Estimating Cost Functions in Differentiated Product Oligopoly Models without Valid Instruments *

Susumu Imai†

Neelam Jain‡

Hiroto Suzuki§and

Miyuki Taniguchi¶

March 12, 2023

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. (2022) 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. We also propose a method to consistently estimate the coefficient of the observed product characteristic in the market share function when valid instruments are not available. We illustrate our methodology with Cobb-Douglas technology and logit demand structure, assuming a multiplicative cost shock. We also conduct Monte Carlo experiments and show that our method works well even when the conventional instruments are invalid.

Keywords: Differentiated Goods Oligopoly, Instruments, Identification, Cost data.

JEL Codes: C13, C14, L13, L41

Preliminary draft. Comments are welcome.

*We thank seminar participants at BI Norway, and City, University of London.

†Corresponding author, Faculty of Economics and Business, Hokkaido University, susimaig@gmail.com

‡Department of Economics, City, University of 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 cost functions in a differentiated products oligopoly 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, and 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, the profit function based approach would also be subject to the endogeneity issue in the case of a differentiated products model since firms also choose prices.

We follow Byrne et al. (2022) 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. (2022) use their two-step nonlinear sieve estimator 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 start with 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 total cost, where we replace the unobservable marginal cost with marginal revenue which is a function of observables and parameters. We then generalize this technique to the non-multiplicative cost shock case. 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. We conduct several Monte-Carlo experiments to demonstrate the validity of our method, assuming either logit demand or BLP demand and either Cobb-Douglas or translog technology.

One issue that Byrne et al. (2022) do not address is the estimation of the coefficients on the observed product characteristics in demand estimation. Using a simple setup, we show how we can verify the validity of instruments and then, if needed, construct valid instruments from the invalid ones to consistently estimate these coefficients. To do so, we observe that in demand estimation, the same instruments are used for the price variable and the observed product characteristics. Therefore, we can use instruments to estimate the price coefficient and compare it with the estimate obtained from the instrument-free approach of Byrne et al. (2022). We can then check if the instruments are valid. If they turn out to be invalid, we can construct valid instruments using parametric functions of the invalid instruments, where the parameters are chosen such that the price coefficient estimated by the constructed instruments is close to the instrument-free estimated one. We use these constructed instruments for estimating the coefficients on the observed product characteristics. We show the validity of our approach both theoretically as well as numerically via Monte-Carlo experiments.

This paper is organized as follows. In Section 2, we review the IV-based estimation of the differentiated products model of demand and of the cost function. Then, we discuss the first order condition 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 examine identification when demand and cost data are available and present our formal identification results, including the example of Cobb-Douglas technology and logit demand. In Section 2.2, we discuss estimation issues. Section 4 contains a Monte-Carlo study that illustrates the effectiveness of our estimator. In Section 5 we conclude.

2 IV estimation of Demand and Cost Functions

The key component of our methodology is the first order condition (F.O.C.) of the firm's profit maximzation, which requires its marginal revenue to equal its 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. (2022) 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 joint 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.

We assume $M > 1$ isolated markets.[1] Market m has $J_m + 1 > 2$ products where aggregate demand for product j across individuals is,

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

where q denotes output, Q denotes market size and s denotes market share. If ϵ_{ijm} is assumed to have a logit distribution as in Berry (1994), then, the aggregate market share for product j in market m is given by,

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

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

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

share in Equation (1) as $s_j(\boldsymbol{\delta}(\boldsymbol{\theta})) \equiv s_j(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta})$ 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 $F_\alpha(.; \boldsymbol{\theta}_\alpha)$, where $\boldsymbol{\theta}_\alpha$ is the parameter vector of the distribution, and similarly, $\boldsymbol{\beta}$ has a distribution function $F_{\boldsymbol{\beta}}(.; \boldsymbol{\theta}_{\boldsymbol{\beta}})$ 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 $\boldsymbol{\beta}$:

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

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]

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 $\mathbf{s}_m = [s_{0m}, s_{1m}, ..., s_{J_m m}]'$ 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 $s_j(\boldsymbol{\delta}_m(\boldsymbol{\theta}_d); \boldsymbol{\theta}_d)$ 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,

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

[2]Note that in BLP, the distribution of the random parameters is the same across markets. That is, there is consumer-level heterogeneity within markets but not across markets.

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 $\delta_{jm}\left(\boldsymbol{\theta}_d\right)$ 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 $\boldsymbol{\theta}_{d0}$:

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

where $\mathbf{z}_{jm}$ is a $T\times 1$ 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. $C\left(\right)$ is assumed to be strictly increasing and continuously differentiable in output and the 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 endogenously 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' Maximization Problem

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[3] 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 input prices i.e., $MC_{jm} = exp\left(\mathbf{w}_{jm}\boldsymbol{\gamma}_w + q_{jm}\gamma_q + \upsilon_{jm}\right)$ (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 the F.O.C. based orthogonality conditions are sufficient for identifying demand parameters.

In this research, we follow the insight of Byrne et al. (2022) 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 estimating price coefficients of the demand function and output coefficient of the cost function.[4]

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

[4] As we explained above, estimating demand and cost sides separately raises endogeneity concerns. Further, we cannot use the MR equation to estimate the parameters because it is not observed in the data. Same is true of MC, even if we were to use inversion to use cost to control for the cost shock. However, jointly estimating has the

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. (2022). Given output, input prices and observed product characteristics, they use marginal revenue to control for the cost shock in the cost function. To do so, they transform the cost function into a pseudo-cost function, which is a function of output, input price, observed product characteristics and marginal revenue. The parameters of the marginal revenue function are chosen to have the best fit of the pseudo-cost function to the cost data.

While their approach achieves consistent estimation of the price coefficients of the demand function without instruments, there is a large bias in the estimate of their nonparametric cost function when we try to recover it from the pseudo-cost function. This is due to the pseudo-cost function being highly nonlinear in output, input prices, observed characteristics and marginal revenue. Further, their approach requires numerical integration of marginal revenue which results in bias.

In this paper, we develop an alternative approach where the parameters of the cost function are estimated directly, that is, without using the semi-parametric pseudo-cost function. We are able to do so by using the first order condition of the firm to derive its cost shock which is the source of the endogeneity issue. This in turn is due to the assumption of a parametric cost function in the first instance. Further, we first simplify the model and assume that the cost shock component enters multiplicatively in the cost function. Such a restriction is frequently imposed in the production function analysis (see Gandhi et al. (2020)). The cost shock is simply the inverse of their productivity shock. It turns out that in this case, we can estimate the demand and cost parameters in a straightforward manner because the pseudo-cost function becomes parametric.

advantage that we can exploit the first order condition, thereby using the variation of prices and market shares on the demand side, as well as the variation of outputs and cost on the supply side to identify all of the parameters of interest. Furthermore, exploiting the properties of particular production functions and demand functions, we find that we do not need any instruments to do so. Note that we rely on the exclusion restriction that outputs do not enter marginal revenue directly and market shares do not enter the cost function directly.

3.1 The Cobb-Douglas Production/Cost Function and Logit Demand

Suppose that output is a function of labor and capital inputs, denoted by L and K, given by

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

Here, x is an observed characteristic and υ is the unobserved cost shock (inverse of the productivity shock). In this subsection, we focus on the identification of α_c and β_c and the component that includes the cost shock, $x\eta + \upsilon$.

We can derive the true cost function from the following cost minimization problem:

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

where w is the wage and r is the rental rate of capital. Then, the cost and the marginal cost functions are as follows:

$$C^*(q, w, r, x, \upsilon) = \left[(\alpha_c + \beta_c) \left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] Bexp\,(x\eta + \upsilon)\, q^{\frac{1}{\alpha_c+\beta_c}}. \tag{11}$$

$$MC^*(q, w, r, x, \upsilon) = \left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)} Bexp\,(x\eta + \upsilon)\, q^{\frac{1}{\alpha_c+\beta_c}-1}. \tag{12}$$

Our methodology is centered around the fact that dividing C^* in Equation (11) by MC^* in Equation (12) eliminates the cost shock υ. To be precise,

$$\frac{C^*(q, w, r, x, \upsilon)}{MC^*(q, w, r, x, \upsilon)} = (\alpha_c + \beta_c)\, q,$$

which implies that

$$C^*(q_{jm}, w_{jm}, r_{jm}, \mathbf{x}_{jm}, \upsilon) = (\alpha_c + \beta_c)\, q_{jm} MC^*(q_{jm}, w_{jm}, r_m, x_{jm}, \upsilon).$$

Next we use the first order condition of the firm's profit maximization problem, namely, that marginal revenue equals marginal cost, to rewrite C^* as

$$C^*(q_{jm}, w_{jm}, r_{jm}, \mathbf{x}_{jm}, \upsilon) = (\alpha_c + \beta_c)\, q_{jm} MR_j\,(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d), \tag{13}$$

We then set the observed cost, C_{jm} to be the true cost C^* plus an i.i.d. measurement error u_{cjm}, that is,

$$C_{jm} = C^*_{jm} + u_{cjm} = (\alpha_c + \beta_c)\, q_{jm} MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d) + u_{cjm}. \tag{14}$$

This is the equation we use to estimate the cost and the demand parameters.[5] Notice that because the cost shock is eliminated from the RHS, the only unobservable in the RHS is the measurement error, and by assumption, it is independent to all the other variables in the RHS. Therefore, we do not face any endogeneity issues in estimating the parameter $(\alpha_c + \beta_c)$. Thus, our estimation methodology does not require any instruments. Note that we are essentially estimating the pseudo-cost function since cost is expressed as a function of marginal revenue instead of the cost shock, just as in Byrne et al. (2022). However, we can recover the cost shock from the first order condition as follows:

$$\begin{aligned} MC^*(q_{jm}, w_{jm}, r_m, x_{jm}, \upsilon) \equiv \widetilde{MC}(q_{jm}, w_{jm}, r_m)\, exp(x\eta + \upsilon) = MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d) \\ exp(x\eta + \upsilon) = \frac{MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d)}{\widetilde{MC}(q_{jm}, w_{jm}, r_m)}. \end{aligned}$$

Taking logs, we derive

$$x\eta + \upsilon = ln(MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d)) - ln(\widetilde{MC}(q_{jm}, w_{jm}, r_m)).$$

We can also estimate the cost function directly (rather than dividing cost by marginal cost and then substituting from the F.O.C.) by substituting for $exp(x\eta + \upsilon)$ into the true cost function, given by Equation (11). This yields:

$$\begin{aligned} C^*(q, w, r, x, \upsilon) &= \left[(\alpha_c + \beta_c)\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)}\left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] B \frac{MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d)}{\widetilde{MC}(q_{jm}, w_{jm}, r_m)} q^{\frac{1}{\alpha_c+\beta_c}} \\ &= (\alpha_c + \beta_c)\, q MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d). \end{aligned}$$

The last equality follows from substituting for $\widetilde{MC}$.

