1402782312

SWP 31/89 DOES LOCATION MATTER? COUNTY-SCALE VARIATIONS IN NEW FIRM FORMATION RATES IN GREAT BRITAIN

PAUL WESTHEAD
Cranfield Entrepreneurship Research Centre
Cranfield School of Management
Cranfield University
Cranfield
Bedford MK43 0AL
United Kingdom

Tel: +44 (0)1234 751122
Fax: +44 (0)1234 781806

and

A.MOYES
Dept of Geography
University College of Wales
Aberystwyth
Dyfed SY23 3DB
Wales

Tel: 0970 623111

Copyright: Moyes and Westhead 1989

Table 2.

Correlation coefficients between start formation rate and selected independent variables (n = 63).

		Pearson Correlation	Adjusted R^2	Significance of 'r' : t values	Slope (b)
Independent variables (x):					
(a)	Rurality				
1.	Area/Population	0.26	0.05	2.14*	0.00
†2.	1971-1981 Population Change (POPCH)	0.49	0.23	4.38***	0.31
(b)	Industrial Structure				
3.	% of total employees employed in easy-entry Industries, 1981 (SICs 46, 47 + 49)	0.19	0.02	1.52	58.21
4.	% of total employees employed in easy-entry industries, 1981 (SICs 31, 32, 46, 47, 49)	0.04	0.00	0.35	5.70
†5.	% of total employees employed in heavy industries, 1981 (SICs 11, 14, 21, 22, 23, 24, 25) (HEAVY)	-0.56	0.30	-5.28***	-44.57
6.	% of total employees employed in heavy industries, 1981 (14, 21, 22, 23, 24, 25)	-0.41	0.15	-3.48***	-44.78
†7.	% of total employees employed in Commerce, Retailing + Wholesaling, 1981 (SICs 61, 64/65, 81, 82, 83, 84) (COMMERCE)	0.51	0.25	4.67***	69.35
†8.	% of total employees employed in mining and quarrying, 1981 (SICs 11, 21 + 23)	-0.41	0.16	-3.53***	-58.03
9.	% of total manufacturing employees (excluding SIC 11) employed in easy entry industries, 1981 (SICs 46, 47, + 49)	-0.08	0.00	-0.59	-0.26
10.	% of total manufacturing employees (excluding SIC 11) employed in easy entry industries, 1981 (SICs 31, 32, 46, 47, 49)	-0.05	0.00	-0.42	-0.11
11.	% of total manufacturing employees (excluding SIC 11) employed in heavy industries, 1981 (SICs 14, 21, 22, 23, 24, 25)	0.02	0.00	-0.12	-0.07
12.	% of production employees employed in easy entry industries, 1981 (SICs 46, 47, 49)	0.08	0.00	0.67	0.93
13.	% of production employees employed in easy entry industries, 1981 (SICs 31, 32, 46, 47, 49)	0.14	0.00	1.09	0.87
14.	% of production employees employed in heavy industries, 1981 (SICs 11, 14, 21, 22, 23, 24, 25)	-0.29	0.07	-2.38*	-0.00
15.	% of total employment in agriculture, 1978	0.23	0.04	1.88	0.30
16.	% of total employment in manufacturing, 1978	-0.44	0.13	-3.85***	-0.22
†17.	% of total employment in production, 1978	-0.55	0.29	-5.14***	-0.25
†18.	% of total employment in services, 1978	0.57	0.32	5.48***	0.29
19.	% of total employment in agriculture, 1981	0.10	0.01	0.77	0.03
20.	% of total employment in manufacturing, 1981	-0.39	0.14	-3.28**	-0.22
21.	% of total employment in production, 1981	-0.55	0.29	-5.10***	-0.27
(c)	Agglomeration.				
22.	Total employment, 1978 as a proportion of total land area (ha.)	0.05	0.00	0.37	0.00
23.	Total production employment, 1978 as a proportion of total land area (ha.)	-0.09	0.00	-0.66	-0.25
(d)	Plant Size of Manufacturing Establishments				
†24.	% of manufacturing employment in establishments, 1979 with 11-19 employees (SP20)	0.64	0.40	6.53***	131.81
25.	% of manufacturing employment in establishments, 1979 with 11-50 employees	0.59	0.34	5.74***	48.11
†26.	% of manufacturing employment in establishments, 1979 with >500 employees (LP500)	-0.42	0.17	-3.64***	-12.13
(e)	Socio-Economic Structure/Occupational Structure				
27.	% of total employees in employment in S.E.G.s 1 & 2 (managerial + professional, 1981)	0.08	0.00	0.66	3.10

Abstract

Information about variations in the rate of new firm formation at the county level in Great Britain has recently become available. An opportunity has therefore arisen to explore the correlates of new firm formation rates. The paper begins by analysing the relative importance of factors thought to underly the new firm formation process. Secondly, correlation and regression analysis are used to winnow out those surrogate variables which seem to be most powerfully associated with variations in new firm formation rates from county to county. Finally, an attempt will be made to define 'ecological incubator environments' for new production business formation in Great Britain. From this it is possible to differentiate and classify the counties according to their incubator characteristics, and identify successful and unsuccessful counties.

