WORKING PAPER SERIES

What Makes a Region Entrepreneurial? Evidence from Britain

Yannis Georgellis and Howard J. Wall

Working Paper 1999-009A
http://research.stlouisfed.org/wp/1999/99-009.pdf

PUBLISHED: Annals of Regional Science, 2000, 34(3), pp. 385-403.

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
411 Locust Street
St. Louis, MO 63102

What Makes a Region Entrepreneurial? Evidence from Britain

Yannis Georgellis*

Brunel University

Howard J. Wall†

Federal Reserve Bank of St. Louis

September 1997, revised June 1999

* Department of Economics and Finance, Brunel University, Uxbridge UB8 3PH, United Kingdom; E-mail: yannis.georgellis@brunel.ac.uk; Tel: (44)-1895-203174; Fax: (44)-1895-203384.

† Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442, United States; E-mail:wall@stls.frb.org; Tel: (314)444-8533; Fax: (314)444-8731.

Acknowledgements: We would like to thank Ron Smith, Simon Parker, and Martin Robson for their comments on earlier drafts. The views expressed are those of the authors and do not necessarily represent official positions of the Federal Reserve Bank of St. Louis nor of the Federal Reserve System.

What Makes a Region Entrepreneurial?
Evidence from Britain

There is a great deal of variation in the levels of entrepreneurship, or rates of self-employment, across the regions of Britain. Over the period 1983-95, average self-employment in the North, Scotland, and the West Midlands was respectively 25%, 15%, and 15% lower than the national average, whereas in the South West, East Anglia, and Wales it was respectively 28%, 23%, and 21% higher. We develop a theoretical model of regional self-employment, and estimate the roles of labour market conditions, labour force characteristics, industry composition, and region-specific factors such as entrepreneurial human capital. Our results suggest that all of these factors are important, and that regional heterogeneity and regionally correlated disturbances must be accounted for when estimating regional self-employment relationships. (JEL J23, R12)

I. Introduction

Entrepreneurship, as measured by self-employment, has become an increasingly prominent characteristic of industrialised economies. Since 1980 all but two EU countries have seen an increase in their rates of non-agricultural self-employment, as have Canada, the US, Japan, and Australia.[1] None, however, has seen as dramatic a change as has occurred in the UK. Following a long period during which the share of British employees who were self-employed hovered around 7%, it rose rapidly throughout the 1980s, reaching over 13% by the end of the decade.[2] Because of these trends, a good deal of research has been done to estimate the determinants of entrepreneurship at the national and individual levels in Britain and elsewhere.[3] However, there has been virtually no research looking at entrepreneurship at the regional level, despite the fact that the spatial variation in self-employment rates is at least as dramatic as the recent temporal variation.

As at the national level, self-employment rates rose in all British regions during the

[1] See Campbell and Daly (1992), Acs, Audretsch, and Evans (1994), and Taylor (1997) for international comparisons.
[2] See Campbell and Daly (1992) and Taylor (1997) for detailed analyses of trends in British self-employment.

1980s. In addition, regional rates of self-employment relative to the national average differed widely across regions, and fluctuated significantly over the period. As illustrated by Figure 1, self-employment rates have tended to be relatively higher in Wales and the south of England, and lower in Scotland and the north of England. Over the period 1983-95, average self-employment in the North, Scotland, and the West Midlands was respectively 25%, 15%, and 15% lower than the national average, whereas in the South West, East Anglia, and Wales it was respectively 28%, 23%, and 21% higher. Although there has been a general North-South divide throughout the period, Figure 1 shows that relative self-employment rates nevertheless fluctuated, indicating that significant differences and changes at the regional level were masked by the continual increase at the aggregate level.

Regional variation in entrepreneurship has been at least as prominent in other countries as it has been in Britain. In the US in 1990, there were ten states with self-employment rates that were more than 15% below the national rate, and sixteen states that were more than 15% above it. The range of state self-employment rates is illustrated by Table 1, which provides self-employment rates for eleven selected states for 1980 and 1990. Further, as Table 1 also shows, there were significant shifts in the states' relative self-employment between 1980 and 1990, particularly in the northern plains.[4] As with Britain, we know of no work that has been done to examine US regional entrepreneurship.[5]

Clearly, because farmers are commonly self-employed, the presence of large agricultural sectors would account for some of the regional variation described above. However, this is only a

[3] Recent studies of British self-employment include Taylor (1996), Parker (1996), Cowling and Mitchell (1997), and Robson, (1998a and b). Earlier studies of the UK include Rees and Shah (1986) and Robson (1991). Acs, Audretsch, and Evans (1994), De Wit (1993), and Blanchflower and Oswald (1998) study elsewhere.
[4] See also Georgellis and Wall (1999), who report how entrepreneurship varies greatly across Germany.
[5] See Blanchflower and Oswald (1998) for a recent individual-level study that uses US panel data.

small part of the story for Britain. For a simple illustration, consider the two regions at the extremes in terms of the prevalence of entrepreneurs, the North and the South West. Over the period covered in this study, agriculture's employment shares averaged 0.9% in the North, and 2% in the South West. Over the same period, the self-employment rates were 7.7% and 13.9%, respectively. So, even if we make the extreme assumption that all of those employed in the agricultural sector are self-employed, the regional variation would still be large.[6]

Given the extent of the regional variation in entrepreneurship described above, it is surprising that so little attention has been paid to its determinants. The purpose of the present paper is to develop a simple theoretical model of regional entrepreneurship, and to estimate the resulting self-employment function for British regions. In particular, we examine the roles of regional differences in (i) labour market conditions, (ii) labour force characteristics, (iii) industry composition, and (iv) unobserved region-specific factors. In doing so, we pay particular attention to the spatial econometric issues of regional heterogeneity and regionally correlated disturbances.

