Estimation of Differentiated Product Models without (Valid) Instruments

Susumu Imai (Hitotsubashi University)
Neelam Jain (City, University of London)
Hiroto Suzuki (Josai International University)
Miyuki Taniguchi (Saga University)

Seminar at Queen's University, Department of Economics
March 26, 2024

Preliminary. Comments are very welcome.

Introduction

- We propose a new methodology for estimating cost functions when data on cost and demand variables is available.
- Our method does not use instruments to deal with the endogeneity issues arising both on the demand side and the cost side, unlike in the literature.
- Our method is direct rather than the pseudo-cost-based method of Byrne et al (2022).
- Industries under (partial) government oversight, such as banks, hospitals, nursing homes etc. report cost data.
- We conduct several Monte-Carlo experiments to show our method works well - assuming either Logit or BLP on demand side and Cobb-Douglas production function on the supply side.

- We also propose an approach to estimating consistently the coefficients of the observed characteristics in demand functions without valid instruments.
- We estimate the price coefficient twice, using cost data, and (constructed) instruments. If the two estimates are close, then we obtain information on the validity of the instruments and use it to estimate the coefficients of the observed product characteristics. Our idea is similar in spirit to the Hausman-Wu test (Wu (1973), Hausman (1978)).
- The above idea also allows us to construct valid instrument from the invalid ones.
- We conduct several Monte-Carlo experiments to show our method works well.

Literature

- There has been active research on the measurement of markup (price minus marginal cost in the monopolistic/oligopolistic industries. Examples are : Berry (1994), Berry Levinsohn and Pakes (BLP 1995), DeLoecker and Warzynski (2012) and the subsequent literature.
- Most of them use instruments (Berry (1994), BLP (1995) etc.) or similar orthogonality conditions (Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg, Caves and Frazer (2015)) for the identification of the key parameters (price and output coefficients, production function parameters etc.)
- The number of instruments required grows with the complexity of the model. However, in most cases, their validity is unknown.

Model: Demand Side

- In each market $m = 1, \ldots, M$, there are J_m products
- Consumer choice set in market m: $\{0, 1, \ldots, J_m\}$.
- Choice 0: no purchase option (outside option).
- Consumer utility from consuming product j:

$$u_{ijm} = \boldsymbol{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha_p + \xi_{jm} + \epsilon_{ijm}.$$

Demand side: Berry (1994) logit model

- If ϵ_{ijm} has a logit distribution then market share of product j in market m is given by:

$$\begin{aligned} s_{jm} &= \frac{exp\left(\boldsymbol{x}_{jm}\beta - p_{jm}\alpha_p + \xi_{jm}\right)}{\sum_{l=0}^{J_m} exp\left(\boldsymbol{x}_{lm}\beta - p_{lm}\alpha_p + \xi_{lm}\right)} \\ &= \frac{exp\left(\boldsymbol{x}_{jm}\beta - p_{jm}\alpha_p + \xi_{jm}\right)}{1 + \sum_{l=1}^{J_m} exp\left(\boldsymbol{x}_{lm}\beta - p_{lm}\alpha_p + \xi_{lm}\right)} \qquad (1) \\ s_{0m} &= \frac{1}{1 + \sum_{l=1}^{J_m} exp\left(\boldsymbol{x}_{lm}\beta - p_{lm}\alpha_p + \xi_{lm}\right)} \qquad (2) \end{aligned}$$

- ξ_{jm}: unobserved product characteristics (demand shock).
- s_{0m}: For outside option, we set $p_{0m} = 0$, $\boldsymbol{x}_{0m} = 0$, $\xi_{0m} = 0$.
- $\sum_{j=0}^{J_m} s_{jm} = 1$.

The cost side

- Output and market size: $q_{jm} = Q_m s_{jm}$.
- Total Cost is given by:

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

- $\boldsymbol{\theta}_c$ is the parameter vector; $\boldsymbol{w}_{jm}$: $L \times 1$ input price vector; υ_{jm} the unobserved cost shock.
- Log linear form:

$$logC_{jm} = c_0 + \left(logq_{jm}\right) c_q + \left(log\,\boldsymbol{w}_{jm}\right) \boldsymbol{c}_w + \boldsymbol{x}_{jm}\boldsymbol{c}_x + \upsilon_{jm}.$$

Variables in the data

Demand Side:

- p_{jm}: price, s_{jm}: market share, $\boldsymbol{x}_{jm}$: observed product characteristics.

Supply Side

- C_{jm}: total cost, q_{jm}: output, $\boldsymbol{w}_{jm}$: input price, $\boldsymbol{x}_{jm}$.
- C_{kjm}: cost of input $l, l = 1, ..., L$.

Market

- Q_m: market size, $q_{jm} = Q_m s_{jm}$, $j = 1, \ldots, J_m$.
- We only need to observe two out of these three variables in the data.

The IV approach

Demand Side

- Dividing Equation (1) by (2), we can eliminate the denominator.
- Then, taking logs, we derive:

$$\log(s_{jm}) - \log(s_{0m}) = \boldsymbol{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha_p + \xi_{jm}.$$

- The endogeneity problem: price likely correlated with the unobserved product characteristics, that is, the error term ξ_{jm}. OLS estimate of α_p is biased.
- The IV approach is to use an instrument for price and then the moment condition $E[\xi_{jm}\boldsymbol{z}_{djm}] = \mathbf{0}$, $\boldsymbol{z}_{djm}$ is an $L_d \times 1$ vector of instruments that is correlated with p_{jm} and uncorrelated with ξ_{jm}.