SHEFFIELD UNIVERSITY LIBRARY

INTRODUCTION

New firm formation has become an important issue in economic policy in Britain and elsewhere in recent years. A high rate of new firm formation in an area is testimony to the area's enterprise culture. Traditional regional economic policies are being supplemented if not even supplanted by measures to popularise such an enterprise culture, and encourage new business formation particularly in regions with the weakest economies.

The personal characteristics of those individuals who do set up businesses, such as motivation and ambition, are difficult to entangle from broader environmental conditions which may have fostered or inhibited their ambitions, such as learned value-systems, formal education and perceived social and business contacts. The concept of environment, though hazy, consists of at least two elements : (i) the web of experiences and contacts provided by the incubator organisation within which the new firm founder is operating prior to start-up, and (ii) a wider socio-economic environment containing those factors which varyingly allow the business to develop. It should be possible to identify territorially-bounded areas where characteristics which are thought to have a bearing on new firm formation are present to a distinctive extent.

better. They broadly support the presumption that certain features of counties' employment and occupational structures are positively associated with high formation rates, particularly self-employment, employment in small firms, and home-ownership. The most closely-related variable was clearly a county's 'stop' rate : the number of businesses of all kinds ceasing to be registered for VAT purposes during 1980-3 per thousand production workers. High 'birth' rates often accompanied (and perhaps stimulated) high 'death' rates. Some expected relationships were not sustained, however; for example the 'enhancing' effect of a high percentage of Managers and Professionals and 'inhibiting' effect of a high percentage of manual workers detected by Gould and Keeble (1984) at a regional scale were not found here.

Because the initial variables deliberately included some near-duplicates within the ten domains of interest, a hard core of sixteen independent variables was picked out which represented those most strongly-associated with the formation rate (Y) in each of the ten domains, at at least the 0.01 level. They are marked (†) in Table 2. Variables discarded were normally instances where the same variable applied to two different time-periods, in which case the strongest relationship was retained. Full listings of the regression equations and test statistics not included here for reasons of space are available from the authors.

Inevitably, the individual variables are by no means independent of each other. A formal backward elimininination method of multiple regression can be used to establish the joint interaction of the variables; the most powerful regression statement to emerge from this procedure was (t values in parenthesis):

$$Y = 2.889 + 0.753STOP + 0.161POPCH + 0.455SP20 - 0.047LOCAL$$
$$(3.933)\quad (16.74)^{***}\quad (6.69)^{***}\quad (5.97)^{***}\quad (-2.72)^{**}$$

There is no shortage of plausible factors representing the personal characteristics referred to in (i) above, and characteristics of the socio-economic environment into which new businesses are born. Compiled from a range of studies, Table 1 shows fourteen factors which have been deemed important in promoting, or impeding new firm formation. These factors can be captured numerically by surrogate variables. It should hence be possible to identify where new firm formation prospects are the most, and least favourable from their distinct blends of environmental conditions.

New firm formation rates vary between countries, because national value-systems, institutional frameworks and economic structures confer a distinct character on the environment within which businesses operate. The environment at this scale can meaningfully be talked of in territorially-bounded terms. Translation of the concept of "environment" into a form appropriate to the regional scale is not easy. The stimuli which might have brought about the formation of many businesses are no longer highly localised. Familial, social and business information networks may well be national in scale, if not broader. Some potential founders may spend much of their lives in one area and move to another in order to set up a new business. But most new firms are presumed to have been conceived where founders were living at the time, and locational choice restricted very much to founders' local areas, certainly the areas within which they would be expected to search for work. Travel-to-work Areas (TTWAs) have been defined to isolate areas within which most people both live and work. Whilst they have a strong rationale, evidence from Wales suggests that TTWAs are not consistently effective in defining areas within which new firm founders both live and are prepared to set up new businesses (Westhead, 1988). Some of this lack of self-containedness arises from the fact that white-collar and blue-collar TTWAs differ

Table 1

Factors identified by the New Firm Research Literature which are associated with either promoting or impeding New Firm Formation

Factor	Surrogate Variables	Promoting/ Impeding
1. Size of 'incubator' firm	% of total employment in plants employing < 10 persons	Promoting
	% of total employment in plants employing > 500 persons	Impeding
2. Occupational experience	% of population in managerial and professional groupings	Promoting
	% of population in manual groupings	Impeding
	% of population self-employed	Promoting
3. Education	% of population with higher degrees	Promoting
4. Access to capital	Savings per head of population	Promoting
	House-owning population	
	Local-authority renting population	Impeding
5. Entry into industry	% of population in industries with low entry barriers	Promoting
	% of population in heavy industry	Impeding
	% of population in mining and quarrying industries	
6. Market demand	Regional income distribution	Promoting
	Rate of change in production employment growth	Promoting
	Rate of change in total employment growth	Promoting
7. Degree of local autonomy	% of total manufacturing employment in indigenous plants	Promoting
8. Age of investment	% of total manufacturing employment in 'young' plants	Promoting
9. Turbulence	Rate of employment loss in manufacturing plant closures	Promoting
10. Industrial specialisation	High Tress specialisation statistic	Promoting
11. Premises	Availability and low cost of premises	Promoting
12. Non-production entrepreneurship	% of total employment in services	Promoting
	% of total employment in commerce, retailing and wholesaling	Promoting
13. Unemployment	% level of unemployment	Promoting/ Impeding
	% change in the rate of unemployment	Promoting/ Impeding
14. Rurality	% of population living in towns of over 5,000 population	Impeding

Sources: adapted from Cross (1981); Gould and Keeble (1984); Lloyd and Mason (1984); O'Farrell and Crouchley (1984); Storey (1982); Storey and Jones (1987).