A previous paper that also examines spatial variations in self-employment rates is Acs, *et al* (1994), who attribute much of the *international* variation to differing stages of economic development. Also, Robson (1998b) uses regional-level data to focus on the time-series properties of the relationship between self-employment and unemployment rates in Britain. Our approach is substantially different from his in that our interest is solely on the regional differences, to the exclusion of any time-series effects.

Before proceeding, we should note that previous research on entrepreneurship has used two different notions of an entrepreneur, both of which have been measured by self-employment. The first notion simply uses the definition of an entrepreneur as a person who undertakes a

[6] See Tables 2 and 3 below for the data for all the regions.

commercial venture. Related to this is the Schumpeterian entrepreneur, an innovator who develops new products and technologies in pursuit of capitalist profit. For the most part, the literature we cite concerns itself with the definitional notion of an entrepreneur, while hinting at the Schumpeterian effects of self-employment. Schiller and Crewson (1997), though, are interested only in Schumpeterian entrepreneurship, but nonetheless proxy for it with self-employment, arguing that self-employment is a "pragmatic if not compelling index of entrepreneurial creativity" (p. 525). In their view, while not all of the self-employed are innovators, self-employment and entrepreneurial creativity are highly correlated. Further along these lines, Noteboom (1994) discusses the link between small businesses and innovation, concluding that 10-20% of small business owners are Schumpeterian entrepreneurs.

In what follows, we restrict ourselves to the definitional notion of entrepreneurship, and try to explain regional variations in self-employment. However, if one agrees with Schiller and Crewson's argument that self-employment is a useful proxy for business innovation, our results are also directly applicable to explaining regional variations in Schumpeterian entrepreneurship.

II. Theoretical Framework for Regional Entrepreneurship

Following Knight (1921), the decision to become an entrepreneur has usually been modelled as an expected-utility-maximising choice between entrepreneurship and the pursuit of paid-employment. Taking account of the financial and non-financial returns on offer, an individual chooses to be an entrepreneur when the expected utility of doing so dominates that of paid employment. Modern extensions of the model include Blau (1987), who considered the general equilibrium aspects; Evans and Jovanovic (1989), who introduced credit constraints; and Parker (1996), who developed an intertemporal model with uncertainty. The model below is not

an addition to the theory of entrepreneurship, but presents a simple framework for the extension of a simple individual-level random-utility model to regional analysis.

Assume that each member of the labour force has a choice of pursuing paid-employment, or of becoming and entrepreneur. The outcome of each of these options is uncertain and depends on the individual's abilities and preferences in each of the activities, and on the prevailing regional market conditions. Define the *mean person* as that member of a labour force who possesses the mix of characteristics and skills expected of a randomly selected person. Denote the utility that the country's mean person would attain if he pursued self-employment in region *i* as U_i^{se}, and that from paid-employment in region *i* as U_i^{pe}.

The utility levels U_i^{se} and U_i^{pe} differ across regions because the regions differ in their suitability (including profitability) for entrepreneurship relative to paid-employment. These differences arise because of regional differences in industrial composition, wages for paid-employment, and risk associated with paid-employment (possible unemployment).

For region *i*'s mean person, the difference in utility between self- and paid-employment in region *i* is $U_i^{se} - U_i^{pe} + \delta_i$, where δ_i differentiates the mean person in region *i* from the mean person of the country as a whole. δ_i differs across regions because of regional difference in average education levels, age, entrepreneurial human capital, and other individual characteristics, some of which may not be observable nor measurable.

Define a random variable s_{ij} such that s_{ij}=1 if individual *j* in region *i* is self-employed, and s_{ij}=0 otherwise. If individual *j* was randomly selected from region *i*, the probability that he will be self-employed is the probability that the difference in utility from the two activities is positive for region *i*'s mean person:

$$\Pr[s_{ij} = 1] = \Pr[U_i^{se} - U_i^{pe} + \delta_i > 0]. \quad (1)$$

Summing (1) across the L_i workers in i, and dividing by L_i,

$$S_i \equiv \frac{1}{L_i} \sum_{j=1}^{L_i} \Pr[s_{ij} = 1] = F(U_i^{se} - U_i^{pe} + \delta_i); \quad (2)$$

where $F' > 0$. Assuming a large L_i, the left-hand-side of (2) is the rate of self-employment in region i, S_i .

Equation (2) says that the self-employment rate of a region is increasing in the relative utility from self-employment that would obtain for the region's mean person. This differs across regions because: (i) Controlling for labour-force characteristics, the relative suitability of regions for self-employment differs regionally (making $U_i^{se} - U_i^{pe}$ differ regionally); and (ii) Regional labour forces differ in their skills and preferences towards self-employment (making δ_i differ regionally).

Variables that capture the first of these reasons are a region's wage for paid-employment, its unemployment rate, and the composition of its industries. Variables that capture the second of these reasons are a regional labour force's educational and age composition. There may also be cultural, historical, geographic, and sociological factors that are difficult to observe or even to measure, but which account for some of the regional variation in self-employment not accounted for by the above variables. If such unobserved regional heterogeneity is important, then even if all other variables listed above are the same across regions, we would still observe regional differences in rates of self-employment. These factors can affect a region's relative-suitability for self-employment, $U_i^{se} - U_i^{pe}$, and/or the regional labour force's suitability for self-employment, δ_i. As is standard in panel data analyses such as this, to account for the possibility of such

unobservable, immeasurable, or intangible heterogeneity, we allow the intercepts of the self-employment function (2) to be region-specific.

