Estimating Cost Functions in Differentiated Product Oligopoly Models without Instruments *

Susumu Imai†
Neelam Jain‡
Hiroto Suzuki§and
Miyuki Taniguchi¶

November 3, 2021

Abstract

We propose a methodology for estimating cost and market share functions of differentiated products oligopoly model when both demand and cost data are available. The method deals with the endogeneity of prices to demand shocks and the endogeneity of outputs to cost shocks without any instruments by using cost data. In contrast to the indirect approach by Byrne et al. (2021) who recover the pseudo-cost function, and then, derive the cost function from it, we propose a method that directly estimates the cost function without the need for the semiparametric pseudo-cost function.

Keywords: Differentiated Goods Oligopoly, Instruments, Identification, Cost data.
JEL Codes: C13, C14, L13, L41

Preliminary draft. Comments are welcome.

*

†Corresponding author, Faculty of Economics and Business, Hokkaido University, susimaig@gmail.com
‡Department of Economics, City University London, Neelam.Jain.1@city.ac.uk
§Faculty of Management, Josai International University h-suziki@jiu.ac.jp
¶Faculty of Economics, Saga University, tanigchi@cc.saga-u.ac.jp

1 Introduction

In this paper, we develop a new methodology for estimating the cost function in a differentiated products model when both prices and outputs are endogenous. Our approach requires cost data in addition to the commonly used demand-side data on products' prices, market shares, observed characteristics. The literature on cost estimation has addressed the endogeneity issues by using either instruments or assuming demand and cost shocks to be orthogonal (See Amsler et al. (2017) and Kutlu et al. (2019) for more details). Another strand of the literature, such as Kumbhakar (2001) has used the profit function, which is a function of output price and input price. If the output and input markets are perfectly competitive, then, those prices can be considered to be exogenous to the firm, and thus, the profit function can be estimated without any instruments. However, profit function based approach would also be subject to the endogeneity issue in the case of differentiated products model since firms also choose prices.

We follow Byrne et al. (2021) and do not use any instruments or other orthogonality conditions, such as orthogonality of demand and cost shocks for dealing with the endogeneity issues.

Byrne et al. (2021) use their two-step nonlinear sieve estimation to recover a semiparametric pseudo-cost function, which is a function of output, input prices, observed characteristics and marginal revenue. They then propose to recover the cost function from the pseudo-cost function by numerically integrating the marginal revenue function. It turns out that this approach is subject to a large bias. Instead, we apply the idea developed by Gandhi et al. (2020) for estimating production functions. They assume that the productivity shock enters in the production function in a multiplicatively separable manner, so that one can eliminate it by using the ratio of the production function and its derivative. We make a similar assumption but for cost functions. That is, we assume a Hicks neutral cost shock, which is the inverse of the Hicks neutral productivity shock. Then, the cost shock can be eliminated as long as we estimate parameters of the cost function by taking the ratio of marginal cost to cost. We then replace the unobservable marginal cost with marginal revenue which is a function of observables and parameters. We thus show that there is a direct approach that sidesteps the pseudo-cost function estimation to estimate the cost and demand function parameters jointly,

The Monte-Carlo results indicate that such a direct approach greatly improves the finite sample accuracy of both the demand and cost function parameter estimates compared to the indirect approach adopted by Byrne et al. (2021).

This paper is organized as follows. In Section 2, we review the IV-based estimation of the differentiated products model of demand and the cost function. Then, we discuss the F.O.C. of the profit maximization of firms that we use for the instrument-free joint identification and estimation of the demand and cost functions. In Section 3, we study identification when demand and cost data are available and present our formal identification results. In Section 4, we propose the direct approach for estimating the cost function. Section 5 contains a Monte-Carlo study that illustrates the effectiveness of our estimator. In Section 6 we conclude.

2 Demand and cost function and their IV estimation

The key component of our methodology is the first order condition (F.O.C.) of the firm's profit maximzation, so that marginal revenue equals marginal cost. Unlike the other approaches in the literature, in our estimation, marginal revenue plays an important role. That is, we follow the control function approach of Byrne et al. (2021) and use marginal revenue to control for the cost shock. Therefore, we first review the standard differentiated products demand model, and derive its marginal revenue function.

In this section, we describe the standard differentiated products model that we adopt including some of the assumptions and provide an overview of IV estimation of the demand and supply side. For more details, see Berry (1994), Berry et al. (1995), Nevo (2001) and others. Most features of the model we discuss here are carried over to the next section where we explain our cost data-based identification strategy.

2.1 Differentiated products discrete choice demand models

In the standard model, consumer i in market m gets the following utility from consuming one unit of product j:

$$u_{ijm} = \mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm} + \epsilon_{ijm},$$

where $\mathbf{x}_{jm}$ is a $1 \times K$ vector of observed product characteristics, p_{jm} is price, ξ_{jm} is the unobserved product quality (or demand shock) that is known to both consumers and firms but unknown to researchers, and ϵ_{ijm} is an idiosyncratic taste shock. The demand parameter vector is denoted by $\boldsymbol{\theta} = \left[\alpha, \boldsymbol{\beta}'\right]'$, where $\boldsymbol{\beta}$ is a $K \times 1$ vector.

It is assumed that there are $M > 1$ isolated markets.[1] Market m has $J_m + 1 > 2$ products

[1]With panel data, the m index corresponds to a market-period.

whose aggregate demand across individuals is,

$$q_{jm} = s_{jm} Q_m,$$

where q_{jm} denotes output, Q_m denotes market size and s_{jm} denotes market share. In the case of the Berry (1994) logit demand model, ϵ_{ijm} is assumed to have a logit distribution. Then, the aggregate market share for product j in market m is,

$$s_{jm}(\boldsymbol{\theta}) \equiv s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}\right) = \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m}\exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)} = \frac{\exp\left(\delta_{jm}\right)}{\sum_{k=0}^{J_m}\exp\left(\delta_{km}\right)}, \tag{1}$$

where $\mathbf{p}_m = [p_{0m}, p_{1m}, ..., p_{J_m m}]'$ is a $(J_m + 1) \times 1$ vector,

$$\mathbf{X}_m = \begin{bmatrix} \mathbf{x}_{0m} \\ \mathbf{x}_{1m} \\ \vdots \\ \mathbf{x}_{J_m m} \end{bmatrix}$$

is a $(J_m + 1) \times K$ matrix, $\boldsymbol{\xi}_m = [\xi_{0m}, \xi_{1m}, ..., \xi_{J_m m}]'$ is a $(J_m + 1) \times 1$ vector, and

$$\delta_{jm} \equiv \mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm} \tag{2}$$

is the "mean utility" of product j in market m. Using this definition, we can express the market share in Equation (1) as $s_j\left(\boldsymbol{\delta}(\boldsymbol{\theta})\right) \equiv s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}\right)$ where $\boldsymbol{\delta}(\boldsymbol{\theta}) = [\delta_{0m}(\boldsymbol{\theta}), \delta_{1m}(\boldsymbol{\theta}), \ldots, \delta_{J_m m}(\boldsymbol{\theta})]'$.

Good $j = 0$ is labeled the "outside good" or "no-purchase option" that corresponds to not buying any of the $j = 1, \ldots, J_m$ goods. This good's product characteristics, price, and demand shock are normalized to zero (i.e., $\mathbf{x}_{0m} = \mathbf{0}$, $p_{0m} = 0$, and $\xi_{0m} = 0$ for all m), which implies

$$\delta_{0m}(\boldsymbol{\theta}) = 0. \tag{3}$$

This normalization, together with the logit assumption for the distribution of ϵ_{ijm}, identifies the level and scale of utility.

In BLP, or equivalently, the random coefficient logit model, one allows the price coefficient and coefficients on the observed characteristics to be different for different consumers. Specifically, α has a distribution function $F_\alpha\left(.; \boldsymbol{\theta}_\alpha\right)$, where $\boldsymbol{\theta}_\alpha$ is the parameter vector of the distribution, and similarly, $\boldsymbol{\beta}$ has a distribution function $F_{\boldsymbol{\beta}}\left(.; \boldsymbol{\theta}_{\boldsymbol{\beta}}\right)$ with parameter vector $\boldsymbol{\theta}_{\boldsymbol{\beta}}$. The probability with which a consumer with coefficients α and $\boldsymbol{\beta}$ purchases product j is identical to that provided by the market share formula in Equation (1). The aggregate market share of product j is obtained by integrating over the distributions of α and $\boldsymbol{\beta}$:

$$s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}\right) = \int_\alpha \int_{\boldsymbol{\beta}} \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)} dF_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}; \boldsymbol{\theta}_{\boldsymbol{\beta}}\right) dF_\alpha\left(\alpha; \boldsymbol{\theta}_\alpha\right), \tag{4}$$

where $\boldsymbol{\theta} = \left[\boldsymbol{\theta}'_\alpha, \boldsymbol{\theta}'_{\boldsymbol{\beta}}\right]'$. Letting μ_α to be the mean of α and $\boldsymbol{\mu}_{\boldsymbol{\beta}}$ the mean of $\boldsymbol{\beta}$, the mean utility is defined to be

$$\delta_{jm} \equiv \mathbf{x}_{jm}\boldsymbol{\mu}_{\boldsymbol{\beta}} - p_{jm}\mu_\alpha + \xi_{jm}, \tag{5}$$

with $\delta_{0m} = 0$ for the outside good.