In the multiplicative case, the two ways differ only in the sequencing of steps used but as we will see more clearly later, the advantage of the direct approach is that the cost shock does

[5] We assume that we have data on outputs, input prices, total cost (including an i.i.d. measurement error), product prices, market shares and observed characteristics. In our model (further details are in the general section below), market shares and outputs are linked via market size (Q) as follows: $s_{jm} = \frac{Q_m}{q_{jm}}$. Thus, we only need two out of these three variables in the data.

not need to enter the cost function multiplicatively. As long as marginal cost is assumed to be an increasing function of the cost shock, we can use the first order condition to derive the cost shock as a function of marginal revenue and then substitute in the cost function directly.

If we further assume that the market share function is logit, i.e. is specified as in Equation (1), then, we can derive the marginal revenue function:

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

Then, the estimating equation becomes:

$$C_{jm} = q_{jm}p_{jm}(\alpha_c + \beta_c) - \left(\frac{q_{jm}}{1 - s_{jm}}\right)\frac{\alpha_c + \beta_c}{\alpha} + u_{cjm} \tag{16}$$

This is a linear regression equation with observed cost being the dependent variable and revenue $q_{jm}p_{jm}$ and $\frac{q_{jm}}{1-s_{jm}}$ being the independent variables. Since the error term of the Equation (16) is the measurement error u_{cjm}, which we assume to be independent to all the the other variables in the RHS (p_{jm}, q_{jm} and s_{jm}), the coefficients $\alpha_c + \beta_c$ and $1/\alpha$ are estimated without any bias via simple OLS. Hence, α is estimated consistently.

However, this equation does not identify α_c or β_c separately. To do so, as in the existing literature, we can use Shephard's Lemma if we have the cost data for each input. To see this, we first derive the log cost function as follows:

$$\begin{aligned} lnC^* &= ln(\alpha_c + \beta_c) - \frac{\alpha_c}{\alpha_c + \beta_c}ln\alpha_c - \frac{\beta_c}{\alpha_c + \beta_c}ln\beta_c \\ &\quad + \frac{\alpha_c}{\alpha_c + \beta_c}lnw + \frac{\beta_c}{\alpha_c + \beta_c}lnr + \frac{1}{\alpha_c + \beta_c}lnq + x\eta + \upsilon \end{aligned} \tag{17}$$

Shephard' Lemma states that:

$$\frac{\partial lnC^*(q_{jm}, w_{jm}, r_{jm}, x_{jm}; \boldsymbol{\theta}_{c0})}{\partial lnw_{jm}} = \frac{\alpha_c}{\alpha_c + \beta_c} = \frac{w_{jm}L_{jm}}{C^*_{jm}}. \tag{18}$$

Denoting the labor input cost by C_{Ljm}, and manipulating this equation, and using Equation (13), we obtain

$$\begin{aligned} C_{Ljm} = w_{jm}L_{jm} + u_{Ljm} &= \alpha_c q_{jm} MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d) + u_{Ljm} \\ &= \alpha_c q_{jm}\left[p_{jm} - \frac{1}{(1 - s_{jm})\alpha}\right] + u_{Ljm} \end{aligned} \tag{19}$$

where u_{Ljm} is an i.i.d. measurement error in input cost. Thus, we identify α_c, and $\beta_c = \alpha_c + \beta_c - \alpha_c$. Since the error term of Equation (19) is the measurement error, assumed to be independent to p_{jm}, q_{jm} and s_{jm}, we can estimate α_c, and thus, β_c consistently.

In our estimation method explained above, we do not need to use any instruments because neither the unobserved product characteristics ξ_{jm}, which correspond to the demand shock, nor the cost shock υ_{jm} enter in Equations (16) and (19). This is due to the properties of the demand and cost function specifications we have used. First, in logit demand, marginal revenue is a function of price and market share, but not of the unobserved product characteristics (see Berry (1994) for more details). Second, we have assumed that the cost shock enters multiplicatively in the cost function. After estimating α_c and β_c, using Equation (17) we can recover the component $x_{jm}\eta + \upsilon_{jm}$.

Note that the above Cobb-Douglas cost function specification has some additional benefits. First, we can see from Equation (16) that the RHS is a function only of the demand side variables, i.e. it is not a function of the input prices. Thus, as long as we can reasonably assume the cost data is generated by the Cobb-Douglas production function, we can estimate the demand parameters and the returns to scale, $\alpha_c + \beta_c$, without any data on inputs and input prices. Furthermore, Equation (19) implies that as long as we have data on total labor cost, we can separately identify α_c and β_c without using any variation in input prices, in this case, wage rate w_{jm} and rental rate of capital r_{jm}. **?** and Gandhi et al. (2020) also estimate the Cobb Douglas production function without relying on the input price data. Instead of the demand side data, they exploit the panel assumption of lagged inputs being uncorrelated with the productivity shock innovation. That is, they use lagged inputs as instruments. In contrast, we use the information from the demand data, such as product prices and market shares as the source of identification, and thus we do not need any instruments for dealing with the endogenous price in the demand function and the endogenous output in the cost function.

3.2 Identification of the coefficients of the observed characteristics - the Logit case

Byrne et al. (2022) propose an instrument-free identification strategy for the coefficient of the endogenous price by using cost data. For observed product characteristics, which may be correlated with the error term, they assume that valid instruments are available. This is somewhat problematic, as they observe, because then the rival firms' observed characteristics (as BLP suggest) could very well be used as instruments for price as well and their cost data based approach may not be needed except as a check. In this paper, we address this issue in the context of logit demand. Our approach is to use the consistent estimate of the price coefficient provided by the methodology of Byrne et al. (2022) (and applied in this paper to specific demand and cost functions) to develop an IV-based identification method for the coefficients of the observed product characteristics. The novelty of our approach is that we are able to verify if the instruments are valid and if not, to construct valid instruments from the invalid ones. To do so, we borrow the idea of the Hausman specification test in which researchers compare the OLS estimate and the IV estimate. Under the assumption that the instruments are valid, if the two estimates are close, then researchers can conclude that the OLS parameter estimate does not suffer from the endogeneity bias, and use its standard error for hypothesis testing.

Our approach is to compare the price coefficient estimated using the IV-free, cost data based approach of Byrne et al. (2022) with the one estimated by the conventional IV approach, where for instruments, we use the commonly used variables such as input prices and their polynomials. If the two estimates are close, then the instruments are valid for price, and we can use them as instruments for the observed product characteristics as well. If the estimates are not close, then we construct valid instruments using a linear combination of some of the variables and use their polynomials as candidate instruments. To ensure validity, we optimize over all such linear combinations and choose the one whose polynomials, when used as instruments, yields a price coefficient estimate that is close to the IV-free estimate.

Our analysis is related to some recent work on consistent estimation of demand parameters when firms in oligopolistic markets choose observed and unobserved product characteristics in addition to price. In such a setting, the commonly used instruments such as current input prices and observed product characteristics of rival firms are no longer valid instruments (See Petrin and Seo (2016) - they use lags of these variables as instruments). However, product characteristics tend not to change as frequently as price in many industries, likely due to substantial costs of doing so. Thus, in order to properly take into account the endogeneity of the product characteristics, we need to use a dynamic oligopoly model, which is outside the scope of this study.

We explain our methodology using a simple model. We specify the unobserved product characteristics as:

$$\xi_{jm} = \mu_{0\xi} + \varrho_{\xi jm} + \varrho_{\xi 3jm} \tag{20}$$

$$\varrho_{\xi jm} \equiv \varrho_{\xi 1m} + \varrho_{\xi 2jm} \tag{21}$$

where $\varrho_{\xi 1m}$ is i.i.d. mean zero distributed with standard deviation $\sigma_{\xi 1}$, and similarly, $\varrho_{\xi 2jm}$ is i.i.d. mean zero distributed with standard deviation $\sigma_{\xi 2}$ and $\varrho_{\xi 3jm}$ is i.i.d. mean zero distributed with standard deviation $\sigma_{\xi 3}$. Thus, we decompose the unobserved product characteristics into a market-specific component and two product specific components.

We next assume that the potential instruments are input prices, (w_{jm}, r_{jm}), specified as follows:

$$w_{jm} = \mu_w + \varrho_{w1jm} + \varrho_{w2jm} + \delta_{w\xi}\varrho_{\xi jm}, \ .r_{jm} = \mu_r + \varrho_{r1jm} + \varrho_{r2jm} + \delta_{r\xi}\varrho_{\xi jm} \tag{22}$$

where ϱ_{w1jm}, ϱ_{w2jm} ϱ_{r1jm}, ϱ_{r2jm} are all i.i.d. mean zero distributed with standard deviations σ_{w1}, σ_{w2} and σ_{r1}, σ_{r2}, respectively. Finally, we specify the observed product characteristic (assumed to be one dimensional, for simplicity) as follows:

$$x_{jm} = \mu_x + \delta_{xw}\varrho_{w1jm} + \delta_{xr}\varrho_{r1jm} + \delta_{x\xi}\varrho_{\xi jm} + \varrho_{xojm}, \tag{23}$$

where ϱ_{xojm} is i.i.d. mean zero distributed with standard deviations σ_{xo} and xo denotes the observed characteristics of other firms. Note that these specifications allow for possible correlation between observed and unobserved characteristics of the same firm; observed characteristics across firms due to the market specific component of ξ and input prices and observed and unobserved characteristics.

