loecker, J., Eeckhout, J., and Unger, G. (2020), "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 135, 561–644. 3

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016), "Prices, Markups, and Trade Reform," *Econometrica*, 84, 445–510. 2, 2

De Loecker, J. and Warzynski, F. (2012), "Markups and Firm-Level Export Status," *American Economic Review*, 102, 2437–71. 1, 2, 7, 3.1, 3.6.3, A.4.3

Doraszelski, U. and Jaumandreu, J. (2018), "Measuring the Bias of Technological Change," *Journal of Political Economy*, 126, 1027–1084. 1

Ekeland, I., Heckman, J. J., and Nesheim, L. (2004), "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112, S60–S109. 1, 6

Feenstra, R. C. (2003), "A homothetic utility function for monopolistic competition models, without constant price elasticity," *Economics Letters*, 78, 79–86. 3.5

Feenstra, R. C. and Weinstein, D. E. (2017), "Globalization, markups, and US welfare," *Journal of Political Economy*, 125, 1040–1074. 3.5

Flynn, Z., Gandhi, A., and Traina, J. (2019), "Measuring Markups with Production Data," Unpublished. 5

Foster, L., Haltiwanger, J., and Syverson, C. (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98, 394–425. 1, 2

Gandhi, A., Navarro, S., and Rivers, D. A. (2020), "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, forthcoming. 1, 4, 3.2, 3.3.3, 3.3.4

Griliches, Z. and Mairesse, J. (1999), "Production Functions: The Search for Identification," in *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, ed. Strøm, S., Cambridge University Press, Econometric Society Monographs, pp. 169–203. 1

Hall, R. E. (1988), "The Relation between Price and Marginal Cost in US Industry," *Journal of Political Economy*, 96, 921–947. 1, 2, 3.6.3, A.4.3

Heckman, J. J., Matzkin, R. L., and Nesheim, L. (2010), "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, 78, 1569–1591. 6

Horowitz, J. L. (1996), "Semiparametric Estimation of a Regression Model with an Unknown Transformation of the Dependent Variable," *Econometrica*, 103–137. 1

Katayama, H., Lu, S., and Tybout, J. R. (2009), "Firm-Level Productivity Studies: Illusions and a Solution," *International Journal of Industrial Organization*, 27, 403–413. 1, 2, 8

Klette, T. J. and Griliches, Z. (1996), "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous," *Journal of Applied Econometrics*, 11, 343–361. 1, 4, 2, 2, 3.1, 16, 3.5

Levinsohn, J. and Petrin, A. (2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 317–341. 1, 3.1

Lu, Y. and Yu, L. (2015), "Trade liberalization and markup dispersion: evidence from China's WTO accession," *American Economic Journal: Applied Economics*, 7, 221–53. 2, 2

Marschak, J. and Andrews, W. (1944), "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12, 143–205. 1, 2, 3.1, 3.1, 4

Matsuyama, K. and Ushchev, P. (2017), "Beyond CES: Three Alternative Cases of Flexible Homothetic Demand Systems," Buffett Institute Global Poverty Research Lab Working Paper 17-109. 1, 3.5, 18, 3.5, A.1.1, A.1, A.1.1, A.1.1

Mrázová, M. and Neary, J. P. (2017), "Not So Demanding: Demand Structure and Firm Behavior," *American Economic Review*, 107, 3835–74. 3.5, 18

— (2019), "IO for Exports(s)," *International Journal of Industrial Organization*, 102561. 3.5, 18

Newey, W. K. and Powell, J. L. (2003), "Instrumental variable estimation of nonparametric models," *Econometrica*, 71, 1565–1578. 3.6.1

Nishioka, S. and Tanaka, M. (2019), "Measuring Markups from Revenue and Total Cost: An Application to Japanese Plant-Product Matched Data," Rieti Discussion Paper Series 19-E-018. 2, 2

Olley, G. S. and Pakes, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1263–1297. 1

Syverson, C. (2011), "What Determines Productivity?" *Journal of Economic Literature*, 49, 326–65. 1

Van Biesebroeck, J. (2003), "Productivity dynamics with technology choice: An application to automobile assembly," *Review of Economic Studies*, 70, 167–198. 1

A Online Appendix (Not for Publication)

A.1 Identification of Demand Function

A.1.1 Proof for Proposition 6

The proof for Proposition 6 uses the following result of Matsuyama and Ushchev (2017).

Theorem A.1. *(Matsuyama and Ushchev, 2017, Remark 3 and Proposition 1). Consider a mapping* $\mathbf{s}(\mathbf{Y}) := (s_1(Y_1), ..., s_N(Y_N))'$ *from* $\mathbb{R}_+^N$ *to* $\mathbb{R}_+^N$*, which is differentiable almost everywhere, is normalized by*

$$\sum_{i=1}^{N} s_i(Y_i^*) = 1, \tag{A.1}$$

for some point $\mathbf{Y}^* := (Y_1^*, ..., Y_N^*)$ *and satisfies the following conditions*

$$\begin{aligned} s_i'(Y_i)Y_i &< s_i(Y_i) \text{ for } i = 1, ..., N, \\ s_i'(Y_i)s_j'(Y_j) &\geq 0 \text{ for } i, j = 1, ..., N, \end{aligned} \tag{A.2}$$

for all $\mathbf{Y}$ *such that* $\sum_{i=1}^{N} s_i(Y_i) = 1$*. Then, (1) for any such mapping, there exists a unique monotone, convex, continuous, and homothetic rational preference that generates the HSA demand system described by*

$$P_i = \frac{\Phi}{Y_i} s_i\left(\frac{Y_i}{A(\mathbf{Y})}\right) \text{ for } i = 1, .., N,$$

where $\Phi := \sum_{i=1}^{N} P_i Y_i$ *and* $A(\mathbf{Y})$ *is obtained by solving*

$$\sum_{i=1}^{N} s_i\left(\frac{Y_i}{A(\mathbf{Y})}\right) = 1.$$

(2) This homothetic preference is described by a utility function U *which is defined by*

$$\ln U(\mathbf{Y}) = \ln A(\mathbf{Y}) + \sum_{i=1}^{N} \int_{c}^{Y_i/A(Y)} \frac{s_i(\xi)}{\xi} d\xi, \tag{A.3}$$

where c *is a constant.*

Matsuyama and Ushchev (2017) proved (1) from the Antonelli's integrability theorem. See their paper for the proof. Matsuyama and Ushchev (2017) provides a proof for (2) for the case of direct demand functions instead of inverse demand functions considered here. So we will provide the proof for (2) in the following proof for Proposition 6 (b).

Proof for Proposition 6

Proof. (a) We construct $S_t(Y_i/A_t(\mathbf{Y}, z_t), z_{it})$ and $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ as is explained in the main text. Fix $\mathbf{z}_t := (z_{1t}, ..., z_{Nt})$ and time t. For $\mathbf{Y} \in \bar{\mathscr{Y}}$, define $A(\mathbf{Y}) := A_t(\mathbf{Y}, \mathbf{z}_t)$ and $s(\mathbf{Y}) := (s_1(Y_1), ..., s_N(Y_N))$ such that $s_i(Y_i) = S_t(Y_i, z_{it})$.

