Cranfield
UNIVERSITY

School of Management

RP 5/07

IS THERE A NORTH-SOUTH DIVIDE IN
SELF-EMPLOYMENT IN ENGLAND?

© Andrew E. Burke, Felix R. FitzRoy and Michael A. Nolan

Research Paper Series

The Cranfield forum for the latest thinking in management research

Available online at www.som.cranfield.ac.uk/som/research/researchpapers

Research Paper no. 5/07

IS THERE A NORTH-SOUTH DIVIDE IN SELF-EMPLOYMENT IN ENGLAND?

Andrew E. Burke [a] [*], Felix R. FitzRoy [b] and Michael A. Nolan [c]

[a] *School of Management, Cranfield University, Cranfield, Bedford MK43 0AL, United Kingdom and Max Planck Institute of Economics, Germany.*
[b] *School of Economics & Finance, Castlecliffe, The Scores, University of St. Andrews, Fife KY16 9AL, United Kingdom and IZA (Bonn), Germany.*
[c] *Hull University Business School, Hull HU6 7RX, United Kingdom.*

July 2007

The Cranfield School of Management Research Paper Series was established in 1987. It is devoted to the dissemination of new insights which advance the theory and practice of business and the management of organisations.

The Research Papers in this Series often represent preliminary work. They are circulated to encourage cross-fertilisation of ideas both between Cranfield scholars and across the wider academic community.

Editor: Catarina Figueira

---

[*] Corresponding author. Tel. +44 1234 754372; Fax + 44 1234 752136
E-mail. Andrew.Burke@cranfield.ac.uk

IS THERE A NORTH-SOUTH DIVIDE IN
SELF-EMPLOYMENT IN ENGLAND? *

Abstract

Using decomposition analysis, the paper investigates why Northern England has fewer but higher performing self-employed individuals than the South. We find the causes are mainly structural differences rather than regional variation in individual characteristics. There are more self employed individuals in the South, but on average they create fewer jobs. Post compulsory education has a strong negative effect on the probability of self employment in the South, probably due to better employment opportunities there, but little influence in the North. Education has some positive effects on job creation by entrepreneurs in both regions. Aggregate studies may thus give misleading results.

Keywords: Self-employment, job creation, North-South divide, decomposition.

1. Introduction

The regional dimension of the British economy has been well documented. Regional variation in economic performance is revealed in a North-South divide where the South has better economic performance than the North with lower unemployment and higher GDP per capita. This persistent divergence in performance has generated various economic policy responses (Lewis and Townsend (1989), Fothergill (2001) and Gudgin (1996)). In particular, a response to divergence in unemployment (Gray (2004)) or non-employment (Anyadike-Danes (2004)) has been to encourage job creation through self-employment (Shutt and Sutherland (2003)). Robson (1998), and Georgellis and Wall (2000) include a theoretical framework to underlie a relationship between the regional self-employment rate and a variety of characteristics of the regional economy, and estimate results for UK regional data over a short time series. Their models are based on Evans and Jovanovic (1989) and Khilstrom and Laffont's (1979) models of choice between wage work and self-employment. The probability of a person choosing self-employment is positively related to both the relative financial and non-pecuniary benefits compared to wage work, so econometric estimation requires variables affecting these components

To our knowledge, there has not been UK regional analysis at the level of the individual, which is one new contribution of this paper. We thus study regional variation in individual characteristics affecting the *ability* and *predisposition* for self-employment. These include skill, experience, education, psychological attitudes and culture. Variation in regional self-employment rates and job creation can therefore be attributed to both the structural disparities (well documented in research on the determinants of aggregate rates of self-employment) and differences in average individual characteristics which are difficult to measure and estimate at an aggregate level.

In this paper we thus include compositional as well as structural influences. With longitudinal data on individuals, we investigate both the probability of self-employment, and also job creation by the self-employed to explore the North-South divide in English self-employment. To the best of our knowledge, this approach is new.

Neglecting these issues, previous work has not identified the impact of certain individual characteristics on self-employment – if, for example, a significant effect in one region is cancelled by an opposing effect in another. Of particular interest to us is the role of post-secondary education which at a national GB level has been found to reduce the number of self-employed but increase job creation by entrepreneurs. Burke, FitzRoy and Nolan (henceforth BFN, 2000) and Cowling et al. (2004) find that the net (aggregate) result of these opposing effects is positive so that education increases the total number of jobs created by the self-employed. Here we disaggregate this result and uncover substantial regional variation. We use data from the National Child Development Study (NCDS) – males and females being considered separately, as was justified previously by BFN (2002), following the 'self-

employed female underperformance hypothesis' of Rosa et al (1996), and Du Rietz and Henrekson (2000).

In addition, we use decomposition analysis, to investigate how variation in self-employment between North and South results from differences in individual characteristics or from differential responses to given characteristics. Our work follows Reimers (1983), Cotton (1988), Neumark (1988), Oaxaca and Ransom (1994), and Fairlie (1999, 2003) – who extended the seminal analysis of decomposition by Blinder (1973) and Oaxaca (1973).

Recent international research on new firm formation shows how it effects long term job creation (Audretsch and Fritsch, 2002, Fritsch and Mueller, 2004 and van Stel and Storey 2004). Fristch and Mueller (2004) claim that this effect evolves across three regimes. Initially, business start-ups have a direct positive effect on job creation which then turns negative as some of these firms grow and compete with incumbents. Later, a third inclusive phase occurs where these ventures have positive spillover effects on other businesses. Fritsch and Mueller's evidence is mainly based on German regional data (although their results have been replicated in other economies, see Fritsch and Schmunde (eds) 2006) and they argue that the positive effects (particularly, the direct effect) are more pronounced in high productivity regions. Mueller, Van Stel and Storey (2006), hereafter MVS (2006), find that in contrast the UK economy does not follow this productivity pattern and show that the direct effect of job creation in the UK is higher in the less productive Northern English regions than in the South.

All of these studies use aggregate data, and our use of individual level data tests whether the UK is indeed an exception to the three regime regional model. Our data is a cross section snapshot of the self-employed and hence only provides a total net effect of self-employment on job creation in the self-employed sector. However, while this tempers the generality of how our findings relate to the displacement and spillover effects, it does not affect the measurement of direct positive effects – a key area where the UK evidence from MVS (2006) diverges from regional effects found in other countries.

Individual level data can also test the impact of education on labour productivity. In general, higher regional levels of education are associated with higher levels of regional productivity (see Lindsay, 2004 for UK evidence). Thus, one might expect that on average the self-employed in more highly educated regions will be more educated than those of less educated regions and therefore have higher productivity levels. While this is often true, and indeed explains the most common patterns observed in the aggregate analyses, our underlying model of self-employment choice (Evans and Jovanovic, 1989) can also generate the opposite effect. Namely, if qualified job opportunities are plentiful then better educated individuals may be drawn into wage work, resulting in lower education and productivity among the self-employed in a region with relatively high average education. By contrast, a less educated region may have higher levels of education among the self-employed if job opportunities are scarce and more of the educated are pushed into self-employment. If this is true and applies to the UK then it could explain why MVS (2006) find that the lower productivity Northern regions have higher positive direct employment effects than the South. Far from undermining

2

the three regime model, the English regions would then be consistent with it. However, the predictive power of the model in relation to productivity and job creation needs to include (dynamic) individual career choice following Evans and Jovanovic (1989), Blanchflower and Oswald (1998) and BFN (2000, 2002). Our use of individual level data allows us to test this hypothesis and we find that it appears to hold.

To summarise, the novelty of our analytic approach is threefold. First, we complement aggregate regional analysis of self-employment in the UK by estimating, at the level of the individual, the probability of choosing self-employment and the determinants of job creation by the self-employed. Secondly, we introduce the regional dimension to existing studies of self-employment choice and performance at the level of the individual. Finally, we use decomposition analysis in order to separate out the impact of compositional from structural regional variation in determining entrepreneurial choice.

