Implentation of the identification and estimation result on Slide 18.

1 Model Specification and Data Assumptions

The production function is of the form

$$\Phi(x, y, \epsilon, \eta) = \beta_{xy} xy + \beta_{x\eta} x\eta + \beta_{y\epsilon} y\epsilon + \epsilon\eta.$$

It is assumed that the production function is monotone in scalar unobservables ϵ and η,

$$\beta_{y\epsilon} y + \eta > 0, \quad \beta_{x\eta} x + \epsilon > 0.$$

Further, it is assumed that ϵ and η are independent from x and y, and

$$\epsilon \sim \text{LogNormal}(\mu_\epsilon, \sigma_\epsilon), \quad \eta \sim \text{LogNormal}(\mu_\eta, \sigma_\eta).$$

The medians of ϵ and η are normalized to one, i.e. $\mu_\epsilon = \mu_\eta = 0$.

The equilibrium matching, upstream and downstream profits are observed in the data. Observation i is the tuple $\left(x_i, y_i, \pi_i^u, \pi_i^d\right)$.

2 Estimation

For each observation i, under the monotonicity assumption, we recover the normalizied $\hat{\epsilon}_i$ as the conditional CDF of upstream profits,

$$\hat{\epsilon}_i = \hat{F}^u(\pi_i^u | x = x_i).$$

Similarly, for the downstream observation i,

$$\hat{\eta}_i = \hat{F}^d\left(\pi_i^d | y = y_i\right).$$

For a given vector of parameters

$$\boldsymbol{\theta} = (\beta_{xy}, \beta_{x\eta}, \beta_{y\epsilon}, \sigma_\epsilon, \sigma_\eta),$$

we can invert the normalized unonbservables into the quantiles of their parameterized distributions given σ_ϵ and σ_η,

$$\tilde{\epsilon}_i = q(\hat{\epsilon}_i | \sigma_\epsilon), \quad \tilde{\eta}_i = q(\hat{\eta}_i | \sigma_\eta),$$

where q is the quantile function for the Log Normal distribution with $\mu = 0$.

The estimator for $\boldsymbol{\theta}$ minimizes

$$\sum_{i=1}^{n} \left[\left(\pi_i^u + \pi_i^d\right) - \left(\beta_{xy} x_i y_i + \beta_{x\eta} x_i \tilde{\eta}_i + \beta_{y\epsilon} y_i \tilde{\epsilon}_i + \tilde{\epsilon}_i \tilde{\eta}_i\right)\right]^2.$$

2.1 Estimating Conditional CDFs

To estimate the conditional CDF We use the Nadaraya-Watson (NW) estimator. For a fixed profit π^u and x, the estimator is given by

$$\hat{F}^u\left(\pi^u|x\right)=\frac{\sum_{i=1}^n \phi\left(\frac{x-x_i}{h_x}\right)1\left(\pi_i^u\leq\pi^u\right)}{\sum_{i=1}^n \phi\left(\frac{x-x_i}{h_x}\right)}.$$

The above estimator is smooth in x but not in y. The only smoothing parameter is the one for x. We use the leave-one-out cross validation method to choose the bandwidth. For each observation, the leave-one-out residual is given by

$$\hat{e}_i\left(\pi^u\right)=1\left(\pi_i^u\leq\pi^u\right)-\hat{F}_{-i}^u\left(\pi^u|x_i\right),$$

where $\hat{F}_{-i}^u\left(\pi^u|x_i\right)$ is the leave-one-out estimator given by

$$\hat{F_{-i}}^u\left(\pi^u|x\right)=\frac{\sum_{j\neq i}\phi\left(\frac{x-x_j}{h_x}\right)1\left(\pi_j^u\leq\pi^u\right)}{\sum_{j\neq i}\phi\left(\frac{x-x_j}{h_x}\right)},$$

that is the observation i is excluded from the sample used to estimate the conditional cdf at observation i.