Define $\bar{\mathscr{Y}}_A := \{\mathbf{Y}/A(\mathbf{Y}) : \mathbf{Y} \in \bar{\mathscr{Y}}\}$. Then, for all $\mathbf{Y} \in \bar{\mathscr{Y}}_A$, $\sum_{i=1}^N s_i(Y_i) = 1$ holds by construction of $A_t(\cdot)$. At the same time, for all $\mathbf{Y}$ that satisfies $\sum_{i=1}^N s_i(Y_i) = 1$, $A(\mathbf{Y}) = 1$ holds so that $\mathbf{Y} \in \bar{\mathscr{Y}}_A$. Therefore, $\bar{\mathscr{Y}}_A = \{\mathbf{Y} \in \bar{\mathscr{Y}} : \sum_{i=1}^N s_i(Y_i) = 1\}$.

Consider $\mathbf{Y} \in \mathscr{Y}_A$. From Assumption 1 (b) and $y := \ln Y$,

$$0 < \frac{\partial \varphi_t(\ln Y, z)}{\partial \ln Y} = 1 + \frac{\partial \psi_t(\ln Y, z)}{\partial \ln Y} < 1$$

holds. The above inequality implies

$$s_i'(Y) > 0 \text{ and } s_i'(Y)Y < s_i(Y) \text{ for all } i \text{ and } Y$$

because

$$\begin{aligned} s_i'(Y)Y &= \frac{\exp(\varphi_t(\ln Y, z_{it}))}{\Phi_t} \frac{\partial \varphi_t(\ln Y, z_{it})}{\partial \ln Y} \\ &= s_i(Y) \frac{\partial \varphi_t(\ln Y, z_{it})}{\partial \ln \tilde{Y}}. \end{aligned}$$

Therefore, $\mathbf{s}(\mathbf{Y})$ satisfies the inequalities in (A.2) for all $\mathbf{Y}$ satisfying $\sum_{i=1}^N s_i(Y_i) = 1$. From Theorem A.1 (1), there exists a unique monotone, convex, continuous, and homothetic rational preference that generates

$$\begin{aligned} P_{it} &= \frac{\Phi_t}{Y_{it}} s_i \left(\frac{Y_{it}}{A(\mathbf{Y}_t)} \right) \\ &= \frac{\Phi_t}{Y_{it}} S_t \left(\frac{Y_{it}}{A(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it} \right), \end{aligned}$$

where Φ_t is the consumer's budget.

(b) The following derivation of the utility function follows the steps in Matsuyama and Ushchev (2017). Let $U_t(\mathbf{Y}_t, \mathbf{z}_t)$ be the utility function that is homogenous of degree one with respect to $\mathbf{Y}_t$. Then, the indirect utility is linear in income Φ_t:

$$V_t(\mathbf{P}_t, \Phi_t) = \max_{\mathbf{Y}_t} \{U_t(\mathbf{Y}_t, \mathbf{z}_t) | \mathbf{P}_t' \mathbf{Y}_t \leq \Phi_t\} = \frac{\Phi_t}{\Pi_t(\mathbf{P}_t)}, \tag{A.4}$$

where $\Pi_t(\mathbf{P}_t)$ is the ideal price index. The first-order condition is given by

$$\frac{\partial U_t(\mathbf{Y}_t, \mathbf{z}_t)}{\partial Y_{it}} = \lambda_t P_{it},$$

where $\lambda_t = 1/\Pi_t(\mathbf{P}_t)$ is the Lagrange multiplier. The Roy's identity derives the demand for firm i as

$$Y_{it} = -\frac{\partial V_t/\partial P_{it}}{\partial V_t/\partial \Phi_t} = \frac{\Phi_t}{P_{it}}\left(\frac{\partial \Pi_t}{\partial P_{it}}\frac{P_{it}}{\Pi_t}\right). \tag{A.5}$$

From (A.4), the expenditure function is written as $e_t(\mathbf{P}_t, U_t) = \Pi_t(\mathbf{P}_t)U_t$. Applying the Shepard's lemma derives the demand for firm i as

$$Y_{it} = \frac{\partial e_t(\mathbf{P}_t, U_t)}{\partial P_{it}} = \frac{\partial \Pi_t}{\partial P_{it}}U_t. \tag{A.6}$$

Using (A.6), $\lambda_t = 1/\Pi_t$ and the first-order condition, we obtain

$$\frac{\partial \Pi_t}{\partial P_{it}}\frac{P_{it}}{\Pi_t} = \frac{Y_{it}}{U_t}\frac{P_{it}}{\Pi_t} = \frac{Y_{it}}{U_t}\lambda_t P_{it} = \frac{\partial U_t}{\partial Y_{it}}\frac{Y_{it}}{U_t}.$$

Therefore, from (A.5), we have

$$S_t\left(\frac{Y_{it}}{A(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it}\right) = \frac{P_{it}Y_{it}}{\Phi_t} = \frac{\partial \Pi_t}{\partial P_{it}}\frac{P_{it}}{\Pi_t} = \frac{\partial U_t}{\partial Y_{it}}\frac{Y_{it}}{U_t},$$

which can be written as

$$\frac{\partial \ln U_t(\mathbf{Y}_t, \mathbf{z}_t)}{\partial Y_{it}} = \frac{1}{Y_{it}}S_t\left(\frac{Y_{it}}{A_t(\mathbf{Y}_t, \mathbf{z}_t)}, z_{it}\right). \tag{A.7}$$

Let $A_t = A_t(\mathbf{Y}_t, \mathbf{z}_t)$. Since $U_t(\mathbf{Y}_t, \mathbf{z}_t)$ is homogeneous of degree one with respect to $\mathbf{Y}_t$, $\partial U_t(\mathbf{Y}_t, \mathbf{z}_t)/\partial Y_{it}$ is homogenous of degree zero with respect to $\mathbf{Y}_t$. Therefore, it holds

$$\begin{aligned}\frac{\partial \ln U_t(\mathbf{Y}_t/A_t, \mathbf{z}_t)}{\partial Y_{it}} &= \frac{\partial U_t(\mathbf{Y}_t/A_t, \mathbf{z}_t)}{\partial Y_{it}}\frac{1}{U_t(\mathbf{Y}_t/A_t, \mathbf{z}_t)}\\ &= \frac{\partial U_t(\mathbf{Y}_t, \mathbf{z}_t)}{\partial Y_{it}}\frac{A_t}{U_t(\mathbf{Y}_t, \mathbf{z}_t)}\\ &= A_t\frac{\partial \ln U_t(\mathbf{Y}_t, \mathbf{z}_t)}{\partial Y_{it}}.\end{aligned}$$