2.1.1 Recovering demand shocks

For each market $m = 1, \ldots M$, researchers are assumed to have data on prices $\mathbf{p}_m$, market shares $\mathbf{s}_m = \left[s_{0m}, s_{1m}, ..., s_{J_m m}\right]'$ and observed product characteristics $\mathbf{X}_m$ for all firms in the market. Given $\boldsymbol{\theta}_d$ and this data, one can solve for the vector $\boldsymbol{\delta}_m$ through market share inversion. That is, if we denote $s_j\left(\boldsymbol{\delta}_m\left(\boldsymbol{\theta}_d\right); \boldsymbol{\theta}_d\right)$ to be the market share of firm j predicted by the model, market share inversion involves obtaining $\boldsymbol{\delta}_m$ by solving the following set of J_m equations,

$$s_j\left(\boldsymbol{\delta}_m\left(\boldsymbol{\theta}_d\right), j; \boldsymbol{\theta}_d\right) - s_{jm} = 0, \text{ for } j = 0, \ldots, J_m, \tag{6}$$

and therefore,

$$\boldsymbol{\delta}_m\left(\boldsymbol{\theta}_d\right) = \mathbf{s}^{-1}\left(\mathbf{s}_m; \boldsymbol{\theta}_d\right). \tag{7}$$

The vector of mean utilities that solves these equations perfectly aligns the model's predicted market shares to those observed in the data.

In the logit model, Berry (1994) shows we can easily recover mean utilities for product j using its market share and the share of the outside good as $\delta_{jm}\left(\boldsymbol{\theta}_d\right) = \log\left(s_{jm}\right) - \log\left(s_{0m}\right)$, $j = 1, \ldots, J_m$. In the random coefficient model, there is no such closed-form formula for market share inversion. Instead, BLP propose a contraction mapping algorithm that recovers the unique $\delta_{jm}\left(\boldsymbol{\theta}_d\right)$ that solves Equation (7) under some regularity conditions. In both cases, δ_{0m} is normalized to 0.

With the mean utilities and parameters in hand, one can recover the structural demand shocks straightforwardly from Equation (2) for the logit demand and Equation (5) for the BLP demand.

2.1.2 IV estimation of demand

A simple regression of Equation (2) or (5) with $\delta_{jm}(\boldsymbol{\theta}_d)$ being the dependent variable and $\mathbf{x}_{jm}$ and p_{jm} being the regressors would yield a biased estimate of the price coefficient. This is because firms likely set higher prices for products with higher unobserved product quality, which creates a correlation between p_{jm} and ξ_{jm}, violating the OLS orthogonality condition $E[\xi_{jm}p_{jm}] = 0$. Researchers use a variety of demand instruments to overcome this issue. In particular, researchers construct a GMM estimator for $\boldsymbol{\theta}$ by assuming the following population moment conditions are satisfied at the true value of the demand parameters $\boldsymbol{\theta}_{d0}$:

$$E[\xi_{jm}(\boldsymbol{\theta}_{d0})\,\mathbf{z}_{jm}] = \mathbf{0}$$

where $\mathbf{z}_{jm}$ is an $L \times 1$ vector of instruments that is correlated with $\mathbf{x}_{jm}$. Also, instruments are required to satisfy the exclusion restriction that at least one variable in $\mathbf{z}_{jm}$ is not contained in $\mathbf{x}_{jm}$.

2.2 Cost Function Estimation

For each product j in market m, in addition to the data related to demand explained above, researchers observe output q_{jm} (hence, market size $Q_m = q_{jm}/s_{jm}$ as well), $L \times 1$ vector of input price $\mathbf{w}_{jm}$ and cost C_{jm}. The observed cost C_{jm} is assumed to be a function of output, input prices $\mathbf{w}_{jm}$, observed product characteristics $\mathbf{x}_{jm}$ and a cost shock υ_{jm}. That is,

$$C_{jm} = C\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right),$$

where $\boldsymbol{\theta}_c$ is the parameter vector. $C()$ is assumed to be strictly increasing and continuously differentiable in output and cost shock.As with demand estimation, one can recover unobserved cost shocks through inversion:

$$C_{jm} = C\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right) \Rightarrow \upsilon_{jm} = \upsilon\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, C_{jm}; \boldsymbol{\theta}_c\right). \tag{8}$$

Like demand estimation, there are important endogeneity concerns with standard approaches to estimating cost functions. Specifically, output q_{jm} is endogenously determined by profit-maximizing firms as in Equation (9), and is potentially negatively correlated with the cost shock

υ_{jm}. That is, all else equal, less efficient firms tend to produce less. In dealing with this issue, researchers have traditionally focused on selected industries where endogeneity can be ignored, or used instruments for output.

The IV approach to cost function estimation typically uses excluded demand shifters as instruments. Denoting this vector of cost instruments by $\widetilde{\mathbf{z}}_{jm}$, one can estimate $\boldsymbol{\theta_c}$ assuming that the following population moments are satisfied at the true value of the cost parameters $\boldsymbol{\theta}_{c0}$: $E\left[\upsilon_{jm}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, C_{jm}; \boldsymbol{\theta}_{c0}\right) \widetilde{\mathbf{z}}_{jm}\right] = \mathbf{0}$. See Wang (2003)

Typical instruments that can be used for price in demand function estimation and output in cost function estimation are the product characteristics of rival firms in the same market:$\mathbf{X}_{-jm}$. However, if firms endogenouly choose their observed characteristics in response to own and other firms' cost shocks, then $\mathbf{X}_{-jm}$ could be correlated with the cost shock υ_{jm} and thus, won't be valid instruments. One way to deal with this problem is to assume that observed characteristics are uncorrelated with the cost shock in the short run.This assumption is similar to the ones often used in panel data estimation: the innovation of the cost shock is uncorrelated with the current observed product characteristics. Petrin and Seo (2016) utilize similar assumptions for estimation of the market share function. They show that innovations in observed product characteristics can be used as instruments for the cost shock.

2.3 Firms' Profit and Profit Maximization.

Assuming that there is one firm for each product, firm j's profit function is as follows:

$$\pi_{jm} \quad = \quad p_{jm}q_{jm} - C\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right).$$

Let MR_{jm}, be the marginal revenue of firm j in market m. BLP assume that firms act as differentiated products Bertrand price competitors. Therefore, the optimal price and quantity of product j in market m are determined by the F.O.C. that equates marginal revenue and marginal cost:

$$MR_{jm} = \underbrace{\frac{\partial p_{jm}q_{jm}}{\partial q_{jm}} = p_{jm} + s_{jm}\left[\frac{\partial s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right)}{\partial p_{jm}}\right]^{-1}}_{MR_{jm}} = MC_{jm} = \underbrace{\frac{\partial C\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial q_{jm}}}_{MC_{jm}}. \tag{9}$$

Note that given the market share inversion in Equation (6), and the specification of mean utility $\boldsymbol{\delta}_m$, $\boldsymbol{\xi}_m$ is a function of $(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m)$ and $\boldsymbol{\theta}_d$. Therefore, marginal revenue of firm j in market m, MR_{jm} in Equation (9) can be written as a function of observables and parameters as follows:

$$MR_{jm} \equiv MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right), \tag{10}$$

Equations (9) and (10) imply that demand parameters can potentially be identified if there is data on marginal cost[2] or even without such data, if the cost function is known or can be estimated and its derivative with respect to output can be taken. Berry et al. (1995) assume that marginal cost is log-linear in output and observed product characteristics, i.e., $MC_{jm} = exp\left(\mathbf{w}_{jm}\gamma_w + q_{jm}\gamma_q + \upsilon_{jm}\right)$ (see their Equation 3.6). They then use instruments to deal with the endogeneity of output with cost shocks and of prices to demand shocks. As long as the parametric specification of the supply side is accurate and there are enough instruments for identification, the demand side and F.O.C. based orthogonality conditions are sufficient for identifying demand parameters.

In this research, we follow the insight of Byrne et al. (2021) that jointly estimating both demand and cost sides of the model can remove the need for any instruments to deal with the endogeneity issue in the estimation of price coefficient of demand function and output coefficient of cost function.

3 Identification of demand and cost functions using cost data and without instruments

In this section, we present our methodology for dealing with the endogeneity issues in identification mentioned above. We propose using cost data in addition to demand data to identify price parameters and the parameters of the cost function. We do so by using the control function approach of Byrne et al. (2021). That is, given output, input prices and observed product characteristics, we use the observed cost to control for the cost shock. We first present our methodology using general marginal revenue and cost functions and then, illustrate it with the logit demand function and Cobb-Douglas and translog cost functions. First, we explain the underlying assumptions of the model.

[2] Genesove and Mullin (1998) use data on marginal cost to estimate the conduct parameters of the homogeneous goods oligopoly model.

3.1 Main Assumptions

We first state all the main assumptions for our methodology. Most of these assumptions are standard as discussed in the previous section or simply describe the environment our methodology is applicable to. For each market in the population, we attach a unique positive real number m as an identifier. Then, we assume $m \in \mathcal{M}$, where $\mathcal{M}$ is the set of all market identifiers, and is an uncountable subset of R_+.

