Problem Set 0 ECON8825

Federico Rodari

September 10, 2022

1. Simulating the 2-equations SEM

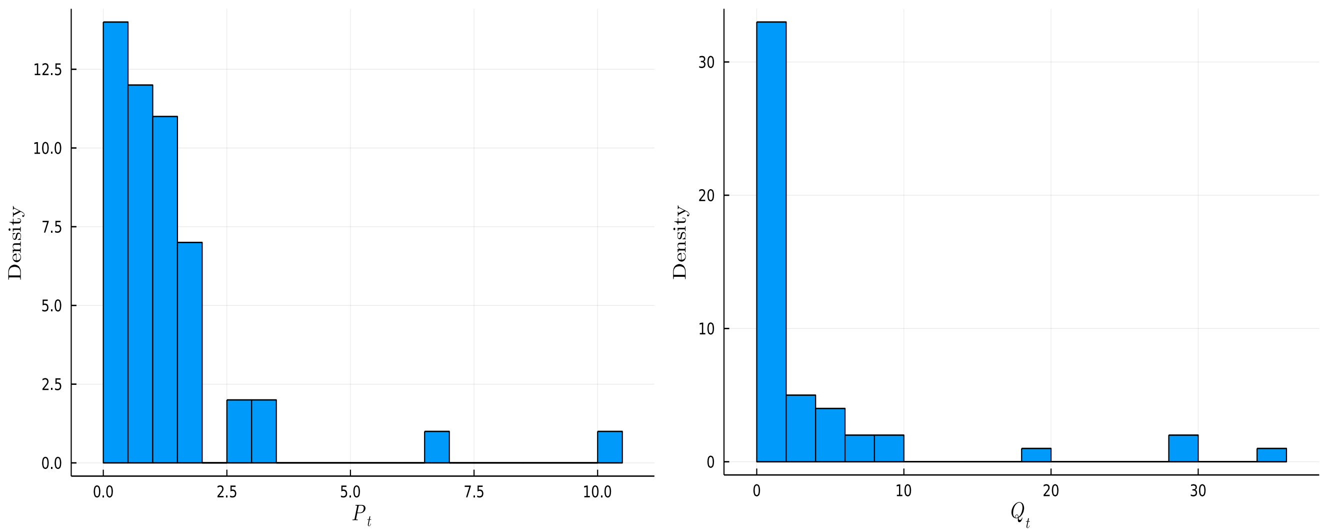

Figure 1: (Left) Price distribution; (Right) Quantity distribution

2. The OLS estimates are

$$\beta^{\text{OLS}} = \begin{bmatrix} -0.03 \\ -1.67 \end{bmatrix}$$

The estimator is neither unbiased nor consistent. To see the direction of the bias:

$$\begin{aligned}
\hat{\beta}_2 &= \frac{\text{Cov}(\log P_t, \log Q_t)}{\text{Var}(\log P_t)} = \frac{\text{Cov}(\log P_t, \beta_1 - \beta_2 \log P_t + \epsilon_t^D)}{\text{Var}(\log P_t)} \\
&= -\beta_2 + \frac{\text{Cov}(\log P_t, \epsilon_t^D)}{\text{Var}(\log P_t)} \\
&= -\beta_2 + \frac{\text{Cov}(\frac{1}{1+\beta_2\delta}\left(\mu + \delta\beta_1 + \delta\epsilon_t^D - \log a_t\right), \epsilon_t^D)}{\text{Var}(\log P_t)} \\
&= -\beta_2 + \frac{\delta}{1+\beta_2\delta}\text{Var}\left(\epsilon_t^D\right)
\end{aligned}$$

which implies the estimator is upward biased. If we are able to find an instrument for $\log P_t$ that satisfies validity and relevance, we can show the OLS estimator is unbiased and consistent.

3. We can use $\log Z_t$ as an instrument and adopt a 2SLS estimation procedure. The first step consists of estimating the following linear model

$$\log P_t = \alpha_0 + \alpha_1 \log Z_t + \upsilon_t$$

and storing the fitted values

$$\widehat{\log P_t} = \hat{\alpha}_0 + \hat{\alpha}_1 \log Z_t.$$

The second stage estimates the demand equation using the fitted values (which are exogenous by construction)

$$\log Q_t = \beta_1 - \beta_2 \widehat{\log P_t} + \epsilon_t^D.$$

We also know that β_2^{2SLS} is a consistent estimator, which can be shown by increasing T. The estimates are

$$\beta^{2SLS} = \begin{bmatrix} 0.053 \\ -2.00 \end{bmatrix}.$$

To show consistency, by simulation

4. The population moments are

$$\widehat{\mathbb{E}}[g(W,\theta)] = \frac{1}{T}\sum_t g(w,\theta) = \begin{bmatrix} \mathbb{E}\left[\log Z_t \log Q_t\right] \\ \mathbb{E}\left[\log Z_t \log P_t\right], \end{bmatrix} = \begin{bmatrix} 1.142 \\ -0.571 \end{bmatrix}$$