Then, (A.7) becomes simplified as

$$\begin{aligned} \frac{\partial \ln U_t(\mathbf{Y}_t,\mathbf{z}_t)}{\partial Y_{it}} &= \frac{1}{Y_{it}} S_t\left(\frac{Y_{it}}{A_t}, z_{it}\right) \\ \Leftrightarrow \frac{\partial \ln U_t(\mathbf{Y}_t/A_t,\mathbf{z}_t)}{\partial Y_{it}} &= \frac{A_t}{Y_{it}} S_t\left(\frac{Y_{it}}{A_t}, z_{it}\right) \\ \Leftrightarrow \frac{\partial \ln U_t(\tilde{\mathbf{Y}}_t,\mathbf{z}_t)}{\partial \tilde{Y}_{it}} &= \frac{S_t\left(\tilde{Y}_{it}, z_{it}\right)}{\tilde{Y}_{it}}, \end{aligned} \tag{A.8}$$

where $\tilde{Y}_{it} := Y_{it}/A_t$ and $\tilde{\mathbf{Y}}_t := \left(\tilde{Y}_{1t}, ..., \tilde{Y}_{Nt}\right)$. Let $c_t(\mathbf{z}_t) := (c_{1t}(\mathbf{z}_t), ..., c_{Nt}(\mathbf{z}_t))$ be defined by $U_t(c_t(\mathbf{z}_t), \mathbf{z}_t) = 1$. Then, integration of (A.8) leads to

$$\ln U_t(\tilde{\mathbf{Y}}_t, \mathbf{z}_t) = \sum_{i=1}^{N} \int_{c_{it}(\mathbf{z}_t)}^{\tilde{Y}_{it}} \frac{S_t\left(\xi, z_{it}\right)}{\xi} d\xi.$$

Since $\ln U_t(\tilde{\mathbf{Y}}_t, \mathbf{z}_t) = \ln U_t(\mathbf{Y}_t/A_t, \mathbf{z}_t) = \ln U_t(\mathbf{Y}_t, \mathbf{z}_t) - \ln A_t$, we obtain the utility function stated in the proposition as follows:

$$\ln U_t(\mathbf{Y}_t, \mathbf{z}_t) = \ln A_t\left(\mathbf{Y}_t, \mathbf{z}_t\right) + \sum_{i=1}^{N} \int_{c_{it}(\mathbf{z}_t)}^{Y_{it}/A_t(\mathbf{Y}_t,\mathbf{z}_t)} \frac{S_t\left(\xi, z_{it}\right)}{\xi} d\xi.$$

(c) The homothetic preference implies that the market share $P_{it}Y_{it}/\Phi_t$ depends only on a price vector and is independent of income. This property requires $A_t\left(\mathbf{Y}_t, \mathbf{z}_t\right)$ to be homogenous of degree one with respect to $\mathbf{Y}_t$ so that for any $k > 0$, it

$$S_t\left(\frac{kY_{it}}{A_t\left(k\mathbf{Y}_t, \mathbf{z}_t\right)}, z_{it}\right) = S_t\left(\frac{kY_{it}}{kA_t\left(\mathbf{Y}_t, \mathbf{z}_t\right)}, z_{it}\right) = S_t\left(\frac{Y_{it}}{A_t\left(\mathbf{Y}_t, \mathbf{z}_t\right)}, z_{it}\right).$$

Let $\varphi_t^{-1}(\bar{r}_{it}, z_{it})$ be the identified log output and $\varphi_t^{*-1}(\bar{r}_{it}, z_{it})$ be its true value. Since $\varphi_t^{-1}(\bar{r}_{it}, z_{it})$ is identified up to location, there is $a \in \mathbb{R}$ such that $\varphi_t^{-1}(\bar{r}_{it}, z_{it}) = a + \varphi_t^{*-1}(\bar{r}_{it}, z_{it})$.

The identified output Y_{it} and the true output Y_{it}^* are related as follows:

$$\begin{aligned} Y_{it} &= \exp(\varphi_t^{-1}(\bar{r}_{it}, z_{it})) \\ &= \exp(a + \varphi_t^{*-1}(\bar{r}_{it}, z_{it})) \\ &= \exp(a) Y_{it}^*. \end{aligned}$$

Since $\varphi_t(y_t, z_t) = \varphi_t^*(y_t - a, z_t)$ for all y_t and z_t,

$$\varphi_t\left(\ln Y_{it}, z_{it}\right) = \varphi_t^*\left(\ln Y_{it} - a, z_{it}\right) = \varphi_t^*\left(\ln Y_{it}^*, z_{it}\right).$$

Then, the market share function $S_t(Y_{it}, z_{it}) := \exp(\varphi_t(\ln Y_{it}, z_{it}))/\Phi_t$ constructed from the identified outputs agrees with the market share function $S_t^*(Y_{it}^*, z_{it}) := \exp\left(\varphi_t^*\left(\ln Y_{it}^*, z_{it}\right)\right)/\Phi_t$ constructed from the true outputs:

$$S_t(Y_{it}, z_{it}) = \frac{\exp(\varphi_t(\ln Y_{it}, z_{it}))}{\Phi_t} = \frac{\exp\left(\varphi_t^*\left(\ln Y_{it}^*, z_{it}\right)\right)}{\Phi_t} = S_t^*(Y_{it}^*, z_{it}).$$

Thus, the identified demand system does not depend on the location normalization of $\varphi^{-1}(\cdot)$.

Since the quantity index $A_t(\mathbf{Y}_t, \mathbf{z}_t)$ is homogenous of degree one with respect to $\mathbf{Y}_t$,

$$\frac{Y_{it}}{A(\mathbf{Y}_t, \mathbf{z}_t)} = \frac{\exp(a)Y_{it}^*}{A(\exp(a)\mathbf{Y}_t^*, \mathbf{z}_t)} = \frac{\exp(a)Y_{it}^*}{\exp(a)A(\mathbf{Y}_t^*, \mathbf{z}_t)} = \frac{Y_{it}^*}{A(\mathbf{Y}_t^*, \mathbf{z}_t)}$$

Let $U_t(\mathbf{Y}_t, \mathbf{z}_t)$ be the identified utility and $U_t^*(\mathbf{Y}_t^*, \mathbf{z}_t)$ be the true utility. Then, they are related as

$$\begin{aligned}
\ln U_t(\mathbf{Y}_t, \mathbf{z}_t) &= \ln A_t(\mathbf{Y}_t, \mathbf{z}_t) + \sum_{i=1}^{N} \int_{c_i(\mathbf{z}_t)}^{Y_{it}/A_t(\mathbf{Y}_t, \mathbf{z}_t)} \frac{S_t(\xi, z_{it})}{\xi} d\xi, \\
&= a + \ln A_t\left(\mathbf{Y}_t^*, \mathbf{z}_t\right) + \sum_{i=1}^{N} \int_{c_i^*(\mathbf{z}_t)}^{Y_{it}^*/A_t\left(\mathbf{Y}_t^*, \mathbf{z}_t\right)} \frac{S_t^*(\xi, z_{it})}{\xi} d\xi \\
&= a + \ln U_t^*(\mathbf{Y}_t^*, \mathbf{z}_t),
\end{aligned}$$

where $c_t^*(\mathbf{z}_t) := \left(c_{1t}^*(\mathbf{z}_t), ..., c_{Nt}^*(\mathbf{z}_t)\right)$ defined by $U^*(c_t^*(\mathbf{z}_t), \mathbf{z}_t) = 1$. Therefore, the log utility function is identified up to the location normalization of $\varphi_t^{-1}(\cdot)$. The identified utility function is a monotonic transformation of the true utility function, which implies both utility functions represent the same consumer preference. □

A.2 Discrete Firm Characteristics z_t

This section proves Propositions 1 and 2 for the case that z_{it} is a discrete variable and have finite support $\mathscr{Z} := \{z^1, ..., z^J\}$.