Assumption 1 *Data Requirements: Researchers have data on outputs, product prices, market shares, input prices, observed product characteristics, and total costs of firms.*

Note that market size can be derived from data on outputs and market shares. Thus, we need to assume observability of only two of these three variables. In contrast to BLP, we require data on total costs of firms. But we do not need data on marginal cost.

Assumption 2 *Isolated Markets: Outputs, market shares, prices and costs in market m are functions of variables in market m .*

Assumption 3 *Logit demand: Market share s_{jm} is specified as in Equation (4) with $\mu_\alpha < 0$.*

Assumption 4 *Equilibrium Concept: Bertrand-Nash equilibrium holds in each market. That is, for any $j = 1, \ldots, J_m$, firm j in market m chooses its price p_{jm} to equalize marginal revenue and marginal cost, given market size Q_m and prices of other firms in the same market $\mathbf{p}_{-j,m}$.*[3]

The next assumption describes the support of variables that determine the equilibrium outcomes in market m. Let the set of these variables be denoted by $\mathbf{V}_m$. Then $\mathbf{V}_m \equiv (Q_m, \mathbf{W}_{jm}, \mathbf{X}_m, \boldsymbol{\xi}_m, \boldsymbol{\upsilon}_m)$, and let $\mathbf{V} \equiv \{V_m\}_{m \in \mathcal{M}}$. Let $\mathbf{V} \setminus w_{lkm}$ to be the set $\mathbf{V}$ without the element $\mathbf{w}_{lm}$ for any $l = 1, 2, \ldots, L$. For other elements of $\mathbf{V}$, the set $\mathbf{V}$ without the element is similarly defined. The assumption imposes substantially weaker restrictions on the support of the variables in $\mathbf{V}$ than is typical in the literature. In particular, it imposes minimal restrictions on the joint distribution of these variables as stated below.

Assumption 5 *Support of $\mathbf{V}$: The support of Q_m conditional on $\mathbf{V} \setminus Q_m$ can be any nonempty subset of R_+ for all m. The support of $\mathbf{w}_{lm}$ conditional on $\mathbf{V} \setminus \mathbf{w}_{lm}$ is R_+ for all l, m; the*

[3] Note that we have assumed this for expositional purposes only. It is not required for identification. MR is a one-to-one function of MC in equilibrium, and not necessarily equal to MC, we can identify the price parameters. This makes our framework applicable to firms that are under government regulation and firms under organizational incentives or behavioral aspects that prevent them from setting $MR = MC$.

support of x_{kjm} conditional on $\mathbf{V} \setminus x_{kjm}$ is either R or R_+ for all k, j, m; and the support of ξ_{jm} conditional on $\mathbf{V} \setminus \xi_{jm}$ is R. Finally, the support of υ_{jm} conditional on $\mathbf{V} \setminus \upsilon_{jm}$ is R_+.

Assumption 5 ensures that the variables in $\mathbf{V}$ are not subject to any orthogonality conditions, which typically restrict the moments of a subset of the unobserved variables $(\boldsymbol{\xi}_m, \boldsymbol{\upsilon}_m)$ conditional on the other variables to be zero. In other words, we do not require them to be econometrically exogenous, and thus, Assumption 5 removes the validity of any conventional instruments.

Note that we do not impose any assumptions on the support of market size other than that it is nonempty and positive. For logit, we require the conditional support to be R_+ since as we show later, market size variation is needed for identifying the price parameters of logit but not for BLP.

We follow Gandhi et al. (2020) and assume that the cost function can be multiplicatively separated into the component that has output, input price and observed product characteristics and the remaining component that only includes observed product characteristics and the cost shock. In addition, we also allow for additive measurement error.

Assumption 6 *The true cost function, denoted as $C^*()$ can be expressed as follows:*

$$C^*(q, \mathbf{w}, \mathbf{x}, \upsilon; \boldsymbol{\theta}_{c0}) = \widetilde{C}(q, \mathbf{w}, \mathbf{x}; \boldsymbol{\theta}_{c0})\, exp(\varphi(\mathbf{x}, \upsilon)). \tag{11}$$

where $\widetilde{C}()$ is only a function of observables, such as output, input prices and observed characteristics and $\varphi(\mathbf{x}, \upsilon)$ is an unspecified smooth function of observed characteristics $\mathbf{x}$ and unobserved characteristics u. Furthermore, the observed cost C_{jm} is given by the sum of the true cost C^*_{jm} and the measuremente error u_{cjm} as follows:

$$C_{jm} = C^*_{jm} + u_{jm} = C^*(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c0}) + u_{cjm},$$

where we assume u_{jm} to be i.i.d. distributed and independent to all other observables in the demand and cost functions and $\boldsymbol{\theta}_{c0}$ is the cost function parameter vector we identify. Similarly, we assume that expenditure on input k, denoted by L_k , whose price is w_{kjm}, is measured with error, i.e.

$$C_{kjm} = w_{kjm} L_{jmk} + u_{kjm},\ \ k = 1, \ldots, K$$

where we assume u_{kjm} to be i.i.d. distributed and independent to other varialbes in the demand and cost functions.

4 Demand and cost function estimation without instruments

4.1 General identification result

In this section, we discuss how to recover the cost function from the data without using any instruments. Byrne et al. (2021) propose to use the pseudo-cost function and marginal cost to recover the cost function by numerically integrating the numerically derived cost function. However, this process results in a large bias in the recovered cost function because numerical errors accumulate during integration. We explain the details later.

In this paper, we propose an alternative approach. More concretely, we eliminate the cost shock from the estimation by exploiting Assumption 6, where we assume the cost shock component to enter multiplicatively in the total cost function. Then, the total cost[4] and the marginal cost functions can be expressed as follows,

$$\begin{aligned} C^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right) &= \widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right) exp\left(\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right). \\ MC^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right) &= \frac{\partial}{\partial q}\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right) exp\left(\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right) \\ &= \widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right) exp\left(\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right). \end{aligned} \tag{12}$$

Therefore, by taking the ratio of margial cost and cost, we obtain

$$\frac{C^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right)}{MC^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right)} = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)}. \tag{13}$$

Note that the RHS does not contain the unobservable cost shock υ_{jm}. Furthermore, from the F.O.C., we obtain

$$MR_{j}\left(\mathbf{p}_{m}, \mathbf{s}_{m}, \mathbf{X}_{m}; \boldsymbol{\theta}_{d}\right) = MC^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right). \tag{14}$$

Using Equation (14) to substitute $MR()$ for $MC^{*}()$ into Equation (13) we derive:

$$\frac{C^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right)}{MR_{j}\left(\mathbf{p}_{m}, \mathbf{s}_{m}, \mathbf{X}_{m}; \boldsymbol{\theta}_{d}\right)} = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)}$$

and by multiplying MR_j on both sides, we get

$$C^{*}_{jm} = C^{*}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c}\right) = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c}\right)} MR_{j}\left(\mathbf{p}_{m}, \mathbf{s}_{m}, \mathbf{X}_{m}; \boldsymbol{\theta}_{d}\right).$$

[4] From now on, we simply call cost for total cost, whenever there is no confusion.

This is how we can express the cost function as a function that does not have the unobservable cost shock υ_{jm}, which was the source of the endogeneity bias.

Our identification strategy is based on the exclusion restriction that there are variables that potentially enter in the marginal revenue function but not in the cost function. These variables are market size Q_m, which enters in the marginal revenue function through $q_{jm} = s_{jm}/Q_m$, prices of firms in market m, $\mathbf{p}_m$, market shares $\mathbf{s}_{-jm}$ and observed characteristics $\mathbf{X}_{-jm}$ of rival firms in the same market. For example, in the logit demand model, marginal revenue is

$$MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right) = p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha} = p_{jm} - \frac{1}{\left(1 - q_{jm}/Q_m\right)\alpha}. \tag{15}$$

Therefore, the exclusion restriction is that price p_{jm} market size Q_m enter in the marginal revenue function, but not in the cost function.

Then, from Assumption 6, observed cost can be specified as follows:

$$C_{jm} = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_{d0}\right) + u_{cjm}. \tag{16}$$

Now, we discuss the difference between the above specification and the conventional one, which is:

$$C_{jm} = C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c0}\right) = \widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right) exp\left(\upsilon_{jm}\right).$$

By taking logs, we obtain

$$lnC_{jm} = ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right) + \upsilon_{jm} \tag{17}$$

The estimation of Equation (17) is subject to the endogeneity bias if output and cost shock are correlated. Since profit maximizing firms tend to increase prices and decrease output as a result of an increase in the cost shock, negative correlation between the cost shock and output is a likely scenario. In contrast, in Equation (16) the cost shock, which is the source of endogeneity is factored out. Since, from Assumption 6, measurement error of cost u_{cjm} is assumed to be independent to all the other variables in the marginal revenue and the marginal cost functions, endogeneity issues do not arise.

