Estimating worker and firm heterogeneity using linked employer-employee data

Richard Upward

December 2008

Leverhulme Centre
for Research on Globalisation and Economic Policy

The University of Nottingham

Introduction

- Since about 1995 there has been an explosion of papers in labour economics which use linked employer-employee data
- Many of these papers cite Abowd, Kramarz & Margolis (1999, Econometrica)
- Are some firms more productive because they hire more productive workers?
- Do some workers earn more because they work for more productive firms?
- Assortative matching?
- In empirical international economics, the use of linked employer-employee data is also spreading, because of the simultaneous availability of measures of globalisation at the firm-level

Some research questions

- Do multinational firms pay higher wages?
- Do exporters pay higher wages?
- Are these higher wages the result of higher firm or worker productivity?
- The effect of imports on wages and skill structure within the firm
- Human capital and firm productivity
- Worker and job turnover

Links to models of trade with heterogeneous firms and workers

- "Workers are differentiated by their skill level ... Z. There are two interpretations of Z. The first is that workers differ in some observable characteristics ... the second interpretation of Z is as a measure of worker quality or ability that can be observed by a firm byt not by an econometrician."
- "... firms that export are larger, used more advanced technology and pay higher wages than those that do not. They also exhibit greater sales per worker, but this apparent productivity advantage disappears when worker heterogeneity is properly controlled." (Yeaple JIE 2005)

A (very) simple illustrative example

- A panel of N individuals: $i = 1, \ldots, N$ individuals, observed T_i times
- Linked to data on J firms: $j = 1, \ldots, J$
- Total sample is $N^* = \sum_{i=1}^{N} T_i$
- y_{it} is the dependent variable (e.g. the wage) and $\mathbf{x}_{it}$ is a vector of time-varying characteristics of individual i at time t (including those of his/her firm)

- So, the data look like this:

i	j	y_{it}
1	A	$y_{1,1}$
1	A	$y_{1,2}$
		$\vdots$
2	C	
2	A	
3	C	
3	C	
4	C	
4	C	
5	A	
5	B	
6	B	
6	B	
		$\vdots$
7	B	
7	B	$y_{7,2}$

- In this example we have a sample of workers and we know every firm they work for
- In other cases, we may have a sample of firms and we know all the workers in those firms
- Or, a sample of firms with a sample of workers in those firms

- θ_i denotes individual fixed effect (ability)
- ψ_j denotes firm fixed effect (productivity?)
- The basic model:

$$\mathbf{y} = \mathbf{D}\boldsymbol{\theta} + \mathbf{F}\boldsymbol{\psi} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- We are ignoring the possibility of "match effects" (Woodcock 2008)
- The matrix $\mathbf{D}$ is the $(N^* \times N)$ matrix of individual dummies:

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- The matrix $\mathbf{F}$ is the $(N^* \times J)$ matrix of firm dummies:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- The usual way to estimate the one-way fixed effects model is to "sweep out" the matrix **D**

$$\mathbf{M}_D\mathbf{y} = \mathbf{M}_D\mathbf{F}\psi + \mathbf{M}_D\mathbf{X}\boldsymbol{\beta} + \mathbf{M}_D\boldsymbol{\epsilon}$$

  and use OLS. The matrix $\mathbf{M}_D = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$ creates deviations from means.

- For $T = 2$, this is equivalent to first-differencing

$$\Delta\mathbf{y} = \Delta\mathbf{F}\psi + \Delta\mathbf{X}\boldsymbol{\beta} + \Delta\boldsymbol{\epsilon}$$

i	$\Delta\mathbf{y}$	$\Delta\mathbf{F}$		
1	Δy_1	0	0	0
2		1	0	-1
3		0	0	0
4	$\vdots$	0	0	0
5		-1	1	0
6		0	0	0
7	Δy_7	0	0	0