The following assumption modifies Assumption 1 for discrete z_{it}.

Assumption A.1. *(a) $f_t(\cdot)$ is continuously differentiable with respect to (m,k,l) on $\mathscr{M} \times \mathscr{K} \times \mathscr{L}$ and strictly increasing in m. (b) For every $z \in \mathscr{Z}$, $\varphi_t(\cdot, z)$ is strictly increasing and invertible with its inverse $\varphi_t^{-1}(\bar{r}, z)$, which is continuously differentiable with respect to $\bar{r}$ on $\bar{\mathscr{R}}$. (c) For every $(k,l,z) \in \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$, $\mathbb{M}_t(\cdot, k, l, z)$ is strictly increasing and invertible with its inverse $\mathbb{M}_t^{-1}(m,k,l,z)$, which is continuously differentiable with respect to (m,k,l) on $\mathscr{M} \times \mathscr{K} \times \mathscr{L}$. (d) ε_t is mean independent of x_t and z_t with $E[\varepsilon_t | x_t, z_t] = 0$.*

The following assumption modifies Assumption 4 for discrete z_{it}.

Assumption A.2. *(a) The distribution $G_\eta(\cdot)$ of η is absolutely continuous with a density function $g_\eta(\cdot)$ that is continuous on its support. (b) η_t is independent of $v_t := (k_t, l_t, z_t, x_{t-1}, z_{t-1})' \in \mathscr{V} := \mathscr{K} \times \mathscr{L} \times \mathscr{Z} \times \mathscr{X} \times \mathscr{Z}$. (c) x is continuously distributed on $\mathscr{X}$. (d) The support Ω of ω is an interval $[\underline{\omega}, \bar{\omega}] \subset \mathbb{R}$ where $\underline{\omega} < 0$ and $1 < \bar{\omega}$. (e) $h(\cdot)$ is continuously differentiable with respect to ω on Ω. (f) The set $\mathscr{A}_{q_{t-1}} := \{(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z} : \partial G_{m_t|v_t}(m_t|v_t)/\partial q_{t-1} \neq 0$ for all $(m_t, k_t, l_t, z_t) \in \mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}\}$ is nonempty for some $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$. (g) For each $(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z}$, it is possible to find $(x_t, z_t) \in \mathscr{X} \times \mathscr{Z}$ such that $\partial G_{m_t|v_t}(m_t|k_t, l_t, z_t, x_{t-1}, z_{t-1})/\partial m_t > 0$.*

A sufficient condition for Assumption A.2 (g) is $g_\eta(\eta) > 0$ for all $\eta \in \mathbb{R}$, under which (A.10) below shows $\partial G_{m_t|v_t}(m_t|k_t, l_t, z_t, x_{t-1}, z_{t-1})/\partial m_t > 0$ holds for all (x_t, z_t).

The following proposition establishes the identification of $\mathbb{M}_t^{-1}(\cdot)$.

Proposition A.1. *Suppose that Assumptions 2, 3, A.1, and A.2 hold. Then, we can identify $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$ up to scale and location, and identify $G_\eta(\cdot)$ up to scale.*

Proof. Choose normalization points (m_{t1}^*, k_t^*, l_t^*) and (m_{t0}^*, k_t^*, l_t^*) in Assumption 3 as well as $x_{t-1}^* \in \mathscr{X}$ such that, for $z_t, z_{t-1} \in \mathscr{Z}$,

$$\mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, l_t^*, z_t) = c_0(z_t),\ \mathbb{M}_t^{-1}(m_{t1}^*, k_t^*, l_t^*, z_t) = c_1(z_t), \text{ and } \bar{h}(x_{t-1}^*, z_{t-1}) = c_2(z_{t-1}), \tag{A.9}$$

where $\{c_0(z_t), c_1(z_t), c_2(z_{t-1})\}_{z_t, z_{t-1} \in \mathscr{Z}}$ are unknown constants. Without loss of generality, let z_t^* in Assumption 3 be $z_t^* = z^1$. Thus, the normalization in Assumption 3 is imposed as

$$c_0(z^1) = 0 \text{ and } c_1(z^1) = 1.$$

From $\eta_t \perp v_t$, the conditional distribution of m_t given v_t satisfies

$$G_{m_t|v_t}(m_t|v_t) = G_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right).$$

Taking the derivatives of $G_{m_t|v_t}(m_t|v_t)$ with respect to $q_t \in \{m_t, k_t, l_t\}$ and $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}\}$. The derivatives of $G_{m_t|v_t}(m|v)$ are

$$\frac{\partial G_{m_t|v_t}(m_t|v_t)}{\partial q_t} = \frac{\partial \mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)}{\partial q_t} g_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right), \tag{A.10}$$

$$\frac{\partial G_{m_t|v_t}(m_t|v_t)}{\partial q_{t-1}} = -\frac{\partial \bar{h}(x_{t-1}, z_{t-1})}{\partial q_{t-1}} g_\eta\left(\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1})\right). \tag{A.11}$$

Using Assumption A.2 (f), we can choose $q_{t-1} \in \{k_{t-1}, l_{t-1}, m_{t-1}, z_{t-1}\}$ and $(\tilde{x}_{t-1}, \tilde{z}_{t-1}) \in \mathscr{A}_{q_{t-1}}$ such that $\partial G_{m_t|v_t}(m_t|k_t, l_t, z_t, \tilde{x}_{t-1}, \tilde{z}_{t-1})/\partial q_{t-1} \neq 0$ for all $(m_t, k_t, l_t, z_t) \in \mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$.

Dividing (A.10) by (A.11), respectively, we obtain for $q_t \in \{m_t, k_t, l_t\}$

$$\frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial q_t} = -\frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}} \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}. \tag{A.12}$$

Then, from (A.9) and (A.14), we have

$$\begin{aligned} 1 &= c_1(z^1) - c_0(z^1) \\ &= \mathbb{M}_t^{-1}(m_{t1}^*, k_t^*, l_t^*, z^1) - \mathbb{M}_t^{-1}(m_{t0}^*, k_t^*, l_t^*, z^1) \\ &= -\frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}} \int_{m_{t0}^*}^{m_{t1}^*} \frac{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z^1,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial m_t}{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z^1,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial q_{t-1}} dm_t \end{aligned}$$

and therefore identify $\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}$ as

$$\frac{\partial \bar{h}(\tilde{x}_{t-1},\tilde{z}_{t-1})}{\partial q_{t-1}} = -\tilde{S}_{q_{t-1}}, \tag{A.13}$$

where

$$\tilde{S}_{q_{t-1}} := \left(\int_{m_{t0}^*}^{m_{t1}^*} \frac{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z^1,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial m_t}{\partial G_{m_t|v_t}\left(m|k_t^*,l_t^*,z^1,\tilde{x}_{t-1},\tilde{z}_{t-1}\right)/\partial q_{t-1}} dm_t \right)^{-1}.$$