The CV criterion for a fixed profit level π^u is

$$\begin{aligned}CV\left(\pi,h_x\right)&=\frac{1}{n}\sum_{i=1}^n\hat{e}_i\left(\pi^u\right)^2 f_x\left(x_i\right)\\&=\frac{1}{n}\left(1\left(\pi_i^u\leq\pi^u\right)-\hat{F}_{-i}^u\left(\pi^u|x_i\right)\right)^2.\end{aligned}$$

The optimal bandwidth minimizes

$$CV\left(h_x\right)=\int CV\left(\pi^u,h_x\right)d\pi^u.$$

We approximate this by a grid over the values of profits, by randomly selecting N profit observations

$$CV\left(h\right)\approx\sum_{i=1}^N CV\left(\pi_i^u,h_x\right).$$

Thus,

$$h_x^*=\arg\min_{h_x}\left\{\sum_{i=1}^N CV\left(\pi_i^u,h_x\right)\right\}.$$

3 Simulation

3.1 Parameterization

$$\Phi\left(x, y, \epsilon, \eta\right) = -3xy + 0.7x\eta + 3.0y\epsilon + \epsilon\eta.$$

$$\sigma_{\epsilon} = 0.2, \sigma_{\eta} = 0.5.$$

Further, I choose a Log Normal distribution for x and y so the support is positive and the monotonicity assumtion is not violated.

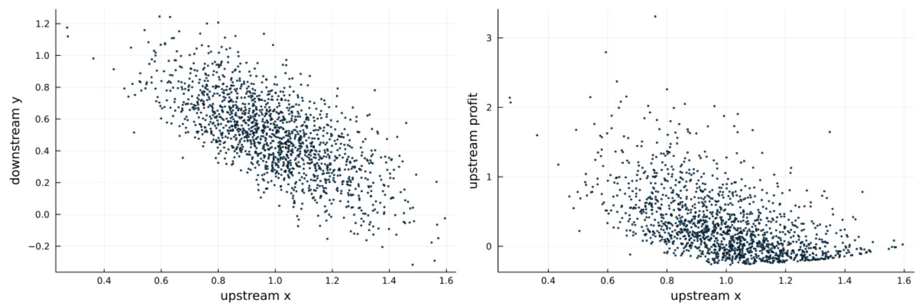

Figure 1: Matching pattern and upstream profits

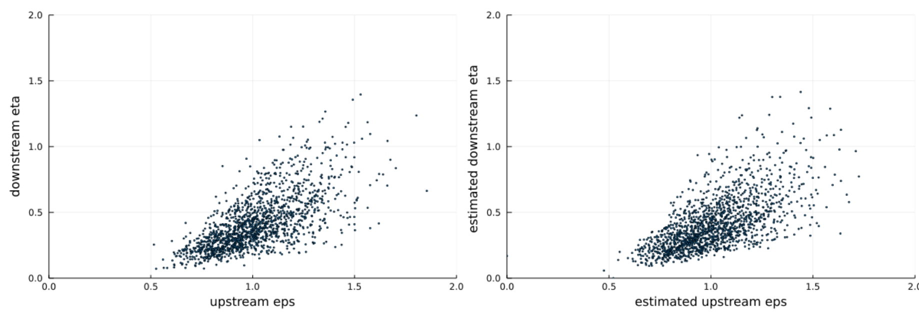

Figure 2: Unobservable realizations (left) versus the estimated unobservables (right)

3.2 Simulation Results

Here are the simulation results from 80 replications.

		Firms 500		**Firms = 1000**		**Firms =1500**	
	truth	**Bias**	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
β_{xy}	-3.0	0.04	0.48	-0.06	0.29	-0.04	0.26
$\beta_{x\eta}$	0.7	0.02	0.11	0.02	0.09	0.02	0.07
$\beta_{y\epsilon}$	3.0	-0.06	0.41	0.02	0.25	0.00	0.22
σ_ϵ	0.2	0.02	0.06	0.01	0.05	0.01	0.05
σ_η	0.7	-0.04	0.30	-0.09	0.36	-0.09	0.36