The Cost side

$$logC_{jm} = c_0 + (logq_{jm})\, c_q + (log\,\boldsymbol{w}_{jm})\,\boldsymbol{c}_w + \boldsymbol{x}_{jm}\boldsymbol{c}_x + \upsilon_{jm}.$$

- The endogeneity problem: log output likely correlated with the cost shock, υ_{jm}, leading to biased estimates of c_q.
- The IV approach is to use an instrument for output and then the moment condition $E[\upsilon_{jm}\boldsymbol{z}_{cjm}] = \mathbf{0}$, $\boldsymbol{z}_{cjm}$ is an $L_c \times 1$ vector of instruments that is correlated with $logq_{jm}$ and uncorrelated with υ_{jm}.

Profit maximization

- Given price and output (market share) of competitor firms, the oligopolistic firm maximizes profit given by:

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

by choosing price. The resulting first order condition equates marginal revenue with marginal cost:

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

Using the FOC

- Suppose we know the marginal cost.
- Then, assuming the logit demand for the sake of simplicity, we derive
$$MR_{jm} = p_{jm} - \frac{1}{(1 - s_{jm})\,\alpha_{p0}} = MC_{jm}$$
We thus identify the true price coefficient as
$$\alpha_{p0} = \frac{1}{(1 - s_{jm})\,(p_{jm} - MC_{jm})}$$
- We usually do not have data on marginal cost, thus, we first need to estimate the cost function to derive the marginal cost.
- Estimating the cost function runs into the endogeneity problem, which requires instruments.

- After estimating the price coefficient using instruments, BLP (1995) use the FOC to estimate cost parameters without cost data.
- They assume that the marginal cost function is a log-linear function of output and input prices. In case of logit demand and log-linear cost function, the FOC would be

$$
\begin{aligned}
& p_{jm} - \frac{1}{(1 - s_{jm})\,\alpha_p} \\
= \quad & mc_0 + (\log q_{jm})\, mc_q + (\log \mathbf{w}_{jm})\, \mathbf{mc}_w + \mathbf{x}_{jm}\mathbf{mc}_x + \upsilon_{jm}.
\end{aligned}
$$

- They need to use instruments for log output.
- We use the FOC and the cost data and estimate parameters without instruments.

Concerns about the IV approach

- Good instruments are difficult to find especially as the demand models have become increasingly complex requiring many instruments and their interactions.
- Commonly-used instruments for demand estimation are cost shifters such as input prices. Firm-level data on wage may reflect product quality or productivity (high product quality or productivity requires highly paid workers).
- Similarly, observed characteristics of the competitor firms as instruments for price may not be valid if they are correlated with unobserved product characteristics or the cost shock.

- Similar arguments apply on the cost side.
  - Commonly-used instruments for cost estimation are demand shifters such as income which may not satisfy the exclusion restrictions.
- Using invalid instruments leads to biased estimates which in turn misinforms policy analysis.
- Even if instruments are available, our methodology can be used as a tool to check the validity of the instruments by comparing estimates.

Instrument-free joint estimation of the price coefficient in demand and output coefficient in cost functions

The Cobb-Douglas example.

- The Cobb-Douglas production function

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

where L and K denote labor and capital respectively and υ is the unobserved cost shock (inverse of productivity shock).

The cost functions

- Then, the cost function is

$$
\begin{aligned}
& C^{*}(q, w, r, x, \upsilon) \\
= \quad & (\alpha_c + \beta_c) \left(\frac{w}{\alpha_c} \right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c} \right)^{\beta_c/(\alpha_c+\beta_c)} \\
& \times A exp\left(x\eta + \upsilon \right) q^{\frac{1}{\alpha_c+\beta_c}}.
\end{aligned}
$$

- And the marginal cost function is

$$
\begin{aligned}
& MC^{*}(q, w, r, x, \upsilon) \\
= \quad & \left(\frac{w}{\alpha_c} \right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c} \right)^{\beta_c/(\alpha_c+\beta_c)} A exp\left(x\eta + \upsilon \right) q^{\frac{1}{\alpha_c+\beta_c}-1}
\end{aligned}
$$

where w and r are the wage rate and rental rate respectively.

Estimating the cost function: key steps of our method

- First, divide true cost by marginal cost:

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

- Substitute MR for $MC^*(q, w, r, x, \upsilon)$ from the FOC:

$$C^*(q_{jm}, w_{jm}, r_{jm}, \mathbf{x}_{jm}, \upsilon) = (\alpha_c + \beta_c)\, q_{jm} MR_j(\boldsymbol{p}_m, \boldsymbol{s}_m, \boldsymbol{X}_m; \boldsymbol{\theta}_d)\,. \tag{3}$$

- Define observed cost to be true cost plus an i.i.d. measurement error:

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

- We use this equation to estimate some of the cost parameters and the demand parameters.

- In the logit case,

$$MR_j\left(\boldsymbol{p}_m, \boldsymbol{s}_m, \boldsymbol{X}_m; \boldsymbol{\theta}_d\right) = p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha_p}.$$

- Substituting in the cost function, we obtain:

$$C_{jm} = \left(\alpha_c + \beta_c\right) q_{jm}\left(p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha_p}\right) + u_{cjm}. \tag{4}$$

That is,

$$C_{jm} = q_{jm} p_{jm}\left(\alpha_c + \beta_c\right) - \frac{q_{jm}}{\left(1 - s_{jm}\right)} \frac{\alpha_c + \beta_c}{\alpha_p} + u_{cjm}. \tag{5}$$

Because the residual u_{cjm} is a measurement error, assumed to be i.i.d., it is uncorrelated with the RHS variables $q_{jm}p_{jm}$ and $q_{jm}/\left(1 - s_{jm}\right)$, thus, there is no endogeneity issue and thus, parameters $\alpha_c + \beta_c$ and $\left(\alpha_c + \beta_c\right)/\alpha_p$ are estimated without any bias via simple OLS. Hence, α_p is estimated consistently.

