Ec141, Spring 2018

Professor Bryan Graham

Problem Set 1

Due: February 5th, 2018 (note this is prior to the final day to add/drop classes for the Spring 2018 semester)

Problem sets are due at 5PM in the GSIs mailbox. You may work in groups, but each student should turn in their own write-up (including a narrated/commented and executed Jupyter Notebook). Please use markdown boxes within your Jupyter notebook for narrative answers to the questions below.

1 The distribution of total factor productivity

The file `mf_firms.out` contains 29,836 firm-by-year observations for a sample of publicly traded manufacturing firms drawn from the *S&P Capital IQ - Compustat* database. The following firm attributes, measured from 1998 to 2014 inclusive, are included:

`gvkey` – Compustat firm identification code

`year` – calendar year

`Y` – total real sales by the firm (in millions of 2009 US$)

`K` – capital stock (in millions of 2009 US$)

`L` – employees (in thousands)

`M` – materials expenditures (in millions of 2009 US$)

`VA` - total real valued added by the firm (in millions of 2009 US$)

`w` - annual wage rate (in 2009 US$)

`i` – real investment (in millions of 2009 US$)

`naics_4digits` – NAICS four digit sector code for the firm

In this problem set you will use this dataset to study the distribution of productivity across firms. A nice introduction to economic research in this area is provided by Syverson (2011).

Preparing and exploring the dataset

1. How many firm-year observations are in the dataset?

2. How many distinct firms?

3. In 2014 what was the aggregate total sales across all manufacturing firms in the dataset?

4. How many employees did these firms employ in total?

5. In 2014 compute the average, standard deviation and 5th, 25th, 50th, 75th and 95th percentiles of total sales, capital stock, employees, materials, and investment across all firms in your dataset. Display this information in a table.

6. Write a few sentences summarizing your dataset.

Profit maximization

Assume that in each period t output, Y_t, is produced using capital, K_t, labor, L_t, and materials, M_t, according to the production technology

$$Y_t = A_t K_t^{1-\alpha-\beta} L_t^{\alpha} M_t^{\beta}, \tag{1}$$

where A_t is a factor neutral shifter or *total factor productivity*. Let R_t and W_t be the prevailing rental rate for capital and wage rate for labor. Assume that firms maximize profits – taking their output price, P_t, as fixed (i.e., perfect competition):

$$\max_{k_t, l_t, m_t} P A_t k_t^{1-\alpha-\beta} l_t^{\alpha} m_t^{\beta} - R_t k_t - W_t l_t - m_t.$$

Let K_t, L_t and M_t denote the profit-maximizing input choices made by the firm. Show that the firm's first order conditions for labor and materials imply that

$$\alpha = \frac{W_t L_t}{Y_t}, \quad \beta = \frac{M_t}{Y_t}.$$

Measuring productivity

Note that `Y` in the dataset corresponds to $P_t Y_t$ in the theoretical model; our theory is about physical units of output, Y_t, but what we observe in the financial filings of firms is generally total sales, $P_t Y_t$. It is this latter quantity which is recorded as `Y` in the dataset.

1. Construct a measure of the firm's wage bill each period, $W_t L_t$, using the formula

$$\text{wage bill} = \frac{(\mathtt{L} \times 1,000) \times \mathtt{w}}{1,000,000}.$$

   Explain the reasoning underlying this formula.

2. Let i index firms and t years. Consider the following estimate of firm i's elasticity of output with respect to labor:

$$\hat{\alpha}_i = \frac{1}{T_i} \sum_t \frac{\text{wage bill}_{it}}{\mathtt{Y}_{it}}$$

where the summation is over all years firm i is in the dataset. Similarly estimate firm i's elasticity of output with respect to materials as:

$$\hat{\beta}_i = \frac{1}{T_i} \sum_t \frac{\mathtt{M}_{it}}{\mathtt{Y}_{it}}.$$

Explain the reasoning underlying these elasticity measures.

3. Compute the average, standard deviation and 5th, 25th, 50th, 75th and 95th percentiles of $\hat{\alpha}_i$ and $\hat{\beta}_i$ across all firms in your dataset. Display this information in a table.

4. Construct the following measure of productivity for each firm-year in your dataset:

$$\mathtt{TFPR}_{it} = \frac{\mathtt{Y}_{it}}{\mathtt{K}_{it}^{1-\hat{\alpha}_i-\hat{\beta}_i} \mathtt{L}_{it}^{\hat{\alpha}_i} \mathtt{M}_{it}^{\hat{\beta}_i}}.$$

How does this measure relate to A_{it} – total factor productivity – as defined in the theoretical model?

5. In 2014 compute the average, standard deviation and 5th, 25th, 50th, 75th and 95th percentiles of $\mathtt{TFPR}_{it}$ across all firms in your dataset. Display this information in a table. Are the productivity differences across firms larger or smaller than you expected?

6. Pick a four digit industry of interest (if you google NAICS 4 digit codes you will find many online look-up tables). As an example 3361 corresponds to motor vehicle manufacturing. In 2014 compute the average, standard deviation and 5th, 25th, 50th, 75th and 95th percentiles of $\mathtt{TFPR}_{it}$ across all firms in your chosen industry. Are the productivity differences across firms in your chosen industry larger or smaller than you expected?

Productivity decomposition

Let $S_t = Y_t / \mathbb{E}[Y_t]$ and show that

$$\begin{aligned}\mathbb{E}[S_t A_t] &= \mathbb{E}[S_t]\,\mathbb{E}[A_t] + \mathbb{C}(S_t, A_t) \\ &= \mathbb{E}[A_t] + \left[\frac{\mathbb{C}(S_t, A_t)}{\mathbb{V}(A_t)}\right] \mathbb{V}(A_t).\end{aligned}$$

Discuss how this expression might be used to understand industry-level change in productivity over time (we will return to this expression in a later problem set).

2 Conditional expectations

Consider the *roof distribution* with probability density function

$$f_{X,Y}(x, y) = x + y$$

if $(x, y) \in [0, 1]^2$ and zero otherwise.

1. Compute the marginal density function of X.
2. Compute the conditional density function of Y given $X = x$.
3. Compute the conditional expectation function of Y given $X = x$.

References

Syverson, C. (2011). What determines productivity. *Journal of Economic Literature*, 49(2), 326 – 365.