Assume that (2) takes the functional form

$$S_i = \alpha_i + \beta_1 w_i + \beta_2 w_i^2 + \gamma_1 u_i + \gamma_2 u_i^2 + \delta' \mathbf{X_i} + \theta' \mathbf{Z_i}; \quad (3)$$

where α_i is the region-specific intercept; w_i is the average real wage for paid-employment in *i*; u_i is the unemployment rate in *i;* $\mathbf{X_i}$ is a vector of variables controlling for the industrial composition of *i*; and $\mathbf{Z_i}$ is a vector controlling for the characteristics of the labour force in *i.*

We have used a quadratic to specify each of the two labour market variables because each has two opposing effects on the self-employment rate. On one hand, a high unemployment rate may 'push' people into self-employment because of a lack of opportunities in paid-employment. On the other hand, they may be 'pulled' into self-employment by a buoyant regional economy, as indicated by a low unemployment rate, because the probability of having a successful entrepreneurial venture is high. By assuming a quadratic form we allow for the possibility that the recession-push effect dominates for some range of unemployment rates, whereas the prosperity-pull effect dominates for another range.[7]

The wage may also have two opposing effects in that it measures the pecuniary benefits of paid-employment (the opportunity cost of being an entrepreneur), but may also act as a measure of the level of income of the customers of the self-employed, and therefore act as a proxy for the level of regional aggregate demand. So, as with the unemployment rate, the quadratic form we assume is flexible enough to handle both of these effects.

Because there is a gap between the time that an individual makes his self-employment decision and the time that he becomes self-employed, assume that current the current self-employment rate depends on the values of the right-hand-side variables for the previous period. Also assume that region-specific effects are fixed over the period we examine below. Adding time subscripts and an error term, the regression equation becomes

$$S_{it+1} = \alpha_i + \beta_1 w_{it} + \beta_2 w_{it}^2 + \gamma_1 u_{it} + \gamma_2 u_{it}^2 + \delta' \mathbf{X}_{it} + \theta' \mathbf{Z}_{it} + \varepsilon_{it}. \quad (4)$$

Because we wish to control for the contemporaneous trends in the RHS variables, we measure the value of each of the variables relative to the average of the regions within a given year. Using relative measures for all the variables means that all level effects are removed from the data, allowing us to focus purely on regional differences. It also eliminates the need have year dummies, which would use up already-scarce degrees of freedom.

III. Data and Variables

The self-employment data are for 1983-1993, and those for the independent variables are for 1982-1992. Our data set is restricted to this period because the industrial classifications were altered in 1981 and 1993. With eleven years of data for the ten standard regions of Britain, we have 110 observations. All data are from relevant issues of *Regional Trends,* and we use the regional consumer price index from Reward Group (1995) to deflate nominal wages to real regional wages. For reference, the regional self-employment rates, real wages, and unemployment rates, all averaged over the sample period, are presented in the first three columns

[7] Recent studies of British self-employment have focused on determining whether it has been the 'recession-push' or the 'prosperity-pull' effect that has been dominant in determining aggregate the self-employment rate. Parker (1996) found a positive relationship between the rates of unemployment and self-employment, although Robson (1998a and b) found the opposite. Cowling and Mitchell (1997) find a negative relationship for short-term unemployment, and a positive one for long-term unemployment.

of Table 2.

The vector of variables representing the characteristics of the regional labour forces, $\mathbf{Z_{it}}$, controls for regional differences in preferences and abilities. The six variables in $\mathbf{Z_{it}}$ are the share of a region's population aged 16-44, the share aged 44 to retirement age (60 for women, 65 for men), the share older than the retirement age, the share with an A-level or higher qualification, the share with no qualification, and the female share of the labour force.[8] Unfortunately, the age groupings provided by *Regional Trends* are not ideal for our purposes as the 16-44 group includes the young, who are the least likely to be self-employed, and those aged 25-44, who are the most likely to be self-employed.[9] Taken together though, the three age groupings should control for the overall age compositions of the regions. For reference, regional averages of the variables included in $\mathbf{Z_{it}}$, measured in absolute terms, are provided by the last six columns of Table 2.

The vector $\mathbf{X_{it}}$ represents a region's suitability for entrepreneurs, as measured by its industrial composition. The variables are the shares of a region's employees who are employed in each of the ten Standard Industrial Classifications: agriculture, forestry, and fishing; energy and water supply; metals, minerals, and chemicals; metal goods, vehicles, and engineering goods; other manufacturing; construction; distribution, hotels, and repairs; transport and communication; financial services; and public and other administration. To prevent perfect collinearity of these variables, we exclude the agriculture, forestry, and fishing variable. For reference, Table 3 provides the regional averages of the employment shares for the ten Standard Industrial Classifications.

[8] Because *Regional Trends* does not provide data for the education variables for 1987 and 1989, we used the average of the previous and subsequent years. Also, prior to 1988, the youngest age group was 15-44.
[9] See Campbell and Daly (1992).

IV. Empirical Results

So as to handle the possibility of non-spherical error terms, we used Generalized Least Squares to estimate equation (4), thus allowing us to correct for within-region heteroskedasticity and cross-region correlation. The latter of these arises when there is cross correlation of regional disturbance terms, due either to spatial autocorrelation or because regions have similar responses to shocks, even if they are not otherwise spatially related.[10] Model I is the 'complete' model, which allows for heteroskedasticity, cross-region correlation, and heterogeneous self-employment functions (region-specific intercepts). Model II differs from Model I only in that it assumes that self-employment functions are homogenous (a common intercept). Model III differs from Model I only in that it assumes that there is no cross-region correlation. Table 4 presents three sets of results, and Table 5 presents the covariance matrix for Model I. Our discussion of our results will focus almost exclusively on Model I, using the results for the other two models for comparison only.