By substituting (A.13) into (A.12), we can identify $\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)/\partial m_t$ and $\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)/\partial q_t$ as

$$\begin{aligned} \frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial m_t} &= \tilde{S}_{q_{t-1}} T_{m_t q_{t-1}}(x_t,z_t), \\ \frac{\partial \mathbb{M}_t^{-1}(m_t,k_t,l_t,z_t)}{\partial q_t} &= \tilde{S}_{q_{t-1}} T_{q_t q_{t-1}}(x_t,z_t), \end{aligned} \tag{A.14}$$

where

$$\begin{aligned} T_{m_t q_{t-1}}(x_t,z_t) &:= \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial m_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}, \\ T_{q_t q_{t-1}}(x_t,z_t) &:= \frac{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_t}{\partial G_{m_t|v_t}(m_t|k_t,l_t,z_t,\tilde{x}_{t-1},\tilde{z}_{t-1})/\partial q_{t-1}}. \end{aligned}$$

From (A.9) and (A.14), $\mathbb{M}_t^{-1}(x_t,z_t)$ is written as

$$\mathbb{M}_t^{-1}(x_t,z_t) = c_0(z_t) + \Lambda_m(x_t,z_t), \tag{A.15}$$

where

$$\Lambda_m(x_t, z_t) := \tilde{S}_{q_{t-1}} \left\{ \int_{m_{t0}^*}^{m_t} T_{m_t q_{t-1}}(s, k_t, l_t, z_t) ds + \int_{k^*}^{k_t} T_{k_t q_{t-1}}(m_{t0}^*, s, l_t, z_t) ds + \int_{l^*}^{l_t} T_{l_t q_{t-1}}(m_{t0}^*, k_t^*, s, z_t) ds \right\}.$$

From Assumption A.2 (g), for a given point $(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z}$, we can find some point $(\tilde{m}_t, \tilde{k}_t, \tilde{l}_t, \tilde{z}_t) \in \mathscr{X} \times \mathscr{Z}$ such that $\partial G_{m_t|v_t}\left(\tilde{m}_t | \tilde{k}_t, \tilde{l}_t, \tilde{z}_t, x_{t-1}, z_{t-1}\right) / \partial m > 0$. Dividing (A.11) by (A.10) identifies $\partial \bar{h}(x_{t-1}, z_{t-1}) / \partial q_{t-1}$ as

$$\frac{\partial \bar{h}(x_{t-1}, z_{t-1})}{\partial q_{t-1}} = -\frac{\partial G_{m_t|v_t}\left(\tilde{m}_t | \tilde{k}_t, \tilde{l}_t, \tilde{z}_t, x_{t-1}, z_{t-1}\right) / \partial q_{t-1}}{\partial G_{m_t|v_t}\left(\tilde{m}_t | \tilde{k}_t, \tilde{l}_t, \tilde{z}_t, x_{t-1}, z_{t-1}\right) / \partial m} \frac{\partial \mathbb{M}_t^{-1}(\tilde{m}_t, \tilde{k}_t, \tilde{l}_t, \tilde{z}_t)}{\partial m}.$$

Repeating this, we can identify $\partial \bar{h}(x_{t-1}, z_{t-1}) / \partial q_{t-1}$ for all $(x_{t-1}, z_{t-1}) \in \mathscr{X} \times \mathscr{Z}$. From (A.9) and (A.13), we can write $\bar{h}_t(x_{t-1}, z_{t-1})$ as

$$\bar{h}_t(x_{t-1}, z_{t-1}) = c_2(z_{t-1}) + \Lambda_{\bar{h}}(x_{t-1}, z_{t-1}) \tag{A.16}$$

with

$$\Lambda_{\bar{h}}(x_{t-1}, z_{t-1}) := \int_{m_{t-1}^*}^{m_{t-1}} \frac{\partial \bar{h}_t(s, k_{t-1}, l_{t-1}, z_{t-1})}{\partial m_{t-1}} ds + \int_{k_{t-1}^*}^{k_{t-1}} \frac{\partial \bar{h}_t(m_{t-1}^*, s, l_{t-1}, z_{t-1})}{\partial k_{t-1}} ds + \int_{l_{t-1}^*}^{l_{t-1}} \frac{\partial \bar{h}_t(m_{t-1}^*, k_{t-1}^*, s, z_{t-1})}{\partial l_{t-1}} ds.$$

Therefore, we can identify $\mathbb{M}_t^{-1}(m, k_t, l_t, z_t)$ and $\bar{h}_t(x_{t-1}, z_{t-1})$ up to $\{c_0(z), c_2(z)\}_{z \in \mathscr{Z}}$.

Define $\widetilde{H}_t(z_t, z_{t-1}) := E[\Lambda_m(m_t, k_t, l_t, z_t) - \Lambda_{\bar{h}}(x_{t-1}, z_{t-1}) | z_t, z_{t-1}]$. To determine $\{c_0(z), c_2(z)\}_{z \in \mathscr{Z}}$, we evaluate

$$\begin{aligned} 0 &= E[\eta_t | z_t, z_{t-1}] \\ &= E\left[\mathbb{M}_t^{-1}(m, k_t, l_t, z_t) - \bar{h}_t(x_{t-1}, z_{t-1}) | z_t, z_{t-1}\right] \\ &= \widetilde{H}_t(z_t, z_{t-1}) + c_0(z_t) - c_2(z_{t-1}) \end{aligned}$$

at different values of $(z_t, z_{t-1}) \in \mathscr{Z}^2$. First, evaluating $E[\eta_t | z_t, z_{t-1}] = 0$ at $z_t = z^1$, and noting that $c_0(z^1) = 0$, we have

$$c_2(z_{t-1}) = \widetilde{H}_t(z^1, z_{t-1}).$$

Therefore, $c_2(z)$ is identified for all $z \in \mathscr{Z}$. Second, evaluating $E[\eta_t | z_t, z_{t-1}] = 0$ at $z_{t-1} = z^1$, we identify $c_0(z)$ as

$$\begin{aligned} c_0(z_t) &= c_2(z^1) - \widetilde{H}_t(z_t, z^1) \\ &= \widetilde{H}_t(z^1, z^1) - \widetilde{H}_t(z_t, z^1). \end{aligned}$$

Given that $\{c_0(z), c_2(z)\}_{z\in\mathscr{Z}}$ are identified, we can identified $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$ and $\bar{h}_t(x_{t-1}, z_t)$ from (A.15) and (A.16).

Each firm's TFP $\omega_{it} = \mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ is identified up to scale and location normalization. From $E[\eta_{it}|x_{t-1}, z_{t-1}] = 0$, we can identify $\bar{h}_t(x_{t-1}, z_{t-1}) = E[\omega_{it}|x_{t-1}, z_{t-1}]$ and $\eta_{it} = \omega_{it} - \bar{h}_t(x_{it-1}, z_{it-1})$. Thus, we obtain the distribution of η_t, $G_{\eta_t}(\eta)$. □

Note that the proofs for Lemma 1 and Proposition 2 do not rely on the continuity of z_t. Therefore, the exactly same proof proves the following proposition.