Now, for the sake of simplicity, let

$$\mu_w = \mu_r = \mu_x = 0. \tag{24}$$

Recall Equation (2), where $\delta_{jm}(\boldsymbol{\theta}_d) = \log(s_{jm}) - \log(s_{0m})$, $j = 1, \ldots, J_m$. That is,

$$\log(s_{jm}) - \log(s_{0m}) = \mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm} \tag{25}$$

Since α_0 is identified from the cost data, we consider it to be given. If $\delta_{x\xi} \neq 0$, then the OLS estimation of β given α_0 is biased because

$$\begin{aligned}
&Cov\left(x_{jm}, \log\left(s_{jm}\right) - \log\left(s_{0m}\right) + p_{jm}\alpha_0\right) \\
&\quad = Cov\left(x_{jm}, x_{jm}\beta_0 + \xi_{jm}\right) &=& \ Var\left(x_{jm}\right)\beta_0 + Cov\left(x_{jm}, \xi_{jm}\right) \\
&&=& \ Var\left(x_{jm}\right)\beta_0 + \delta_{x\xi}\sigma_\xi^2 \\
&\quad \frac{Cov\left(x_{jm}, x_{jm}\beta_0 + \xi_{jm}\right)}{Var\left(x_{jm}\right)} &=& \ \beta_0 + \frac{\delta_{x\xi}\sigma_\xi^2}{Var\left(x_{jm}\right)} \neq \beta_0.
\end{aligned}$$

3.2.1 Verifying validity of instruments

We now consider the problem of estimating the parameters (α, β) consistently. First, note that:

$$Cov\left(w_{jm}, x_{jm}\right) = \delta_{xw}\sigma_{w1}^2 + \delta_{w\xi}\delta_{x\xi}\sigma_\xi^2,\ Cov\left(w_{jm}, \xi_{jm}\right) = \delta_{w\xi}\sigma_\xi^2 \tag{26}$$

$$Cov\left(r_{jm}, x_{jm}\right) = \delta_{xr}\sigma_{r1}^2 + \delta_{r\xi}\delta_{x\xi}\sigma_\xi^2,\ Cov\left(r_{jm}, \xi_{jm}\right) = \delta_{r\xi}\sigma_\xi^2 \tag{27}$$

Input prices as instruments for x_{jm} are valid if they satisfy the following two conditions: 1) they are correlated with x_{jm}, and 2) they are uncorrelated with ξ_{jm}. The first condition can be checked with the data. Here we assume it holds for both instruments, i.e. $Cov\left(w_{jm}, x_{jm}\right) \neq 0$ and $Cov\left(r_{jm}, x_{jm}\right) \neq 0$. The second condition corresponds to $\delta_{w\xi} = 0$ and $\delta_{r\xi} = 0$ in our specification. In applications where the 2nd condition is hard to verify, researchers assume it to be satisfied. In particular, we focus on the case where $\delta_{x\xi} \neq 0$, $\delta_{w\xi} \neq 0$ as well as $\delta_{r\xi} \neq 0$ so that the OLS estimate of β is biased and input prices are invalid instruments.

The first step in our method to estimate β consistently is to compare the instrument-free estimate of α with the IV based method. We construct the following variable as an instrument for price:

$$z_{pjm} = w_{jm} - \frac{Cov\left(w_{jm}, x_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)} r_{jm}. \tag{28}$$

By construction, $Cov\left(z_{pjm}, x_{jm}\right) = 0$, thus, we can remove the term with x_{jm} from the IV estimation equation. Thus, the instrument z_{pjm} identifies the true price coefficient α_0 if and only if

$$Cov\left(z_{pjm}, p_{jm}\right) \neq 0 \tag{29}$$

$$\begin{aligned}
\frac{Cov\left(z_{pjm}, ln\left(s_{jm}\right) - ln\left(s_{0m}\right)\right)}{Cov\left(z_{pjm}, p_{jm}\right)} &= \frac{Cov\left(z_{pjm}, -p_{jm}\alpha_0 + x_{jm}\beta_0 + \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} \\
&= -\alpha_0 + \frac{Cov\left(z_{pjm}, \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} = -\alpha_0,
\end{aligned} \tag{30}$$

Note Equation (29) corresponds to Condition 1) of the IV validity, and Equation (30), given Equation (29) is equivalent to $Cov\left(z_{pjm}, \xi_{jm}\right) = 0$, which corresponds to Condition 2) of the IV validity. Next we show that Equations (29) and (30) are equivalent to the following two equations:

$$\frac{Cov\left(w_{jm}, p_{jm}\right)}{Cov\left(w_{jm}, x_{jm}\right)} \neq \frac{Cov\left(r_{jm}, p_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)} \tag{31}$$

$$\frac{Cov\left(w_{jm}, \xi_{jm}\right)}{Cov\left(w_{jm}, x_{jm}\right)} = \frac{Cov\left(r_{jm}, \xi_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)}. \tag{32}$$

First, from Equation (28), we can see that Equation (29) holds if and only if Equation (31) is satisfied. Furthermore, given $Cov\left(z_{pjm}, p_{jm}\right) \neq 0$, Equation (30) holds if and only if $Cov\left(z_{pjm}, \xi_{jm}\right) = 0$, which results in Equation (32). Note that the two conditions, (29) and (30), and thus, Equations (31) and (32) can be verified, the first from the data and the second using our instrument-free identification of α_0 using cost data.

Similarly, we construct the following candidate as the instrument for x_{jm}:

$$z_{xjm} = w_{jm} - \frac{Cov\left(w_{jm}, p_{jm}\right)}{Cov\left(r_{jm}, p_{jm}\right)} r_{jm}.$$

Identification of β_0 requires:

$$\begin{aligned} \frac{Cov\left(z_{xjm}, ln\left(s_{jm}\right) - ln\left(s_{0m}\right)\right)}{Cov\left(z_{xjm}, x_{jm}\right)} &= \frac{Cov\left(z_{xjm}, -p_{jm}\alpha_0 + x_{jm}\beta_0 + \xi_{jm}\right)}{Cov\left(z_{xjm}, x_{jm}\right)} \\ &= \beta_0 + \frac{Cov\left(z_{xjm}, \xi_{jm}\right)}{Cov\left(z_{xjm}, x_{jm}\right)} = \beta_0, \end{aligned}$$

where we have used the fact that $Cov\left(z_{xjm}, p_{jm}\right) = 0$. This implies that Equations (31) and

$$\frac{Cov\left(w_{jm}, \xi_{jm}\right)}{Cov\left(w_{jm}, p_{jm}\right)} = \frac{Cov\left(r_{jm}, \xi_{jm}\right)}{Cov\left(r_{jm}, p_{jm}\right)} \tag{33}$$

must hold. Then, Equations (31), (32) and (33) imply instrument validity, i.e.

$$Cov\left(w_{jm}, \xi_{jm}\right) = Cov\left(r_{jm}, \xi_{jm}\right) = 0.$$

While we can verify Equation (31) from the data and Equation (32) using our instrument-free consistent estimator $\widehat{\alpha}$ of α_0, we cannot verify (33).

We, therefore, consider $\left\{w^l_{jm}, r^l_{jm}, l = 2, \ldots\right\}$ as additional instruments for price, as is frequently done in the literature. Then, because of our assumption of mean zero independence of all the ϱ terms, we obtain

$$E\left[\varrho_{wjm}^{k_w}\varrho_{x1jm}^{k_x}\varrho_{\xi jm}^{k_\xi}\right]=0$$

if at least one of (k_w, k_x, k_ξ) is one. Then,

$$Cov\left(w_{jm}^l,\xi_{jm}\right)=E\left[\left(\varrho_{w1jm}+\varrho_{wojm}+\delta_{w\xi}\varrho_{\xi jm}\right)^l\varrho_{\xi jm}\right]$$

$$=\sum_{k_1+k_2+k_3=l}\left(\begin{array}{c}l\\k_1,k_2,k_3\end{array}\right)\delta_{w\xi}^{k_3}E\left[\varrho_{w1m}^{k_1}\right]E\left[\varrho_{wojm}^{k_2}\right]E\left[\varrho_{\xi jm}^{k_3+1}\right]$$

$$Cov\left(w_{jm}^l,x_{jm}\right)=E\left[\left(\varrho_{w1jm}+\varrho_{wojm}+\delta_{w\xi}\varrho_{\xi jm}\right)^l\right]\left(\delta_{xw}\varrho_{w1jm}+\delta_{xr}\varrho_{r1jm}+\delta_{x\xi}\varrho_{\xi jm}+\varrho_{xojm}\right)$$

$$=\sum_{k_1+k_2+k_3=l}\left(\begin{array}{c}l\\k_1,k_2,k_3\end{array}\right)\delta_{xw}\delta_{w\xi}^{k_3}\delta_{x\xi}E\left[\varrho_{w1jm}^{k_1+1}\right]E\left[\varrho_{wojm}^{k_2}\right]E\left[\varrho_{\xi jm}^{k_3+1}\right]$$

As before, we can construct another instrument for price as follows:

$$z_{pjm}^{(l)}\equiv w_{jm}^l-\frac{Cov\left(w_{jm}^l,x_{jm}\right)}{Cov\left(r_{jm},x_{jm}\right)}r_{jm}.$$