Labour market variables

Our results indicate that the relationship between relative self-employment and relative unemployment is hill-shaped, with a peak at a relative unemployment rate of 1.06. So, for a region with low unemployment, relative self-employment should rise along with relative unemployment, indicating that push effects are dominant. But for a region with already high unemployment, relative self-employment should fall if relative unemployment rises, indicating the dominance of prosperity-pull effects.

We find a convex, although always negative, relationship between self-employment and the real wage variable. This is consistent with the notion that the wage in paid-employment represents the opportunity cost of self-employment.

Labour force characteristics

As mentioned earlier, our age variables are not ideally suited for our purposes as the age groupings do not match well with the propensity for people of different ages to be self-employed. Nonetheless, they appear to explain some of the regional variation in self-employment rates, although only the coefficient on the age group 44-retirement is statistically significant. The education variables also appear to be capturing some of the variation in regional self-employment rates, although their signs are not as one might expect, as both groups would be expected to be more likely than the rest of the population to be entrepreneurs. The negative sign on the female share of the labour force is as expected because women are less likely to be entrepreneurs than are men. However, one should be cautious when interpreting the coefficients of these variables because they may also represent availability of customers for entrepreneurs, as well as the availability of people likely to be entrepreneurs.

Industry composition

The coefficients of the industry composition variables are estimated relative to that of the excluded industry; agriculture, fisheries, and forestry, whose coefficient is set to zero. For four industrial groupings: energy and water supply, construction, transport, and financial services; the effect of a higher employment share is statistically no different from a higher employment share

[10] See Greene (1997, ch. 15) for a detailed description of the estimation procedure, and Anselin (1988) for the econometric consequences of spatially correlated disturbances.

for agriculture, etc. In contrast, for those industries with positive and statistically significant coefficients: metals, etc., distribution, etc., and public administration; a higher employment share would have a larger effect than would a higher share for agriculture, etc. For the remaining industries, a higher employment share will have a smaller effect than would a higher share for agriculture, etc.

One should be cautious in interpreting the signs on the coefficients because, as with the labour force characteristics, they are likely to be capturing supply-side and demand-side effects of entrepreneurship. In other words, a region's suitability for entrepreneurs is not indicated solely by the levels of activity in industries in which entrepreneurs are common, but also for the activities of other industries that may be the customers or suppliers of entrepreneurs. For example, although the proportion of employees in public administration who are self-employed is low, shifts of a region's economy towards this sector may increase entrepreneurial opportunities for firms in industries where self-employment rates are relatively high (such as construction, financial services, hotels and distribution). The empirical question is therefore whether inclusion of these variables is statistically important in estimating the self-employment relationship.

By necessity the nine included industry variables are measured relative to the one that is excluded, which unfortunately precludes us from saying much about their importance for explaining variations in regional self-employment. To remedy this we re-estimated Model I under the assumption that, as with the coefficient on the agriculture, etc. employment share, those for the other nine industry employment shares are zero. We do not report these results here, but a likelihood ratio test rejects the null hypothesis that the exclusion of these nine industry variables has no statistical effect on the results. This is with $\chi^2(9) = 19.19$ and a critical value of 16.92 at 95% confidence level. We therefore conclude that industrial composition is statistically

important in explain regional differences in self-employment rates.

Region-specific effects

Recall that we have allowed for regional differences in cultural, historical, geographic, and sociological factors that are difficult to measure or observe. Such factors can affect a region's relative-suitability for entrepreneurial activity and/or the regional labour force's suitability for being entrepreneurs. If these factors are correlated with the other independent variables included in the regression, then the estimated coefficients on these other variables will be biased when they are not accounted for. Because they cannot be controlled for by measuring them with actual variables, we have instead done so by allowing for region-specific intercepts.

In Model I, the region-specific effects are all statistically significant at the 5% level, and they differ substantially across the regions. A more important question though is the effect on the results of relaxing the restriction that the regions have a common intercept. We test this by estimating Model II, which applies this restriction to the complete model. As is clear from a comparison of the results for Models I and II, this restriction on the regional intercepts biases the estimation, as the coefficients on the other variables differ greatly between the two models. Further, a likelihood-ratio test with $\chi^2(10) = 53.61$ and a critical value of 18.31 at the 95% confidence level rejects the null that Models I and II are statistically the same. This implies that the region-specific effects are correlated to a statistically important extent with one or more of the right-hand-side variables, as well as with the self-employment rate.

Refer to Table 6, which describes the extent to which the region-specific effects from Model I differ from each other, and quantifies their importance in explaining regional variations in self-employment rates. The first column gives the differences in relative self-employment rates

from unity (the mean relative self-employment rate), averaged over the period 1983-93, in order of lowest to highest. The second column provides differences in the estimated region-specific effects from their average, 3.476. The third column presents the portion of the regional self-employment rate differentials not accounted for by differences in region-specific effects. Note that there is a wide divergence in self-employment rates, region-specific effects, and the effects of other factors.

The last column of Table 6 provides the ratio of the region-specific effect to the effect of other factors, and is a measure of the relative importance of region-specific effects in explaining regional variations in self-employment. The region-specific effects explain relatively large portions of the regional variation in entrepreneurship, although only that of the North is greater than the sum of the effects of the other factors. The North is an outlier in that the region-specific effect alone can account for nearly all of the difference in its self-employment rate.