Use of Shephard's Lemma

- To estimate α_c and β_c, we can use the Shephard's lemma if input cost data is available:

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

where $C^*_{jm} = \left(\alpha_c + \beta_c\right) q_{jm} MR_{jm}$ as in Equation (3).

- Then denoting the measurement error in the labor cost data by u_{Ljm}, we obtain the estimate of α_c from the following:

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

$$C_{Ljm} = \alpha_c q_{jm}\left(p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha_p}\right) + u_{cjm}.$$

Once again, this is a simple OLS estimation exercise with no endogeneity issue and thus the estimate of α_c is unbiased.

Monte Carlo Experiments for BLP Demand

- We conducted Monte Carlo experiments for the random coefficient logit model or BLP where market share function is as follows:
- $$\begin{aligned} & s_j\left(\boldsymbol{p}_m, \boldsymbol{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}\right) \\ = & \int_{\alpha_p}\int_{\boldsymbol{\beta}} \frac{\exp\left(\boldsymbol{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha_p + \xi_{jm}\right)}{\sum_{k=0}^{J_m}\exp\left(\boldsymbol{x}_{km}\boldsymbol{\beta} - p_{km}\alpha_p + \xi_{km}\right)} dF_{\boldsymbol{\beta}}\left(\boldsymbol{\beta}; \boldsymbol{\theta}_{\boldsymbol{\beta}}\right) dF_{\alpha_p}\left(\alpha_p; \boldsymbol{\theta}_{\alpha_p}\right), \end{aligned}$$

  where $F_{\alpha_p}\left(.; \boldsymbol{\theta}_{\alpha_p}\right)$ denotes the distribution function of the parameter vector $\boldsymbol{\theta}_{\alpha_p}$ and similarly for β.

Monte Carlo experiments for BLP demand

- Letting μ_{α_p} to be the mean of α_p and $\boldsymbol{\mu}_\beta$ the mean of $\boldsymbol{\beta}$, the mean utility is given by:

$$\delta_{jm} \equiv \boldsymbol{x}_{jm}\boldsymbol{\mu}_\beta - p_{jm}\mu_{\alpha_p} + \xi_{jm}.$$

- The parameter set up is the same as for logit demand except μ_{α_p} replaces α_p and $\boldsymbol{\mu}_\beta$ replaces $\boldsymbol{\beta}$ and in addition, we need to estimate the standard deviations of these parameters, denoted by σ_{α_p} and σ_β respectively.
- We assume four firms in each market so that the sample size, denoted by T equals $4M$, where M denotes the number of markets.
- We report statistics from 100 Monte Carlo simulation/estimation exercises.

Table: Monte Carlo Parameter Values

Parameter	Description	Value
(a) Demand-side parameters		
μ_{α_p}	Price coef. mean	2.0
σ_{α_p}	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

Table: Supply Side Parameter Values

Parameter	Description	Value
(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 cost function	1.0
$\sigma_{\nu+\varsigma}$	Measurement error std. dev.	0.4

Other Parameters

- Correlation Parameters
  - Correlation between ξ_{jm} and
    Own Observed Characteristics (δ_x) $= 0$.
    Other firms' observed characteristics (δ_{xo}) $=$ wages (δ_w) $=$
    Rental Rate (δ_r) $=$ Market Size (δ_Q) $=$0.0833.
    The cost shock (δ_v) $= -0.0833$
  - Correlation between the cost shock and market size (v_{jm} and Q_m) $= \zeta_Q = 0.0833$
- Cobb-Douglas Production function parameters:
  Labor Coefficient (α_c) $= 0.5$
  Capital Coefficient (β_c) $= 0.3$

Table: Parameter estimates based on Shephard's Lemma

(x_{jm} correlated with ξ_{jm} and υ_{jm})

(a) Demand side parameters

			$\hat{\mu}_{\alpha_p}$			$\hat{\sigma}_{\alpha_p}$	
	Sample		Std.			Std.	
Markets	Size	Mean	Dev.	RMSE	Mean	Std.	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

			$\hat{\mu}_{\beta}$			$\hat{\sigma}_{\beta}$	
	Sample		Std.			Std.	
Markets	Size	Mean	Dev.	RMSE	Mean	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

Markets	Sample Size	$\hat{\alpha}_c$ Mean	$\hat{\alpha}_c$ Std. Dev.	$\hat{\alpha}_c$ RMSE	$\hat{\beta}_c$ Mean	$\hat{\beta}_c$ Std. Dev.	$\hat{\beta}_c$ 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		

Markets	Sample Size	$\hat{\eta}$ Mean	$\hat{\eta}$ Std. Dev.	$\hat{\eta}$ RMSE
50	200	0.1613	0.0141	0.0412
100	400	0.1642	0.0096	0.0370
200	800	0.1628	0.0073	0.0379
400	1600	0.1619	0.0045	0.0384
True Value		0.2		

Discussion of results and next step

- In the experiment, observed characteristics are correlated with the demand shock and the cost shock. Results show that while μ_β and η are biased due to the correlation, all other parameter estimates continue to be close to the true values.
- To address the consistent estimation of the observed characteristics coefficient, we next examine the logit case.
- We focus on the following linear regression equation.

$$log\left(s_{jm}\right) - log\left(s_{0m}\right) = -p_{jm}\alpha_{p0} + x_{jm}\beta_0 + \xi_{jm}.$$

- Currently, we focus on the single characteristic case so that x_{jm} is a scalar.