Then,

$$\frac{Cov\left(z_{pjm}^{(l)},ln\left(s_{jm}\right)-ln\left(s_{0m}\right)\right)}{Cov\left(z_{pjm}^{(l)},p_{jm}\right)}=\frac{Cov\left(z_{pjm}^{(l)},-p_{jm}\alpha_0+x_{jm}\beta_0+\xi_{jm}\right)}{Cov\left(z_{pjm}^{(l)},p_{jm}\right)}=-\alpha_0+\frac{Cov\left(z_{pjm}^{(l)},\xi_{jm}\right)}{Cov\left(z_{pjm}^{(l)},p_{jm}\right)}$$

Proposition 1 $\alpha_0=\alpha_{IV}\Rightarrow\delta_{w\xi}=\delta_{r\xi}=0.$

Proof. The IV $z_{pjm}^{(l)}$ is valid if and only if:

$$\frac{Cov\left(z_{pjm}^{(l)},\xi_{jm}\right)}{Cov\left(z_{pjm}^{(l)},p_{jm}\right)}=0,l=1,2..$$

Now, as before, $Cov\left(z_{pjm}^{(l)},\xi_{jm}\right)=0,l=1,\ldots$ if and only if

$$\frac{Cov\left(w_{jm}^l,\xi_{jm}\right)}{Cov\left(w_{jm}^l,x_{jm}\right)}=\frac{Cov\left(r_{jm},\xi_{jm}\right)}{Cov\left(r_{jm},x_{jm}\right)}\equiv B,\ l=1,\ldots \tag{34}$$

Thus, the following holds

$$Cov\left(w_{jm}^{l},\xi_{jm}\right)=BCov\left(w_{jm}^{l},x_{jm}\right),\ l=1,\ldots.$$

which implies,

$$Cov\left(w_{jm}^{l},\xi_{jm}-Bx_{jm}\right)=0,\ l=1,\ldots$$

Given our assumption in Equation (24), we obtain

$$Cov\left(w_{jm}^{l},\xi_{jm}-Bx_{jm}\right)=E\left[w_{jm}^{l}\left(\xi_{jm}-Bx_{jm}\right)\right]-E\left(w_{jm}^{l}\right)E\left(\xi_{jm}-Bx_{jm}\right)=0,\ l=1,\ldots$$

Hence,

$$E\left[w_{jm}^{l}\left(\xi_{jm}-Bx_{jm}\right)\right]=0\ l=1,\ldots \tag{35}$$

which implies, given our assumption that the conditional moment is a continuous function of w_{jm}, that the conditional moment condition below is satisfied:

$$E\left[\xi_{jm}-Bx_{jm}|w_{jm}\right]=0. \tag{36}$$

To show that Equation (35) implies Equation (36) more formally, we consider the following conditional moment condition:

$$E\left[\xi_{jm}-Bx_{jm}|w_{jm}\geq w\right]=\frac{E\left[\left(\xi_{jm}-Bx_{jm}\right)I\left(w_{jm}\geq w\right)\right]}{E\left[I\left(w_{jm}\geq w\right)\right]}=E\left[\left(\xi_{jm}-Bx_{jm}\right)\frac{I\left(w_{jm}\geq w\right)}{E\left[I\left(w_{jm}\geq w\right)\right]}\right].$$

Then, from the Dominated convergence theorem, suppose f_n is a uniformly bounded sequence of differentiable function such that

$$f_n\left(w\right)\rightarrow\frac{I\left(w_{jm}\geq w\right)}{E\left[I\left(w_{jm}\geq w\right)\right]}$$

almost everywhere. Then,

$$\lim_{n\rightarrow\infty}E\left[\left(\xi_{jm}-Bx_{jm}\right)f_n\left(w\right)\right]=E\left[\left(\xi_{jm}-Bx_{jm}\right)\frac{I\left(w_{jm}\geq w\right)}{E\left[I\left(w_{j'm'}\geq w\right)\right]}\right]$$

and, from the Weierstrass Theorem, each continuous function f_n can be expressed as the infinite sum of polynomials $\left\{w_{jm}^l\right\}_{l=1}^{\infty}$. Therefore,

$$\begin{aligned} E\left[\xi_{jm}-Bx_{jm}|w_{jm}\geq w\right] &= \lim_{n\to\infty} E\left[f_n\left(w_{jm}\right)\left(\xi_{jm}-Bx_{jm}\right)\right] \\ &= \lim_{n\to\infty}\sum_{l=1}^{\infty}\psi_{n,l}E\left[w_{jm}^l\left(\xi_{jm}-Bx_{jm}\right)\right]=0. \end{aligned}$$

By taking the derivative of the above equation with respect to w, we obtain Equation (36).

Next we show that Equation (36) implies that $B=0$. Suppose otherwise. First, consider the case of $\delta_{xw}>0$ and $B>0$. Note that both w_{jm} and x_{jm} contain the random term ϱ_{w1jm}, but ξ_{jm} does not. Then, a random incease in ϱ_{wjm}, increases w_{jm} and Bx_{jm} as well, but does not affect ξ_{jm}. Therefore, a random increase in ϱ_{w1jm} increases w_{jm} and decreases $\xi_{jm}-Bx_{jm}$, implying that w_{jm} and Bx_{jm} are stochastically related via ϱ_{w1jm}. The same logic applies when $B<0$. Therefore, Equation (36) implies that $B=0$, and thus, we have, from Equation (34),

$$\frac{Cov\left(r_{jm},\xi_{jm}\right)}{Cov\left(r_{jm},x_{jm}\right)}=\frac{Cov\left(w_{jm},\xi_{jm}\right)}{Cov\left(w_{jm},x_{jm}\right)}=0\Rightarrow\delta_{r\xi}=0=\delta_{w\xi}.$$

Thus, we can use w_{jm}, r_{jm} as valid instruments for identifying β_0 as well.

The analysis of $\delta_{xw}<0$ is similar. Finally, consider $\delta_{xw}=0$. Then, if w_{jm} satisfies Condition 2) of instrument validity, i.e. $\delta_{w\xi}=0$, we have

$$Cov\left(w_{jm},x_{jm}\right)=0$$

which violates Condition 1) of instrument validity. Thus, the Proposition is proved. ■

So far we have shown that our instrument-free estimate of the price coefficient can be used to verify validity of potential instruments for observed characteristics, for example, input prices. If these instruments satisfy the two conditions for validity, we are able to obtain consistent estimates of the parameter β in addition to the price coefficient. Next, we consider the possibility that these instruments turn out to be invalid. As demand models become increasingly complex, the problem of finding valid instruments becomes accordingly more severe. In the next subsection, we propose a new method to create valid instruments using the invalid ones. The logic used is quite similar to what we used for verification of the existing instruments.

3.2.2 Constructing valid instruments

First, we define the following two candidates for instruments, using input prices (w and r) and observed characteristic (x) and parameters D_w and D_r:

$$z_{wjm} = x_{jm} - D_w w_{jm},\ z_{rjm} = x_{jm} - D_r r_{jm}.$$

Next, let

$$D_{w0} \equiv \frac{\delta_{\xi x}}{\delta_{w\xi}},\ D_{r0} \equiv \frac{\delta_{\xi x}}{\delta_{r\xi}}$$

and let z_{w0jm}, z_{r0jm} be defined as follows:

$$z_{w0jm} \equiv x_{jm} - D_{w0} w_{jm},\ z_{r0jm} \equiv x_{jm} - D_{r0} r_{jm}.$$

Then,

$$\begin{aligned}
z_{w0jm} &= \mu_x - D_{w0}\mu_w + \left(\delta_{xw} - D_{w0}\right)\varrho_{w1jm} - D_{w0}\varrho_{wojm} + \delta_{xr}\varrho_{rjm} + \varrho_{xojm} \\
z_{r0jm} &= \mu_x - D_{r0}\mu_r + \left(\delta_{xr} - D_{r0}\right)\varrho_{r1jm} - D_{r0}\varrho_{rojm} + \delta_{xw}\varrho_{wjm} + \varrho_{xojm},
\end{aligned}$$

Note that z_{w0jm} and z_{r0jm} do not contain unobserved product characteristics ξ_{jm}, which is the source of endogeneity. Hence,

$$Cov\left(z_{w0jm}, \xi_{jm}\right) = Cov\left(z_{r0jm}, \xi_{jm}\right) = 0.$$

However, we cannot use these variables as instruments because D_{w0} and D_{r0} are unknown.