The paper proceeds with Section 2, which focuses on the data. Section 3 discusses methodology, concentrating on application of decomposition to this topic. Section 4 presents results. It focuses initially on analysis of the probability of self-employment, and decomposing differences between South and North into compositional and structural parts, and then proceeds to male self-employment job creation. There then follows a concluding section.

2. Data Description

The National Child Development Study (NCDS) surveys a cohort born in the week 3rd March 1958 to 9th March 1958 inclusive and living in Great Britain. These surveys (in 1958, 1965, 1969, 1974, 1981 and 1991) were used by Blanchflower and Oswald (1998) and BFN (2000, 2002). The first of these papers only considers the self-employment decision, while the second and third also analyse measures of entrepreneurial performance, and the third disaggregates by gender.

The precise extent of self-employment indicated by the fifth sweep NCDS data from 1991 (NCDS5 hereafter) depends upon the exact definition that is chosen. Blanchflower and Oswald (1998) choose to define as self-employed the 1,279 (out of 11,369) individuals who indicate self-employment to be their main economic activity – some of whom are only part-time self-employed (but they exclude those part-time self-employed whose main economic activity category is not self-employment). We follow the broader definition used by BFN (2000, 2002) – including some people for whom part-time self-employment is not their main economic activity – by adding those with self-employment as their main economic activity to those reporting a self-employment income whose economic activity category is specified, and is not 'sick or disabled'. This yields a total of 1,558 self-employed. The maximum sample size, for our analysis of the probability of choosing self-employment rather than being solely in some other form of economic activity, is 11,113. We include the unemployed and, for

example, housewives – many of whom are often considered to be economically inactive. In principle, these individuals could make a decision to become self-employed – provided appropriate incentives are offered so that self-employment provides them with greater utility than any feasible alternative. This indicates a self-employment probability of 0.140 across males and females. The self-employment proportions given for 1991 in Blanchflower and Oswald (1998), calculated on quite a different basis, are very similar at 14.2% of employment within the NCDS cohort, and 15% across Great Britain.

The 11,113 individuals in our sample include 5,432 males and 5,681 females. The self-employment probability for males is 0.195, while that for females is only 0.088. However, this puts the share of self-employment for women at about 32.0% – which is above the 24.8% indicated in OECD (2000) for the whole UK across 1990-96; while the same source indicates similar female shares of self-employment across 1990-97 for Italy (23.4%), Sweden (25.7%), France (26.0%) and Germany (28.3%), but a larger share in the USA (37.0%). Parker and Robson (2004) report a male self-employment rate of 17.7% of the workforce for the UK in 1990, and 7.4% for females. Corresponding figures for other countries include 12.7% and 4.7% for Sweden; and 10.4% and 6.2% for the USA (indicating that the greater female share of self-employment there exists in the context of a rather low overall rate of self-employment).

Of course, national self-employment rates mask significant regional variation. Table 1 (below) shows in more detail how the self-employment probability differs by region in the NCDS5 data. Although our paper does not discuss how the UK North-South divide has developed over time, Georgellis and Wall (2000) show that while self-employment has changed across time, regional variation has not fluctuated much. This poses a particular challenge for policy makers because altering regional differences in self-employment is perceived as a key component of advancing regional economic development. Thus, understanding the causes of these differences in self-employment is a key input to policy aimed at alleviating the North-South divide in England. The last four rows of the table calculate self-employment probabilities for composite regions comprising at least two postcode-based Standard Regions (SRs). This paper focuses particularly on Southern England (a combination of Greater London, South East England and South West England) and Northern England (a combination of Yorkshire & The Humber, North West England and the North of England SR). The table shows clearly that, for the NCDS5 data, there is a noticeably higher self-employment rate in Southern England than in Northern England – 23% against 17% among males, and 11% against 8% among females. It is also clear – as expected – that self-employment rate is substantially higher for males than females (about 19½% against 9%).

Table 1: The self-employment probability by region – males and females separately.

	MALES			FEMALES		
Region	*Total*	*S/E*	*Prob*	*Total*	*S/E*	*Prob*
Greater London	358	72	0.201	397	53	0.134
South East England	1302	274	0.210	1351	140	0.104
South West England	464	133	0.287	499	50	0.100
East Anglia	182	32	0.176	219	14	0.064
East Midlands	295	55	0.186	301	28	0.093
West Midlands	512	101	0.197	501	41	0.082
Wales	417	82	0.197	352	29	0.082
Yorkshire & The Humber	555	101	0.182	582	39	0.067
North West England	581	107	0.184	628	53	0.084
North of England (NCDS)	285	34	0.119	287	23	0.080
Scotland	462	64	0.139	529	25	0.047
Unknown	*19*	*4*	*0.211*	*35*	*4*	*0.114*
GREAT BRITAIN	5432	1059	0.195	5681	499	0.088
Southern England	2124	479	0.226	2247	243	0.108
Central England	989	188	0.190	1021	83	0.081
Northern England	1421	242	0.170	1497	115	0.077
Wales & Scotland	879	146	0.166	881	54	0.061

The NCDS5 data does not offer detailed information about the industry and occupation categories of the self-employed. However, examination of Labour Force Survey (LFS) data for 1991 demonstrates only modest differences in the industry and occupation distributions of self-employment between Northern England and Southern England. Since the LFS is sample-based, and covers the entire age range, it is far from straightforward to perform a reliable and useful projection onto NCDS cohort data: improved comparability comes at the expense of higher relative sampling variation.

Turning from self-employment rates to performance, we use a measure of employment by the self-employed which is also provided by NCDS5 – where each self-employed cohort member indicates how many employees he/she has. Some summary statistics are shown in Table 2 below. Table 2 deals with 1526 self-employed individuals that report a value (nil in the majority of cases) for job creation. The higher job creation rate by the self-employed in Northern England is an interesting feature to emerge from the table. This is true both for males – 3.529 jobs per self-employed individual on average, compared to 2.652 jobs for Southern England – and for females (3.477 jobs per self-employed individual, versus 3.079). However, while average male job creation in all northern SRs is above that for any southern

SR, the highest female job creation average is for Greater London (7.820), well ahead of the highest northern SR (North West England, at 4.538). Another insight offered by Table 2 is that there is less of a job creation gap between self-employed men and self-employed women than there is a gap between the gender-specific self-employment probabilities. Indeed, there is almost no difference by gender in the job creation rate for the self-employed of Northern England.

We now turn to factors likely to determine self-employment choice and performance. These NCDS variables are used in our estimation of self-employment probability logits, and/or job creation tobits. They are motivated by previous papers on self-employment using this dataset, (Blanchflower and Oswald (1998) and BFN (2000, 2002)), which themselves use the self-employment versus wage work theoretical framework developed by Evans and Jovanovic (1989).   Thus, the variables are those likely to affect self-employment income or non pecuniary satisfaction relative to wage work. They include the following:

1. *Ability, education and training* – more able individuals are likely to secure higher income in both self-employment and wage work. Therefore, relative impact will determine ability's role in influencing self-employment. In terms of job creation by the self-employed, the more able are likely to be more successful – but there is an income effect, which may allow such individuals to try less hard and settle for a given performance level. We use dummies to indicate whether the highest academic qualification achieved is O level (or equivalent), A level, first degree or higher degree; up to four pairs of dummies capture performance in reading and maths tests at age seven (NCDS2) and age sixteen (NCDS3). For each test, a dummy is used to indicate a score definitively (not tied) in the top quintile of the cohort and another indicates a score in the bottom quintile – leaving the middle 60% (plus ties) of each ability distribution as the base case. A dummy captures apprenticeship by 1981; another denotes receipt of a vocational qualification by 1991.