Proposition A.2. *Suppose that Assumptions 2, 3, A.1, A.2, and 5 hold. Then, we can identify $\varphi_t^{-1}(\cdot)$ and $f_t(\cdot)$ up to scale and location and each firm's markup $\partial \varphi_t^{-1}(\bar{r}_{it}, z_{it})/\partial r_t$ up to scale.*

A.3 Demand Function with Unobservable Demand Shifter

We derive the demand function (55) form a representative consumer's maximization problem. Suppose there are I products. Let $Y_{it} = \exp(y_{it})$ and $P_{it} = \exp(p_{it})$ be the output and price levels of firm i. Consider a representative consumer's utility maximization problem:

$$\max_{\{Y_{it}\}_{i=1}^I} U\left(u\left(\exp(\xi_{1t})Y_{1t}, z_{1t}\right), \ldots., u\left(\exp(\xi_{It})Y_{It}, z_{It}\right)\right) \text{ s.t. } \sum_{i=1}^{I} P_{it} Y_{it} = Y_t,$$

where Y_t is income, the upper tier utility $U(\cdot)$ is symmetric in its arguments and the lower tier $u(\cdot)$ is common for all products. Using $p_{it}^{\dagger} := p_{it} - \xi_{it}$ and $y_{it}^{\dagger} := y_{it} + \xi_{it}$, the utility maximization problem is rewritten as

$$\max_{\{y_{it}^{\dagger}\}_{i=1}^I} U\left(u\left(\exp\left(y_{1t}^{\dagger}\right), z_{1t}\right), \ldots., u\left(\exp\left(y_{It}^{\dagger}\right), z_{It}\right)\right) \text{ s.t. } \sum_{i=1}^{I} \exp\left(p_{it}^{\dagger}\right) \exp\left(y_{it}^{\dagger}\right) = Y_t.$$

The first-order condition for maximization is

$$U' \frac{\partial u\left(\exp\left(y_{it}^{\dagger}\right), z_{it}\right)}{\partial \exp\left(y_{it}^{\dagger}\right)} = \lambda_t \exp\left(p_{it}^{\dagger}\right),$$

where λ_t is the Lagrange multiplier and each firm takes λ_t and U' as given under monopolistic competition. The inverse demand function for firm i is written as:

$$p_{it}^{\dagger} = \psi_t(y_{it}^{\dagger}, z_{it}).$$

A.4 IID Productivity Shock

A firm receives an i.i.d. shock e_{it} to output after choosing inputs:

$$y_{it} = f_t(x_{it}) + \omega_{it} + e_{it}.$$

We suppose that firm's revenue r_{it} is given by

$$r_{it} = \varphi_t(y_{it}, z_{it}) = \varphi_t(f_t(x_{it}) + \omega_{it} + e_{it}, z_{it}). \tag{A.17}$$

A firm chooses m_{it} at time t by maximizing the expected profit conditional on the information available at the time denoted by $\mathscr{I}_{it}$ that includes all past variables and all time t variables except e_{it}:

$$\begin{aligned} m_{it} &= \mathbb{M}_t(\omega_{it}, k_{it}, l_{it}, z_{it}) \\ &:= \underset{m \in \mathscr{M}}{\operatorname{argmax}}\ E\left[\exp\left(\varphi_t(f_t(m, k_{it}, l_{it}) + \omega_{it} + e_{it}, z_{it})\right) | \mathscr{I}_{it}\right] - \exp(p_t^m + m) \\ &= \underset{m \in \mathscr{M}}{\operatorname{argmax}}\ E_e\left[\exp\left(\varphi_t(f_t(m, k_{it}, l_{it}) + \omega_{it} + e_{it}, z_{it})\right)\right] - \exp(p_t^m + m), \end{aligned} \tag{A.18}$$

where E_e is the expectation operator with respect to e.

The identification of $\varphi_t^{-1}(\cdot)$ and $f_t(\cdot)$ in the second step uses the conditional *distribution* of r_t given $w_t := (x_t, z_t)$, beyond the conditional *expectation* in Assumption 2.

Assumption A.3. *The following information at time t is known: (a) the conditional distribution $G_{m_t|v_t}(\cdot)$ of m_t given v_t; (b) the conditional distribution $G_{r_t|w_t}(r|w)$ of r_t given $w_t := (x_t, z_t)$; (c) firm's expenditure on material* $\exp(p_t^m + m_{it})$.

A.4.1 Identification of Control Function and TFP

Since $\mathbb{M}_t^{-1}(m_{it}, k_{it}, l_{it}, z_{it})$ remains a function of the same set of variables, Proposition 1 holds with the same proof.

Proposition A.3. *Suppose that Assumptions 1, 3, 4, and A.3 hold. Then, we can identify $\mathbb{M}_t^{-1}(m_t, k_t, l_t, z_t)$ up to scale and location for all $(m_t, k_t, l_t, z_t) \in \mathscr{M} \times \mathscr{K} \times \mathscr{L} \times \mathscr{Z}$ and identify $G_\eta(\cdot)$ up to scale.*

A.4.2 Identification of Production Function

We make the following assumption that corresponds to Assumption A1–A3 and A5–A6 in Chiappori et al. (2015). (Assumption 1 (b) corresponds to Assumption A4 in Chiappori et al. (2015).)

Assumption A.4. *(a) The distribution $G_{e_t}(\cdot)$ of e_t is absolutely continuous with a density function $g_{e_t}(\cdot)$ that is continuous on its support. (b) e_t is independent of $w_t := (x_t, z_t)'$ with $med(e_t|w_t) = 0$. (c) w_t is continuously distributed on $\mathscr{W} := \mathscr{X} \times \mathscr{Z}$. (d) The support $\mathscr{Y}$ of y_t is an interval on $\mathbb{R}$ that contains 0. (e) The set $\mathscr{B}_{q_t} := \{x_t \in \mathscr{X} : \partial G_{r_t|w_t}(r|w_t)/\partial q_t \neq 0$ for every $(r_t, z_t) \in \mathscr{R} \times \mathscr{Z}\}$ is nonempty for some $q_t \in \{m_t, k_t, l_t\}$.*

The conditional median restriction in Assumption A.4(b) is location normalization. We continue to use the first-order condition with respect to material as a restriction for identification.

Assumption A.5. *The first-order condition with respect to material for the profit maximization problem (A.18) holds for all firms as follows:*

$$E_e\left[\exp\left(\varphi_t(\tilde{y}_{it} + e_{it}, z_{it})\right) \frac{\partial \varphi_t(\tilde{y}_{it} + e_{it}, z_{it})}{\partial \tilde{y}_{it}}\right] \frac{\partial f_t(x_{it})}{\partial m_{it}} = \exp(p_t^m + m_{it}), \tag{A.19}$$

where $\tilde{y}_{it} := f_t(x_{it}) + \omega_{it}$ and the expectation E_e is taken with respect to e_{it}.