Thus, we start with:

$$\begin{aligned}
z_{wjm} &= \mu_x - D_w\mu_w + \left(\delta_{xw} - D_w\right)\varrho_{w1jm} - D_w\varrho_{wojm} + \delta_{xr}\varrho_{rjm} + \varrho_{xojm} + \left(D_{w0} - D_w\right)\delta_{w\xi}\varrho_{\xi jm} \\
z_{rjm} &= \mu_x - D_r\mu_r + \left(\delta_{xr} - D_r\right)\varrho_{r1jm} - D_r\varrho_{rojm} + \delta_{xw}\varrho_{wjm} + \varrho_{xojm} + \left(D_{r0} - D_r\right)\delta_{r\xi}\varrho_{\xi jm}.
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
Cov\left(z_{wjm}, x_{jm}\beta_0 + \xi_{jm}\right) &= Cov\left(z_{wjm}, x_{jm}\right)\beta_0 + \left(D_{w0} - D_w\right)\delta_{w\xi}\sigma_\xi^2, \\
Cov\left(z_{wjm}, x_{jm}\right) &= Var\left(x_{jm}\right) - D_w Cov\left(w_{jm}, x_{jm}\right).
\end{aligned}$$

Hence,

$$\beta_w \equiv \frac{Cov\left(z_{wjm}, x_{jm}\beta_0 + \xi_{jm}\right)}{Cov\left(z_{wjm}, x_{jm}\right)} = \beta_0 + \frac{\left(D_{w0} - D_w\right)\delta_{w\xi}\sigma_\xi^2}{Var\left(x_{jm}\right) - D_w Cov\left(w_{jm}, x_{jm}\right)}.$$

Similarly,

$$\beta_r \equiv \frac{Cov\left(z_{rjm}, x_{jm}\beta_0 + \xi_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)} = \beta_0 + \frac{\left(D_{r0} - D_r\right)\delta_{r\xi}\sigma_\xi^2}{Var\left(x_{jm}\right) - D_r Cov\left(r_{jm}, x_{jm}\right)}.$$

As in the previous subsection, the instruments that identifies α_0 are the ones with D_w and D_r that satify

$$\frac{Cov\left(z_{wjm}, p_{jm}\right)}{Cov\left(z_{wjm}, x_{jm}\right)} \neq \frac{Cov\left(z_{rjm}, p_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)} \tag{37}$$

$$\frac{Cov\left(z_{wjm}, \xi_{jm}\right)}{Cov\left(z_{wjm}, x_{jm}\right)} = \frac{Cov\left(z_{rjm}, \xi_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)}. \tag{38}$$

These two equations imply:

$$\begin{aligned} \frac{Cov\left(z_{wjm}, \xi_{jm}\right)}{Cov\left(z_{wjm}, x_{jm}\right)} &= \frac{\left(D_{w0} - D_w\right)\delta_{w\xi}\sigma_\xi^2}{Var\left(x_{jm}\right) - D_w Cov\left(w_{jm}, x_{jm}\right)} \\ &= \frac{\left(D_{r0} - D_r\right)\delta_{r\xi}\sigma_\xi^2}{Var\left(x_{jm}\right) - D_r Cov\left(r_{jm}, x_{jm}\right)} = \frac{Cov\left(z_{rjm}, \xi_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)}. \end{aligned}$$

Using the same logic as in the previous subsection, we can see that if

$$\frac{Cov\left(z_{\mathrm{w}jm}, \xi_{jm}\right)}{Cov\left(z_{wjm}, x_{jm}\right)} = \frac{Cov\left(z_{rjm}, \xi_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)} \neq 0,$$

then Equations (37) and (38) imply

$$\frac{Cov\left(z_{wjm}, \xi_{jm}\right)}{Cov\left(z_{wjm}, p_{jm}\right)} \neq \frac{Cov\left(z_{rjm}, \xi_{jm}\right)}{Cov\left(z_{rjm}, p_{jm}\right)}, \tag{39}$$

violating the condition that is equivalent to Equation (33), which is the condition for instrument validity for x_{jm}. That is, identification of the true parameter estimate α_0 does not imply identification of β_0 when instruments are invalid.

As before, we next use z_{wjm}^l as instrument as well. To do so, we define

$$\begin{aligned}
& Cov\left(z_{wjm}^l, \xi_{jm}\right) \\
= \; & E\left[\left(\left(\delta_{xw} - D_w\right)\varrho_{w1jm} - D_w\varrho_{wojm} + \delta_{xr}\varrho_{rjm} + \varrho_{xojm} + \left(D_{w0} - D_w\right)\delta_{w\xi}\varrho_{\xi}\right)^l \varrho_{\xi jm}\right] \\
= \; & \sum_{k_1+k_2+k_3+k_4+k_5=l} \begin{pmatrix} l \\ k_1, k_2, k_3, k_4, k_5 \end{pmatrix} \left(\delta_{wx} - D_w\right)^{k_1} \left(-D_w\right)^{k_2} \delta_{xr}^{k_3} \left[\left(D_{w0} - D_w\right)\delta_{w\xi}\right]^{k_5+1} \\
& \times E\left[\varrho_{w1jm}^{k_1}\right] E\left[\varrho_{wojm}^{k_2}\right] E\left[\varrho_{xrjm}^{k_3}\right] E\left[\varrho_{xojm}^{k_4}\right] E\left[\varrho_{\xi jm}^{k_5+1}\right]. \qquad (40)
\end{aligned}$$

Then, for identifying the true price coefficient α_0, we obtain equations that are equivalent to the conditions (32), if, we use the newly defined z_{wjm}^l and z_{rjm} instead of z_{wjm} and z_{rjm} as instruments. Thus, the following holds.

$$\frac{Cov\left(z_{wjm}^l, \xi_{jm}\right)}{Cov\left(z_{wjm}^l, x_{jm}\right)} = \frac{Cov\left(z_{rjm}, \xi_{jm}\right)}{Cov\left(z_{rjm}, x_{jm}\right)} \equiv B, \; l = 1, \ldots$$

In order for the above equality to hold for any $l = 1, \ldots,$

$$Cov\left(z_{wjm}^l, \xi_{jm}\right) = BCov\left(z_{wjm}^l, x_{jm}\right).$$

Therefore,

$$Cov\left(z_{wjm}^l, \xi_{jm} - Bx_{jm}\right) = 0, \; l = 1, \ldots$$

which, given appropriate assumptions, implies

$$E\left[\xi_{jm} - Bx_{jm} | z_{wjm}\right] = 0. \qquad (41)$$

Next, we prove that $B = 0$. Recall $z_{wjm} = x_{jm} - D_w w_{jm}$. Then, given all other random components in z_{wjm}, higher ϱ_{xojm} implies higher z_{wjm}, which changes Bx_{jm}, but not ξ_{jm}. Therefore, if $B \neq 0$, Equation (41) does not hold. It then follows that

$$Cov\left(z_{wjm}^l, \xi_{jm}\right) = Cov\left(z_{rjm}, \xi_{jm}\right) = 0.$$

From Equation (40), we can see that the above equality holds if

$$D_w = D_{w0}.$$

Therefore, identifying the true price coefficient α_0 allows researchers to construct the valid instrument z_{rjm} and using it to identify the true coefficient β_0.

3.2.3 IV Estimation of observed characteristic coefficient without orthogonality conditions

Next, we present the estimation algorithm that is based on the identification results in this subsection. Let $\widehat{\alpha}$ be the price coefficient estimated by using the cost data, using Equation (16). Then, the estimation algorithm is based on the 2SLS estimator $(\alpha_{IV}(D_w, D_r), \beta_{IV}(D_w, D_r))$ where $z_{wjm}(D_w)$, $z_{rjm}(D_r)$ are the instruments for price and observed characteristics, given the candidate parameters (D_w, D_r). That is,

$$\begin{bmatrix} -\alpha_{IV}(D_w, D_r) \\ \beta_{IV}(D_w, D_r) \end{bmatrix} = \left[\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}\right]^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$$

where

$$\begin{aligned}
\mathbf{x}_{jm} &= [p_{jm}, x_{jm}] \\
\mathbf{X} &= \left[\mathbf{x}'_{11}, \ldots \mathbf{x}'_{J_1 1}, \ldots \mathbf{x}'_{1M}, \ldots, \mathbf{x}'_{J_M M}\right]' \\
\mathbf{z}_{jm}(D_w, D_r) &= \Big[z_{wjm}(D_w), z_{rjm}(D_r), z_{wjm}(D_w)^2 \\
&\quad , z_{wjm}(D_w) \times z_{rjm}(D_r), z_{rjm}(D_r)^2, \ldots, z_{rjm}(D_r)^3\Big] \\
\mathbf{Z} &= \left[\mathbf{z}'_{11}, \ldots \mathbf{z}'_{J_1 1}, \ldots \mathbf{z}'_{1M}, \ldots, \mathbf{z}'_{J_M M}\right]' \\
\mathbf{y} &= (y_{11}, \ldots, y_{J_1 1}, \ldots, y_{1M}, \ldots, y_{J_M M}), \; y_{jm} = ln(s_{jm}) - ln(s_{0m}).
\end{aligned}$$

In the estimation algorithm, we choose (D_w, D_r) so that the difference between the 2SLS estimator of the price coefficient using the constructed instruments $(z_{wjm}(D_w), z_{rjm}(D_r))$, $\alpha_{IV}(D_w, D_r)$, and the price coefficient $\widehat{\alpha}$ estimated by using the cost data is minimized. That is,

$$(D_w^*, D_r^*) = argmin_{\{D_w, D_r\}}\left[\alpha_{IV}(D_w, D_r) - \widehat{\alpha}\right]^2.$$

Then, the estimator of the coefficient on the observed product characteristics is:

$$\widehat{\beta} = \beta_{IV}(D_w^*, D_r^*)$$

Next, we provide a more general discussion of our approach focusing on the price coefficient.

In Subsection 3.3, we present the assumptions.

3.3 Main Assumptions of the General Model

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, total costs and individual input costs of firms.*

Note that market size can be derived from data on outputs and market shares because output of a firm in any market equals its market share times the market size. 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 as well as individual input costs. 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 or BLP demand structure: Market share s_{jm} is specified either as in Equation (1) with $\alpha > 0$ or 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 $\mathbf{p}_{-j,m}$.*

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 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.

