Job Creation Supplement

Bob Butcher, NIESR and BIS; and Matt Bursnall, BIS

A. Data Annex, 2 August 2013
B. Other supplementary items may be added in due course.

A. Data Annex for NIESR article: "How dynamic is the private sector? Job creation and insights from workplace-level data", by Bob Butcher and Matt Bursnall, published August 2013

1. Labour Force Survey

LFS data were obtained from the ESRC Data Archive[1]. We validated our analysis including use of weights, by replicating the Labour Market Statistics Bulletins for non-seasonally adjusted figures (eg table 8 in the Bulletins).

Other than the primary classification of economic activity, the main set of variables used have been the economic status one year ago (OYCIRC and related variables), asked in Q2 each year (April to June). Many of the variables started in 2006, the first year reported in table 3 in the article. We use these variables to measure churn over time, as they reveal not only whether someone moved from unemployment or inactivity to work or vice versa, but also whether they changed job in the year. The LFS also provides longitudinal data, which measures directly the economic activity of individuals in each of 5 consecutive quarters, and hence their position 12 months apart, but that does not reveal whether people moved from one job to another, only whether they moved from employment to non-employment or vice versa.

Issues. The main issues in using the one-year-ago variables are a) the relatively large volume of 'DNAs' (people coded as Do Not Apply), and b) the reliance on respondent memory over a 12 month period. The DNA codes arise when data are brought forward from the previous quarter, ie people who have been interviewed in the previous survey wave but were not contactable in the wave that falls in Q2. Analysis shows that these DNA codes are evenly spread across most key variables - region, sector, economic activity. We have used the percentages derived ignoring the DNAs, which assumes the missing values are distributed proportionately. There is variation across age, with younger people more likely to be classed as DNA, presumably because young people tend to have higher non-contact rates. On memory, labour market changes are likely to be significant enough to be remembered but precise timing may be mistaken, with an average tendency to report events as more recent than they were, leading to a possible overstatement of total churn.

2. Business Structure Database

[1] Office for National Statistics. Social Survey Division and Northern Ireland Statistics and Research Agency. Central Survey Unit, *Quarterly Labour Force Survey, April - June, 2012* [computer file]. Colchester, Essex: UK Data Archive [distributor], September 2012. SN: 7108 , http://dx.doi.org/10.5255/UKDA-SN-7108-1

The Business Structure Database[2] (BSD) is the source for most of the analysis in this paper. As well as the key points on structure and timing in the main text, the following paragraphs describe the coverage and the approach we take to using the BSD to derive estimates of job creation and destruction.

Coverage. BSD is an extract from the Interdepartmental Business Register (IDBR) taken each year in March. IDBR is constructed from firms that are registered for Value Added Tax or for PAYE tax. BSD includes about 2 million firms. An annual publication by the Department for Business Innovation and Skills[3] estimates how many extra businesses there are if one includes the small ones that do not have employees and have a low level of trading below the threshold for VAT. While there is a large number of such businesses, they contribute at most a few percentage points of total UK turnover. This paper focuses solely on the businesses on the BSD.

As noted in the main article, the date given to the BSD data is a little misleading, since BSD 2012 primarily refers to the situation and activity in 2011, and our tables show the actual reference year.

Measuring job creation and destruction.

Job creation occurs when a workplace expands or newly enters the economy, and destruction when a workplace contracts or closes down and exits. An entry is a workplace (or firm) that was present on the BSD at year t but not at year t-1 and an exit is a workplace that was present in year t-1 but not in year t[4]. Expanding and contracting workplaces and firms are identified from amongst continuing firms by comparing the number of employees between the two years.

Net job creation for workplace i in year t is

$$NJ_{it} = J_{it} - J_{it-1}$$

where J_{it} is the number of jobs in workplace i in year t (ie employment), and is defined as zero where the workplace does not exist in year t (eg for entering workplaces, $J_{it-1} = 0$).

We define job creation for the firm, C_{it} , as equal to NJ_{it} where it is positive, and zero otherwise. Job destruction similarly. This can be written

$$C_{it} = \tfrac{1}{2}.(NJ_{it} + |NJ_{it}|) \text{ , a non-negative number, and}$$

$$D_{it} = \tfrac{1}{2}.(NJ_{it} - |NJ_{it}|) \text{ .}$$

---

[2] Office for National Statistics, *Business Structure Database, 1997-2011: Secure Access* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], October 2012. SN: 6697.

[3] Business Population Estimates statistical release, BIS. Latest publication 17 October 2012.

[4] Note that some studies use different time horizons, eg 5 years, and define entry as firms present at the second time point but not at the first.