2. *Non-cognitive attributes* – self-employment is often associated with unique psychological characteristics but empirical support for these propositions is limited (see Parker, 2004). Therefore, several psychological measures are included as discrete scores. Creativity comes from NCDS1 (1965) – a zero value denoting no creativity, and other values rescaled to a maximum of 0.4; while unforthcomingness, withdrawal, depression, anxiety acceptance and hostility towards (other) children are taken from NCDS2 (1969) – each with a zero minimum; and caution, flexibility, moodiness, timidity, sociability and laziness measures are derived from NCDS3 (1974) – varying in the range [–2,+2]. There is a dummy for fear of new situations (1974).. A number of dummies indicate what the cohort member regarded, in 1981 (NCDS4), as being most important when choosing a job. Included are promotion, being in charge, being one's own boss, lack of responsibility, job security and good pay. Cohort members responding with some other job characteristic form the base group. This set of dummies serves to capture the individual's primary motivation across pecuniary and non-pecuniary dimensions emphasised in BFN (2000).

3. *Family background* – included to capture the impact of role model, mentor and network effects on entrepreneurial choice and performance. A dummy reflects family financial difficulties (NCDS1); another denotes use of the English Language at home in 1969 (NCDS2); a series of dummies are used to indicate occupation of the cohort member's father in 1969 – including employee manager of small firm, employee manager of large firm, professional self-employed, professional employee, foreman (manual work), skilled manual, worker with own account (a type of self-employment), farmer employee-manager and farmer with own account; two grouped variables from NCDS3 indicate the age at which the cohort member's father and mother left full-time education; another grouped variable indicates, for the cohort member's 1974 school, the percentage of male parents in a non-manual job.

4. *Current family* – a dummy variable captures having no children by 1991; we also investigate the interaction of this dummy with higher level qualification (at least A level). Being childless may be relevant because people with children face extra obligations and, thus, time constraints. BFN (2002) found that both males, and less qualified females, with children are more likely to be self-employed (perhaps due to self-employment's potential flexibility in working time). That paper also found that highly qualified males with children, once self-employed, hired more workers than otherwise similar childless entrepreneurs. To capture exogenous finance, three variables are also constructed (NCDS5) to capture the size (linearly and quadratically[1]) and timing (year) of any inheritance – to capture liquidity constraints.

5. *Region* – similarly to BFN (2002), we construct four regions from the 11 SRs of Great Britain, of which two (Southern England and Northern England) are our focus here. Even within the composite regions, there may be some variation in costs (particularly housing) and demand conditions. In our logits of self-employment, we include two SR dummies – with South-West England SR being the base part of our Southern England composite region, and North of England SR being the base within our broader definition of Northern England. In the self-employment tobit equations, we include the average SR unemployment rate as a control, rather than pairs of SR dummies.

6. *Aspects of self-employment* – we include a control for the length in years by 1991 of a spell of self-employment ongoing at NCDS5. We also use a dummy for non-full-time self-employed.

7. *Missing value dummies* – for some individual regressors, and some groups of regressors, an extra dummy is used to indicate missing data, and as a (rather limited) control for this fact. This approach is quite common.

---

[1] These enter in present value (1991) form. The linear term is divided by 10000 (yielding a mean, across all cases with specified region and gender, of 0.5321) and the quadratic term is divided by 1.0×10^{10} (which gives a mean across all cases of 0.4996). The timing control indicates the year in which the inheritance was received (subtracting 1900 from the year in question, and then dividing by 100).

Table 2: Job creation by the self-employed by region – males and females separately.

	MALES			FEMALES		
Region	*Jobs*	*S/E*	*Mean*	*Jobs*	*S/E*	*Mean*
Greater London	164	67	2.448	391	50	7.820
South East England	762	265	2.875	265	139	1.906
South West England	302	131	2.305	80	50	1.600
East Anglia	119	32	3.719	23	13	1.769
East Midlands	315	55	5.727	61	27	2.259
West Midlands	305	100	3.050	60	40	1.500
Wales	179	82	2.183	155	29	5.345
Yorkshire & The Humber	398	101	3.941	109	37	2.946
North West England	349	106	3.292	236	52	4.538
North of England (NCDS)	100	33	3.030	41	22	1.864
Scotland	669	62	10.790	87	25	3.480
Unknown	*2*	*4*	*0.500*	*6*	*4*	*1.500*
GREAT BRITAIN	3664	1038	3.530	1514	488	3.102
Southern England	1228	463	2.652	736	239	3.079
Central England	739	187	3.952	144	80	1.800
Northern England	847	240	3.529	386	111	3.477
Wales & Scotland	848	144	5.889	242	54	4.481

3. Empirical methodology

The well-known logit model provides a straightforward method of estimating the individual's probability of self-employment. We perform basic decomposition analysis to determine whether North-South differences in self-employment are primarily a result of 'compositional' or 'structural' differences. As in previous literature, compositional differences reflect between-region differences (on average) in the characteristics of individuals; and structural differences refer to between-region differential responses to given characteristics.

The initial work on decomposition by Oaxaca (1973) and Blinder (1973) pertained to linear regressions of the logarithm of wages. However, since probit and logit specifications also typically involve a linear specification, Gomulka and Stern (1990) and Fairlie (1999, 2003) have noted that decomposition is feasible for these models. When appropriately transformed, estimated coefficients from probit and logit models are typically similar – which is unsurprising, given the shapes of the normal and logistic distributions that underlie the respective models. For our decompositions, a useful feature of the logit model is that the predicted probability of a given outcome is identical to the actual probability, not only for the

whole sample, but also for sub-samples where a given dummy variable takes a particular value.

The basic format of the decomposition is as follows:

$$\overline{Y}_S - \overline{Y}_N = \left[\sum_{i=1}^{n_S} \frac{F\left(\mathbf{X}_{iS}^{/}\hat{\boldsymbol{\beta}}_S\right)}{n_S} - \sum_{i=1}^{n_N} \frac{F\left(\mathbf{X}_{iN}^{/}\hat{\boldsymbol{\beta}}_S\right)}{n_N} \right] + \left[\sum_{i=1}^{n_N} \frac{F\left(\mathbf{X}_{iN}^{/}\hat{\boldsymbol{\beta}}_S\right)}{n_N} - \sum_{i=1}^{n_N} \frac{F\left(\mathbf{X}_{iN}^{/}\hat{\boldsymbol{\beta}}_N\right)}{n_N} \right] \qquad (1)$$

where $\overline{Y}$ indicates the mean of the dependent variable in a logit model, $F(.)$ is the Cumulative Distribution Function of the logistic distribution, the S and N subscripts indicate (throughout) Southern and Northern England respectively, n denotes the number of individuals in a particular area, $\mathbf{X}_i^{/}$ is an individual's vector of characteristics and $\hat{\boldsymbol{\beta}}$ is the vector of estimated coefficients from the logit model. The estimated coefficients vectors have the subscript S or N attached because they are generated through separate estimations of the logit model for those from Southern and Northern England. The version of the decomposition shown in equation (1) is split into a compositional effect (the term in the first square bracket) and a structural effect. The compositional term looks at the average predicted probability of self-employment that would be generated if individuals from Northern England responded according to the logit estimates for Southern England, relative to the average predicted probability of self-employment in Southern England. The structural term measures the influence on the self-employment probability of the difference between the response coefficients for Southern England and Northern England, together with the impact of unobserved regional differences in characteristics (which, by definition, cannot be captured within the logit estimation – and, in our case, include industry and occupation category).