- Petrin Ponder and Seo (2023) use the F.O.C. of optimal choice of $\boldsymbol{x}$ assuming lagged observed characteristics as instruments for identification of β.
- However, since firms do not change product characteristcs each period, they face discrete-continuous dynamic choice problem, which cannot be estimated simply by F.O.C.
- Furthermore, To properly estimate the oligopoly model based on optimal choice, the problems of multiple equilibria and equilibrium selection need to be addressed. (Ciliberto, Murry and Tamer (2023)).

IV estimation

- Market share equation is;

$$log\left(s_{jm}\right) - log\left(s_{0m}\right) = -p_{jm}\alpha_{p0} + x_{jm}\beta_0 + \xi_{jm}.$$

- Taking covariance with IV $z_{jm} \in \{w_{jm}, r_{jm}\}$ results in:

$$\begin{aligned} Cov\left(z_{jm}, lns_{jm} - lns_{0m}\right) &= -Cov\left(z_{jm}, p_{jm}\right)\alpha_{p0} \\ &\quad + Cov\left(z_{jm}, x_{jm}\right)\beta_0 + Cov\left(z_{jm}, \xi_{jm}\right) \end{aligned} \tag{6}$$

- Using the cost data, we already identified $\alpha_{pC} = \alpha_{p0}$. Then, putting what we know on the RHS,

$$\begin{aligned} & Cov\left(z_{jm}, lns_{jm} - lns_{0m}\right) + Cov\left(z_{jm}, p_{jm}\right)\alpha_{p0} \\ = \quad & Cov\left(z_{jm}, x_{jm}\right)\beta_0 + Cov\left(z_{jm}, \xi_{jm}\right) \end{aligned} \tag{7}$$

- If we assume $Cov\left(z_{jm}, \xi_{jm}\right) = 0$, then the IV estimation above identifies β_0.

Identification of the valid instruments for β_0

- We allow for the violation of the instrument orthogonality condition.

$$E\left[\xi_{jm} | z_{jm}\right] \neq 0, \ \ or, \ \ Cov\left(z_{jm}, \xi_{jm}\right) \neq 0$$

- Suppose $Cov\left(z_{jm}, x_{jm}\right) = 0$. Then, from Equation (6), we derive

$$-\alpha_{IV} = \frac{Cov\left(z_{jm}, Ins_{jm} - Ins_{0m}\right)}{Cov\left(z_{jm}, p_{jm}\right)} = -\alpha_{p0} + \frac{Cov\left(z_{jm}, \xi_{jm}\right)}{Cov\left(z_{jm}, p_{jm}\right)}$$

Since $\alpha_{pC} = \alpha_{p0}$ in the population, $\alpha_{IV} = \alpha_{pC}$ implies $Cov\left(z_{jm}, \xi_{jm}\right) = 0$, and thus, the IV orthogonality condition holds.

- We estimate the price coefficient α_p twice: $\widehat{\alpha}_{pC}$ using the cost data and $\widehat{\alpha}_{pIV}$ using the instruments (This is similar to the Hausman test).
- Since we can identify the true price coefficient from the cost data, $(\alpha_{pC} = \alpha_{p0})$ we check whether $\alpha_{pIV} = \alpha_{pC}$.
- If yes, then we know $Cov\left(z_{jm}, \xi_{jm}\right) = 0$, the orthogonality condition for valid IV holds.
- That is, instead of the conventional IV orthogonality condition $Cov\left(z_{jm}, \xi_{jm}\right) = 0$, which we cannot verify from the data, we use the moment condition: $\alpha_{pIV} = \alpha_{pC}$, which we can verify from the data.
- However, $Cov\left(z_{jm}, x_{jm}\right) = 0$, thus, the IV relevance condition is violated. Therefore, z_{jm} cannot be used as an instrument.

Simple specification of instrument invalidity

- We allow for endogeneity of x_{jm} as well as invalidity of IV orthogonality condition:
  - the following specification for w, r and x:

$$x_{jm} = x_{jm}^* + \delta_{x\xi}\xi_{jm}, \ w_{jm} = w_{jm}^* + \delta_{w\xi}\xi_{jm}, \ r_{jm} = r_{jm}^* + \delta_{r\xi}\xi_{jm} \quad (8)$$

  - ξ_{jm} is i.i.d. mean zero with standard deviation σ_ξ.
  - $\left(x_{jm}^*, w_{jm}^*, r_{jm}^*\right)$ is a vector of mean zero random variables, independent to ξ_{jm}.
- We focus on the case where $\delta_{x\xi} \neq 0$, and $\delta_{w\xi} \neq 0$ or $\delta_{r\xi} \neq 0$ or both, so that the OLS estimation of β given α_{p0} is biased and the input prices are invalid instruments for x_{jm}.

Bias of the IV identification?