The next assumption is regarding the cost function. Let C^*_{jm} denote true cost. Then,

Assumption 6

$$C^*_{jm} \equiv C\left(q_{jm}, \mathbf{w}_m, \mathbf{x}_{jm}, \upsilon_{jm}\right), \tag{42}$$

which is a continuous function of q*,* $\mathbf{w}$*,* $\mathbf{x}$ *and* υ*, strictly increasing, and continuously differentiable in* q *and* υ*, and marginal cost is strictly increasing in* υ*; Further, for any* $q > 0$*,* $w_l > 0$*,* $l = 1, \ldots, L$ *and* $\mathbf{x} \in \mathcal{X}$*, where* $\mathcal{X}$ *is the support of* $\mathbf{x}$*,*

$$lim_{\upsilon \searrow 0} \frac{\partial C^{\upsilon}\left(q, \mathbf{w}, \mathbf{x}, \upsilon\right)}{\partial q} = 0, \quad lim_{\upsilon \nearrow \infty} \frac{\partial C^{\upsilon}\left(q, \mathbf{w}, \mathbf{x}, \upsilon\right)}{\partial q} = \infty.$$

We also consider a special case, as in 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. That is,

$$C^*\left(q, \mathbf{w}, \mathbf{x}, \upsilon; \boldsymbol{\theta}_{c0}\right) = \widetilde{C}\left(q, \mathbf{w}, \mathbf{x}; \boldsymbol{\theta}_{c0}\right) exp\left(\varphi\left(\mathbf{x}, \upsilon\right)\right). \tag{43}$$

where $\widetilde{C}\left(\right)$ *is the deterministic component of cost and* $\varphi\left(\mathbf{x}, \upsilon\right)$ *is an unspecified smooth function of observed characteristics* $\mathbf{x}$ *and unobserved characteristics* υ*. Furthermore, the observed cost* C_{jm} *is given by the sum of the true cost* C^*_{jm} *and the measurement error* u_{cjm} *as follows:*

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

where we assume u_{cjm} 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 l whose price is w_{ljm}, is measured with error, i.e.

$$C_{ljm} = C^*_{ljm} + u_{kjm} = w_{ljm} l_{jm} + u_{kjm},\ l = 1, \ldots, L,$$

where we assume u_{ljm} to be i.i.d. distributed and independent to other variables in the demand and cost functions.[6]

3.4 General identification result

We first explain the simpler case of the multiplicative cost shock (see Assumption 6). Then, using Equation (43), the marginal cost function can be expressed as follows,

$$\begin{aligned} MC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right) &= \frac{\partial}{\partial q} \widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right) exp\left(\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right) \\ &= \widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right) exp\left(\varphi\left(\mathbf{x}_{jm}, \upsilon_{jm}\right)\right). \end{aligned} \tag{44}$$

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

$$\frac{C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{MC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)} = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}. \tag{45}$$

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

$$MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right) = MC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right). \tag{46}$$

Using Equation (46) to substitute $MR\left(\right)$ for $MC^*\left(\right)$ into Equation (45) we derive:

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

[6] Total cost equals the sum of input costs and thus, the measurement error in total cost equals the sum of measurement errors in individual input costs.

and by multiplying MR_j on both sides, we get

$$C_{jm}^* = C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right) \quad = \quad \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_c\right)} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d\right). \quad (47)$$

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

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

$$C_{jm} = \frac{\widetilde{C}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)}{\widetilde{MC}\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0}\right)} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_{d0}\right) + u_{cjm}. \quad (48)$$

Note that since u_{cjm} is the measurement error, and is assumed to be independent to all other observable variables in the demand and cost functions, the variables on the RHS and the error term u_{cjm} are independent, and thus, we don't face the endogeneity issue. Thus, this equation can be used to estimate both the demand parameters $\boldsymbol{\theta}_{d0}$ and some or all of the cost parameters $\boldsymbol{\theta}_{c0}$ without any endogeneity issues once we have the functional forms for the cost function and the demand function.

To see which of the cost parameters can be estimated using Equation (48), note that

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

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

Since we have data on the cost of each input, we can identify some of the remaining parameters from Shephard's Lemma, which states that

$$\frac{\partial C^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_{c0}\right)}{\partial w_{kjm}} = L_{kjm}. \quad (49)$$

Or equivalently,

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

and thus, we can identify those parameters that determine the input price elasticity of cost. Then, using Equation (43), the above equation can be modified as follows:

$$\frac{\partial ln\widetilde{C}\left(q_{jm},\mathbf{w}_{jm},\mathbf{x}_{jm};\boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}}=\frac{\partial lnC^{*}\left(q_{jm},\mathbf{w}_{jm},\mathbf{x}_{jm},\upsilon_{jm};\boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}}=\frac{w_{kjm}L_{kjm}}{C_{jm}^{*}},k=1,\ldots,K \quad (50)$$

Then, Shephard's Lemma and Assumption 6 together imply

$$C_{kjm}=w_{kjm}L_{kjm}+u_{kjm}=C_{jm}^{*}\frac{\partial ln\widetilde{C}\left(q_{jm},\mathbf{w}_{jm},\mathbf{x}_{jm};\boldsymbol{\theta}_{c0}\right)}{\partial lnw_{kjm}}+u_{kjm},k=1,\ldots,K-1, \quad (51)$$

which we use for estimation together with the true cost function $C^{*}()$, which we recover from Equation (48). As before, since the measurement error in input cost u_{kjm} is assumed to be independent to the variables in the marginal revenue and cost functions, the above equation is not subject to any endogeneity issue. Note that in estimation, as we can see in Equation (51), only $K-1$ input share equations are used. This is because total cost C_{jm} equals the sum of the K input costs. Similarly as before, let $\boldsymbol{\theta}_{cw}$ be the vector of parameters that are not in $\boldsymbol{\theta}_{cq}$ but can be estimated by applying Shephard's Lemma. In the Cobb-Douglas production function example, $\boldsymbol{\theta}_{cw}=\alpha_{c}$. The parameters of the cost function that remain unidentified, denoted by $\boldsymbol{\theta}_{c,(-q,-w)}$, can be identified from the remaining cost component because it can be expressed as a function only of $\mathbf{x}_{jm}$ and υ_{jm}, i.e., $ln\varphi\left(\mathbf{x}_{jm},\upsilon_{jm};\boldsymbol{\theta}_{c,(-q,-w)}\right)$. Since there are no terms involving output, the presence of the cost shock does not lead to any endogeneity issue.

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

$$MR_{j}\left(\mathbf{p}_{m},\mathbf{s}_{m},\mathbf{X}_{m};\boldsymbol{\theta}_{d}\right)=p_{jm}-\frac{1}{\left(1-s_{jm}\right)\alpha}=p_{jm}-\frac{1}{\left(1-q_{jm}/Q_{m}\right)\alpha}. \quad (52)$$

Therefore, the exclusion restriction is that price p_{jm} and market size Q_{m} enter in the marginal revenue function, but not in the cost function. However, note that in contrast to the conventional identification arguments, such exclusion restrictions do not lead to any orthogonality conditions for instruments.

Note that in the case of the multiplicative cost shock, we can recover the cost shock component from the first order condition as follows:

$$MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d) = MC^*(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c) = \widetilde{MC}(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0})\, exp(\varphi(\mathbf{x}_{jm}, \upsilon_{jm}))$$

and thus,

$$exp(\varphi(\mathbf{x}_{jm}, \upsilon_{jm})) = \frac{MR_j(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_d)}{\widetilde{MC}(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}; \boldsymbol{\theta}_{c0})} \tag{53}$$

or in log specification,

$$\varphi(\mathbf{x}_{jm}, \upsilon_{jm}) = ln(MR_{jm}) - ln\left(\widetilde{MC}_{jm}\right)$$

That is, the cost shock component can be recovered from the difference between the log of marginal revenue and the deterministic component of the marginal cost function, both of which do not include any unobservable variables.[7] This is in contrast to the conventional literature that identifies the cost shock as the difference between the observed cost and the deterministic component of the cost.

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 total cost as the sum of various cost components. 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.

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.

3.5 Estimation issues

We estimate Equations (48) and (51) jointly. The conventional methods are feasible generalized least squares (FGLS) and maximum likelihood (ML) methods. We use FGLS in our Monte-Carlo experiments. FGLS estimates the parameters by minimizing the following objective function

[7] This is similar to Byrne et al (2022) in that they used the MR=MC condition to identify the cost shock. However, unlike here, they assume a nonparametric cost function and thus, cannot explicitly solve for the cost shock. Further, they use the FOC to define their pseudo-cost function, where marginal revenue can be used in place of the cost shock.

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

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 $\mathbf{W}=\mathbf{I}$.

Note that since the RHS of Equation (48) 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 Equation (44) and the F.O.C. in Equation (46),

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

which results in

$$\widehat{\varphi}\left(\mathbf{x}_{jm},\upsilon_{jm}\right)=lnMR_j\left(\mathbf{p}_m,\mathbf{s}_m,\mathbf{X}_m;\widehat{\boldsymbol{\theta}}_d\right)-ln\widetilde{MC}\left(q_{jm},\mathbf{w}_{jm},\mathbf{x}_{jm};\widehat{\boldsymbol{\theta}}_c\right). \tag{54}$$

In the conventional approach, the cost shock component 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 component 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 component.

Next, we discuss how to separate the observed and unobserved cost components, , i.e., $\mathbf{x}_{jm}$ and υ_{jm} in the estimated function $\widehat{\varphi}\left(\mathbf{x}_{jm},\upsilon_{jm}\right)$. We can either assume that potential instruments for them, i.e. input prices $\mathbf{w}_{jm}$, and observed characteristics of rival firms, are valid, or adopt the procedure of instrument construction discussed in Subection 3.2.

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

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

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.