As long as we assume that the variables in $\widetilde{C}\left(\right)$ are independent to the productivity shock υ_{jm}, we can estimate the parameter vector $\boldsymbol{\theta}_c$ consistently. Even then, it is customary to jointly estimate the model with equations that can be derived using Shephard's Lemma for the sake of efficiency. Shephard's Lemma states that

$$\frac{\partial C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial w_{kjm}} = L_{kjm}. \tag{18}$$

Then, using Equation (17), the above equation can be modified as follows:

$$\frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}} = \frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}} = \frac{w_{kjm}L_{kjm}}{C^*_{jm}}, k = 1, \ldots, K \tag{19}$$

and after adding an error term ϵ_{kjm}, $k = 1, \ldots, K$, for estimation purpose, which also indicates allocative inefficiency, we obtain

$$\frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}} = \frac{w_{kjm}L_{kjm}}{C^*_{jm}} + \epsilon_{kjm}, \ k = 1, \ldots, K-1, \tag{20}$$

which, together with Equation (17) forms a system that is estimated joinly by Maximum Likelihood, where researchers construct the log likelihood by assuimg υ_{jm} and ϵ_{jm} to be jointly normally distributed. Note that in estimation, as we can see in Equation (20), only $K-1$ input share equations are used. This is becauce of the linear dependence of the K input shares since they add to one. The existence of the allocative inefficiency in Equation (20) implies that it should also be included in the cost function. Kumbhakar (1997) derived the modified translog cost function that explicitly includes the allocative inefficiency, i.e., $C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}, \boldsymbol{\epsilon}_{jm}; \boldsymbol{\theta}_c\right)$, where $\boldsymbol{\epsilon}_{jm} \equiv (\epsilon_{1jm}, \ldots, \epsilon_{Kjm})$, and Kumbhakar and Tsionas (2005) provided a Bayesian estimation method for estimating the translog cost function when such allocative inefficiencies exist.

Instead, we use Assumption 6. That is, Shephard's Lemma and Assumption 6 together imply

$$C_{kjm} = w_{kjm}L_{kjm} + u_{kjm} = C^*_{jm}\frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}} + u_{kjm}, \tag{21}$$

which we use for estimation. As before, since the measurement error in input cost u_{kjm} are assumed to be independent to the variables in the marginal revenue and cost functions, the above equation is not subject to an endogeneity issue. Furthermore, note that the difference in how we and Kumbhakar (1997) and the subsequent literature interpret the error term that corresponds to the discrepancy between the input price elasticity of cost function and the input cost share. Kumbhakar (1997) specified them as the allocation error and pointed out the need for including them in the cost function as unobservable variables. In our specification of cost function in Assumption 6, measurement error does not enter in the true cost function C^*.

So far, we have specified measurement errors as additive to the total cost as well as the components of the cost. We believe that it is more realistic to specify to the total cost as the sum of various cost components, and so are the measurement error of the total cost. Therefore, if we specify the measurement errors as additive to cost, then the measurement error of the total cost can be simply expressed as the sum of all the measurement errors of the individual cost components.

Next, we discuss in more detail the identification. First, we discuss an example with the cost function that is based on the following Cobb-Douglas production function.

$$q = \left[Bexp\left(x\eta + \upsilon\right)\right]^{-(\alpha_c+\beta_c)} L^{\alpha_c} K^{\beta_c}.$$

We can derive the cost function from the below cost minimization problem.

$$\begin{aligned} C^*\left(q, w, r, x, \upsilon\right) &= max_{K,L} rK + wL \\ s.t.\ q &\leq \left(Bexp\left(x\eta + \upsilon\right)\right)^{-(\alpha_c+\beta_c)} L^{\alpha_c} K^{\beta_c}. \end{aligned}$$

Then, given wage w and the capital rental rate r, the cost and the marginal cost functions are as follows:

$$C^*\left(q, w, r, x, \upsilon\right) = \left[\left(\alpha_c + \beta_c\right)\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)}\left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] Bexp\left(x\eta + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}}.$$

$$MC^*\left(q, w, r, x, \upsilon\right) = \left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)}\left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)} Bexp\left(x\eta + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}-1}.$$

Thus, the log cost function is specified as

$$\begin{aligned} lnC^* &= ln\left(\alpha_c + \beta_c\right) - \frac{\alpha_c}{\alpha_c + \beta_c} ln\alpha_c - \frac{\beta_c}{\alpha_c + \beta_c} ln\beta_c \\ &\quad + \frac{\alpha_c}{\alpha_c + \beta_c} lnw + \frac{\beta_c}{\alpha_c + \beta_c} lnr + \frac{1}{\alpha_c + \beta_c} lnq + \eta x + \upsilon \end{aligned} \tag{22}$$

By taking derivative of the above cost function with respect to output,

$$\frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon\right)}{\partial lnq} = \frac{1}{\alpha_c + \beta_c} = \frac{MR\left(\right) q_{jm}}{C^*_{jm}}.$$

Then,

$$C^* \left(q_{jm}, w_{jm}, r_{jm}, \mathbf{x}_{jm}, \upsilon\right) = \left(\alpha_c + \beta_c\right) q_{jm} MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right), \tag{23}$$

and the estimating equation is:

$$C_{jm} = C^*_{jm} + u_{cjm} = \left(\alpha_c + \beta_c\right) q_{jm} MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right) + u_{cjm}, \tag{24}$$

which identifies $\alpha_c + \beta_c$ and the demand parameters $\boldsymbol{\theta}_d$. If we assume the market share function to be logit, then, we obtain

$$C_{jm} = C^*_{jm} + u_{cjm} = \left(\alpha_c + \beta_c\right) q_{jm} \left[p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha}\right] + u_{cjm},$$

which identifies $\alpha_c + \beta_c$ and α. However, Equation (24) does not identfy α_c or β_c separately. If we have the cost data for each input item, then using the Shephard' Lemma, we derive

$$\frac{\partial ln\widetilde{C}\left(q_{jm}, w_{jm}, r_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial ln w_{jm}} = \frac{\alpha_c}{\alpha_c + \beta_c} = \frac{w_{jm} L_{jm}}{C^*_{jm}}.$$

Manipulating the above equation, we obtain

$$C_{Ljm} = w_{jm} L_{jm} + u_{Ljm} = \alpha_c q_{jm} MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right) + u_{Ljm}.$$

and thus, we identify α_c, and $\beta_c = \alpha_c + \beta_c - \alpha_c$. In case we do not have data on input cost, given $\alpha_c + \beta_c$, we can identify the remaining parameters by putting the terms with them on the RHS as follows:

$$lnC^*_{jm} - \left(\alpha_c + \beta_c\right) lnq_{jm} = A + \frac{\alpha_c}{\alpha_c + \beta_c} lnw_{jm} + \frac{\beta_c}{\alpha_c + \beta_c} lnr_{jm} + x_{jm}\eta + \upsilon_{jm}.$$

where

$$A \equiv ln\left(\alpha_c + \beta_c\right) - \frac{\alpha_c}{\alpha_c + \beta_c} ln\alpha_c - \frac{\beta_c}{\alpha_c + \beta_c} ln\beta_c$$

and C^*_{jm} can be obtained from Equation (23) given both $\alpha_c + \beta_c$ and θ_d are identified. If we assume x_{jm}, w_{jm} and r_{jm} to be orthogonal to υ_{jm}, the above equation can be esitmated consistently by OLS.

Next, we summarize the above results. That is, estimation that is based solely on Equation

(16) can only identify the parameters in $\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right) / \widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)$. Because

$$
\begin{aligned}
\frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial lnq} &= \frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \boldsymbol{\theta}_c\right)}{\partial lnq} \\
= \quad & \frac{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)} q_{jm} = \frac{MR\left(\mathbf{p}_m, \mathbf{x}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right)}{C^*_{jm}} q_{jm},
\end{aligned}
$$

this implies that what we identify from the F.O.C. of profit maximization is the demand parameters $\boldsymbol{\theta}_d$ and the output elasticity of cost. In other words, F.O.C. only identifies the paramters of the cost function that affect the output elasticity of cost, which we denote as the vector $\boldsymbol{\theta}_{cq}$. In the above example with the cost function that is based on the Cobb-Douglas production function, $\boldsymbol{\theta}_c = (\alpha_c + \beta_c, \alpha_c)$, and $\boldsymbol{\theta}_{cq} = \alpha_c + \beta_c$.

If we have data on cost of each input, we can identify some of the remaining parameters from the Shephard's Lemma, which identifies the elasticity of cost with respect to input price, i.e.

$$
\frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}} = \frac{w_{kjm} L_{kjm}}{C^*_{jm}},
$$

and thus the parameters that determine the input price elasticity of cost. Similarly as before, let $\boldsymbol{\theta}_{cw}$ be the vector of such parameters that is not in the element of $\boldsymbol{\theta}_{cq}$. In the example with the cost function based on the Cobb-Douglas production function, $\boldsymbol{\theta}_{cw} = \alpha_c$. The parameters of the cost function that remains unidentified is not a function of either log output or log input price, and thus, can be identified from the remaining cost component that can be expressed as a function of $\mathbf{x}_{jm}$ and υ_{jm}, i.e., $ln\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c,(-q,-w)}\right)$.