- Suppose the IV orthogonality conditions do not hold. Then, the price coefficient identifed from the population being $\alpha_{pIV} = \alpha_{pC} = \alpha_{p0}$ does not imply $\beta_{IV} = \beta_0$.
- Consider the following violation of the IV moment condition:

$$\frac{Cov\left(w_{jm}, \xi_{jm}\right)}{Cov\left(w_{jm}, x_{jm}\right)} = \frac{Cov\left(r_{jm}, \xi_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)} \neq 0. \tag{9}$$

- Let

$$\beta_{IV} = \beta_0 + \frac{Cov\left(w_{jm}, \xi_{jm}\right)}{Cov\left(w_{jm}, x_{jm}\right)} = \beta_0 + \frac{Cov\left(r_{jm}, \xi_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)} \neq \beta_0.$$

  Then, $(\alpha_{p0}, \beta_{IV})$ still satisfies Equation (6) without the IV orthogonality conditions in Equation (7).

Another way to look at the bias

- We can discuss the non-identification equivalently using the following potential instrument for price:

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

  The above IV satisfies $Cov\left(z_{pjm}, x_{jm}\right) = 0$.

- We can check from the observed variables whether $\alpha_{pIV} = \alpha_{pC} \left(= \alpha_{p0}\right)$, which implies,

$$\begin{aligned} -\alpha_{pIV} &= \frac{Cov\left(z_{pjm}, ln\left(s_{jm}\right) - ln\left(s_{0m}\right)\right)}{Cov\left(z_{pjm}, p_{jm}\right)} \\ &= \frac{Cov\left(z_{pjm}, -p_{jm}\alpha_{p0} + x_{jm}\beta_0 + \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} \\ &= -\alpha_{p0} + \frac{Cov\left(z_{pjm}, \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} = -\alpha_{pC} = -\alpha_{p0} \end{aligned}$$

- Hence, $\alpha_{pIV} = \alpha_{pC} = \alpha_{p0}$ implies

$$\frac{Cov\left(z_{pjm}, \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} = 0, \tag{10}$$

which validates the instrument z_{pjm}. However, that does not imply validity of w_{jm}, r_{jm}.

- Equation (10) implies

$$\frac{Cov\left(w_{jm}, \xi_{jm}\right)}{Cov\left(w_{jm}, x_{jm}\right)} = \frac{Cov\left(r_{jm}, \xi_{jm}\right)}{Cov\left(r_{jm}, x_{jm}\right)} \tag{11}$$

but cannot exclude the possibility of

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

which is equivalent to Equation (9).

- Thus, $\alpha_{pIV} = \alpha_{pC}$ doesn't identify the validity of w_{jm} and r_{jm} as instruments for x_{jm}.

- That is, the instruments can be constructed even if they identify the true price coefficient if their biases are the same.
- Instead, we use the conditional expectation to derive the IV estimates. That is, let $A_w \equiv [\underline{w}, \overline{w}]$ be a closed intervals of w, and let $A_r = [\underline{r}, \overline{r}]$ be a closed interval of r. Let

$$z_{jm}(A_w, A_r) \equiv \frac{I(w_{jm} \in A_w, r_{jm} \in A_r)}{Pr(w_{jm} \in A_w, r_{jm} \in A_r)}$$

and let

$$z_{pjm} \equiv \frac{z_{jm}(A_w, A_r)}{E[z_{jm}(A_w, A_r) x_{jm}]} - \frac{z_{jm}(A'_w, A'_r)}{E[z_{jm}(A'_w, A'_r) x_{jm}]}$$

Then,

$$E[z_{pjm} x_{jm}] = 0$$

and therefore,

$$E[z_{pjm}(lns_{jm} - lns_{0m})] = -E[z_{pjm} p_{jm}] \alpha_{p0} + E[z_{pjm} \xi_{jm}].$$

- Then, $\alpha_{pIV} = \alpha_{pC}$ implies
$$E\left[z_{pjm}\xi_{jm}\right] = 0 \; \forall \left(A_w, A_r, A'_w, A'_r\right)$$
which implies
$$\frac{E\left[\xi_{jm} | w_{jm} \in A_w, r_{jm} \in A_r\right]}{E\left[x_{jm} | w_{jm} \in A_w, r_{jm} \in A_r\right]} = \frac{E\left[\xi_{jm} | w_{jm} \in A'_w, r_{jm} \in A'_r\right]}{E\left[x_{jm} | w_{jm} \in A'_w, r_{jm} \in A'_r\right]}$$
$$\forall \left(A_w, A_r, A'_w, A'_r\right) \quad (12)$$
- Equation (12) is equivalent to the below equation for a constant B.
$$\frac{E\left[\xi_{jm} | w_{jm} = w, r_{jm} = r\right]}{E\left[x_{jm} | w_{jm} = w, r_{jm} = r\right]} \equiv B \; \forall (w, r) \in R^2_+.$$
- Then, we can rewrite the above as:
$$E\left[\xi_{jm} - Bx_{jm} | w_{jm} = w, r_{jm} = r\right] = 0. \quad (13)$$
for all $(w, r) \in R^2_+$.
- Instruments w_{jm}, r_{jm} are valid if $B = 0$. Next, we prove validity.

Outline

- We use the decomposition of w_{jm}, r_{jm} into the component orthogonal to ξ_{jm} and the rest.
- Because of the linear specification of endogeneity, there exists a linear combination of w_{jm} and r_{jm} that is independent of ξ_{jm}, hence a valid instrument.
- Then, rather than conditioning on w_{jm} and r_{jm}, we can condition on this linear combination and r_{jm}.
- By integrating out the other component r_{jm}, we obtain a contradiction to the assumption that the input prices are not valid - that is, we prove validity of the input prices as instruments.