3.6 Estimation using translog cost function.

We next apply our methodology to the translog cost function while still assuming logit demand. The translog cost function is specified as follows:

$$\begin{aligned} lnC^*_{jm} &= \gamma_0 + \gamma_q lnq_{jm} + \frac{1}{2}\gamma_{qq}\left(lnq_{jm}\right)^2 + \sum_{k=1}^{K}\gamma_k lnw_{kjm} \\ &+ \frac{1}{2}\sum_{k=1}^{K}\sum_{k'=1}^{K}\gamma_{kk'} lnw_{kjm} lnw_{k'jm} + \sum_{k=1}^{K}\gamma_{kq} lnw_q lnq_{jm} + \mathbf{x}_{jm}\boldsymbol{\gamma}_x + \upsilon_{jm}. \end{aligned} \quad (55)$$

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:

$$\begin{aligned} &\frac{\partial lnC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial lnq_{jm}} \\ = &\frac{q_{jm}MC^*\left(q_{jm}, \mathbf{w}_{jm}, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{C^*_{jm}} = \gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k=1}^{K}\gamma_{kq} lnw_k. \end{aligned} \quad (56)$$

Substituting Equation (56) into Equation (48), we obtain

$$C_{jm} = C^*_{jm} + u_{cjm} = \frac{q_{jm}}{\gamma_q + \gamma_{qq} lnq_{jm} + \sum_{k=1}^{K}\gamma_{kq} lnw_k} MR_j\left(\mathbf{p}_m, \mathbf{s}_m, \mathbf{X}_m; \boldsymbol{\theta}_{d0}\right) + u_{cjm}. \quad (57)$$

Thus, in the first step, we can estimate parameters $\gamma_q, \gamma_{qq}, \gamma_{kq}$ without using any instruments, and $\boldsymbol{\theta}_{d0}$ as explained above. Identification of those parameters on the RHS allows us to identify the true cost C^*_{jm} as well. The remaining parameters γ_k and $\gamma_{kk'}$, $k = 1, \ldots, K$, $k' = 1, \ldots, K'$ can be identified from Shephard's Lemma as follows:

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

Finally, the intercept term γ_0 can be identified from Equation (55), and

$$\begin{aligned} \gamma_0 \quad = \quad & lnC^*_{jm} - \gamma_q lnq_{jm} - \frac{1}{2}\gamma_{qq}\left(lnq_{jm}\right)^2 - \sum_{k=1}^{K} \gamma_k lnw_{kjm} \\ & -\frac{1}{2}\sum_{k=1}^{K}\sum_{k'=1}^{K} \gamma_{kk'} lnw_{kjm} lnw_{k'jm} - \sum_{k=1}^{K} \gamma_{kq} lnw_q lnq_{jm} - E\left[\mathbf{x}_{jm}\boldsymbol{\gamma}_x + \upsilon_{jm}\right]. \end{aligned} \tag{58}$$

where C^*_{jm} and the parameters on the RHS are identified and we set $E\left[\upsilon_{jm}\right]$ to be zero.

This approach is related to Kumbhakar et al. (2012), who estimate markups using the output elasticity of translog cost function. They start with the assumption that the markup is strictly positive, that is:

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

which implies

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

and thus,

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

In our study, we additionally focus on the endogeneity of output with respect to the shock $u_{jm} + \upsilon_{jm}$. That is, if firms tend to reduce output to increase markup, then the shock and the output are negatively correlated, resulting in a downward bias of the estimate of γ_q and thus, the residual $u_{jm} + \upsilon_{jm}$ could misrepresent the true markup. In our approach, we deal with it by using marginal revenue 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. These functions can be fairly flexible, except that there are variables in the market share function but not in the cost function.

4 Monte-Carlo experiments

This section presents results from a series of Monte-Carlo experiments that highlight the finite sample performance of our estimator.

4.1 BLP demand and Cobb-Douglas Technology

We first design the Monte-Carlo setup to focus on the issue of endogenous price in the demand function and endogenous output in the cost function. To generate samples, we use the following random coefficients logit demand model:

$$s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right) = \int_\alpha \int_{\boldsymbol{\beta}} \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)} dF_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}; \boldsymbol{\theta}_{\boldsymbol{\beta}}\right) dF_\alpha\left(\alpha; \boldsymbol{\theta}_\alpha\right), \qquad (59)$$

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 in Equations (11) and (12), respectively. Notice that the cost function is homogeneous of degree one in input prices.

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

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

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

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

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

$TN\left(0, 1\right)$ is the truncated standard normal distribution, where we truncate both upper and lower 0.82 percentiles. $U\left(Q_L, Q_H\right)$ 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:

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

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

where $\varrho_{\xi jm}$ and $\varrho_{\upsilon jm}$ are the idiosyncratic components of the demand and supply shocks. We assume that $\varrho_\xi \sim i.i.d.TN\,(0,1)$ and $\varrho_\upsilon \sim i.i.d.TN\,(0,1)$.

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, input prices and the observed product characteristics of own product or products of other firms cannot be used as instruments since they are designed to be correlated with the cost shock. Furthermore, the 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. 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 possibility of any conventional instruments in either the demand or the supply equation.

To solve for the equilibrium price, quantity, and market share for each oligopoly firm, we use the golden section search on price.[8]

We estimate the parameters using GLS, where, in this case, we first set $\mathbf{W} = \mathbf{I}$, and then, given the residual estimated from the first step, derive $\mathbf{W} = \widehat{\mathbf{\Sigma}}^{-1}$, where $\widehat{\mathbf{\Sigma}} = Varcov\,(\widehat{u}_{jm})$, $m = 1, \ldots, M$, $j = 1, \ldots, J_m$.

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

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

cost function parameters under the assumption that the own observed characteristics are uncorrelated with the own unobserved product characteristics and the cost shock. 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 particular, we set $corr\left(x_{jm}, \xi_{jm}\right) = 0.0833 > 0$ and $corr\left(x_{jm}, \upsilon_{jm}\right) = -0.0833 < 0$. 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.

4.2 Logit Demand and Cobb-Douglas Technology

We then present results from a series of Monte-Carlo experiments that highlight the finite sample performance of our estimator that consistently estimates the coefficient on the observed product characteristics in the demand equation even if the instruments are invalid. To generate samples, we use the following random coefficients logit demand model:

$$s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right) = \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)}, j = 1, \ldots, 4 \tag{60}$$

$$s_0\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right) = \frac{1}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)} \tag{61}$$

where $\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m$ are defined the same as before.

We create our Monte-Carlo samples, following Equations (20), (21), (22) and (23). We assume that

$$\varrho_{\xi 1m} \sim \sigma_{\xi 1} \times i.i.d.TN\left(0,1\right), \; \varrho_{\xi 2jm} \sim \sigma_{\xi 2} \times i.i.d.TN\left(0,1\right), \; \varrho_{\xi ojm} \sim \sigma_{\xi o} \times i.i.d.TN\left(0,1\right)$$

$$\varrho_{w1jm} \sim \sigma_{w1} \times i.i.d.TN\left(0,1\right), \ \varrho_{wojm} \sim \sigma_{wo} \times i.i.d.TN\left(0,1\right)$$

$$\varrho_{r1jm} \sim \sigma_{r1} \times i.i.d.TN\left(0,1\right), \ \varrho_{rojm} \sim \sigma_{ro} \times i.i.d.TN\left(0,1\right)$$

$$\varrho_{xojm} \sim \sigma_{xo} \times i.i.d.TN\left(0,1\right), \ \varrho_{\upsilon} \sim \sigma_{\upsilon} \times i.i.d.TN\left(0,1\right)$$

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

where $TN\left(0,1\right)$ is the truncated standard normal distribution, used as before. For the supply shock, we assume that $\varrho_{\upsilon} \sim i.i.d.TN\left(0,1\right)$.

Table 4 summarizes the parameter setup of the Monte-Carlo experiments. The setup assumes that wage, rental rate and observed product characteristics are correlated with the unobserved product characteristics. Furthermore, by including the market level fixed effects in the demand shock, the observed characteristics of rival firms are also correlated with the demand shock. Hence, on the demand side all conventional instruments are invalid. In order to focus on the endogeneity on the demand side, we assume that the wage and rental rate, and observed product characteristics are uncorrelated with the cost shock.

In Table 5, we present the Monte-Carlo results of the direct estimator that estimates the parameters of the demand and cost functions. We can see that on average, the parameter estimates are close to the true ones even when sample size is as small as 200. The results indicate that the IV procedure discussed in Subsection 3.2 successfully removes the bias in the β coefficients. With the exception of the coefficient estimate β on the observed product characteristics, the standard errors and the root mean square errors decrease with sample size. The standard error of β estimates are higher than of the other parameter estimates, and decrease slower, and increases in with sample size from 400 to 800. There are three sources of variation for the high standard errors of the β estimate. First component is the variation of the IV estimation. The second one is the variation from the estimation of the price coefficient. The third one is the variation coming from the construction of the instrument.

In Table 6, we present the simple OLS results for comparison. That is, we estimate the following equation:

$$log\left(s_{jm}\right) - log\left(s_{0m}\right) = -p_{jm}\alpha + x_{jm}\beta + \xi_{jm},$$

where ξ_{jm} is the error term. Then, we estimate the log cost function in Equation (17), where the cost shock υ_{jm} is the error term. That is, we follow the convention and assume away the measurement error. We can see that the OLS estimated price coefficient $\widehat{\alpha}$ is biased downwards, and the coefficient on the observed product characteristics, $\widehat{\beta}$, biased upwards. This is due to the positive correlation between the price and the demand shock on one hand, and between the observed product characteristics and the demand shock on the other in the Monte-Carlo setup. Furthermore, the Cobb-Douglas production function coefficients α_c and β_c are estimated with an upward bias. This is due to the downward bias of the output coefficient, which is the estimate of $1/(\alpha_c + \beta_c)$. The source of this bias is the negative correlation between the cost shock and the output, chosen to maximize profits. Since output is also a function of x, the bias in the output coefficient also leads to the bias of $\widehat{\eta}$, the coefficient on the observed product characteristics x_{jm}. Overall, the bias of the coefficients is due to the correlation between price, observed product characteristics and the demand shock, and the correlation between output and the cost shock. These are well known sources of bias that arise when we estimate the demand and cost functions using OLS.

Finally, in Table 7, we present the results where we use the conventional instruments, which are wage, rental rate and the average observed product characteristics of rival firms. Note that since all the instruments are specified to be correlated with the demand shock ξ_{jm}, they are invalid. We can confirm this by observing that all the demand parameters are estimated to be quite different from the true values. On the other hand, the Cobb-Douglas production function parameters α_c and β_c tend to move closer to the true values as sample size increases. This confirms the validity of the instruments for cost function estimation since they are not correlated with the cost shocks. On the other hand, we see a large bias in the η estimate.