Consider the other case where we only have data on total cost, and not the cost data for each input. Given that $\boldsymbol{\theta}_{cq}$ are the parameters that determines the output elasticity of cost, and thus identified from $\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon; \boldsymbol{\theta}_c\right) / \partial lnq$, we identify the remaining parameters $\boldsymbol{\theta}_{c,-q}$ from the component that is not affected by log output from the regression.

$$
lnC^* = ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \boldsymbol{\theta}_{cq}, \boldsymbol{\theta}_{c,-q}\right) + \upsilon_{jm} \tag{25}
$$

Since only the parameters $\boldsymbol{\theta}_{cq}$ are subject to endogeneity bias, the remaining parameters $\boldsymbol{\theta}_{c,-q}$ can be estimated via regression consistently.

4.2 Estimation issues

Next, we discuss the estimation when we have cost data for each input. We would then estimate jointly the equations (16) and (21). The conventional methods are feasible generalized least squares (FGLS) and maximum likelihood (ML) methods. If the measurement errors are specified as i.i.d., jointly normally distributed, then in large sample, FGLS and ML in this particular cases are very similar in the sense that they both result in consistency and they have the same asymptotic distribution. FGLS estimates the parameters by minimizing the following objective function

$$\mathbf{u'}_{jm}\mathbf{W}\mathbf{u}_{jm},$$

where the efficient weight would be $\mathbf{W} = \mathbf{\Sigma}^{-1}$, and $\mathbf{\Sigma} \equiv Var\left(\mathbf{u}_{jm}\right)$ is the variance-covariance matrix. The variance-covariance matrix is estimated by using the residual of the first stage parameter estimate which is estimated by initially setting the weighting matrix to $\mathbf{W} = \mathbf{I}$.

Similarly, we can use the Maximum Likelihood procedure and construct the likelihood based on the joint distribution of the measurement errors. Since u_{kjm} is specified to be a measurement error, it is a component of the measurement error of total cost u_{cjm}. Hence, we need to allow for correlation between u_{kjm} and u_{cjm}.

We assume that the measurement errors $(u_{cjm}, u_{1jm}, \ldots, u_{K-1,jm})$ are jointly normally distributed. That is,

$$\left(u_{cjm}, u_{kjm}\right)' \sim N\left(\mathbf{0}, \mathbf{\Sigma}_0\right)$$

where

$$\begin{aligned}
\Sigma_{0,11} &= \sigma_{c0}^2 \\
\Sigma_{0,k+1,k+1} &= \sigma_{k0}^2,\ k = 1, \ldots, K-1 \\
\Sigma_{0,1,k+1} &= \rho\sigma_{c0}\sigma_{k0},\ k = 1, \ldots, K-1 \\
\Sigma_{0,k+1,l+1} &= 0,\ k, l > 0,\ k \neq l.
\end{aligned}$$

σ_{c0} is the standard deviation of the measurement error of total cost, σ_{k0} is the standard deviation of the measurement error of the cost of kth input, and ρ is the correlation between the measurement error of the overall cost data and the measurement error of the cost of kth input.

Let l_{jm} be the log likelihood increment of firm j in market m. Since the measurement errors are jointly normally distributed with mean zero and variance-covariance matrix $\mathbf{\Sigma}$:

$$l_{jm} = -\frac{K}{2} ln\pi - \frac{1}{2} ln\left|\mathbf{\Sigma}\right| - \frac{1}{2}\mathbf{U}'_{jm}\mathbf{\Sigma}^{-1}\mathbf{U}_{jm}$$

where

$$\begin{aligned} U_{1jm} &= u_{cjm} \equiv C_{jm} - \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right), \\ U_{k+1jm} &= u_{kjm} \equiv C_{kjm} - C^*_{jm}\frac{\partial ln\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\partial ln w_{kjm}}, \ k = 1, \ldots, K-1. \end{aligned}$$

We choose the parameters so as to maximize the sum of log likelihood increments over all firms. That is,

$$\left(\widehat{\boldsymbol{\theta}}_c, \widehat{\boldsymbol{\theta}}_d\right) = argmax_{\{(\boldsymbol{\theta}_c, \boldsymbol{\theta}_d) \in \Theta_c \times \Theta_d\}} \sum_{j,m} l_{jm}$$

In case we only have data on total cost, we estimate Equation (16) by either nonlinear OLS, where u_{cjm} is the residual, or Maximum Likelihood by assuming u_{cjm} to be normally distributed with mean zero and variance σ_c^2. Next, given the parameter estimates $\widehat{\boldsymbol{\theta}}_{cq}$, the remaining parameters $\boldsymbol{\theta}_{c,-q}$ can be estimated by Equation (25).

Note that since the RHS of Equation (16) does not contain either the cost shock or the demand shock, our estimation procedure does not suffer from any endogeneity issues and thus, we do not need to impose any orthogonality conditions using instruments.

We next show that from the Equation (12) and the F.O.C. in Equation (14),

$$MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \widehat{\boldsymbol{\theta}}_d\right) = \widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \widehat{\boldsymbol{\theta}}_c\right) exp\left(\widehat{\varphi}\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right),$$

which results in

$$\widehat{\varphi}\left(\mathbf{x}_{jm}, \upsilon_{jm}\right) = lnMR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \widehat{\boldsymbol{\theta}}_d\right) - ln\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \widehat{\boldsymbol{\theta}}_c\right). \tag{26}$$

In the conventional approach, the cost shock is identified as part of the residual of the cost function estimates, i.e., it is the difference between the cost data and the cost predicted by the cost function. In contrast, we identify the cost shock as the difference between log marginal revenue and log of the deterministic component of the marginal cost. The economic logic behind the above result is as follows: the logit model predicts that firms with larger market shares have higher monopoly power. It then follows that a firm with high price and small market share does not have much monopoly power, and thus, its marginal cost should be close to its price. Then, we can infer that it has high marginal cost, and thus, a high cost shock.

It is important to note that in general, we need to impose additional assumptions for separately identifying the cost shock υ_{jm} from the observed product characteristics $\mathbf{x}_{jm}$. For example, if we assume linearity, and cost shock to be uncorrelated with observed characteristics, then

$$\varphi_{jm} = \eta_0 + \mathbf{x}_{jm}\boldsymbol{\eta}_{\mathbf{x}} + \upsilon_{jm}\eta_\upsilon.$$

where $\upsilon_{jm}\eta_\upsilon$ is the residual. We can normalize the cost shock and set $\eta_\upsilon = 1$. The less restrictive approach could be to assume that cost shock is orthogonal to the observed product characteristics and specify φ_{jm} as follows:

$$\varphi_{jm} = \eta\left(\mathbf{x}_{jm}\right) + \upsilon_{jm}.$$

where $\eta\left(\right)$ can be estimated as the polynomial of $\mathbf{x}_{jm}$. Even less restrictive would be to assume that υ_{jm} is independent of $\mathbf{x}_{jm}$, and $\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)$ is increasing in υ_{jm} by looking at the quantiles of $\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)$ given $\mathbf{x}_{jm}$. To do so, we first normalize the cost shock υ so that it is uniformly distributed in $(0,1)$. Then, τth quantile of υ is τ. Therefore, if we define the τth quantile of $y$conditional on $\mathbf{x}$ as $y_{\tau|\mathbf{x}}$, then,

$$y_{\tau|\mathbf{x}} = \varphi\left(\mathbf{x}, \tau\right)$$

Thus, in order to recover the cost shock, we do not require orthogonality between the cost shock and the output. Furthermore, the price coefficient on the demand side can be estimated without any orthogonality restrictions between the demand shock (or unobserved product characteristics) and other variables such as input price, own and rival firm observed product characteristics, and market size.

After estimation, we can analyze properties of the cost shock in various ways. For example, efficiency of a firm can be obtained by decomposing the cost shock as follows:

$$\upsilon_{jm} = -\zeta_{jm} + \eta_{jm}$$

where ζ_{jm} and η_{jm} are assumed to be independent and ζ_{jm} is specified to be half normally distributed, and η_{jm} to be mean zero normal. We can make additional decompositions as below as well

$$\upsilon_{jm} = \omega_j + \chi_m + u_{jm} + \eta_{jm}$$

where ω_j is the firm specific and χ_m is the market specific fixed effect, if we define the market as time. For more details on the estimation of firm specific fixed effects, see Greene (2005), Wang and Ho (2010) and others.

Diewert and Fox (2008) also use the F.O.C. of profit maximization to estimate markup and the cost function parameters. We extend their approach by including the cost shock into the cost function, and thereby explicitly deal with the endogeneity issues, but at the same time, without the use of instruments. We also jointly estimate the parameters of the cost function and the demand function.