Our purpose in including region-specific fixed effects is to control for regional differences in cultural, historical, geographic, and sociological factors that are difficult to measure or observe. If we have controlled for all market factors that determine a region's self-employment rate, such as the value of paid-employment, the age and educational profile of the labour force, and the structure of the regional economy, then what remains may be called the 'entrepreneurial human capital' of a region's average person. One potential source of this is suggested by individual-level studies which find that a person's probability of being self-employed is higher if his/her parent was self-employed.[11] At the regional level then, the self-employment rate can be higher than that suggested by current market factors because *past* market factors were favourable to self-employment. The region-specific fixed effects would then be

determined in part by entrepreneurial inertia due to intergenerational transfers of entrepreneurial human capital. Of course, this would be changing throughout our sample period as levels of entrepreneurship changes, but eleven years is likely insufficient for significant changes to have occurred.

We control for missing variables that are unobservable or unmeasurable by assuming that they are fixed throughout the sample period, and can therefore be captured by the time-invariant regional intercepts. By definition, the regional intercepts would also pick up any variable that is measurable and observable, but which happens to be fixed, such as geographic variables. However, they would also pick up any variables that may have been excluded from the analysis either inadvertently or by necessity. Variables that we were unable to include because of the unavailability of a sufficiently long series of region-level observations, are measures of the ethnic composition of regional populations.

As Campbell and Daly (1992) report, Britons of West Indian, Guyanese, and African descent are less likely than whites to be self-employed, whereas South Asians are more likely.[12] As these variables were not available for years prior to 1991, we were unable to include them in our estimation of the complete model. Although these groups make up just over 5% of the British population, as summarised by Table 7, there is considerable variation across regions. This may account for some of the variation in self-employment rates, and if so, would be captured by the variation in the regional intercepts. To test this, we regressed our ten estimated fixed effects on the regional population shares of these ethnic groups for 1991, presented for reference by Table 7. Although the ethnic composition of Britain changed over the period of our study, the relative

[11] See Lentz and Laband (1990), De Wit and Van Winden (1990), Dunn and Holtz-Eakin (1996), and Georgellis and Wall (1999).

composition of regions was likely unchanged to any significant extent. The Ordinary Least Squares results are

$$\hat{\alpha}_i = \underset{(41.62)}{4.4912} + \underset{(0.563)}{0.0468}\,(\text{Black share}) - \underset{(0.605)}{0.0297}\,(\text{South Asian share});$$

where t-statistics are in parentheses, and R^2=0.05. As the coefficients are statistically no different from zero, we find no evidence that the differences in the regional intercepts are due to the omission of these variables from Model I.

It is not wholly satisfying to attribute large portions of the variation in regional entrepreneurship to region-specific fixed effects without identifying their sources. The standard method for disentangling fixed effects is to estimate simple cross-sectional regressions of the estimated intercepts against any number of fixed sociological, geographic, economic, and/or cultural variables. However, because of the small number of cross-sectional units in our present study, we are very limited in this regard, although this is more or less what we have done in testing whether racial composition plays a part.

Cross-region correlation

Recall that Model I allows for the disturbance terms to be correlated across regions, which can be due to, among other factors, spatial autocorrelation and a shared response to shocks. So as to test the statistical importance of this, we estimated Model III, which restricts the cross-regional correlation to zero, as in classical regressions. A visual inspection of the results of Models I and III in Table 4 reveals substantial differences in their estimates. Further, a likelihood ratio test with $\chi^2(45) = 182.28$ and a critical value of 61.37 at the 95% confidence level rejects

[12] Table 18 in Campbell and Daly (1992) reports a self-employment rate of 13% for Whites, 7.2% for West

the null that the models are statistically the same. We therefore conclude that cross-region correlation should be accounted for in estimating regional self-employment relationships. Table 8 presents the matrix of regional correlations from Model I. So as to test whether the primary source of these correlations is spatial autocorrelation, we regressed their absolute values against a contiguity dummy variable that takes the value of 1 if two regions are contiguous, and zero if they are not.[13] As Anselin (1988) describes, the assignment of spatial weights is unavoidably arbitrary, and there are many acceptable methods. We have chosen the simplest of these methods, which assumes that spatial relationships do not extend beyond contiguous regions.[14] Given the relatively large size of our spatial units, this is not unreasonable. Besides, as our purpose is limited to a simple test of whether the cross-region correlation is due primarily to spatial autocorrelation, this is likely to be sufficient. The Ordinary Least Squares results are

$$\begin{array}{lcll} |\text{cross-region corr. coeff.}| & = & 0.4315 & - \ 0.0611 \text{ Contiguity;} \\ & & (9.631) & (0.838) \end{array}$$

where t-statistics are in parentheses, and $R^2 = 0.016$. Because the coefficient on the contiguity dummy is statistically no different from zero, we conclude that the cross-region correlation coefficients are not primarily due to spatial autocorrelation. Of course, this is not to say that spatial autocorrelation is not present or important, only that there are other sources of cross-region correlation that are statistically more important.

V. Conclusions

Indian/Guyanese, 20.2% for Indians, and 21.9% for Pakistani/Bangladeshi, averaged over 1989-91.

[13] Of the 45 regional pairings, there are 17 instances for which regions are contiguous.

[14] The results are unchanged when the contiguity vector is replaced with a vector of the lengths of regional borders.

We developed a theoretical model of regional entrepreneurship, as measured by regional self-employment rates, and estimated the roles of labour-market conditions, labour-force characteristics, industry composition, and fixed region-specific factors. Our results suggest that all of these factors are important in explaining differences in regional entrepreneurship, and that regional heterogeneity and regional cross-correlation must be accounted for when estimating regional self-employment relationships. In particular, we find that regional factors such as regional entrepreneurial human capital play a significant role in determining why some British regions are more entrepreneurial than others. For most regions their effect is between 20% and 80% the size of the influence of all other factors combined, but for the North it is 14 times as important.