The following expression for the decomposition is equivalent:

$$\overline{Y}_S - \overline{Y}_N = \left[\sum_{i=1}^{n_S} \frac{F\left(X_{iS}^{/}\hat{\boldsymbol{\beta}}_N\right)}{n_S} - \sum_{i=1}^{n_N} \frac{F\left(X_{iN}^{/}\hat{\boldsymbol{\beta}}_N\right)}{n_N} \right] + \left[\sum_{i=1}^{n_S} \frac{F\left(X_{iS}^{/}\hat{\boldsymbol{\beta}}_S\right)}{n_S} - \sum_{i=1}^{n_N} \frac{F\left(X_{iS}^{/}\hat{\boldsymbol{\beta}}_N\right)}{n_S} \right]. \qquad (2)$$

The interpretation of equation (2) is the opposite to equation (1), where now individuals from the buoyant South react to Northern coefficients.

Even and Macpherson (1990, 1993) noted that the decomposition component attributable to differences in characteristics can be split on a variable by variable basis (or for a group of variables, where this is more appropriate). The numerator for the required ratio is given by the size of the difference in sample means across the two groups (regions in our case) for the single variable, weighted by its estimated coefficient. The denominator is the difference in sample means across the two groups for *all* variables (each difference being weighted by the corresponding estimated coefficient). Hence the contribution to the probability gap by regressor r is as follows:

$$\left[\sum_{i=1}^{n_S}\frac{F\left(\boldsymbol{X}'_{iS}\hat{\boldsymbol{\beta}}_S\right)}{n_S}-\sum_{i=1}^{n_N}\frac{F\left(\boldsymbol{X}'_{iN}\hat{\boldsymbol{\beta}}_S\right)}{n_N}\right]\left[\frac{\left(\overline{X}_{rS}-\overline{X}_{rN}\right)\hat{\beta}_{rS}}{\left(\overline{X}'_S-\overline{X}'_N\right)\hat{\boldsymbol{\beta}}_S}\right]. \tag{3}$$

Before we proceed to our estimation results, we discuss the appropriate decomposition. In its applications to wages, the question of what constitutes the 'no discrimination' distribution is often discussed. In our case too, we might expect that norm coefficients should be something other than either those for the South (as in equation (1)) or those for the North (equation (2)). While Reimers (1983) suggested the use of the arithmetic mean of the two sets of coefficients, and Cotton (1988) proposed the use of a weighted mean (the relative weights being determined by relative sample sizes) it has become more usual to consider the coefficients resulting from pooled estimation across the groups under examination – as in Neumark (1988) and Oaxaca and Ransom (1994), where it is demonstrated that the coefficients from a pooled regression can be written as a weighted sum of the regression coefficients for the two groups. In the case of the logit model, the decomposition can be written as follows:

$$\left[\sum_{i=1}^{n_S}\frac{F\left(\mathbf{X}'_{iS}\hat{\boldsymbol{\beta}}_P\right)}{n_S}-\sum_{i=1}^{n_N}\frac{F\left(\mathbf{X}'_{iN}\hat{\boldsymbol{\beta}}_P\right)}{n_N}\right]+\left[\sum_{i=1}^{n_S}\frac{F\left(\mathbf{X}'_{iS}\hat{\boldsymbol{\beta}}_S\right)}{n_S}-\sum_{i=1}^{n_N}\frac{F\left(\mathbf{X}'_{iS}\hat{\boldsymbol{\beta}}_P\right)}{n_S}\right]$$
$$+\left[\sum_{i=1}^{n_S}\frac{F\left(\mathbf{X}'_{iN}\hat{\boldsymbol{\beta}}_P\right)}{n_N}-\sum_{i=1}^{n_N}\frac{F\left(\mathbf{X}'_{iN}\hat{\boldsymbol{\beta}}_N\right)}{n_N}\right], \tag{4}$$

where the *P* subscript refers to the pooled sample of Southern plus Northern England. The first term reflects the compositional effect (viewing the pooled coefficients as applying to all individuals). The second and third terms each reflect how the regional coefficients depart from the pooled norm.

Job creation by the self-employed is estimated using a Tobit model. This affords easier comparison with the previous work of BFN (2000, 2002) – and broadly similar statistical significance results to the negative binomial model for count data. Ordinary Least Squares (OLS) estimation is rejected on the basis that, while job creation might not be formally censored at zero, it seems highly unlikely that all cases of zero job creation by the self-employed (more than half of the group) reflect similar propensities to create jobs. If that suspicion is correct, OLS estimates are biased. An intuitively unappealing feature of the OLS fitted values for job creation is that some (in this case, 25-30%) are negative.

4. Estimation results

We use Limdep (Greene (2002)). Firstly, we consider the factors which influence the probability of an individual being self-employed and Table 3, below, shows logit maximum likelihood estimates for males – in Southern and Northern England separately. The regressors

in Table 3 are those remaining after a general-to-specific process based principally on at least some weak evidence of statistical significance in one or both regions[2]. A number of differences are apparent, beyond the higher mean self-employment in Southern England shown in Table 1. Some of these differences are shown by the simple means of the regressors in the third and sixth columns of numbers within Table 3. For example, among NCDS cohort males, a higher proportion reach first degree level in Southern England (16%, versus 10% in Northern England). Nor is this North-South divide on education confined to the NCDS generation itself – since the regressor means for the (grouped) variables on parental education are both noticeably higher for the South than the North.

Comparing our results to previous work on this dataset (see Table 4), the negative link between post-compulsory qualifications and the probability of male self-employment (BFN, 2002) shows clearly for Southern England – particularly for those with children – but, for the North, it is notable by its absence (especially for those with children). This may be a result of the less buoyant economy having fewer employment opportunities for the relatively well educated (since these opportunities would normally tend to draw them away from self-employment). There is a similar result for vocational qualifications, but apprenticeship has a positive association with male self-employment probability in both regions. The significant positive coefficient (Southern England) on the dummy identifying low reading ability at age 16 might indicate that these individuals have relatively poor employment prospects given the generally higher levels of education in the South and are pushed into self-employment.

Creativity, found by BFN (2000) to be positively – though weakly – linked to self-employment across both genders, is shown above to be almost significant for males in Southern England only. We also find that 'relatively depressed at age 11' Northerners are more likely to take up self-employment. The fact that 'lazy at 16' and 'sociable at 16' are each positively linked to self-employment only in the South may indicate a North-South gap in ways of working and self-employment activities. There is a negative relation between 'timid at 16' and self-employment at 33, for Northern England only, and a lower mean for timidity compared to the South (perhaps a timid individual is more unsuitable for self-employment in the North). A desire to be one's own boss at 23 is associated, as expected, with generally higher self-employment – and there is also the expected negative link of self-employment with the desire for job security (although a higher proportion of males in the North rated job security as the most important job characteristic in 1981 – when unemployment was particularly high, especially in the North).

---

[2] The two region dummies for which estimates are reported near the foot of Table 3 refer respectively to the Greater London and South-East England SRs in the case on Southern England (the South-West England SR forming the base); and to Yorkshire and the Humber and the North-West England SR in the case of Northern England (the North of England SR being the base here).

Table 3: Male logits of the self-employment probability – South versus North.