After comparing the various Monte-Carlo results, we conclude that our estimation methodology based on the instrument-free identification approach consistently estimates the price and output coefficients, and the IV methods we propose tend to remove bias of the coefficient estimates of the observed product characteristics in the demand function, even though the instruments are invalid.

5 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 as well as possible correlation between the observed product characteristics and the demand shock (in the logit case), and the cost function in the presence of output endogeneity without any valid instruments. Moreover, our method can accommodate measurement error.

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, and V. Sarafides (2022): "Instrument-free Identification and Estimation of Differentiated Products Models without Instruments Using Cost Data," *Journal of Econometrics*, 228(2), 278–301.

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. (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.

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.

6 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.
(x_{jm} uncorrelated with ξ_{jm} and υ_{jm}.)

(a) Demand side parameters

		$\hat{\mu}_\alpha$			$\hat{\sigma}_\alpha$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.033	0.2075	0.2091	0.4972	0.1232	0.1226
100	400	2.000	0.1395	0.1388	0.5037	0.0947	0.0943
200	800	1.993	0.1105	0.1102	0.4865	0.0575	0.0588
400	1600	2.001	0.0711	0.0715	0.4995	0.0453	0.0451
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	1.013	0.1457	0.1455	0.4100	0.0845	0.0846
100	400	0.9944	0.0818	0.0816	0.4025	0.0514	0.0512
200	800	1.000	0.0732	0.728	0.4033	0.0384	0.0384
400	1600	1.004	0.0480	0.0479	0.4041	0.0283	0.0284
True Value		1.0			0.4		

(b) Production function parameters

		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.5051	0.0441	0.0442	0.3041	0.0280	0.0282
100	400	0.5058	0.0295	0.0299	0.3023	0.0178	0.0179
200	800	0.4992	0.0204	0.0203	0.2995	0.0145	0.0145
400	1600	0.4998	0.0141	0.0141	0.3001	0.0102	0.0101
True Value		0.5			0.3		

		$\hat{\eta}$			
Markets	Sample Size	Mean	Std. Dev.	RMSE	Obj. fct
50	200	0.1990	0.0094	0.0094	1.993
100	400	0.2012	0.0067	0.0068	1.997
200	800	0.1996	0.0042	0.0042	1.998
400	1600	0.1995	0.0048	0.0040	1.999
True Value		0.2			

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

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	2.038	0.1939	0.1967	0.5007	0.1121	0.1115
100	400	1.999	0.1391	0.1384	0.5005	0.0734	0.0730
200	800	2.000	0.1095	0.1090	0.4964	0.0559	0.0558
400	1600	2.006	0.0698	0.0700	0.4981	0.0367	0.0365
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	1.192	0.1121	0.2221	0.4185	0.0600	0.0625
100	400	1.173	0.0787	0.1900	0.4036	0.0437	0.0436
200	800	1.175	0.0651	0.1866	0.4013	0.0322	0.0320
400	1600	1.179	0.0421	0.1835	0.4052	0.0221	0.0226
True Value		1.0			0.4		
		(b) Production function parameters					
		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.5025	0.0347	0.0346	0.3023	0.0220	0.0220
100	400	0.5034	0.0230	0.0231	0.3007	0.0130	0.0129
200	800	0.5011	0.0189	0.0189	0.3006	0.0135	0.0134
400	1600	0.4992	0.0115	0.0115	0.2998	0.0081	0.0081
True Value		0.5			0.3		
		$\hat{\eta}$					
Markets	Sample Size	Mean	Std. Dev.	RMSE			
50	200	0.1613	0.0141	0.0412	1.994		
100	400	0.1642	0.0096	0.0370	1.997		
200	800	0.1628	0.0073	0.0379	1.999		
400	1600	0.1619	0.0045	0.0384	1.999		
True Value		0.2					

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

Table 4: Monte Carlo Parameter Values

Parameter	Description	Value
(a) Demand-side parameters		
α	Price coef. mean	2.0
β	Product characteristic coef. mean	1.0
μ_X	Product characteristic mean	0.4
σ_{Xo}	Product characteristic std. dev.	0.3
δ_0	Unobserved product quality mean	4.0
	Unobserved product quality std. dev.	
$\sigma_{\xi 1}$		0.2
$\sigma_{\xi 2}$		0.2
σ_{xio}		0.3
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	wage mean	1.0
	wage std. dev.	
σ_{w1}		0.2
σ_{w2}		0.2
μ_w	rental rate mean	1.0
	rental rate std. dev.	
σ_{r1}		0.2
σ_{r2}		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) Covariance parameters		
δ_{xw}	x_{jm} and ϱ_{w1jm}	0.4
δ_{xr}	x_{jm} and ϱ_{r1jm}	0.4
$\delta_{x\xi}$	x_{jm} and $\varrho_{\xi jm}$	0.4
$\delta_{w\xi}$	w_{jm} and $\varrho_{\xi jm}$	0.4
$\delta_{r\xi}$	r_{jm} and $\varrho_{\xi jm}$	0.4
δ_v	ξ_{jm} and v_{jm} correlation	0.0
δ_Q	ξ_{jm} and Q_m correlation	0.0
ζ_Q	v_{jm} and Q_m correlation	0.0
(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 5: Parameter estimates using cost data
(Logit demand. Product characteristic x_{jm} and unobserved product quality ξ_{jm} are correlated;
instruments (w_{jm}, r_{jm}, average of x_{-jm}) are correlated with ξ_{jm}
x_{jm}, instruments and cost shock υ_{jm} are uncorrelated)

(a) Demand side parameters

		$\hat{\alpha}$			$\hat{\beta}$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.0004	0.0447	0.0444	0.9413	0.1676	0.1768
100	400	1.9950	0.0364	0.0366	0.9735	0.1307	0.1328
200	800	1.9963	0.0194	0.0197	0.9632	0.1329	0.1372
400	1600	2.0023	0.0157	0.0158	0.9646	0.1044	0.1098
True Value		2.0			1.0		

(b) Production function parameters

		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.4991	0.0088	0.0088	0.3001	0.0089	0.0089
100	400	0.5010	0.0064	0.0064	0.3000	0.0061	0.0060
200	800	0.5008	0.0037	0.0038	0.3002	0.0043	0.0043
400	1600	0.4998	0.0029	0.0029	0.2996	0.0034	0.0034
True Value		0.5			0.3		

			$\hat{\eta}$	
Markets	Sample Size	Mean	Std. Dev.	RMSE
50	200	0.2059	0.1058	0.1055
100	400	0.1943	0.0713	0.0712
200	800	0.1969	0.0468	0.0467
400	1600	0.2032	0.0396	0.0395
True Value		0.2		

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

Table 6: OLS estimation under endogeneity
(Logit demand. Product characteristic x_{jm} and unobserved product quality ξ_{jm} are correlated;
instruments (w_{jm}, r_{jm}, average of x_{-jm}) are correlated with ξ_{jm}
x_{jm}, instruments and cost shock υ_{jm} are uncorrelated)

(a) Demand side parameters

		$\hat{\alpha}$			$\hat{\beta}$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	1.8198	0.0474	0.1862	1.1290	0.0848	0.1541
100	400	1.8197	0.0384	0.1843	1.1350	0.0615	0.1483
200	800	1.8225	0.0244	0.1792	1.1441	0.0431	0.1503
400	1600	1.8240	0.0160	0.1767	1.1465	0.0290	0.1493
True Value		2.0			1.0		

(b) Production function parameters

		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.6768	0.0764	0.1925	0.6275	0.0771	0.3365
100	400	0.6792	0.0549	0.1874	0.6246	0.0531	0.3289
200	800	0.6815	0.0366	0.1851	0.6212	0.0368	0.3233
400	1600	0.6820	0.0255	0.1838	0.6203	0.0248	0.3212
True Value		0.5			0.3		

			$\hat{\eta}$	
Markets	Sample Size	Mean	Std. Dev.	RMSE
50	200	0.1155	0.0396	0.0932
100	400	0.1240	0.0298	0.0815
200	800	0.1225	0.0220	0.0805
400	1600	0.1259	0.0157	0.0758
True Value		0.2		

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

Table 7: IV estimation

(Logit demand. Product characteristic x_{jm} and unobserved product quality ξ_{jm} are correlated; instruments (w_{jm}, r_{jm}, average of x_{-jm}) are correlated with ξ_{jm} x_{jm}, instruments and cost shock υ_{jm} are uncorrelated)

		(a) Demand side parameters					
		$\hat{\alpha}$			$\hat{\beta}$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	1.5505	0.9095	1.1047	1.1802	1.8331	1.8328
100	400	1.9407	0.9981	0.9948	1.9853	2.1384	2.3447
200	800	1.8045	1.0073	1.0211	1.7207	2.1221	2.2311
400	1600	2.2570	2.1861	2.1903	2.6725	4.2931	4.5873
True Value		2.0			1.0		
		(b) Production function parameters					
		$\hat{\alpha}_c$			$\hat{\beta}_c$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	0.6083	0.1926	0.2201	0.5253	0.2909	0.3668
100	400	0.6164	0.1590	0.1965	0.4804	0.1991	0.2679
200	800	0.5811	0.1509	0.1706	0.4261	0.1876	0.2253
400	1600	0.5355	0.0604	0.0698	0.3504	0.1140	0.1242
True Value		0.5			0.3		
		$\hat{\eta}$					
Markets	Sample Size	Mean	Std. Dev.	RMSE			
50	200	-0.2863	0.5322	0.7190			
100	400	-0.2269	0.4555	0.6307			
200	800	-0.1422	0.3854	0.5140			
400	1600	0.0598	0.3731	0.3968			
True Value		0.2					

Notes: Monte-carlo experiment results based on calibration described in Table 4. Instruments used are wage, rental rate and the mean of observed product characteristics of rival firms.