Case 1: assume $\delta_{r\xi} \neq 0$

- Because $x_{jm} = x_{jm}^{*} + \delta_{x\xi}\xi_{jm}$, Equation (15) implies

$$E\left[\left(1 - B\delta_{x\xi}\right)\xi_{jm} - Bx_{jm}^{*} | w_{jm} = w, r_{jm} = r\right] = 0 \qquad (14)$$

- Let

$$\nabla\left(w_{jm}, r_{jm}, D\right) \equiv w_{jm} - Dr_{jm}.$$

- Then, for

$$D_0 \equiv \frac{\delta_{w\xi}}{\delta_{r\xi}},$$

  ξ_{jm} cancels out, i.e.,

$$\nabla\left(w_{jm}, r_{jm}, D_0\right) = w_{jm} - D_0 r_{jm} = w_{jm}^{*} - D_0 r_{jm}^{*}$$

  which is independent of ξ_{jm}.

- Then, Equation (16) implies

$$E\Big[- Bx^*_{jm} + (1 - B\delta_{x\xi})\,\xi_{jm} | r_{jm} = r, \nabla\left(w_{jm}, r_{jm}, D_0\right) = \overline{D}\Big]$$
$$= 0$$

- We take expectation with respect to r_{jm} given $\nabla\left(w_{jm}, r_{jm}, D_0\right) = \overline{D}$ and obtain

$$E\Big[- Bx^*_{jm} + (1 - B\delta_{x\xi})\,\xi_{jm} | \nabla\left(w_{jm}, r_{jm}, D_0\right) = \overline{D}\Big] = 0 \qquad (15)$$

- Note that

$$E\Big[\xi_{jm} | \nabla\left(w_{jm}, r_{jm}, D_0\right)\Big] = 0$$

due to independence of $\nabla\left(w_{jm}, r_{jm}, D_0\right)$ to ξ_{jm}. Hence,

$$(18) \quad : \quad -BE\left[x^*_{jm} | \nabla\left(w_{jm}, r_{jm}, D_0\right) = \overline{D}\right] = 0 \qquad (16)$$

- Suppose we know D_0. Then as long as there exists $\overline{D}$ such that $E\left[x^*_{jm}|\nabla(w_{jm}, r_{jm}, D_0) = \overline{D}\right] \neq 0$. Then, Equation (18) implies $B = 0$. However, we do not know D_0.
- We still can establish $B = 0$ if we make stronger assumption: If we assume that for any D, there exists $\overline{D}$ such that $E\left[x_{jm}|\nabla(w_{jm}, r_{jm}, D) = \overline{D}\right] \neq 0$, then the same holds for $D = D_0$ and thus, Equation (18) implies $B = 0$.
- This is similar to the instrument relevance condition.

- Then, from Equation (15),

$$E\left[\xi_{jm} - Bx_{jm} | w_{jm} = w, r_{jm} = r\right] = E\left[\xi_{jm} | w_{jm} = w, r_{jm} = r\right] = 0 \quad \forall (w, r) \in R_+^2$$

and thus, we obtain a contradiction to the assumption that $\delta_{r\xi} \neq 0$. Hence $\delta_{r\xi} = 0$ holds. Thus, we have shown that the instrument r_{jm} is valid.

Case 2:assume $\delta_{w\xi} \neq 0$

The same arguments hold if we let

$$\nabla\left(w_{jm}, r_{jm}, D\right) \equiv r_{jm} - Dw_{jm}$$

and then, set

$$D_0 \equiv \frac{\delta_{r\xi}}{\delta_{w\xi}},$$

and follow the same procedure as above. Then we have proved that $\delta_{w\xi} = 0$ and thus, w is a valid instrument. Combining the two cases, it follows that both w and r are valid instruments. Thus, $\delta_{r\xi} = \delta_{w\xi} = 0$ hold.

Constructing valid instruments

- While the above procedure helps verify validity of existing instruments, it may not be useful if there are't any valid instruments.
- Using procedures similar to above, we can construct valid instruments from invalid ones.
- Define the following two candidates for instruments for x_{jm}.

$$z_{wjm} = w_{jm} - D_{xw} x_{jm}, \; z_{rjm} = r_{jm} - D_{xr} x_{jm}.$$

- Let $D_{xw0} \equiv \frac{\delta_{w\xi}}{\delta_{x\xi}}, \; D_{xr0} \equiv \frac{\delta_{r\xi}}{\delta_{x\xi}}$. Then, ξ_{jm} cancels out:

$$z_{w0jm} = w^*_{jm} - D_{xw0} x^*_{jm}, z_{r0jm} = r^*_{jm} - D_{xr0} x^*_{jm}.$$

(z_{w0jm}, z_{r0jm}) do not contain unobserved product characteristics. Hence,

$$E\left[\xi_{jm} | z_{w0jm}, z_{r0jm}\right] = 0$$

- On the other hand, for $(D_{xw}, D_{xr}) \neq (D_{xw0}, D_{xr0})$,

$$
\begin{aligned}
z_{wjm} &= w_{jm}^{*} - D_{xw} x_{jm}^{*} + (D_{xw0} - D_{xw})\, \delta_{x\xi} \xi_{jm} \\
z_{rjm} &= r_{jm}^{*} - D_{xr} x_{jm}^{*} + (D_{xr0} - D_{xr})\, \delta_{x\xi} \xi_{jm}
\end{aligned}
$$

  contain ξ_{jm}. Hence, at least one of the instruments is invalid.
- We cannot simply derive and use z_{w0jm}, z_{r0jm} as instruments because we do not know the coefficients $\delta_{w\xi}$, $\delta_{r\xi}$ and $\delta_{x\xi}$.
- But for every $\delta_{w\xi}$, $\delta_{r\xi}$ and $\delta_{x\xi}$, we can construct instruments (z_{wjm}, z_{rjm}) and check their validity using the method discussed above.