Proposition A.4. *Suppose that Assumptions 1, 3,4, A.3, A.4, and A.5 hold. Then, we can identify $\varphi_t^{-1}(\cdot)$, $f_t(\cdot)$, and $G_{e_t}(\cdot)$ up to scale and location.*

Proof. Because φ_t is strictly increase in its first argument, from $med(e_t|w_t) = 0$, we can identify

$$\begin{aligned} \phi_t(x_t, z_t) &:= \varphi_t(f_t(x_t) + \mathbb{M}_t^{-1}(x_t, z_t), z_t) \\ &= med(r_t|x_t, z_t). \end{aligned}$$

From

$$\varphi_t^{-1}(\phi_t(x_t, z_t), z_t) = f_t(x_t) + \mathbb{M}_t^{-1}(x_t, z_t), \tag{A.20}$$

the error term e_t is expressed as

$$e_t = \varphi_t^{-1}(r_t, z_t) - \varphi_t^{-1}(\phi_t(x_t, z_t), z_t). \tag{A.21}$$

From $e_t \perp w_t$ and $w_t := (x_t, z_t)$, the conditional distribution function $G_{r_t|w_t}(r_t|w_t)$ satisfies

$$\begin{aligned} G_{r_t|w_t}(r_t|w_t) &= G_{e_t|w_t}(\varphi_t^{-1}(r, z_t) - f_t(x_t) - \mathbb{M}_t^{-1}(x_t, z_t)\,|w_t) \\ &= G_{e_t}\left(\varphi_t^{-1}(r, z_t) - f_t(x_t) - \mathbb{M}_t^{-1}(x_t, z_t)\right). \end{aligned} \tag{A.22}$$

For $q_t \in \{m_t, k_t, l_t\}$, the derivatives of (A.22) are

$$\frac{\partial G_{r_t|w_t}(r_t|w_t)}{\partial r} = \frac{\partial \varphi_t^{-1}(r_t, z_t)}{\partial r} g_{e_t}\left(\varphi_t^{-1}(r_t, z_t) - f_t(x_t) - \mathbb{M}_t^{-1}(x_t, z_t)\right), \tag{A.23}$$

$$\frac{\partial G_{r_t|w_t}(r_t|w_t)}{\partial q_t} = -\left(\frac{\partial f_t(x_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial q_t}\right) g_{e_t}\left(\varphi_t^{-1}(r_t, z_t) - f_t(x_t) - \mathbb{M}_t^{-1}(x_t, z_t)\right), \tag{A.24}$$

$$\frac{\partial G_{r_t|w_t}(r_t|w_t)}{\partial z_t} = \left(\frac{\partial \varphi_t^{-1}(r_t, z_t)}{\partial z_t} - \frac{\partial \mathbb{M}_t^{-1}(x_t, z_t)}{\partial z_t}\right) g_{e_t}\left(\varphi_t^{-1}(r_t, z_t) - f_t(x_t) - \mathbb{M}_t^{-1}(x_t, z_t)\right). \tag{A.25}$$

Using Assumption A.4(e), choose $q_t \in \{m_t, k_t, l_t\}$ and $\tilde{x}_t \in \mathscr{B}_{q_t}$ such that $\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial q_t \neq 0$ for all $(r_t, z_t) \in \mathscr{R} \times \mathscr{Z}$. Dividing (A.23) by (A.24) and (A.25) by (A.24), respectively, we obtain

$$\frac{\partial \varphi_t^{-1}(r_t, z_t)}{\partial r} = -\left(\frac{\partial f_t(\tilde{x}_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(\tilde{x}_t, z_t)}{\partial q_t}\right) \frac{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial r}{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial q_t}, \tag{A.26}$$

$$\frac{\partial \varphi_t^{-1}(r_t, z_t)}{\partial z_t} - \frac{\partial \mathbb{M}_t^{-1}(\tilde{x}_t, z_t)}{\partial z_t} = -\left(\frac{\partial f_t(\tilde{x}_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(\tilde{x}_t, z_t)}{\partial q_t}\right) \frac{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial z_t}{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial q_t}, \tag{A.27}$$

for all $r_t \in \mathscr{R}$.

Let $x_{t0}^* := (m_{t0}^*, k_t^*, l_t^*)$ and $r_t^* := \phi_t(x_{t0}^*, z_t^*)$. Then, the normalization Assumption 3 implies:

$$\begin{aligned} \varphi_t^{-1}(r_t^*, z_t^*) &= \varphi_t^{-1}(\phi_t(x_{t0}^*, z_t^*), z_t^*). \\ &= f_t(x_{t0}^*) + \mathbb{M}_t^{-1}\left(x_{t0}^*, z_t^*\right) \\ &= 0. \end{aligned}$$

Integrating (A.26) with respect to r and using $\varphi_t^{-1}(r_t^*, z_t^*) = 0$, we obtain

$$\begin{aligned} \varphi_t^{-1}(r_t, z_t^*) &= \int_{r_t^*}^{r_t} \frac{\partial \varphi_t^{-1}(s, z_t^*)}{\partial r} ds \\ &= \left(\frac{\partial f_t(\tilde{x}_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}\left(\tilde{x}_t, z_t^*\right)}{\partial q_t}\right) S_{q_t}(r_t), \end{aligned} \tag{A.28}$$

where

$$S_{q_t}(r_t) := -\int_{r_t^*}^{r_t} \frac{\partial G_{r_t|w_t}(s|\tilde{x}_t, z_t^*)/\partial r}{\partial G_{r_t|w_t}(s|\tilde{x}_t, z_t^*)/\partial q_t} ds > 0 \tag{A.29}$$

is well-defined under Assumption A.4(e).

Define

$$c_m := \frac{\partial f_t(\tilde{x}_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}\left(\tilde{x}_t, z_t^*\right)}{\partial q_t}. \tag{A.30}$$

From (A.28) and (A.21), $\varphi_t^{-1}(r_t, z_t^*)$ and e_t are identified up to c_m as:

$$\varphi_t^{-1}(r_t, z_t^*) = c_m S_{q_t}(r_t) \tag{A.31}$$

$$e_t = c_m\left[S_{q_t}(r_t) - S_{q_t}(\phi(x_t, z_t^*))\right]. \tag{A.32}$$

Because e_t is independent of z_t and x_t, we can identify the distribution of $\tilde{e}_t := e_t/c_m$ as $G_{\tilde{e}_t}(t) = \Pr(S_{q_t}(r_t) - S_{q_t}(\phi(x_t, z_t^*)) \leq t | x_t, z_t^*)$ from (A.32).