	Southern England			Northern England		
Variable	*Estimate*	*Est./S.E.*	*Mean*	*Estimate*	*Est./S.E.*	*Mean*
A level is highest	-0.912	-3.71	0.121	0.152	0.47	0.080
First degree is highest	-0.682	-2.87	0.164	-0.204	-0.58	0.104
Higher degree is highest	-1.385	-2.60	0.028	-1.113	-1.39	0.019
Professional qualification	0.167	0.86	0.123	0.347	1.34	0.108
Vocational qualification	-0.394	-2.94	0.460	-0.211	-1.20	0.483
Apprenticeship 1981	0.298	3.68	0.466	0.347	3.32	0.645
Maths High Aged 7	0.047	0.30	0.201	0.292	1.41	0.182
Maths Low Aged 7	-0.300	-1.53	0.107	0.139	0.59	0.142
Reading High Aged 16	-0.376	-1.87	0.162	0.0003	0.001	0.121
Reading Low Aged 16	0.505	2.67	0.097	-0.387	-1.52	0.141
Creativity	1.539	1.84	0.165	0.259	0.24	0.165
Depression	-0.0003	-0.01	0.877	0.145	3.04	0.986
Caution	-0.139	-1.55	0.178	0.062	0.54	0.173
Laziness	0.217	3.34	-0.167	0.125	1.44	-0.091
Moodiness	0.096	1.41	-0.454	0.074	0.85	-0.432
Sociability	0.175	2.49	0.449	0.058	0.66	0.404
Timidity	-0.013	-0.13	0.025	-0.339	-2.47	0.004
Own boss important 1981	0.765	4.47	0.102	0.552	2.35	0.092
Job security important 1981	-0.406	-2.70	0.217	-0.376	-1.94	0.289
Dad manager of small firm	0.258	1.46	0.114	0.652	2.35	0.067
Dad professional employee	-0.158	-0.55	0.056	0.678	1.62	0.033
Dad worker own account	0.268	1.02	0.038	0.606	1.44	0.025
Dad farmer employee-manager	0.732	1.50	0.010	1.866	2.69	0.007
Dad farmer own account	1.586	2.70	0.007	4.213	3.53	0.004
Dad's years of education	0.064	1.36	2.912	-0.062	-0.78	2.637
Mum's years of education	0.016	0.29	2.952	0.175	1.99	2.653
No children	-0.349	-2.43	0.353	-0.046	-0.23	0.274
No children * higher quals	0.486	1.77	0.142	-0.368	-0.85	0.072
Inheritance	0.082	3.15	0.687	0.179	1.42	0.243
Inheritance squared	-0.023	-0.61	0.184	-0.480	-1.01	0.018
Year of inheritance	-0.423	-2.62	0.268	-0.190	-0.74	0.192
Sub-region 1	-0.469	-2.53	0.169	0.523	2.25	0.391
Sub-region 2	-0.438	-3.28	0.613	0.551	2.37	0.409
Constant	-1.294	-4.12	1.000	-3.066	-6.98	1.000
Log-likelihood	-1023.262			-590.672		
Sample size	2124			1421		
Mean of dependent variable	0.22552			0.17030		

Table 4: Self-employment probability – statistical significance of regressors.

Variable	S. England	N. England	BFN (00)	BFN(02)	BO(98)
A level is highest	****	+	*		N/A
First degree is highest	****	*	****	****	N/A
Higher degree is highest	****	*	*		N/A
Professional qualification	+	+	++	+	N/A
Vocational qualification	****	*	**	***	N/A
Apprenticeship 1981	++++	++++	++++	++++	N/A
Maths High Aged 7	+	+	N/A	N/A	N/A
Maths Low Aged 7	*	+	N/A	N/A	N/A
Reading High Aged 16	**	+	N/A	N/A	N/A
Reading Low Aged 16	++++	*	N/A	N/A	N/A
Unforthcomingness	(+)	(+)	*	N/A	*
Hostility towards (other) children	(+)	(+)	*	N/A	+
Anxiety acceptance	(+)	(+)	++++	N/A	***
Creativity	++	+	+	N/A	N/A
Depression	*	++++	N/A	N/A	N/A
Caution	*	+	N/A	N/A	N/A
Laziness	++++	+	N/A	N/A	N/A
Moodiness	+	+	N/A	N/A	N/A
Sociability	+++	+	N/A	N/A	N/A
Timidity	*	***	N/A	N/A	N/A
Own boss important 1981	++++	+++	++++	++++	N/A
Job security important 1981	****	**	****	****	N/A
Dad manager of small firm	+	+++	++++	N/A	++++
Dad professional employee	*	+	++	N/A	N/A
Dad worker own account	+	+	++++	N/A	++++
Dad farmer employee-manager	+	++++	++++	N/A	N/A
Dad farmer own account	++++	++++	++++	N/A	++++
Dad's years of education	+	*	N/A	N/A	N/A
Mum's years of education	+	+++	N/A	N/A	N/A
No children	***	*	N/A	***	N/A
No children * higher quals	++	*	N/A	+	N/A
Inheritance	++++	+	++++	++++	++++
Inheritance squared	*	*	****	*	*
Year of inheritance	****	*	***	**	N/A

13

1. Asterisks indicate a negative sign – one for statistical insignificance at the 10% level, and an extra asterisk for significance at each of the 10%, 5% and 1% levels.
2. Plus signs indicate a positive sign. The number of plus signs follows the same rules as above.
3. Of the five columns summarising signs, columns 1, 2 and 4 refer to results for males only. Columns 3 and 5 aggregate males and females.
4. For three regressors, a plus sign has been enclosed in brackets – since these regressors were excluded from our preferred specification.
5. N/A indicates the absence of that regressor.
6. Blanchflower and Oswald (1998) did also include 12 parental social class dummies, and reported results for a variety of probit specifications. BFN (2000, 2002) also used a probit model, whereas this paper employs the logit model.

Having a male parent who was the manager of a small firm is less common in Northern England, but this only appears to have a significant positive impact on self-employment in the North (in contrast to Blanchflower and Oswald (1998) and BFN (2000)). Parental education lasts about 0.3 years longer on average for each parent in the South, and its links with self-employment also suggests a North-South divide – with a weak positive link from father's education (only) in Southern England, and a stronger positive link with mother's education (only) in the North.

The effect of inheritance on the self-employment probability found by Blanchflower and Oswald (1998) and BFN (2000, 2002) shows evidence of a non-linear component. Only in the last paper are males considered separately. The effect of inheritance on self-employment there – a statistically strong positive linear effect and a rather weak negative quadratic effect, plus a weak link with timing whereby recent inheritance comes with less chance of self-employment – is altered when the regional dimension is considered. This paper finds a positive linear effect of inheritance for Southern England only, and no evidence of a quadratic effect. There is also a statistically significant link between recent inheritance and lower self-employment probability in the South, but no evidence of any effect of the magnitude or timing of inheritance for Northern England.

Decomposition results:

To calculate our decompositions, we have dropped all dummies to capture Standard Region (SR) or a wider region – using the rather severe assumption that, under the 'no discrimination' distribution, there is no purely spatial aspect to variation in the self-employment probability. For the logit model, predicted probabilities reflect actual frequencies precisely. Decompositions are shown in Table 5, below – for five alternative approaches. The compositional term is negative in three out of five instances, equation (2) yielding the most notable exception. The choice of 'no discrimination' distribution clearly does matter. In each case, however, the compositional term is dwarfed by the structural part(s) of the overall gap

in self-employment probability – so male self-employment appears to be subject to a substantial North-South divide unaccounted for by the characteristics of individuals.

Table 5: Male logit – decomposition of the self-employment probability gap.