Aggregate expansion is $Exp_t = \sum_{i \in G} C_{it}$, where G is the set of ongoing workplaces

Aggregate entry is estimated by

$Ent_t = \sum_{i \in N} C_{it}$, where N is the set of workplaces that enter in year t.

Aggregate contraction, $Cont_t$, and exit, Ex_t, are written similarly.

Gross job creation, C_t , and job destruction D_t, are

$$C_t = Exp_t + Ent_t = \sum_{i \in G} C_{it} + \sum_{i \in N} C_{it}$$, the sum of entry and expansion, and

$$D_t = Cont_t + Ex_t = \sum_{i \in G} D_{it} + \sum_{i \in X} D_{it}$$.

where X is the set of exiting workplaces

The *rate* of job creation for year t is calculated as $CRt = Ct / Jt\text{-}1$, so a job creation rate of say 18% in 2007 means that 18 new jobs are created in 2007 for every 100 jobs that existed in 2006.

Aggregate net job creation $NJ_t = C_t + D_t$.; aggregate *churn* is $Ch_t = |C_t| + |D_t|$

Workplace vs firm

Gross job creation derived at workplace level is a larger number than at firm level. This is because job creation in some firms is netted off by there also being job *destruction* within the same firms. For example there might be one workplace expanding or entering and another workplace within the same firm that contracts or exits. If we use the superscript F for a firm level figure and W for workplace, we have net job creation for firm s being

$NJ^F_s = \sum C^W_{sv} . + \sum D^W_{sv}$, summed over workplaces v within firm s.

Note that $NJ^F_s \leq \sum C^W_{sv}$ since the D's are always negative or zero

The estimates of job creation and destruction for a firm s are equal to

$C^F_s = ½.(NJ^F_s + | NJ^F_s |)$, and $D^F_s = ½.(NJ^F_s - | NJ^F_s|)$.

Gross job creation derived from workplace level is the sum over firms s of $\sum C^W_{sv}$ while job creation at firm level is the sum over s of C^F_s ., which in general will be smaller than the workplace level figure. Table 6 in the paper shows there are 4.0m

jobs created per year in 2004-07 when measured at workplace level, and 3.1m when measured at firm level.

Definitions

Size. We define the size of a firm at enterprise group level as the number in employment in the base year, and for a new entrant group the number in employment in its first year.

Alternative time points are used for defining size in the literature. Options include defining size equal to the average of the first and second time point, with zero given to entrants at time 1. The enables the rate of increase to be defined for entrants. Also considered is defining size at the second time point. There is no uniformly preferred option. The approach we adopt is simple, and easy to explain. Whatever definition is used, care is needed in interpreting the volume of job creation by size.

Age. We define the age at workplace level by the number of years since it first appeared in the BSD. New entrant workplaces have age zero.

Robust existence of workplaces over time

We had been concerned that some job creation or destruction may be over-stated if workplace reference numbers disappeared or were created following acquisition or disposal of the workplace by another firm while the same jobs with the same incumbents continue to exist simply with different workplace references. We were reassured on this point following discussion with ONS and with Dr Michael Anyadike-Danes of Aston University, and we now judge that in general workplace references will continue to exist through change of ownership. Based on his study investigating a cohort of workplaces born in 1998 and following them to 2008, and based on his long experience of working with the ONS business data and the collection processes, Dr Anyadike-Danes observes that workplaces with employees do not seem to disappear when acquired. ONS report that in the past there have been occasions when a firm might change the workplace grouping of employees one year compared to the last one, but in recent years (at least the last five), they resist this kind of change and it is rarely if ever known to occur.

Nevertheless, there will be times when a plant or outlet moves from one location to a nearby one, for example to expand into new premises. In such cases many of the existing jobs and people might transfer, and estimates of total job creation will include them, even though in many respects such jobs would not generally be regarded as job creation in the same way as completely new jobs with people doing different work from their previous occupation. It would be helpful to have an enquiry to estimate the magnitude of different types of job creation in this sense.

.

Missing data - analysis starts at 2004.

In looking at time trends in the workplace (Local Unit) version of the BSD, we found a problem with BSD2003. There appears to be a shortfall in employment volume of about 3m and in the volume of workplaces of about 300,000. This shortfall aligns

with a problem that has been identified by Dr Michael Anyadike-Danes in the cohort of workplaces he has studied over time. He has found a significant proportion disappearing in about 2002 or 2003 and then re-appearing with different workplace reference numbers. For this reason we start our BSD analyses with BSD2004, and hence the first year of job creation that we measure is BSD2005 (which refers to calendar year 2004 in terms of activity).