where

$$\begin{aligned} \mathbb{E}[g^1(W,\theta)] &= \mathbb{E}\left[\log Z_t \log Q_t \mid \theta\right] \\ &= \mathbb{E}\left[\log Z_t \left(\frac{1}{1+\beta_2\delta}\left(\beta_1 - \beta_2\mu + \beta_2 \log a_t + \epsilon_t^D\right)\right) \mid \theta\right] \\ &= \frac{1}{1+\beta_2\delta}\left((\beta_1 - \beta_2\mu)\,\mathbb{E}(\log Z_t) + \beta_2 \mathbb{E}(\log Z_t \log a_t)\right) \\ &= \frac{1}{1+\beta_2\delta}\left((\beta_1 - \beta_2\mu)\,\mathbb{E}(\log Z_t) + \beta_2\gamma\sigma^Z\right) \end{aligned}$$

$$\begin{aligned} \mathbb{E}[g^2(W,\theta)] &= \mathbb{E}\left[\log Z_t \log P_t \mid \theta\right] \\ &= \mathbb{E}\left[\log Z_t \left(\frac{1}{1+\beta_2\delta}\left(\mu + \delta\beta_1 + \delta\epsilon_t^D - \log a_t\right)\right) \mid \theta\right] \\ &= \frac{1}{1+\beta_2\delta}\left((\mu + \delta\beta_1)\,\mathbb{E}(\log Z_t) - \mathbb{E}(\log Z_t \log a_t)\right) \\ &= \frac{1}{1+\beta_2\delta}\left((\mu + \delta\beta_1)\,\mathbb{E}(\log Z_t) - \gamma\sigma^Z\right). \end{aligned}$$

The empirical counterpart is

$$\widehat{\mathbb{E}}[g(W,\theta)] = \frac{1}{T}\sum_t g(w,\theta) = \begin{bmatrix} \frac{1}{T}\sum_t \log Z_t \log Q_t \\ \frac{1}{T}\sum_t \log Z_t \log P_t \end{bmatrix} = \begin{bmatrix} 1.375 \\ -0.691 \end{bmatrix}$$

5. After simulating the model S times

$$\mathbb{E}(g(W,\theta)) - \mathbb{E}^{\text{sim}}(g(W,\theta)) = \begin{bmatrix} -0.0521 \\ 0.0196 \end{bmatrix}$$

   where $\widetilde{\mathbb{E}}^{\text{sim}}[g(W,\theta)] = \frac{1}{S}\sum_s \frac{1}{T}\sum_t g(w_s,\theta)$

6. We perform a naive-grid search over β_2. We could use numerical optimization, but since we know the true value, we decided to plot the loss function over a neighborhood of the true parameter. We expect that on average, the minimum of the loss function reached across all simulations will be attained at $\beta_2 = 2$

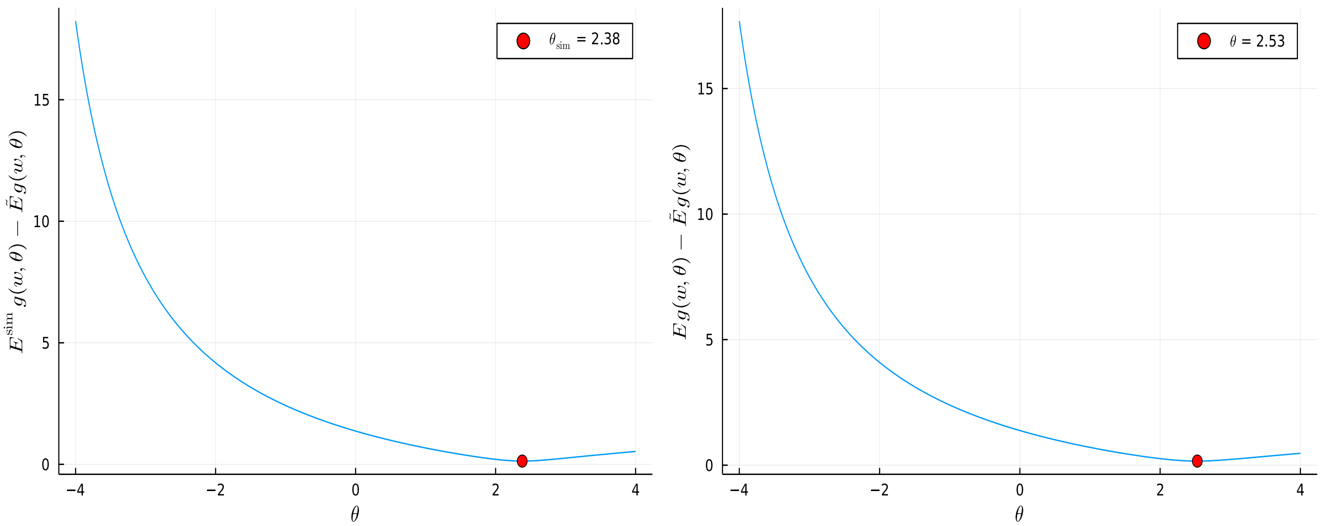

Figure 2: (Left) Simulated Moments; (Right) Population Moments

7. Different norms will change the weight we give to each loss point, so it will change the result.

8. We cannot identify δ as the SEM is under-identified. We do not have an instrument we can use in the supply equation.