References

Acs, Z.J., D.B. Audretsch, and D.S. Evans, 1994, 'Why Does the Self-Employment Rate Vary Across Countries and Over Time?', CEPR Discussion Paper No. 871.

Anselin, L., 1988, *Spatial Econometrics: Methods and Models*, Kluwer.

Blanchflower, D.G. and A.J. Oswald, 1998, 'What Makes and Entrepreneur?', *Journal of Labour Economics 16,* 26-60.

Blau, D.M., 1987, 'A Time-Series Analysis of Self-Employment in the United States', *Journal of Political Economy 95,* 445-467.

Campbell, M. and Daly, M., 1992, 'Self-Employment into the 1990s', *Employment Gazette,* 269-292.

Cowling, M. and P. Mitchell, 1997, 'Trends in Self-Employment: Government Policy and the Role of the Macroeconomy', *The Manchester School 65,* 427-442.

De Wit, G., 1993, 'Models of Self-Employment in a Competitive Market', *Journal of Economic Surveys 7,* 367-397.

De Wit. G. and F. Van Winden, 1990, 'An Empirical Analysis of Self-Employment in the Netherlands', *Economics Letters 32,*97-100.

Dunn, T. and D. Holtz-Eakin, 1996, 'Financial Capital, Human Capital, and the Transition into Self-Employment: Evidence from Intergenerational Links', Metropolitan Studies Program Series Occasional Paper No. 181, Syracuse University.

Evans, D.S. and B. Jovanovic, 1989, 'An Estimated Model of Entrepreneurial Choice under Liquidity Constraints', *Journal of Political Economy 97,* 808-827.

Georgellis, Y. and H.J. Wall, 1999, 'Gender Differences in Self-Employment: Panel Evidence from the Former West Germany', *Federal Reserve Bank of St. Louis Working Paper 99-*

008.

Greene, W.H., 1997, *Econometric Analysis*, Prentice-Hall.

Knight, F.H., 1921, *Risk, Uncertainty, and Profit,* Houghton Mifflin.

Lentz, B.S. and D. Laband, 1990, 'Entrepreneurial Success and Occupational Inheritance among Proprietors', *Canadian Journal of Economics 23,* 563-579.

Nooteboom, B., 1994, 'Innovation and Diffusion in Small Firms: Theory and Evidence', *Small Business Economics 6*, 327-347.

Parker, S.C., 1996, 'A Time Series Model of Self-Employment under Uncertainty', *Economica 63*, 469-475.

Rees, H. and A. Shah, 1986, 'An Empirical Analysis of Self-Employment in the UK', *Journal of Applied Econometrics 1,* 95-108.

Reward Group, 1995, *Cost of Living Report: Regional Comparisons*, Reward Group.

Robson, M.T., 1991, 'Self-Employment and New Firm Formation', *Scottish Journal of Political Economy 38*, 352-368.

_____, 1998a, 'The Rise in Self-Employment Amongst UK Males', *Small Business Economics 10*, 199-212.

_____, 1998b, 'Self-Employment in the UK Regions', *Applied Economics 30,* 313-322.

Schiller, B.R. and P.E. Crewson, 1997, 'Entrepreneurial Origins: A Longitudinal Inquiry', *Economic Inquiry 35*, 523-531.

Taylor, M.P., 1996, 'Earnings, Independence or Unemployment: Why Become Self-Employed?', *Oxford Bulletin of Economics and Statistics 58,* 253-266.

_____, 1997, 'The Changing Picture of Self-Employment in Britain', Working Paper 97-9, ESRC Research Centre on Micro-Social Change, University of Essex.