- In practice, $\Delta\mathbf{F}$ is a large, non-patterned matrix
- How large? It is $N \times J$, and so we would require a dummy variable for each firm
- Analytical results are not available ... we think
- If we estimate the one-way fixed effects model on our example data:

$$\hat{\boldsymbol{\psi}} = (\Delta\mathbf{F}'\Delta\mathbf{F})^{-1}(\Delta\mathbf{F}'\Delta\mathbf{y})$$

$$\begin{pmatrix} \hat{\psi}_B - \hat{\psi}_A \\ \hat{\psi}_C - \hat{\psi}_A \end{pmatrix} = \begin{pmatrix} \Delta y_5 \\ -\Delta y_2 \end{pmatrix}$$

- In this example, the firm-effect for firm B (normalised relative to firm A) is identified by the individual who moves from firm A to firm B in the data
- The firm-effect for firm C (relative to firm A) is identified by the individual who moves from firm C to firm A
- This relates to the pattern of movement in the data

i	j	y_{it}
1	A	$y_{1,1}$
1	A	$y_{1,2}$
2	C	$\vdots$
2	A	
3	C	
3	C	
4	C	
4	C	
5	A	
5	B	
6	B	
6	B	
7	B	$\vdots$
7	B	$y_{7,2}$

- Having estimated $\boldsymbol{\psi}$, one can go back and estimate $\boldsymbol{\theta}$, because

$$\begin{aligned}
\bar{y}_{1.} &= \theta_1 + \psi_A \\
\bar{y}_{2.} &= \theta_2 + \frac{\psi_A + \psi_C}{2} \\
\bar{y}_{3.} &= \theta_3 + \psi_C \\
\bar{y}_{4.} &= \theta_4 + \psi_C \\
\bar{y}_{5.} &= \theta_5 + \frac{\psi_A + \psi_B}{2} \\
\bar{y}_{6.} &= \theta_6 + \psi_B \\
\bar{y}_{7.} &= \theta_7 + \psi_B
\end{aligned}$$

(Again a normalisation is required: $\psi_A = 0$ or $\psi_A + \psi_B + \psi_C = 0$)

- For each firm *j* the firm average person effect $\bar{\theta}_j$ is the duration-weighted average of the θ_i for all individuals who work in that firm
- Abowd *et al* refer to ψ_j as measuring high or low wage firms, while $\bar{\theta}_j$ measures high or low wage workers
- Symmetrically, we also have an effect $\bar{\psi}_i$ which is the average firm quality for individual *i* over their employment history

- Suppose we did not have linked data
- From a panel of individual workers we could estimate

$$\mathbf{y} = \mathbf{D}\boldsymbol{\theta} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- From a panel of firms we could estimate (individual data is averaged to firm level)

$$\mathbf{y} = \mathbf{F}\boldsymbol{\psi} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- The bias for individuals is

$$\tilde{\theta}_i - \theta_i = \sum_{t=1}^{T} \frac{\psi_j}{T_i}$$

  This is just an employment-weighted average of the firm effects for every firm that individual i worked for
- For example, the bias for $i = 2$ in our example is just

$$\frac{\psi_C + \psi_A}{2}$$

  because that individual works for one period for A and for one period for C

- Similarly, the bias for firms would be

$$\tilde{\psi}_j - \psi_j = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{\theta_i}{N_j}$$

  which is the employment-duration weighted average of the person effects θ_i

- So for example the bias from ignoring θ for firm A is

$$\frac{2\theta_1 + \theta_2 + \theta_5}{4}$$

  because firm A employs individual 1 for two periods and individuals 2 and 5 for one period each

- In addition, we would also get biased estimates of $\boldsymbol{\beta}$ if there is any correlation between $\mathbf{X}$ and the omitted individual dummies $\mathbf{D}$ or the omitted firm dummies $\mathbf{F}$

What about match effects?

- Woodcock (2008) argues that match quality is important
- Can be thought of as the degree of complementarity between workers' and firms' productivities
- The good news is that one can estimate $\boldsymbol{\beta}$ consistently in the presence of match effects by using "within match" deviations or differences
- But it is much more difficult to subsequently recover the estimates of θ_i, ψ_j and the match effect without making additional assumptions

Links to trade models and measures of globalisation

- Reduced form wage equations
  - Do multinationals or exporters "really" pay higher wages?
  - Is there still a wage premium after controlling for firm and worker fixed effects?
- The reaction of wages to changes in trade costs, separated by worker type (θ_i)