Version	*Compositional Term*	*Structural 1*	*Structural 2*	*Overall gap*
Equation (1)	0.2255 – 0.2366 = -0.0111	NIL	0.0663	0.0552
Equation (2)	0.1789 – 0.1703 = +0.0086	0.0467	NIL	0.0552
Equation (4)	0.2035 – 0.2032 = +0.0003	0.0220	0.0329	0.0552
Reimers	0.1956 – 0.1968 = -0.0012	0.0299	0.0265	0.0552
Cotton	0.2004 – 0.2037 = -0.0033	0.0251	0.0334	0.0552

Notwithstanding the small compositional term found above, we now investigate – using equation (3) following Even and Macpherson (1993) – the contributions of the various categories of regressor laid out previously[3] in Section 2. The second term in equation (3) can be used to break down the coefficient-weighted difference in regressor sample means, in accordance with each reported method of decomposition in turn. The results are shown in Table 6, below. Each row has six columns of numbers. Each of the first five give the numerator for the second term in equation (3) for that particular category of regressor. The last column (which is the sum of the first five) is the denominator for the second term in equation (3). The second row illustrates an important problem in the use of equation (3) – where the denominator of the second term is very small relative to four of the numerators. A general point should be noted about the relationship between rows 1, 2 and 4 in Table 6 – namely, that the definition of the second term of equation (3) combines with the Reimers (1983) definition of the 'no discrimination' coefficients (as the simple arithmetic mean of the coefficient vectors for Southern and Northern England) so that each element in row 4 is equal to the simple average of the corresponding elements from rows 1 and 2. Similarly, there is a relationship between rows 1, 2 and 5 – whereby each element in row 5 is equal to the weighted average of the corresponding elements from rows 1 and 2 (as in the Cotton (1988) definition of the 'no discrimination' coefficients).

One of the key features of Table 6 is that the first category (ability, education and training) of regressors has a robust effect across four of the five forms of the decomposition – whereby the extra ability, education and training of an average individual in Southern England leads to a lower probability of self employment than in Northern England (probably through the extra opportunities for employees with better education and training). However, the gap in family background characteristics (category 3) is shown in Table 6 to have a robust effect in the opposite direction for the same four forms of decomposition – so that the type of family background enjoyed by the average individual in Southern England differs from that of his

---

[3] Categories 5 (regions) and 6 (characteristics of self-employment) do not apply to our decomposition of the probability of self-employment.

counterpart in Northern England in ways that, ceteris paribus, make self-employment more likely.

Table 6: Male logit of the self-employment probability – breakdown of term 2 in (3).

	Regressor category (see Section 2)						
Version	*1*	*2*	*3*	*4*	*7*		*Overall*
Equation (1)	-0.1610	0.0237	0.0392	0.0047	0.0217		-0.0718
Equation (2)	-0.0508	0.0015	0.1083	-0.0487	-0.0055		+0.0048
Equation (4)	-0.1075	0.0183	0.0647	0.0022	0.0160		-0.0063
Reimers	-0.1059	0.0126	0.0737	-0.0220	0.0081		-0.0335
Cotton	-0.1168	0.0148	0.0669	-0.0167	0.0108		-0.0411

Female self-employment:

Table 7 shows self-employment logit estimates for females – separately for Southern England and Northern England. In addition to the higher self-employment rate in Southern England already noted in Table 1, the sample means shown in Table 7 indicate some interesting regional differences – although the fact that there are fewer statistically significant regressors means that this table is shorter than Table 3.

As for males, females in the South are more highly educated on average than those in the North. For females, in contrast to males, there is little evidence of academic qualifications having an impact on self-employment. However, females in the South with A-levels as their highest qualification may be more likely to be in self-employment at age 33. No corresponding effect is evident for females in Northern England, although high reading ability at age 7 is associated with more likely self-employment in NCDS5.

We briefly summarise some other results on self-employment probability that differ from BFN (2000, 2002), or exhibit a North-South divide. Previous depression (positive) and timidity (negative) have impact only in the South. So too does the desire at age 23 to be one's own boss (positive) – this was not apparent in earlier work. The link between having a father working with his own account and subsequent self-employment of the child is now shown to stem from females in Northern England. The positive association between childlessness and female self-employment is significant only at the 10% level, and only disappears among highly qualified females for Northern England. Magnitude of inheritance is only significant for females in the South – where the linear effect is statistically strong, but the quadratic (negative) effect is only significant at the 10% level.

Table 7: Female logits of the self-employment probability – South versus North.

	Southern England			Northern England		
Variable	*Estimate*	*Est./S.E.*	*Mean*	*Estimate*	*Est./S.E.*	*Mean*
O level equivalent is highest	0.304	1.59	0.409	0.164	0.64	0.444
A level is highest	0.556	2.12	0.123	-0.148	-0.32	0.081
First degree is highest	0.297	1.11	0.149	-0.260	-0.56	0.101
Professional qualification	0.150	0.61	0.081	0.275	0.73	0.066
Vocational qualification	-0.072	-0.46	0.338	-0.013	-0.06	0.304
Apprenticeship 1981	0.415	2.60	0.083	0.452	2.00	0.068
Reading High Aged 7	0.188	1.06	0.212	0.518	2.19	0.219
Reading Low Aged 16	-0.306	-1.00	0.093	0.445	1.43	0.145
Depression	0.108	2.07	0.701	-0.014	-0.17	0.687
Caution	-0.174	-1.53	0.152	-0.009	-0.06	0.229
Flexibility	0.065	0.61	0.186	-0.254	-1.72	0.217
Timidity	-0.296	-2.15	0.069	-0.138	-0.79	0.100
Promotion important 1981	-0.434	-1.18	0.054	0.275	0.62	0.041
Own boss important 1981	1.627	6.16	0.033	0.842	1.67	0.021
Job security important 1981	-0.357	-1.38	0.123	-0.299	-0.94	0.145
Family financial difficulties	-0.028	-0.07	0.045	-0.743	-1.39	0.068
Dad manager of small firm	0.289	1.35	0.107	0.313	0.95	0.084
Dad worker own account	0.111	0.28	0.034	1.099	2.20	0.023
Dad farmer employee-mngr	1.342	2.27	0.007	0.925	1.12	0.008
No children	-0.482	-1.90	0.249	-0.631	-1.75	0.206
No children * higher quals	0.461	1.33	0.112	1.028	1.83	0.078
Inheritance	0.096	3.00	0.748	0.435	1.28	0.585
Inheritance squared	-0.076	-1.98	0.410	-7.865	-1.07	2.351
Sub-region 1	0.247	1.09	0.177	-0.188	-0.66	0.389
Sub-region 2	0.086	0.47	0.601	0.131	0.48	0.420
Constant	-2.794	-11.19	1.000	-2.566	-7.61	1.000
Log-likelihood	-710.91			-378.53		
Sample size	2247			1496		
Mean of dependent variable	0.10814			0.07687		

Decomposition results:

The decompositions are again shown for the same five approaches. In contrast to the results for males, the compositional term for females is positive in every case – although the overall gap is of the same sign as it was for males. For equations (1) and (4), the compositional term accounts for about half the overall gap – and, for all five decomposition approaches, there is a

greater relative importance for the regional differences in average characteristics in determining female probability of self-employment than was found for male self-employment probability. See Table 8, below:

Table 8: Female logit – decomposition of the self-employment probability gap.

Version	Compositional Term	Structural 1	Structural 2	Overall gap
Equation (1)	0.1081 – 0.0931 = +0.0150	NIL	0.0163	0.0313
Equation (2)	0.0828 – 0.0769 = +0.0059	0.0254	NIL	0.0313
Equation (4)	0.1013 – 0.0871 = +0.0142	0.0068	0.0102	0.0313
Reimers	0.0891 – 0.0810 = +0.0081	0.0190	0.0042	0.0313
Cotton	0.0915 – 0.0827 = +0.0088	0.0167	0.0059	0.0313

Viewing the respective balances of the compositional and structural elements in rows 1 and 2 as those for two opposite extremes in terms of the form of decomposition, the element balances for the other ('intermediate') forms of decomposition do fall in between. Although this ordering was widely assumed by previous authors, Oaxaca and Ransom (1994) gave a counter-example.