Figure 1: *Self-employment relative to average of regions*,1978-95

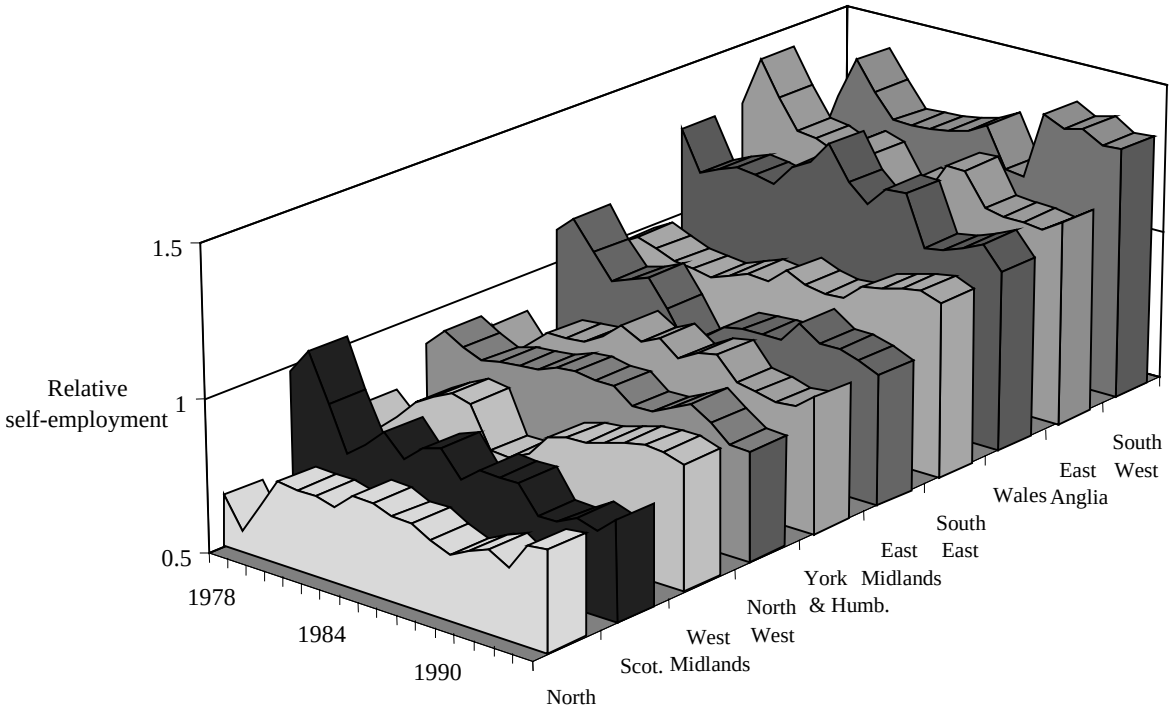

Table 1: *Relative self-employment rates for selected US states*,1980 and 1990

	1980	*1990*
Delaware	0.69	0.73
Nevada	0.77	0.80
New York	0.81	0.84
North Carolina	0.98	0.96
Tennessee	1.03	1.01
Kentucky	1.25	1.10
New Mexico	1.10	1.21
Kansas	1.55	1.36
Iowa	1.86	1.59
Montana	1.77	1.89
South Dakota	2.65	2.19

Source: US Census Bureau, *Labour force, Employment, and Journey to Work CENSUS*

Table 2: *Average self-employment rates, labour market variables, and labour force characteristics*

	self-emp. rate	*unemp. rate*	*real wage*	*share age 16-44*	*share age 45-retire*	*share retire+*	*female share*	*share A-level+*	*share no qualif.*
North	7.65	13.45	1.45	42.38	19.70	18.13	0.29	34.04	41.23
Yorks & Hum'side	9.93	10.76	1.42	42.82	19.06	18.24	0.30	34.87	41.41
East Midlands	10.26	8.72	1.39	43.30	19.14	17.69	0.32	34.05	41.30
East Anglia	13.14	6.87	1.41	42.53	18.67	19.17	0.31	32.66	40.71
South East	11.18	7.12	1.69	44.03	18.79	17.98	0.37	27.75	42.41
South West	13.85	7.80	1.40	41.34	19.12	21.06	0.34	28.83	42.12
West Midlands	9.10	10.96	1.40	43.02	19.47	17.19	0.30	37.21	40.49
North West	9.59	11.97	1.41	42.64	19.03	18.06	0.32	32.58	42.47
Wales	12.72	11.44	1.40	41.51	19.32	19.32	0.30	34.81	40.52
Scotland	7.73	10.83	1.27	39.64	17.40	15.83	0.33	31.19	38.41
Mean of Regions	10.52	9.99	1.42	42.32	18.97	18.27	0.32	32.80	41.11

Table 3: *Average employment shares: Standard Industrial Classifications*

	ag, for, fishing	*energy, water*	*metals, chems*	*vehicles, engineer*	*other manuf*	*cons-truction*	*distrib, hotels*	*trans, comm*	*finance services*	*public admin*
North	0.9	3.5	4.4	8.4	7.2	4.2	14.8	4.1	5.4	24.3
Yorks & Hum'side	1.1	3.4	3.9	7.4	9.9	3.8	16.5	4.4	5.9	22.4
East Midlands	1.5	3.5	3.1	9.6	13.2	3.3	15.4	4.0	5.3	21.2
East Anglia	3.5	1.2	2.2	8.0	9.2	3.7	17.3	5.5	6.8	22.5
South East	0.7	1.2	1.8	7.8	5.9	3.4	16.9	6.3	12.4	25.6
South West	2.0	1.3	1.9	8.9	6.8	3.4	18.1	4.1	7.7	24.0
West Midlands	1.1	1.6	4.2	15.8	6.9	3.4	14.9	3.6	6.6	20.9
North West	0.5	1.7	3.4	9.2	9.2	3.5	16.1	4.6	6.9	23.2
Wales	1.7	3.1	4.4	7.3	6.4	3.6	14.5	3.8	5.3	25.4
Scotland	1.3	2.3	1.8	6.6	7.2	4.7	14.6	4.2	5.9	23.0
Mean of Regions	1.4	2.3	3.1	8.9	8.2	3.7	15.9	4.5	6.8	23.3

Table 4: *GLS results*, Dependent variable: log of relative self-employment rate; All variables relative to average of regions