We now explain further the difference between the approach taken in this paper, and the one in Byrne et al. (2021). They take a more general approach, where they show that the cost function can be expressed as a function of output, input price vector, observed product characteristics and marginal revenue as follows:

$$C^* \left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right) = \psi \left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right)\right)$$

They call the function $\psi\left(\right)$ the pseudo-cost function and show that it can be estimated from the data using the following equation:

$$C_{jm} = \psi \left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right)\right) + u_{jm}$$

where u_{jm} is the residual term that is assumed to be independent to all the variables in the pseudo-cost function and the marginal revenue function. They show that the pseudo-cost function ψ can be estimated nonparametrically by using polynomials of q_{jm}, $\mathbf{w}_{jm}$, $\mathbf{x}_{jm}$ and MR_{jm}. They then propose an algorithm to obtain the cost function from the pseudo-cost function by using the F.O.C. The procedure involves starting from a point $\left(q_{j^{(0)}m^{(0)}}, \mathbf{w}_{j^{(0)}m^{(0)}}, \mathbf{x}_{j^{(0)}m^{(0)}}, \mathbf{p}_{m^{(0)}}, \mathbf{s}_{m^{(0)}}, \mathbf{X}_{m^{(0)}}\right)$ $=\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m\right)$, and then finding another point that satisfies

$$\begin{aligned}
& MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right) = \frac{\partial C^* \left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial q} \\
= \; & (\Delta q)^{-1} \Big[\psi \left(q_{jm} + \Delta q, \mathbf{w}_{jm^{(1)}}, \mathbf{x}_{j^{(1)},m^{(1)}}, MR_{j^{(1)}} \left(\mathbf{p}_{m^{(1)}}, \mathbf{s}_{m^{(1)}}, \mathbf{X}_{m^{(1)}}; \boldsymbol{\theta}_d\right)\right) \\
& - \psi \left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, MR_j \left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right)\right)\Big].
\end{aligned}$$

That is, researchers have to find a point $\left(q_{j^{(1)}m^{(1)}}, \mathbf{w}_{j^{(1)}m^{(1)}}, \mathbf{x}_{j^{(1)}m^{(1)}}, \mathbf{p}_{m^{(1)}}, \mathbf{s}_{m^{(1)}}, \mathbf{X}_{m^{(1)}}\right)$ in the population which satisfies $q_{j^{(1)}m^{(1)}} = q_{jm} + \Delta q$ and the 2nd equality. Then, the two points

$(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m)$ and $\left(q_{j^{(1)}m^{(1)}}, \mathbf{w}_{j^{(1)}m^{(1)}}, \mathbf{x}_{j^{(1)}m^{(1)}}, \mathbf{p}_{m^{(1)}}, \mathbf{s}_{m^{(1)}}, \mathbf{X}_{m^{(1)}}\right)$ have approximately the same cost shock. By continuing this procedure, one can obtain the set of points in the population $\left\{\left(q_{j^{(l)}m^{(l)}}, \mathbf{w}_{j^{(l)}m^{(l)}}, \mathbf{x}_{j^{(l)}m^{(l)}}, \mathbf{p}_{m^{(l)}}, \mathbf{s}_{m^{(l)}}, \mathbf{X}_{m^{(l)}}\right)\right\}_{l=0}^{\infty}$, $q_{j^{(l)}m^{(l)}} = q^{(0)} + l\Delta q$, $l > 0$, that have approximately the same cost shock. That is,

$$
\begin{aligned}
& MR_j\left(\mathbf{p}_{m^{(l)}}, \mathbf{s}_{m^{(l)}}, \mathbf{X}_{m^{(l)}}; \boldsymbol{\theta}_d\right) \\
= \; & (\Delta q)^{-1}\left[\psi\left(q_{j^{(l+1)}m^{(l+1)}}, \mathbf{w}_{m^{(l+1)}}, \mathbf{x}_{j^{(l+1)}m^{(l+1)}}, MR_{j^{(l+1)}}\left(\mathbf{p}_{m^{(l+1)}}, \mathbf{s}_{m^{(l+1)}}, \mathbf{X}_{m^{(l+1)}}; \boldsymbol{\theta}_d\right)\right)\right. \\
& \left. -\psi\left(q_{j^{(l)}m^{(l)}}, \mathbf{w}_{m^{(l)}}, \mathbf{x}_{m^{(l)}}, MR_{j^{(l)}}\left(\mathbf{p}_{m^{(l)}}, \mathbf{s}_{m^{(l)}}, \mathbf{X}_{m^{(l)}}; \boldsymbol{\theta}_d\right)\right)\right].
\end{aligned}
$$

As we can see, this procedure relies on the numerical derivative which is subject to inaccuracies. These inaccuracies accumulate if we iterate on l.

Note, in out set up where the cost shock enters multiplicatively in the cost function, the pseudo-cost function can be expressed as follows.

$$
\psi\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_{d0}\right)\right) = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_{d0}\right).
$$

Then, as long as the functional form of the cost function is specified in advance, direct estimation of $\widetilde{C}\,()$ is more efficient.

4.3 Identification of market size.

Given the data on output, market share can be derived by dividing the output by the market size. However, market size may not be directly observable, unless we have a variable that unambiguously corresponds to it. Here, we follow Byrne et al. (2021) and estimate the market size as a function of observables $\mathbf{z}_m$, which typically would be demographic variables, as follows:

$$
Q_m = \lambda_0 + \boldsymbol{\lambda}_1 \mathbf{z}_m.
$$

Here, we use the example of logit demand function to show that market size can be identified.

$$
\begin{aligned}
C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right) &= \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}\left[p_{jm} - \frac{1}{\left(1 - q_{jm}/Q_m\right)\alpha}\right] \\
&= \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}\left[p_{jm} - \frac{\lambda_0 + \boldsymbol{\lambda}_1 \mathbf{z}_m}{\left(\lambda_0 + \boldsymbol{\lambda}_1 \mathbf{z}_m - q_{jm}\right)\alpha}\right].
\end{aligned}
$$

Then, we can identify the market size function separately from the cost function. First, the parenthesis that includes the price p_{jm} should only have variables that explains the market size, $\mathbf{z}_m$.

4.4 Estimation using translog cost function.

We next discuss the estimation of the parameters of the translog cost function and the logit market share function. The translog cost function is specified as follows:

$$\begin{aligned} lnC^*_{jm} &= \gamma_0 + \gamma_q lnq_{jm} + \frac{1}{2}\gamma_{qq}\left(lnq_{jm}\right)^2 + \sum_{k=1}^{K}\gamma_k lnw_{kjm} \\ &+ \frac{1}{2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\gamma_{kk'} lnw_{kjm} lnw_{k'jm} + \sum_{k=1}^{K}\gamma_{kq} lnw_q lnq_{jm} + ln\upsilon_{jm}. \end{aligned}$$

We impose the following restrictions on the cost function parameters so that the cost function has homogeneity of degree one in input prices:

$$\sum_{k=1}^{K}\gamma_k = 1,\ \sum_{k=1}^{K}\gamma_{kk'} = 0,\ \sum_{k'=1}^{K}\gamma_{kk'} = 0,\ \sum_{k=1}^{K}\gamma_{kq} = 0.$$

Then, taking the derivative of the log cost function with respect to log output, we obtain:

$$\begin{aligned} &\frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial lnq_{jm}} \\ = &\frac{q_{jm}MC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{C^*_{jm}} = \gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k=1}^{K}\gamma_{kq} lnw_k. \end{aligned} \tag{27}$$

The remaining parameters γ_k and $\gamma_{kk'}$, $k = 1, \ldots, K$, $k' = 1, \ldots, K'$ can be identified from the Shephard's Lemma as follows:

$$\frac{w_{kjm}L_{kjm}}{C^*_{jm}} = \frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial lnw_{kjm}} = \gamma_k + \sum_{k'=1}^{K}\gamma_{kk'} lnw_{k'jm} + \gamma_{kq} lnq_{jm}.$$

Thus, the log likelihood increment of firm j in market m is

$$l_{jm} = -\frac{K}{2}ln\pi - \frac{1}{2}ln\left|\boldsymbol{\Sigma}\right| - \frac{1}{2}\mathbf{U}'_{jm}\boldsymbol{\Sigma}^{-1}\mathbf{U}_{jm}$$

where

$$
\begin{aligned}
U_{1jm} &= C_{jm} - \left(\gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k=1}^{K} \gamma_{kq} lnw_k\right)^{-1} \left(p_{jm} - \frac{1}{(1-s_{jm})\,\alpha}\right) q_{jm} \\
U_{k+1jm} &= C_{kjm} - \frac{\gamma_k + \sum_{k'=1}^{K} \gamma_{kk'} lnw_{k'jm} + \gamma_{kq} lnq_{jm}}{\gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k'=1}^{K} \gamma_{k'q} lnw_{k'}} \left(p_{jm} - \frac{1}{(1-s_{jm})\,\alpha}\right) q_{jm},\ k = 1, \ldots, K-1
\end{aligned}
$$

and

$$
\begin{aligned}
\Sigma_{11} &= \sigma_c^2, \\
\Sigma_{k+1,k+1} &= \sigma_{kk}^2, k = 1, \ldots, K-1 \\
\Sigma_{k+1,1} &= \rho \sigma_c \sigma_{kk}, k = 1,\ \ldots K-1, \\
\Sigma_{kl} &= 0,\ for\ k \neq l
\end{aligned}
$$

The log likelihood is the sum of the log likelihood increment:

$$
l = \sum_{m=1}^{M} \sum_{j=1}^{J_m} l_{jm}
$$

As we can see above, we do not need to use any instruments for the estimation of either the demand parameters or the cost function parameters.