That is, valid instruments can be identified by finding D_{xr}, D_{xr} in

$$z_{jm}\left(A_w, A_r\right) \equiv \frac{I\left(z_{wjm}\left(D_{xw}\right) \in A_w, z_{rjm}\left(D_{xr}\right) \in A_r\right)}{Pr\left(z_{wjm}\left(D_{xw}\right) \in A_w, z_{rjm}\left(D_{xr}\right) \in A_r\right)}$$

and let

$$z_{pjm} \equiv \frac{z_{jm}\left(A_w, A_r\right)}{E\left[z_{jm}\left(A_w, A_r\right) x_{jm}\right]} - \frac{z_{jm}\left(A'_w, A'_r\right)}{E\left[z_{jm}\left(A'_w, A'_r\right) x_{jm}\right]}$$

that satisfy $\alpha_{pIV} = \alpha_{pC}$, i.e.,

$$\begin{aligned} -\alpha_{pIV} &= \frac{Cov\left(z_{pjm}, ln\left(s_{jm}\right) - ln\left(s_{0m}\right)\right)}{Cov\left(z_{pjm}, p_{jm}\right)} \\ &= -\alpha_{p0} + \frac{Cov\left(z_{pjm}, \xi_{jm}\right)}{Cov\left(z_{pjm}, p_{jm}\right)} = -\alpha_{pC} = -\alpha_{p0} \end{aligned}$$

for any $(z_w, z_r) \neq (z'_w, z'_r)$.

Estimation of valid instruments and consistent β

- We have the following market share equation:

$$lns_{jm} - lns_{0m} = -p_{jm}\alpha_0 + x_{jm}\beta_0 + \xi_{jm}$$

- Use the indicator function of sets of overlapping rectangles as instruments.

$$\begin{aligned} & R\left(\Delta z_{wk}, \Delta z_{rl}\right) \\ \equiv \; & \left\{\left(z_{wjm}, z_{rjm}\right) : \left(z_{wjm}, z_{rjm}\right) \in \left[z_{wk}, z_{w,k+\Delta}\right] \times \left[z_{rk}, z_{r,k+\Delta}\right]\right\} \end{aligned}$$

- Use the sets of overlapping rectangles $R\left(\Delta z_{wk}, \Delta z_{rl}\right)$ that cover the domain of $\left(z_w, z_r\right)$, we derive the following conditional expectations.

$$\begin{aligned} & E\left[lns_{jm} - lns_{0m} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right] \\ = \; & -E\left[p_{jm} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right]\alpha_0 \\ & +E\left[x_{jm} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right]\beta_0 \\ & +E\left[\xi_{jm} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right] \end{aligned}$$

Furthermore, let

$$
\begin{aligned}
\widehat{y}_{kl}\left(D_{xw}, D_{xr}\right) &= E\left[Ins_{jm} - Ins_{0m} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right] \\
\widehat{p}_{kl}\left(D_{xw}, D_{xr}\right) &= E\left[p_{jm} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right] \\
\widehat{x}_{kl}\left(D_{xw}, D_{xr}\right) &= E\left[x_{jm} \middle| \left(z_{wjm}\left(D_{xw}\right), z_{rjm}\left(D_{xr}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right]
\end{aligned}
$$

- If $D_{xw} = D_{xw0}$ and $D_{xr} = D_{xr0}$, then

  $$E\left[\xi_{jm} | \left(z_{wjm}\left(D_{xw0}\right), z_{rjm}\left(D_{xr0}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right] = 0$$

  Therefore,

  $$\widehat{y}_{kl} = b_0 + \widehat{p}_{kl} b_p + \widehat{x}_{kl} b_x.$$

  $b_p = -\alpha_{p0}$, $b_x = \beta_x$.

- Furthermore, similar arguments as before can be used to show that $b_p = -\alpha_{p0}$ implies validity of the instruments $I\left(\left(z_{wjm}\left(D_{xw0}\right), z_{rjm}\left(D_{xr0}\right)\right) \in R\left(\Delta z_{wk}, \Delta z_{rl}\right)\right)$, and thus, $b_x = \beta_{x0}$.
- However, in the actual finite sample, the sample analog of the expectation E is the sample average, and is subject to the sample error, which we denote to be u_{kl}.

- We obtain the estimates $\left(\widehat{b}_0, \widehat{b}_p, \widehat{b}_x\right)$ by minimizing the following loss function:

$$\left(\widehat{b}_0, \widehat{b}_p, \widehat{b}_x\right) = argmin_{(b_0, b_p, b_x, D_{xw}, D_{xr})}\left(\sum_{k,l} u_{kl}^2 + \phi\left(\left|b_p - \widehat{\alpha}_{pC}\right|\right)\right)$$

  where u_{kl} is the residual, i.e.

$$u_{kl} \equiv y_{kl} - \widehat{y}_{kl}$$

  and $\phi() \geq 0$ is the loss function, i.e., $\phi(v) > 0$ if $v \neq 0$ and $\phi(v) = 0$ if and only if $v = 0$. The example of a loss function is

$$\phi(v) = v^2$$

- Set $\widehat{\beta} = \widehat{b}_x$.