Private sector and self-employment

Private sector employment in official statistics is derived by subtracting the public sector total that is obtained direct from public sector employers, from the LFS figure for total employment. It is possible to obtain very similar figures just from the LFS data using the definition of public sector employment that the person is an employee and works for one of central government, local government (including police, fire, state schools), the NHS or armed forces[5]. The estimates are about 300k lower than the published (seasonally-adjusted) estimates for June each year. The trend is very similar with no clear pattern for the slight fluctuation from year to year. There are several reasons for the official public sector estimates to be smaller than those reported in the LFS. One possibility is that there is a higher response rate from public sector employees for which the LFS weighting is not fully able to correct.

On the BSD, we define private sector by removing items described as central government or local authority[6]. This provides the series in table 1 of the main article (6th column of numbers) which is an average of 300,000 below the LFS column over the 2004-11 period.

The overall levels and annual change figures (columns 3-4 and 6-7) show sufficient similarity to suggest they are measuring essentially the same phenomenon.

There could be a misalignment in 2009-10, when the LFS shows -0.8 change in 2009, and 0.0 in 2010, while BSD shows -0.5 in each year. Part of the explanation could be the time reference point to which each dataset refers. The LFS data is for quarter 2 each year, centred on May. Much of the BSD employment data refers to September, though some will be more historic.

To what extent does self-employment explain the differences in levels, trends and cyclical fluctuation in the private sector volumes from the different sources?

Self-employment explains some of the difference in level between the LFS and BSD, and accounts for a portion of the trend increase in private sector employment, but appears to have little effect on the fluctuation over time. In brief the position is as follows. From LFS, the volume of self-employment has been steadily increasing from 3.6m in 2004 to 4.2m in 2012 at about 70,000 per year, with perhaps an acceleration in 2012, after the main period of our analysis.

As with the LFS, the BSD includes self-employment where relevant, whether they are sole traders or have employees. However, the BSD only includes businesses that are

[5] Variable SECTRO03 from the LFS.
[6] Using the STATUS variable. Other codes include company, partnership, public corporation, sole proprietor and non-profit making body.

registered for VAT or PAYE or with companies house, so it will omit a proportion of sole traders with low turnover. In addition, for our analysis we exclude firms that have no employees (ie that only have a working proprietor or no employment at all) as one person can own several firms and the category of those with no employees is definitely incomplete in the BSD. Thus there is a greater volume of self-employment in the LFS than the BSD, which will therefore explain some of the gap between the two.

On broad trends, self-employment increases as a proportion of private sector employment gradually over the period, from 16% to 18% according to LFS. Thus self-employment makes its proportionate contribution to the overall trend in net job creation, but including or excluding self-employment would make little difference in the fluctuation over time.

Hidden job creation

The estimates of creation and destruction from the BSD and similar datasets in this country and other countries only measure job creation at the workplace or firm level. Some of these workplaces are large, even several thousand employees. Whether small or large there will be job creation in some workplaces that is 'hidden' by counterbalancing job destruction. Firms create new posts and teams all the time and will close others. While the net position will reveal the creation of some jobs the gross position will be larger in some cases. The literature notes this (for example Hijzen et al 2011) and considers the current measures to provide useful insights, even if only partial.

We provide an example of the issue in the paper by showing that a million jobs a year are created at the workplace level which are hidden if the analysis is undertaken at firm level. It is likely that observed job creation would be larger still if we could measure within workplaces.

This measurement deficiency could lead to a misinterpretation of the data, in particular it is likely that the current estimates under-state the job creation by large workplaces, leading to a greater emphasis on the role of small workplaces and firms than is justified.

A fuller assessment would need data collection from firms on the extent to which they are creating jobs, not simply on their total number of job in one year compared to the next, which would seem feasible at least as a one off exercise.

Unit of analysis. The literature varies in which unit has been used for analysis. ONS report that identifiers and reference numbers for all three levels are robust (workplace, firm and enterprise group), at least for recent years - and that they are reliable for longitudinal analysis. The smallest units, the workplaces, provide a greater accuracy in revealing job creation and loss than higher levels where there can be *hidden* job creation (for example an expanding workplace in a firm that has another workplace that is contracting), and also *over-stated* job creation (for example firm expansion due to the takeover of an existing workplace). For this reason we use workplace as our unit of analysis. We are still interested in the higher levels though, as the literature

that only considers lower levels has a tendency to imply that the churn is all to do with market forces, and is external to firms, when a considerable part is in fact within the firm, and therefore subject to rather different forces. We find roughly half of job creation which arises from the entry of new workplaces, is accounted for by existing firms setting up new workplaces within themselves. We are not aware of other analyses of administrative business data that have used the levels in this combined way, and even using survey data it is rare, one very good example being Bryson and Nurmi 2011..

Bob Butcher
Matt Bursnall
19 July 2013