This approach is related to Kumbhakar et al. (2012), who estimate markups using the output elasticity of translog cost function. They start from the price being above markup

$$
p_{jm} > MC_{jm} \equiv \frac{\partial C_{jm}}{\partial q_{jm}},
$$

which implies

$$
\frac{p_{jm} q_{jm}}{C_{jm}} > \frac{\partial lnC_{jm}}{\partial lnq_{jm}},
$$

and thus, specify the revenue divided by cost by

$$
\frac{p_{jm} q_{jm}}{C_{jm}} = \frac{\partial lnC_{jm}}{\partial lnq_{jm}} + u_{jm} + v_{jm} = \gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k=1}^{K} \gamma_{kq} lnw_{kjm} + u_{jm} + v_{jm},\ u_{jm} \geq 0.
$$

We additionally focus on the endogeneity of output with respect to the shock $u_{jm} + v_{jm}$. That is, if firms tend to reduce output to increase markup, then the shock and the output are negatively correlated, resulting in downward bias of the estimate of γ_q and thus, the residual $u_{jm} + v_{jm}$

could misrepresent the true markup. In our approach, we deal with it by using the marginal revenue, more explicitly, which we derive from the demand side.

While we can deal with the endogeneity issue of both demand and supply side without using instruments, we impose some functional form assumptions on both the demand and cost functions. They can be fairly flexible, except that there are variables in the market share function but not in the cost function.

5 Monte-Carlo experiments

This section presents results from a series of Monte-Carlo experiments that highlight the finite sample performance of our estimator. To generate samples, we use the following random coefficients logit demand model:

$$s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right) = \int_\alpha \int_{\boldsymbol{\beta}} \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)} dF_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}; \boldsymbol{\theta}_{\boldsymbol{\beta}}\right) dF_\alpha\left(\alpha; \boldsymbol{\theta}_\alpha\right), \qquad (28)$$

where $\mathbf{x}_{jm}$ is the $1 \times K$ vector of observed product characteristics. We set the number of product characteristics K to be 1. We assume that each market has four firms, each producing one product (e.g., $J_m = J = 4$). Hence consumers in each market have a choice of $j = 1, \ldots, 4$ differentiated products or not purchasing any of them ($j = 0$).

On the supply-side, we assume firms compete on prices a la differentiated products Bertrand competition, use labor and capital inputs in production and have a Cobb-Douglas production function. Given output, input prices $\mathbf{w} = [w, r]'$ (w is the wage and r is the rental rate of capital), total cost and marginal cost functions are specified as

$$C^*\left(q, w, r, x, \upsilon\right) = \left[\left(\alpha_c + \beta_c\right)\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)}\left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] Bexp\left(x\eta + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}}$$

$$MC^*\left(q, w, r, x, \upsilon\right) = \left[\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)}\left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] Bexp\left(x\eta + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}-1}.$$

Notice that in the above specification, the cost function is homogeneous of degree one in input prices.

To create our Monte-Carlo samples, we generate wage, rental rate, variable cost shock, market size Q_m, and observable product characteristics x_{jm} as follows:

$$ln\left(w_{jm}\right) \sim i.i.d.TN\left(\mu_w, \sigma_w\right), \quad e.g., ln\left(w_{jm}\right) = \mu_w + \sigma_w \varrho_{wm}, \quad \varrho_{wm} \sim i.i.d.TN\left(0, 1\right).$$

$$ln\left(r_{jm}\right) \sim i.i.d.TN\left(\mu_r, \sigma_r\right), \quad e.g., ln\left(r_{jm}\right) = \mu_r + \sigma_r \varrho_{rm}, \quad \varrho_{rm} \sim i.i.d.TN\left(0,1\right).$$

$$Q_m \sim i.i.d.U\left(Q_L, Q_H\right).$$

$$x_{jm} \sim i.i.d.TN\left(\mu_x, \sigma_x\right), \quad e.g., x_{jm} = \mu_x + \sigma_x \varrho_{xjm}, \quad \varrho_{xjm} \sim i.i.d.TN\left(0,1\right).$$

$TN(0,1)$ is the truncated standard normal distribution, where we truncate both upper and lower 0.82 percentiles. $U(Q_L, Q_H)$ is the uniform distribution with lower bound of Q_L and upper bound of Q_H.

We also specify the unobserved characteristics and the cost shock so as to allow for correlation between ξ_{jm} and input prices, the cost shock, market size and the observed characteristics of the products other than j in market m denoted by $xo_{jm} \equiv (1/3)\sum_{l\neq j} \varrho_{xlm}$. Specifically, we set:

$$\begin{aligned} \xi_{jm} \quad = \quad & \delta_{0\xi} + \delta_{1\xi}\varrho_{\xi jm} + \delta_w \varrho_{wm} + \delta_r \varrho_{rm} + \delta_\upsilon \varrho_{\upsilon jm} \\ & + \delta_Q \Phi^{-1}\left(\delta + (1.0 - 2\delta)\frac{Q_m - Q_L}{Q_H - Q_L}\right) + \delta_x \varrho_{xlm} + \delta_{xo} xo_{jm}, \end{aligned}$$

and

$$\begin{aligned} \upsilon_{jm} \quad = \quad & \delta_{0\upsilon} - \delta_{1\upsilon}\varrho_{\upsilon jm} - \delta_w \varrho_{wm} - \delta_r \varrho_{rm} - \delta_\xi \varrho_{\xi jm} \\ & - \delta_Q \Phi^{-1}\left(\delta + (1.0 - 2\delta)\frac{Q_m - Q_L}{Q_H - Q_L}\right) - \delta_x \varrho_{xlm} - \delta_{xo} xo_{jm}, \end{aligned}$$

where ϱ_ξ is the idiosyncratic component of the demand shock.

For transforming the uniformly distributed market size shock to truncated normal distribution, we use small positive $\delta = 0.025$ for truncation. We truncate the distribution of the shocks to ensure that the true cost function is positive and bounded given the parameter values of the cost function we set (which will be discussed later).

By construction, neither market size nor observed product characteristics of own product or products of other firms can be used as instruments since they are designed to be correlated with the cost shock. Furthermore, cost shock is set to be correlated with the demand shock, and thus, demand side variables such as prices and market shares cannot be used as instruments either. The only sources of exogenous variation we allow for the model are input prices. We assume competitive markets for inputs and thus, they are exogenous to the firm. In other words, we do not consider monopsony or oligopolistic behavior of firms in the input markets. In sum, we exclude the possibilit of any conventional instruments in either demand or supply equation. Since no instruments for demand estimation is valid, we won't be able to estimate first the parameters of the market share equation, and then, derive the marginal revenue to derive the equivalent marginal cost, and then, from the marginal cost, calculate the cost function.

To solve for the equilibrium price, quantity, and market share for each oligopoly firm, we use the golden section search on price.[5]

We estimate the paramters using GLS, where, in this case, we assume we know the true variance covariance matrix of the measuremente errors, and set it to $\mathbf{W} = \mathbf{I}$.

Table 1 summarizes the parameter setup of the Monte-Carlo experiments.

In Table 2, we present the Monte-Carlo results of the direct estimator that estimates the cost function parameters without using the pseudo-cost function. We report the average, standard deviation, and square root of the mean squared errors (RMSE) of the parameter estimates of the BLP market share function and the Cobb-Douglas cost function from 100 Monte-Carlo simulation/estimation replications. As we can see, the averages of the parameter estimates are close to the true values, even for the cases with sample size of only 200. Furthermore, the standard errors and RMSEs of the estimates decrease with the sample size, demonstrating the validity of our approach.

In Table 3, we present the results where we also allow the observed product characteristics to be correlated with the unobserved product characteristics and the cost shock. In paticular, we set $corr\left(x_{jm}, \xi_{jm}\right) > 0$ and $corr\left(x_{jm}, \upsilon_{jm}\right) < 0$. Then, we can see that the parameter estimates $\widetilde{\mu}_{\beta}$ and $\widehat{\eta}_c$, which are the coefficients of the observed product characteristics are both biased, indicating the bias due to the correlation mentioned above. Nonetheless, we can see that all the other parameter estimates are close to the true values, and the standard errors and the RMSEs decrease with sample size.

We next present the results of the scenario where quantity of inputs are not observable. To obtain consistent parameter estimates, we require that log wage, log rental rate are uncorrelated with the cost shock. The results are shown in Table 4. Except for the parameter estimates with sample size of 200, all the other sample averages of the parameter estimates are close to the true values. Furthermore, the standard deviations and the RMSEs decrease with sample size. Overall, if we compare the standard deviations and RMSEs with the ones in Table 2, we can see that they are higher. Therefore, we conclude that the use of input quantity improves the efficiency of the estimation procedure.

[5] The algorithm for finding equilibria in oligopoly markets is available upon request.

6 Conclusion

We have developed a new methodology for estimating the cost parameters of a differentiated products oligopoly model. The method uses data on prices, market shares, and product characteristics, and some data on firms' costs. Using these data, our approach identifies demand parameters in the presence of price endogeneity, and the cost function in the presence of output endogeneity without any instruments. Moreover, our method can accommodate measurement error and fixed cost in cost data that do not have to be randomly distributed.

References

Amsler, C., A. Prokhorov, and P. Schmidt (2017): "Endogenous environmental variables in stochastic frontier models," *Journal of Econometrics*, 199, 131–140.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

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

Byrne, D. P., S. Imai, N. Jain, V. Sarafides, and M. Hirukawa (2021): "Identification and Estimation of Differentiated Products Models without Instruments using Cost data," *mimeo.*

Diewert, E. W. and K. J. Fox (2008): "On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups," *Journal of Econometrics*, 145, 174–193.