Monte-Carlo results

(a) Demand side estimates

Markets	Sample Size	$\widehat{\alpha}_p$ Mean	$\widehat{\alpha}_p$ Std. Dev.	$\widehat{\alpha}_p$ RMSE	$\widehat{\beta}$ Mean	$\widehat{\beta}$ Std. Dev.	$\widehat{\beta}$ RMSE
50	200	1.9984	0.0628	0.0625	0.7285	0.1246	0.2984
100	400	2.0191	0.0474	0.0508	0.9116	0.0993	0.1326
200	800	2.0016	0.0296	0.0295	0.9799	0.0861	0.0880
400	1600	2.0011	0.0226	0.0225	1.0497	0.0750	0.0897
True Value:		2.0			1.0		

(b) OLS estimates

Markets	Sample Size	$\widehat{\alpha}_{pOLS}$ Mean	$\widehat{\alpha}_{pOLS}$ Std. Dev.	$\widehat{\alpha}_{pOLS}$ RMSE	$\widehat{\beta}_{OLS}$ Mean	$\widehat{\beta}_{OLS}$ Std. Dev.	$\widehat{\beta}_{OLS}$ RMSE
50	200	1.7255	0.0701	0.2832	0.9272	0.0649	0.0973
100	400	1.7246	0.0491	0.2797	0.9274	0.0453	0.0855
200	800	1.7181	0.0313	0.2836	0.9244	0.0280	0.0806
400	1600	1.7205	0.0216	0.2804	0.9238	0.0218	0.0793
True value:		2.0			1.0		

(c) IV estimates: (w_{jm}, r_{jm})

			$\widehat{\alpha}_{pIV}$			$\widehat{\beta}_{IV}$	
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	2.1180	1.3817	1.3799	1.3051	1.3712	1.3980
100	400	1.9777	0.2132	0.2133	1.1656	0.2115	0.2678
200	800	2.0477	0.2198	0.2237	1.2379	0.2163	0.3209
400	1600	2.0165	0.1128	0.1134	1.2045	0.1106	0.2322
True Value:		2.0			1.0		

(d) Estimates given the price coefficient $\widehat{\alpha}_{pC}$

		$\widehat{\beta}_{OLS}\,(\widehat{\alpha}_{pC})$			$\widehat{\beta}_{IV}\,(\widehat{\alpha}_{pC})$		
Markets	Sample Size	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
50	200	1.1505	0.0600	1.1619	1.1827	0.0704	0.1960
100	400	1.1683	0.0476	0.1749	1.2052	0.0582	0.2132
200	800	1.1558	0.0267	0.1580	1.1917	0.0318	0.1943
400	1600	1.1533	0.0209	0.1548	1.1894	0.0245	0.1910
True Value		1.0			1.0		

		Hausman-Wu			Sargan		
Markets	Sample Size	test stat.	p-value		test stat.	p-value	
50	200	3.4522	0.2387		1.2913	0.4595	
100	400	7.9305	0.0786		0.7833	0.5447	
200	800	13.466	0.0224		1.2794	0.4683	
400	1600	25.754	0.0001		0.8729	0.4870	

- IV validity parameter setup: $\delta_{x\xi} = 0.8$, $\delta_{w\xi} = \delta_{r\xi} = 0.4$: instruments are invalid, and $\beta_{IVw} \approx \beta_{IVr}$.
- Our procedure: the price coefficient ($\widehat{\alpha}_p$), the coefficient on the observed characteristics ($\widehat{\beta}$) are close to the true values.
- OLS and IV estimates of (α_p, β): closeness of the IV estimated price coefficient to cost-based estimate does not correspond to the validity of the IV for β.
  - OLS: $\widehat{\alpha}_{pOLS}$ (IV estimate with instruments: p_{jm} and x_{jm}) has large downward bias, downward bias of $\widehat{\beta}_{OLS}$ is small.
  - IV: $\widehat{\alpha}_{pIV}$ only has small asymptotic bias, large upward asymptotic bias in $\widehat{\beta}_{IV}$
- Estimate β given $\widehat{\alpha}_p$: Upward asymptotic bias of β_{IV} larger than the upward bias of β_{OLS}. On the other hand,
  - Sargan test insignificant, IV validity is not rejected.
  - Hausman-Wu test: significant: OLS has downward bias.
- Hence, commonly used test results indicate the IVs are valid, and thus, OLS has downward bias. Our procedure implies that IVs are not valid and OLS has upward bias because the true beta is 1, not 1.18 as these tests conclude.

Identification of demand with multivariate $\boldsymbol{x}$.

- Both in the logit and the random coefficient aggregate demand (BLP) model, cost data identifies the marginal revenue without instruments.
- Logit model: marginal revenue only identifies the price coefficient α_p. Insufficient for the identification of the coefficient of multivariate $\boldsymbol{x}$.
- BLP model: marginal revenue identifies the price coefficient α_p and the vector random coefficient parameters $\boldsymbol{\sigma}_\beta$: identify the multivariate $\boldsymbol{x}$.

Conclusion

- We show that we can consistently estimate the demand coefficients and the key coefficients of the cost function of a differentiated product oligopoly model by using cost data and without instruments for output and price.
- We develop a way to verify the validity of existing instruments and constructing valid instruments from invalid ones for observed product characteristics.
- In our Monte-Carlo experiments, we show that our method works well even when all the commonly used instruments are invalid.

- Benefit for industries with cost data: we can estimate the key parameters of the cost function and all parameters of the logit or BLP demand parameters.
- For industries without cost data:information on instruments that are identified or constructed as valid instruments using cost data could be useful.
  - Researchers can use the validated or constructed instruments.
  - Or, such information on instruments can be helpful in constructing the bounds or prior distributions of the IV moment conditions, when researchers allow for moment conditions to not exactly equal to zero (see Conley, Hansen and Rossi (2012)).

Thank you for your attention.