The breakdown of the coefficient-weighted difference in regressor sample means is shown in Table 9, below – although it should be noted that the effects of inheritance on female self-employment in Northern England (measured by statistically insignificant, but quite large coefficients) seem to distort the results for category 4 and the overall breakdown in rows 2, 4 and 5. One key difference between these results for females and the corresponding results for males in Table 6 is the effect of regressors from category 1. Although, there is evidence of more education and training on average for females in Southern England, this has rather limited impact on the self-employment probability (and the effect is positive for most decompositions). There also seems to be more of a role for differences in non-cognitive characteristics than was evident for men. The effect of family background regressors from category 3, on the other hand, is rather smaller.

Table 9: Female logit of the self-employment probability – breakdown of term 2 in (3).

	Regressor category (see Section 2)					
Version	1	2	3	4	7	Overall
Equation (1)	0.0479	0.0447	0.0062	0.0197	0.0174	+0.1358
Equation (2)	-0.0400	0.0504	0.0174	-2.7111	0.0328	-2.6506
Equation (4)	0.0222	0.0502	0.0090	0.0229	0.0310	+0.1352
Reimers	0.0039	0.0475	0.0118	-1.3457	0.0251	-1.2574
Cotton	0.0127	0.0470	0.0107	-1.0717	0.0235	-0.9778

Job creation by self-employed males:

Table 10, below, shows estimates from censored (Tobit) regression for males – again, comparing Southern England and Northern England. Marginal effects can be readily calculated, via a scale factor (the Standard Normal Cumulative Distribution Function, evaluated – using the parameter estimates – at a chosen point, which is often the sample mean of the regressors). Among the male self-employed, job creation is positively associated with academic qualifications – but while having an A-level as highest is significant in the South, it is the more advanced first degree that is (weakly) significant for Northern England. Meanwhile, the positive effect of a professional qualification – found elsewhere in more aggregated samples – is preserved across the two separate regions. Unforthcomingness was found by BFN (2000) to have a statistically very weak negative link to job creation across all the self-employed. Table 10 illustrates that this relationship is statistically significant at the 1% level for males in Southern England. Timidity is negatively linked to job creation in the South – whereas, in the North, it is negatively linked to the self-employment probability. The classification of a job characteristic as most important does not appear to be a good indicator of self-employment job creation – although there is a very weak positive link for the promotion characteristic in the North. Having a father who was a professional employee is associated with greater job creation by self-employed males in Southern England, but not for the North (negative but insignificant here) – whereas BFN (2000) found a statistically weak positive effect aggregated across self-employed males and females. Another notable effect found in this family background category is the strong positive relationship between father's education and job creation in the South only.

The estimates on the inheritance regressors superficially appear different between the two regions – but the differences are not statistically significant. The positive linear coefficients is statistically significant at the 10% level for the South only. The negative quadratic estimate for the South is significant at the 10% level, while that for the North is insignificant. The other noticeable distinction is that the regressor means are quite different between the two regions: those for the South are substantially higher, and given the greater gap for the quadratic regressor, this is an indication of some rather large inheritances having been received among the South's self-employed males. Unsurprisingly, for both regions, self-employed males that operate their business from home create fewer jobs than those based elsewhere.

Table 10: Male tobits of self-employment job creation – South versus North.

Variable	Southern England			Northern England		
	Estimate	*Est./S.E.*	*Mean*	*Estimate*	*Est./S.E.*	*Mean*
O level equivalent is highest	2.306	1.06	0.380	1.731	1.01	0.375
A level is highest	10.662	2.71	0.076	-0.340	-0.11	0.079
First degree is highest	-3.263	-0.79	0.110	5.888	1.62	0.092
Professional qualification	10.753	3.24	0.093	6.415	2.36	0.117
Maths High Aged 7	-6.362	-2.38	0.184	0.545	0.27	0.204
Reading High Aged 16	5.927	1.76	0.102	8.102	2.69	0.117
Unforthcomingness	-1.647	-2.63	1.238	0.155	0.39	1.446
Anxiety acceptance	-1.104	-1.08	0.408	1.112	1.64	0.492
Caution	3.694	2.53	0.022	-1.685	-1.37	0.088
Moodiness	-1.778	-1.74	-0.268	0.362	0.51	-0.325
Timidity	-3.046	-1.71	-0.076	0.817	0.54	-0.133
Promotion important 1981	-0.980	-0.29	0.076	4.842	1.41	0.050
Dad professional employee	13.951	2.77	0.039	-0.550	-0.13	0.046
Dad professional self-emp	-12.317	-1.44	0.011	36.118	3.00	0.004
Dad worker own account	-9.589	-2.02	0.058	-0.353	-0.09	0.038
Dad's years of education	1.725	2.59	2.765	0.472	0.66	2.692
No children	-0.083	-0.03	0.289	-0.573	-0.28	0.254
No children * higher quals	-5.734	-1.20	0.089	-1.641	-0.37	0.058
Inheritance	2.029	1.88	1.290	3.038	1.52	0.397
Inheritance squared	-15.497	-1.84	0.651	-33.526	-1.11	0.044
NCDS region unemployment rate	0.095	0.06	6.527	0.157	0.19	10.995
Years self-employed	0.420	1.80	4.708	0.093	0.47	4.375
Not full time	1.121	0.40	0.168	-1.255	-0.60	0.225
Operated from home	-9.096	-4.67	0.516	-9.040	-5.33	0.483
Constant	-11.688	-1.05	1.000	-5.071	-0.49	1.000
Sigma	-15.614	18.50	N/A	9.363	14.18	N/A
Log-likelihood	-890.94			-473.77		
Sample size	463			240		
Mean of dependent variable	2.65227			3.52917		

5. Conclusions

This paper focuses on the North-South divide in England in self-employment and job creation by the self-employed, against the background of known differences in the regional economies of the two areas. The paper uses individual level data which so far has only been used to analyse UK self-employment at a national level. Males and females are studied separately, as this has been shown to be important previously. We also use decomposition analysis to clarify the distinction between regional structural effects and the effects of regional differences in individual characteristics. We have several new findings and implications.

For male self-employment, we find a negative effect of post-compulsory education only for Southern England – probably due to better employment opportunities there. Education levels are also higher in the south , so less educated individuals may be disadvantaged in the labour market and pushed into self-employment. We also find self-employment is associated with lower ability and motivation in the South. In the North, fewer job opportunities may push marginal well educated individuals into self-employment.

Blanchflower and Oswald (1998) note that inheritance is a good proxy for exogenous availability of finance – and should have a positive and significant effect on self-employment if finance constraints exist. Our results for Southern England include a positive effect of inheritance, and of the time since receipt – but neither of these effects appear for Northern England. This result is interesting because one would expect more finance availability in the South. The regional variation we find may indicate that more business opportunities and competition among greater numbers of entrepreneurs for finance in the South means constraints are more likely there. Thus, the analysis uncovers some interesting regional differences compared to aggregate effects of inheritance on self-employment in the previous work of Blanchflower and Oswald (1998), and BFN (2000, 2002). Obviously, further research is needed but the results do raise issues of regional policy.

Decomposition analysis shows that higher male self-employment in Southern England occurs in spite of higher education, ability and training; and instead partly through an opposing effect resulting from differences in family background. These findings are inevitably missing in traditional aggregated studies of male self-employment.

Our results on male job creation indicate there are differences by region in which qualifications are associated with the creation of more jobs. Having a father who was a professional employee only seems to help job creation for those in the South; and job creation in this region (only) is also linked to paternal education.

For female self-employment, there is much less evidence of education having an impact – although A levels may influence self-employment in Southern England. Inheritance has no effect on self-employment for females in the North – another regional contrast. Higher female

self-employment in Southern England is affected by both education and family background, but the situation is complicated by inheritance, which is only significant in the South.