Gandhi, A., S. Navarro, and D. Rivers (2020): "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, 128, 2973–3016.

Genesove, D. and W. P. Mullin (1998): "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914," *RAND Journal of Economics*, 29, 355–377.

Greene, W. (2005): "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model," *Journal of Econometrics*, 126, 269–303.

Kumbhakar, S. C. (1997): "Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share equations: An Exact rRelationship," *Journal of Econometrics*, 76, 351–356.

——— (2001): "Estimation of profit functions when profit is not maximum," *American Journal of Agricultural Economics*, 83, 1–19.

Kumbhakar, S. C., S. Baardsen, and G. Lien (2012): "A New Method for Estimating Market Power with an Application to Norwegian Sawmilling," *Review of Industrial Organization*, 40(2), 109–129.

Kumbhakar, S. C. and E. G. Tsionas (2005): "Measuring Technical and Allocative Inefficiency in the Translog Cost System: a Bayesian Approach," *Journal of Econometrics*, 126, 355–384.

Kutlu, L., K. C. Tran, and M. G. Tsionas (2019): "A time-varying true individual effects model with endogenous regressors," *Journal of Econometrics*, 211, 539–559.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

Petrin, A. and B. Seo (2016): "Identification and Estimation of Discrete Choice Demand Models when Observed and Unobserved Characteristics are Correlated," *University of Minnesota Working Paper*.

Wang, C. J. (2003): "Productivity and Economies of Scale in the Production of Bank Service Value Added," *FRB Boston Working Papers Series*, 03-7.

Wang, H.-J. and C.-W. Ho (2010): "Estimating Fixed-Effect Panel Stochastic Frontier Models by Model Transformation," *Journal of Econometrics*, 157, 286–296.

7 Tables and Figures

Table 1: Monte Carlo Parameter Values

Parameter	Description	Value
(a) Demand-side parameters		
μ_α	Price coef. mean	2.0
σ_α	Price coef. std. dev	0.5
μ_β	Product characteristic coef. mean	1.0
σ_β	Product characteristics coef. std. dev.	0.2
μ_X	Product characteristic mean	3.0
σ_X	Product characteristic std. dev.	1.0
δ_0	Unobserved product quality mean	2.0
δ_ξ	Unobserved product quality std. dev.	0.5
Q_L	Lower bound on market size	5.0
Q_H	Upper bound on market size	10.0
(b) Supply-side parameters		
η	coef. on observed product characteristics	0.2
μ_w	log wage mean	1.0
σ_w	log wage std. dev.	0.2
μ_r	log rental rate mean	1.0
σ_r	Rental rate std. dev.	0.2
μ_v	log cost shock mean	-5.0
σ_v	log cost shock std. dev.	0.1
J	Number of firms in each market	4
B	Scaling factor for output in the cost function	1.0
(c) Cost measurement error		
$\sigma_{\nu+\varsigma}$	Measurement std. dev.	0.4
(d) Correlation parameters with unobservables ξ_{jm} and v_{jm}		
δ_x	ξ_{jm} and $\mathbf{x}_{jm}$ correlation	0
δ_{xo}	ξ_{jm} and $\mathbf{X}_{-jm}$ correlation	0.0833
δ_w	ξ_{jm} and w_m correlation	0.0833
δ_r	ξ_{jm} and r_m correlation	0.0833
δ_v	ξ_{jm} and v_{jm} correlation	0.0833
δ_Q	ξ_{jm} and Q_m correlation	0.0833
ζ_Q	v_{jm} and Q_m correlation	0.0833
(e) Cobb-Douglas Production Function Parameters		
α_c	Labor coef. in Cobb-Douglas prod. fun.	0.5
β_c	Capital coef. in Cobb-Douglas prod. fun.	0.3

Table 2: Parameter estimates based on Shephard's Lemma

(a) Demand side parameters

		$\hat{\mu}_\alpha$			$\hat{\sigma}_\alpha$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.031	0.3687	0.3682	0.5102	0.1260	0.1258
100	400	1.999	0.2740	0.2726	0.4982	0.1056	0.1051
200	800	1.982	0.1799	0.1798	0.4908	0.0751	0.0753
400	1600	2.001	0.1556	0.1548	0.4965	0.0600	0.0598
True Value		2.0			0.5		

(a) Demand side parameters

		$\hat{\mu}_\beta$			$\hat{\sigma}_\beta$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.9913	0.2412	0.2401	0.2167	0.1155	0.1161
100	400	0.9859	0.1880	0.1876	0.2160	0.0973	0.0981
200	800	0.9896	0.1289	0.1287	0.2048	0.0595	0.0594
400	1600	1.0016	0.1034	0.1029	0.2027	0.0404	0.0402
True Value		1.0			0.2		

(b) Production function parameters

		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.5233	0.0679	0.0715	0.3111	0.0412	0.0425
100	400	0.5144	0.0497	0.0515	0.3057	0.0316	0.0320
200	800	0.5071	0.0419	0.0423	0.3051	0.0281	0.0284
400	1600	0.5034	0.0295	0.0296	0.3014	0.0180	0.0180
True Value		0.5			0.3		

		$\hat{\eta}$		
Markets	Sample Size	Mean	Std. Dev.	RMSE
50	200	0.1998	0.0130	0.0129
100	400	0.2002	0.0083	0.0082
200	800	0.2007	0.0057	0.0057
400	1600	0.1995	0.0048	0.0040
True Value		0.2		

Notes: Monte-carlo experiment results based on calibration described in panels (a)-(d) of Table 1.

Table 3: Parameter estimates based on Shephard's Lemma
(Product characteristic x_{jm} and unobserved product quality ξ_{jm},
cost shock υ_{jm} are correlated)

		(a) Demand side parameters					
		$\hat{\mu}_\alpha$			$\hat{\sigma}_\alpha$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.059	0.3790	0.3816	0.5095	0.1398	0.1394
100	400	2.011	0.2512	0.2501	0.5021	0.1053	0.1048
200	800	1.994	0.1631	0.1623	0.4931	0.0793	0.0792
400	1600	2.005	0.1568	0.1560	0.4979	0.0564	0.0562
True Value		2.0			0.5		

		(a) Demand side parameters					
		$\hat{\mu}_\beta$			$\hat{\sigma}_\beta$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	1.1614	0.2294	0.2795	0.2335	0.1314	0.1349
100	400	1.1542	0.1608	0.2222	0.2151	0.0873	0.0882
200	800	1.1535	0.1087	0.1878	0.2054	0.0594	0.0593
400	1600	1.1606	0.0967	0.1872	0.2040	0.0431	0.0431
True Value		1.0			0.2		

		(b) Production function parameters					
		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.5210	0.0654	0.0684	0.3098	0.0407	0.0416
100	400	0.5130	0.0500	0.0514	0.3046	0.0303	0.0305
200	800	0.5044	0.0357	0.0358	0.3035	0.0238	0.0239
400	1600	0.5031	0.0278	0.0279	0.3014	0.0170	0.0170
True Value		0.5			0.3		

			$\hat{\eta}$	
Markets	Sample Size	Mean	Std. Dev.	RMSE
50	200	0.1694	0.0188	0.0358
100	400	0.1688	0.0145	0.0343
200	800	0.1679	0.0098	0.0336
400	1600	0.1666	0.0083	0.0344
True Value		0.2		

Notes: Monte-carlo experiment results based on calibration described in panels (a)-(d) of Table 1.

Table 4: Parameter estimates without using input data
(Product characteristic x_{jm} and unobserved product quality ξ_{jm},
cost shock υ_{jm} are uncorrelated)

		(a) Demand side parameters					
		$\hat{\mu}_\alpha$			$\hat{\sigma}_\alpha$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.698	2.787	2.860	0.4935	0.2994	0.2990
100	400	2.115	0.5304	0.5403	0.4540	0.1877	0.1923
200	800	2.066	0.3504	0.3546	0.4593	0.1387	0.1439
400	1600	2.035	0.2242	0.2258	0.4837	0.1141	0.1147
True Value		2.0			0.5		
		(a) Demand side parameters					
		$\hat{\mu}_\beta$			$\hat{\sigma}_\beta$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	1.3284	1.5293	1.5567	0.3348	0.4509	0.4684
100	400	1.0734	0.4003	0.4050	0.2235	0.1437	0.1449
200	800	1.1057	0.2603	0.2652	0.2102	0.0978	0.0979
400	1600	1.1027	0.1593	0.1608	0.2027	0.0755	0.0752
True Value		1.0			0.2		
		(b) Production function parameters					
		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.5216	0.0733	0.0761	0.3048	0.0917	0.0914
100	400	0.5101	0.0610	0.0615	0.3101	0.0724	0.0728
200	800	0.4997	0.0443	0.0441	0.2983	0.0530	0.0528
400	1600	0.4992	0.0342	0.0340	0.3024	0.0379	0.0378
True Value		0.5			0.3		
		$\hat{\eta}$					
Markets	Sample Size	Mean	Std. Dev.	RMSE			
50	200	0.1968	0.0168	0.0170			
100	400	0.1998	0.0106	0.0105			
200	800	0.2007	0.0068	0.0068			
400	1600	0.1996	0.0060	0.0059			
True Value		0.2					

Notes: Monte-carlo experiment results based on calibration described in panels (a)-(d) of Table 1.