A Note of Professor Stone's Comment on Dustmann and Frattini
by Bob Rowthorn

3 January 2014

In the following I shall ignore social housing and use the term "benefits" to include tax credits. Dustmann and Frattini (henceforth D&F) start from the following basic equation for the probability that a random individual in quarter q will receive benefits:

$$D = \beta I + \tau_q \tag{1}$$

where I indicates migration status ($= 0$ for a native and $= 1$ for an immigrant). According to D & F the coefficient β indicates a difference in the probability of receiving benefits between migrants and natives observed at the same moment in time. The above equation assumes that this difference in probability remains constant through time.

The observed differential probability that a migrant will receive benefits at time q is given by

$$\Delta_q = \frac{a_q}{a_q + c_q} - \frac{b_q}{b_q + d_q}$$

where $a_q(b_q)$ is the number of migrants (natives) who receive benefits, and $c_q(d_q)$is the number of migrants (natives) who do not receive benefits. Stone points out that the OLS estimator of β is given as follows

$$\beta^* = w_1\Delta_1 + ... + w_Q\Delta_Q \tag{2}$$

where $Q = 44$ is the total number of quarters and the weights w_q sum to unity and

$$w_q \propto \frac{(a_q + c_q)(b_q + d_q)}{n_q} \tag{3}$$

where

$$n_q = a_q + c_q + b_q + d_q$$

Equation (2) shows how the OLS estimator is the weighted sum of the differential probabilities in each quarter. The weights can be expressed in a simple form if we make the following assumptions:

$$\begin{aligned} b_q + d_q &= N \\ a_q + c_q &= \mu q N \end{aligned}$$

The first assumption implies that the number of natives in the sample remains constant through time. The second assumption implies the stock of migrants in this sample increases by an amount μN each year. These assumptions imply that

$$n_q = (1 + \mu q)N$$

Substituting in (3) yields

$$w_q \propto \frac{\mu q N}{(1+\mu q)N} \tag{4}$$

If μq is small we can approximate the above proportionality formula as follows

$$w_q \propto \mu q \tag{5}$$

Note that in the D&F the ratio of post-2001 EEA (non-EEA) migrants to the stock of natives in the final year 2011 is approximately 0.03 (0.04).

Since the weights sum to unity it follows that

$$\begin{aligned} w_q &= \frac{\mu q}{\sum_1^Q \mu q} \\ &= \frac{2q}{Q(Q+1)} \end{aligned} \tag{6}$$

0.1 A linear example

It is reasonable to assume that the longer migrants remain in the country, the greater the amount of benefits they will receive, in part because they have children. As an example, assume that the differential probability of receiving benefits for a migrant who has been in the country for j quarters is given by

$$f_j = r + sj \tag{7}$$

This assumes, contrary to D&F, that the differential probability changes through time. D&F assume that the probability differential for migrants remains constant through time (subject only to random fluctuations). In the present example this is equivalent to assuming that $s = 0$.In quarter q there are μN immigrants of vintage j. Hence, the observed average probability for migrants present in the country in quarter q is given by

$$\begin{aligned} \Delta_q &= \frac{\sum_{j=1}^{j=q} \mu N f_j}{\sum_{j=1}^{j=q} \mu N} \qquad (8) \\ &= \frac{\sum_{j=1}^{j=q}(r+sj)}{q} \\ &= r + \frac{s\sum_{j=1}^{j=q} j}{q} \\ &= r + \frac{s(q+1)}{2} \\ &= r + \frac{s}{2} + \frac{sq}{2} \end{aligned}$$

Substituting in (2) yields

$$\begin{aligned}
\beta^* &= \sum_1^Q w_q \Delta_q \\
&= \sum_1^Q \left(\frac{2q}{Q(Q+1)}\right)\left(r + \frac{s}{2} + \frac{sq}{2}\right) \\
&= r + \frac{s}{2} + \frac{s}{Q(Q+1)} \sum_1^Q q^2 \\
&= r + \frac{s}{2} + \frac{s}{Q(Q+1)} \frac{Q(Q+1)(2Q+1)}{6} \\
&= r + \frac{s}{2} + \frac{(2Q+1)s}{6} \\
&= r + \frac{(2Q+4)s}{6} \\
&= r + \frac{(Q+2)s}{3} \qquad (9)
\end{aligned}$$

Note that

$$\Delta_1 = r + s \qquad (10)$$

$$\Delta_Q = r + s\left(\frac{Q+1}{2}\right) \qquad (11)$$

Thus,

$$\beta^* = \frac{1}{3}\Delta_1 + \frac{2}{3}\Delta_Q \qquad (12)$$

The above formula illustrates clearly how the estimator of β^* weights later observations more heavily than earlier ones.

0.1.1 Numerical example.

Suppose that $\Delta_1 = -0.28$ and $\Delta_Q = -0.16$. Then, $\beta^* = -0.20$, which is roughly what D&F get for post-2000 migrants in the equations without controls. Equations (10) and (11) imply that

$$\begin{aligned}
s &= \frac{2*(\Delta_Q - \Delta_1)}{Q-1} = \frac{0.24}{43} = +0.0055814 \\
r &= \Delta_1 - s = -0.28 - 0.0055814 = -0.2855814
\end{aligned}$$

If $\Delta_1 = -0.28$ and $\Delta_Q = -0.16$ the differential probability for a migrant who has been in the country for j quarters is equal to $-0.2855814 + 0.0055814j$. For a migrant who has been in the country for the entire observation period $j = 44$

quarters, the differential probability is equal to $-0.2855814 + 0.0055814 \times 44 = -0.04$.

Thus, if $\Delta_1 = -0.28$ and $\Delta_Q = -0.16$ the estimated coefficient is $\beta^* - 0.20$. However, by the last quarter of the observation period, the long-established migrants in this example have a probability differential equal to -0.04.Thus, long-established migrants and natives have a very similar propensity to claim benefits. By assuming that the differential probability for migrants remains constant through time, D&F obtain a large negative OLS coefficient. This result disguises what may be the virtual disappearance of the migrant differential for a given cohort after eleven years presence in the country. The virtual dispearance of the migrant differential after eleven years is extremely important when considering the future evolution of the net fiscal contribution of migrants.