	Model I			**Model II**			**Model III**		
	coefficient	*s.e.*	*t statistic*	*coefficient*	*s.e.*	*t statistic*	*coefficient*	*s.e.*	*t statistic*
Labour market and labour force variables									
Unemployment rate	0.595	0.077	7.721	0.680	0.073	9.321	0.709	0.231	3.070
Square of unemployment rate	-0.281	0.033	-8.583	-0.418	0.028	-15.023	-0.334	0.104	-3.212
Real wage	-1.674	0.981	-1.706	-13.789	1.596	-8.642	-1.685	2.832	-0.595
Square of real wage	0.651	0.482	1.351	6.594	0.788	8.365	0.596	1.351	0.442
% aged 16-44	-0.295	0.186	-1.590	-1.671	0.049	-34.088	-0.654	0.422	-1.551
% aged 44-retirement	-1.095	0.329	-3.326	-1.960	0.282	-6.950	-1.474	0.701	-2.102
% older than retirement	0.231	0.177	1.305	0.967	0.060	16.117	0.520	0.389	1.337
% A-level or higher	-0.139	0.045	-3.062	-0.047	0.023	-2.045	-0.209	0.106	-1.976
% no qualification	0.070	0.037	1.871	0.161	0.068	2.376	0.108	0.154	0.700
% female	-0.299	0.135	-2.208	-1.918	0.125	-15.383	-0.253	0.474	-0.534
Industry composition									
Energy and water supply	-0.020	0.016	-1.295	-0.115	0.012	-9.631	-0.002	0.044	-0.049
Metals, minerals, chemicals	0.042	0.021	1.978	-0.065	0.017	-3.805	0.081	0.046	1.755
Metal goods, vehicles, etc	-0.175	0.037	-4.671	-0.148	0.019	-7.780	-0.220	0.087	-2.523
Other manufacturing	-0.301	0.046	-6.524	-0.013	0.021	-0.629	-0.383	0.091	-4.212
Construction	-0.019	0.022	-0.859	-0.238	0.012	-19.335	-0.030	0.056	-0.532
Distribution, hotels, repairs	0.122	0.051	2.375	0.125	0.037	3.402	0.150	0.101	1.483
Transport and communications	-0.029	0.031	-0.933	-0.131	0.018	-7.430	-0.014	0.068	-0.202
Financial services	0.043	0.039	1.112	-0.116	0.033	-3.486	0.032	0.086	0.373
Public administration	0.110	0.044	2.490	0.388	0.060	6.423	0.109	0.102	1.066
Region-specific intercepts									
North	3.218	0.515	6.251	-	-	-	3.679	1.667	2.207
Yorkshire & Humberside	3.428	0.506	6.780	-	-	-	3.901	1.660	2.350
East Midlands	3.640	0.496	7.344	-	-	-	4.183	1.667	2.509
East Anglia	3.663	0.495	7.403	-	-	-	4.154	1.648	2.521
South East	3.455	0.490	7.054	-	-	-	4.000	1.690	2.367
South West	3.656	0.497	7.362	-	-	-	4.110	1.631	2.520
West Midlands	3.410	0.489	6.978	-	-	-	3.932	1.669	2.355
North West	3.425	0.499	6.868	-	-	-	3.927	1.665	2.358
Wales	3.555	0.512	6.944	-	-	-	3.970	1.636	2.426
Scotland	3.307	0.502	6.584	-	-	-	3.832	1.681	2.280
Common intercept	-	-	-	12.711	0.675	18.832	-	-	-
log-likelihood	308.216			281.409			217.076		

Table 5: *Covariance matrix,* Model I

	North	*Yorks & Humb*	*East Midlands*	*East Anglia*	*South East*	*South West*	*West Midlands*	*North West*	*Wales*	*Scotland*
North	1.9E-04									
York & Humb	4.7E-05	1.1E-03								
East Midlands	-2.4E-04	-9.2E-04	1.9E-03							
East Anglia	-6.2E-04	7.3E-04	-1.7E-04	3.1E-03						
South East	6.1E-05	8.6E-05	-3.6E-04	-2.2E-04	2.9E-04					
South West	2.9E-04	-1.4E-03	5.0E-04	-2.3E-03	7.0E-04	5.3E-03				
West Midlands	-4.4E-04	-1.4E-03	1.5E-03	5.5E-04	-3.5E-04	7.4E-04	2.9E-03			
North West	1.3E-04	-1.6E-06	-1.7E-04	-5.1E-04	1.2E-04	4.5E-04	-3.3E-04	1.7E-04		
Wales	6.2E-05	1.6E-03	-1.6E-03	1.1E-03	-7.0E-05	-2.7E-03	-1.8E-03	-1.3E-04	3.4E-03	
Scotland	-5.8E-05	4.8E-04	9.6E-05	7.3E-04	-4.1E-04	-1.9E-03	-3.6E-04	-2.8E-04	9.6E-04	1.5E-03

Table 6: *The relative importance of region-specific effects*

	average relative self-emp. - 1 (1)	region intercept - average intercept (2)	effect of other factors (1) - (2)	relative import of region effect (2)/[(1)-(2)]
North	-0.276	-0.258	-0.018	14.08
Scotland	-0.192	-0.048	-0.144	0.33
West Midlands	-0.144	0.164	-0.308	-0.53
North West	-0.094	0.187	-0.281	-0.67
Yorks & Humb	-0.062	-0.021	-0.041	0.50
East Midlands	-0.032	0.180	-0.212	-0.85
South East	0.053	-0.066	0.119	-0.55
Wales	0.202	-0.051	0.253	-0.20
East Anglia	0.240	0.079	0.161	0.49
South West	0.306	-0.169	0.475	-0.36

Table 7: *Regional minority population shares,* 1991

	Black Caribbean, Black African, Black other	Indian, Pakistani, Bangladeshi
North	0.1	0.7
York & Humb	0.7	3.0
East Midlands	1.0	3.0
East Anglia	0.7	0.7
South East	3.5	4.0
South West	0.5	0.4
West Midlands	2.0	5.4
North West	0.7	2.3
Wales	0.3	0.5
Scotland	0.2	0.6

Source: *Regional Trends, 1993.*

Table 8: *Matrix of regional correlations,* Model I

	North	*Yorks & Humb*	*East Midlands*	*East Anglia*	*South East*	*South West*	*West Midlands*	*North West*	*Wales*	*Scotland*
North	1.000									
York & Humb	0.102	1.000								
East Midlands	-0.389	-0.630	1.000							
East Anglia	-0.805	0.392	-0.068	1.000						
South East	0.261	0.154	-0.482	-0.232	1.000					
South West	0.287	-0.595	0.156	-0.570	0.571	1.000				
West Midlands	-0.598	-0.799	0.628	0.185	-0.382	0.190	1.000			
North West	0.748	-0.004	-0.307	-0.702	0.541	0.482	-0.478	1.000		
Wales	0.076	0.846	-0.629	0.346	-0.071	-0.637	-0.566	-0.165	1.000	
Scotland	-0.107	0.373	0.056	0.338	-0.618	-0.663	-0.172	-0.555	0.422	1.000