The male results also give some new insights into why self-employment is higher in Southern England but the average entrepreneur creates more jobs in Northern England. Briefly, this seems to be due to differences in the regions and not the characteristics of the average individual in each region. Thus lower post-compulsory education in Northern England would actually generate higher levels of self-employment and similar job creation if the Northern economy was like the South.

We also find that total effects on self employment job creation differ in various ways between North and South. Thus, for example, first degrees have a negative total effect in the South by reducing self employment and not raising jobs per entrepreneur. In the North, however, first degrees have a positive effect on job creation and no influence on self-employment, giving a positive total contribution. It follows that aggregate effects for the UK deduced in previous studies may be the net result of opposing tendencies in the different regions, and hence provide quite misleading guides for policy-makers.

This result can also explain why MVS (2006) find that UK regions do not match the pattern in Europe (Fritsch and Mueller, 2004), where direct job creation by new ventures is higher in high productivity/educated regions. Our results show that more highly educated regions do not necessarily have more highly educated self-employed sectors. Our use of individual level data and regional analysis show that since a lack of job opportunities can push well educated people into self-employment, then if push factors vary by region, less buoyant regions can have higher levels of education among the self employed than even regions with more highly educated workforces. Our results indicate that this appears to be the case in the UK, with the less educated North having higher direct job creation among the self-employed than the better educated South – as in MVS (2006).

In summary, the new results in this paper show that the explanations of self-employment prevalence and job creation are sufficiently different between North and South England as to require corresponding regional variation in enterprise policy – particularly in relation to education and finance. Our disaggregated analysis also indicates how unreliable predictions for the performance of regional self-employment are likely to be under the usual assumption of regional homogeneity This is particularly important for education. For example, based on Southern estimates, lower post compulsory education in the North should boost the probability of Northern self-employment. Yet our regional analysis shows a different result – hypothetically higher post compulsory education in the North should not reduce self employment but could increase job creation by the self employed. The paper, therefore, highlights the importance of identifying differences across regional economies and provides new evidence that the North-South economic divide is not merely a traditional industrial phenomenon but also a hallmark of self-employment too.

References

Anyadike-Danes, M. (2004) The Real North-South Divide? Regional Gradients in UK Male Non-employment, *Regional Studies* 38(1): 85-95.

Audretsch, D.B. and Fritsch, M. (2002) Growth Regimes over Time and Space, *Regional Studies* 36: 113-124.

Blanchflower, D. and Oswald, A. (1998) What Makes an Entrepreneur? *Journal of Labor Economics* 16: 26-60.

Blinder, A. S. (1973) Wage Discrimination: Reduced-Form and Structural Estimates, *The Journal of Human Resources* 8(4): 436-455.

Burke, A.E., FitzRoy, F.R. and Nolan, M.A. (2000) When Less is More: Distinguishing Between Entrepreneurial Choice and Performance, *Oxford Bulletin of Economics and Statistics* 62: 565-587.

Burke, A.E., FitzRoy, F.R. and Nolan, M.A. (2002) Self-employment Wealth and Job Creation: The Roles of Gender, Non-pecuniary Motivation and Entrepreneurial Ability, *Small Business Economics* 19: 255-270.

Cotton, J. (1988) On the Decomposition of Wage Differentials, *Review of Economics and Statistics* 70(2): 236-243.

Cowling, M., Taylor, M. and Mitchell, P. (2004) Job Creators, *The Manchester School* 72(5): 601-617.

Du Rietz, A. and Henrekson, M. (2000) Testing the Female Underperformance Hypothesis, *Small Business Economics* 14(1): 1-10.

Evans, D. and Jovanovic, B. (1989) An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints, *Journal of Political Economy* 97: 808-827.

Even, W.E. and Macpherson, D.A. (1990) Plant Size and the Decline of Unionism, *Economics Letters* 32(4): 393-398.

Even, W.E. and Macpherson, D.A. (1993) The Decline of Private-Sector Unionism and the Gender Wage Gap, *Journal of Human Resources* XXVIII(2): 279-296.

Fairlie, R.W. (1999) The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-employment, *Journal of Labor Economics* 17(1): 80-108.

Fairlie, R.W. (2003) An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models, *Economic Growth Center Discussion Paper No. 873, Yale University*.

Fritsch, M. and Mueller, P. (2004) The Effects of New Business Formation on Regional Development over Time, *Regional Studies* 38: 961-975.

Fritsch, M. and Mueller, P. (2006) The Effects of New Business Formation on Regional Development over Time: the Case of Germany, Max Planck Institute for Economics,

Entrepreneurship Group Discussion Paper 1906 and *forthcoming in Small Business Economics*.

Fritsch, M. and Shmunde, J. (eds) (2006) *Entrepreneurship in the Region*. New York: Springer.

Fothergill, S. (2001) The true scale of the regional problem in the UK, *Regional Studies* 35(3): 241-246.

Georgellis, Y. and Wall, H.J. (2000) What Makes a Region Entrepreneurial? Evidence from Britain, *Annals of Regional Science* 34(3): 385-403.

Gomulka, J. and Stern, N. (1990) The Employment of Married Women in the United Kingdom 1970-83, *Economica* 57(2): 171-199.

Gray, D. (2004) Persistent Regional Unemployment Differentials Revisited, *Regional Studies* 38(2): 167-176.

Greene, W.H. (2002) *LIMDEP, version 8.0 for the PC*, Econometric Software, Inc.

Gudgin, G. (1996) Regional problems and policy in the UK, *Oxford Review of Economic Policy* 11(2).

Khilstrom, R. and Laffont, J.J. (1979) A general equilibrium entrepreneurial theory of the firm formation based on risk aversion. *Journal of Political Economy* 87: 719-748.

Lewis, J and Townsend, A. (1989) *The North-South Divide: Regional Change in Britain in the 1980s*, edited volume, Paul Chapman, London.

Lindsay, C. (2004) Labour Productivity, Special Feature, Labour Market Division, Office for National Statistics, UK.

Mueller, P., Van Stel, A. and Storey, D.J. (2006) The Effects of New Firm Formation on Regional Development Over Time: the Case of Great Britain, Max Planck Institute for Economics, Entrepreneurship Group Discussion Paper 2406 and *forthcoming in Small Business Economics*.

Neumark, D. (1988) Employers' Discriminatory Wage Behavior and the Estimation of Wage Discrimination, The Journal of Human Resources 23: 279-295.

Oaxaca, R. (1973) Male-Female Wage Differentials in Urban Labor Markets, *International Economic Review* 14(3): 693-709.

Oaxaca, R.L. and Ransom, M.R. (1994) On discrimination and the decomposition of wage differentials, *Journal of Econometrics* 61: 5-21.

OECD (2000) *Employment Outlook*, Organisation for Economic Co-operation and Development.

Parker, S.C. (2004) The Economics of Self-Employment and Entrepreneurship, Cambridge University Press, Cambridge, England.

Parker, S.C. and Robson, M.T. (2004) Explaining International Variations in Self-Employment: Evidence from a Panel of OECD Countries, *Southern Economic Journal* 71(2): 287-301.

Reimers, C.W. (1983) Labor Market Discrimination Against Hispanic and Black Men, *Review of Economics and Statistics* 65(4): 570-579.

Robson, M.T. (1998) Self-employment in the UK Regions, *Applied Economics* 30: 313-322.

Rosa, P., Carter, S. and Hamilton, D. (1996) Gender as a Determinant of Small Business Performance: Insights from a British Study, *Small Business Economics* 8(4): 463-478.

Shutt, J. and Sutherland, J. (2003) Encouraging Transition into Self-employment, *Regional Studies* 37(1): 97-103.

Van Stel, A. and Storey, D.J. (2004) The Link Between Firm Births and Job Creation: Is there and Upas Tree Effect? *Regional Studies* 38: 893-909.