Let $y_t := \varphi_t^{-1}(r_t, z_t^*) = f(x_t) + \mathbb{M}_t^{-1}\left(x_t, z_t^*\right) + e_t$. Then, (A.31) implies

$$\frac{y_t}{c_m} = \frac{\varphi_t^{-1}(r_t, z_t^*)}{c_m} = S_{q_t}(r_t). \tag{A.33}$$

Since $S_{q_t}(\cdot)$ is an increasing function, there exists its inverse function $D(\cdot) := S_{q_t}^{-1}(\cdot)$ such that:

$$r_t = \varphi_t(y_t, z_t^*) = D_t\left(\frac{y_t}{c_m}\right) \text{ and } \frac{\partial \varphi_t(y_t, z_t^*)}{\partial y_t} = \frac{1}{c_m} D_t'\left(\frac{y_t}{c_m}\right) \tag{A.34}$$

From $y_t - e_t = f(x_t) + \mathbb{M}_t^{-1}\left(x_t, z_t^*\right) = \varphi_t^{-1}(\phi_t(x_t, z_t^*), z_t^*)$, (A.33) implies

$$\frac{y_t}{c_m} - \tilde{e}_t = \frac{\varphi_t^{-1}(\phi_t(x_t, z_t^*), z_t^*)}{c_m} = S_{q_t}(\phi_t(x_t, z_t^*)). \tag{A.35}$$

From (A.34) and (A.35), the expectation term in the first-order condition (A.31) for a firm with (x_t, z_t^*) times c_m can be written as:

$$
\begin{aligned}
& c_m E_e\left[\exp\left(\varphi_t(y_t, z_t^*)\right)\frac{\partial \varphi_t(y_t, z_t^*)}{\partial y_t}\right] \\
=& c_m E_e\left[\exp\left(D_t\left(\frac{y_t}{c_m}\right)\right)\frac{1}{c_m}D_t'\left(\frac{y_t}{c_m}\right)\right] \text{ from } (A.34) \\
=& E_e\left[\exp\left(D_t\left(S_{q_t}(\phi(x_t, z_t^*)) + \tilde{e}_t\right)\right)D_t'\left(S_{q_t}(\phi(x_t, z_t^*)) + \tilde{e}_t\right)\right] \text{ from } (A.35) \\
=& \int \exp\left(D_t\left(S_{q_t}(\phi(x_t, z_t^*)) + \tilde{e}_t\right)\right)D_t'\left(S_{q_t}(\phi(x_t, z_t^*)) + \tilde{e}_t\right) dG_{\tilde{e}_t}(s) \\
=:& \Upsilon(x_t) \qquad (A.36)
\end{aligned}
$$

where $\Upsilon(x_t)$ is identified because $D_t(\cdot)$, $S_{q_t}(\cdot)$, $\phi(\cdot)$, and $G_{\tilde{e}_t}(\cdot)$ are already identified.

From (A.36), the first-order condition (A.31) for a firm with (x_t, z_t^*) becomes

$$
\frac{\Upsilon(x_t)}{c_m}\frac{\partial f_t(x_t)}{\partial m_t} = \exp(p_t^m + m_t). \qquad (A.37)
$$

Evaluating (A.37) at $(\tilde{x}_t, z_t^*)$ and substituting it into (A.30), we identify c_m as

$$
c_m = \frac{\Upsilon(\tilde{x}_t)}{\Upsilon(\tilde{x}_t) - \exp(p_t^m + \tilde{m}_t)}\frac{\partial \mathbb{M}_t^{-1}\left(\tilde{x}_t, z_t^*\right)}{\partial m_t}.
$$

Given that c_m is identified, we identify $\varphi_t^{-1}(r_t, z_t^*)$ from (A.31), e_t from (A.32) and $f_t(\cdot)$ as

$$
f(x_t) = \varphi_t^{-1}(\phi(x_t, z_t^*), z_t^*) - \mathbb{M}_t^{-1}\left(x_t, z_t^*\right).
$$

Finally, we identify $\partial \varphi_t^{-1}(r_t, z_t)/\partial z_t$ from (A.27) as

$$
\frac{\partial \varphi_t^{-1}(r_t, z_t)}{\partial z_t} = -\left(\frac{\partial f_t(\tilde{x}_t)}{\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(\tilde{x}_t, z_t)}{\partial q_t}\right)\frac{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial z_t}{\partial G_{r_t|w_t}(r_t|\tilde{x}_t, z_t)/\partial q_t} + \frac{\partial \mathbb{M}_t^{-1}(\tilde{x}_t, z_t)}{\partial z_t}.
$$

and $\varphi_t^{-1}(r_t, z_t)$ as:

$$
\varphi_t^{-1}(r_t, z_t) = \varphi_t^{-1}(r_t, z_t^*) + \int_{z_t^*}^{z_t}\frac{\partial \varphi_t^{-1}(r_t, s)}{\partial z_t}ds.
$$

□

A.4.3 Identification of Markup

Because of the i.i.d. shock e_{it}, the first-order condition (A.19) includes the expectation with respect to e_{it}. Thus, the identified value of $\partial \varphi_t^{-1}(r_{it}, z_{it})/\partial r_{it}$ no longer equals the markup. Instead, we obtain the markup from the cost minimization, following Hall (1988) and De Loecker and Warzynski (2012).

Consider a cost minimization problem of producing $\exp(\tilde{y}_{it})$ unit of output:

$$C_t(\tilde{y}_{it}, k_{it}, l_{it}) := \min_m \exp(p_t^m + m) \text{ s.t. } \exp\left(f_t(m, k_{it}, l_{it}) + \omega_{it}\right) \geq \exp\left(\tilde{y}_{it}\right). \tag{A.38}$$

The first-order condition is

$$\lambda_{it} \exp\left(\tilde{y}_{it}\right) \frac{\partial f_t(x_t)}{\partial m_t} = \exp(p_t^m + m_{it}) \tag{A.39}$$

where λ_{it} is the Lagrange multiplier and interpreted as the marginal costs. Using the cost function (A.38), we write the profit maximization problem:

$$\max_{\tilde{y}_{jt}} E\left[\exp\left(\varphi_t(\tilde{y}_{it} + e_{it}, z_{it})\right) | \mathcal{I}_{it}\right] - C_t(\tilde{y}_{it}, k_{it}, l_{it}). \tag{A.40}$$

The first-order condition for (A.40) is

$$E_e\left[\exp\left(\varphi_t(\tilde{y}_{it} + e_{it}, z_{it})\right) \frac{\partial \varphi_t(\tilde{y}_{it} + e_{it}, z_{it})}{\partial \tilde{y}_{it}}\right] = \frac{\partial C_t(\tilde{y}_{it}, k_{it}, l_{it})}{\partial \tilde{y}_{it}} = \lambda_{it} \exp\left(\tilde{y}_{it}\right). \tag{A.41}$$

Substituting (A.41) into (A.39) obtains the first-order condition (A.19) for the profit maximization problem (A.18). Therefore, the problem (A.40) and the problem (A.18) achieve the identical maximized profit.

From (A.19) and (A.41), the marginal cost λ_{it} is expressed as

$$\begin{aligned}
\lambda_{it} &= \frac{E_e\left[\exp\left(\varphi_t(\tilde{y}_{it} + e_{it}, z_{it})\right) \frac{\partial \varphi_t(\tilde{y}_{it} + e_{it}, z_{it})}{\partial \tilde{y}_{it}}\right]}{\exp\left(y_{it} - e_{it}\right)} \\
&= \frac{\exp(p_t^m + m_{it}) / \frac{\partial f_t(x_{it})}{\partial m_{it}}}{\exp\left(y_{it} - e_{it}\right)}.
\end{aligned}$$

Then, the markup becomes

$$\frac{\exp(p_{it})}{\lambda_{it}} = \frac{\partial f_t(x_{it}) / \partial m_{it}}{\exp(p_t^m + m_{it}) / \exp\left(r_{it} - e_{it}\right)},$$

which is identified given our identification of $\partial f_t(x_{it}) / \partial m_{it}$ and e_{it}.