(Coombes et al., 1988). Counties appear to be an appropriate areal framework, more self-contained than TTWAs, and have the advantage of being generally perceived more clearly as territorial units by those in business. The 63 sub-areas to be used in this study are mapped in Fig. 1.

ANALYTICAL FRAMEWORK

As suggested in Table 1, based on industrial location theory and on previous empirical analysis of new firm founders and formation, many factors potentially bearing on new firm formation can be listed and most can be conveyed in the form of surrogate variables. But the more are considered, the less easy it may be to generalise about their individual and collective importance, particularly as they vary across geographic space. Given a consistent and logical selection of initial variables, univariate and multivariate correlation techniques allow the numbers of variables that ought to be considered to be reduced to manageable dimensions.

DATA

The analysis is based on data on the registration of businesses for VAT purposes between 1980 and 1983 (Ganguly, 1985), using data relating to production industries (dominantly businesses in manufacturing). For the purposes of this study, the recording of a new production business for VAT purposes was treated as the equivalent of the formation of a new manufacturing firm. The spatial scale of the analysis is of the counties of Great Britain, with the exception that the Highlands and Islands have been aggregated to form a single unit (see Fig. 1). Variations in he number of VAT 'starts' standardised to a rate per thousand production workers per county, 1981 are indicated in Figure 2; the rates are the dependent variable to be analysed.

The independent surrogate variables used in the following regression analysis apply to a variety of time-periods. However, as far as possible they relate logically to the 1980-83 period to which the dependent variable applies. The main data sources were the 1981 Population Censuses of England, Wales and Scotland, particularly the County Reports, County Monitors and Economic Activity Leaflets; the Census of Production (Business Monitor, PA1003, 1979), and the Central Statistical Office publication Regional Trends. A total of 57 variables was extracted covering the following themes: "rurality" (population density and change); industrial structure (pointing towards barriers to entry in activities characteristic of each county); "push" factors concerned with employment decline; market demand (mass) factors; agglomeration; personal income; plant size; socio-economic and occupational structures, and educational qualifications. In some cases, information for the same characteristic but for different dates could be included, since there was no a priori case to prefer data for one date than another. Because of the technical requirements of some of the multivariate procedures subsequently used, many variables have been converted into the form of ratios and/or percentages.

RESULTS OF THE CORRELATION AND REGRESSION ANALYSES

Table 2 shows the relationship between the dependent variable (the new firm formation rate) and each of 57 independent variables, across the 63 counties; they are grouped into ten headings consonant with new firm literature proposals. They were obtained using the statistical analysis package SPSSX, and show the strength (r), direction (±) and statistical significance of the association, with the regression parameter (b). Most of the correlation coefficients were in the direction hypothesised in Table 1. Twenty six of the independent variables were associated with the dependent variable at the 0.05 level of statistical significance or

Table 2.

Correlation coefficients between start formation rate and selected independent variables (n = 63).

		Pearson Correlation	Adjusted R^2	Significance of 'r' : t values	Slope (b)
Independent variables (x):					
(a)	Rurality				
1.	Area/Population	0.26	0.05	2.14*	0.00
†2.	1971-1981 Population Change (POPCH)	0.49	0.23	4.38***	0.31
(b)	Industrial Structure				
3.	% of total employees employed in easy-entry Industries, 1981 (SICs 46, 47 + 49)	0.19	0.02	1.52	58.21
4.	% of total employees employed in easy-entry industries, 1981 (SICs 31, 32, 46, 47, 49)	0.04	0.00	0.35	5.70
†5.	% of total employees employed in heavy industries, 1981 (SICs 11, 14, 21, 22, 23, 24, 25) (HEAVY)	-0.56	0.30	-5.28***	-44.57
6.	% of total employees employed in heavy industries, 1981 (14, 21, 22, 23, 24, 25)	-0.41	0.15	-3.48***	-44.78
†7.	% of total employees employed in Commerce, Retailing + Wholesaling, 1981 (SICs 61, 64/65, 81, 82, 83, 84) (COMMERCE)	0.51	0.25	4.67***	69.35
†8.	% of total employees employed in mining and quarrying, 1981 (SICs 11, 21 + 23)	-0.41	0.16	-3.53***	-58.03
9.	% of total manufacturing employees (excluding SIC 11) employed in easy entry industries, 1981 (SICs 46, 47, + 49)	-0.08	0.00	-0.59	-0.26
10.	% of total manufacturing employees (excluding SIC 11) employed in easy entry industries, 1981 (SICs 31, 32, 46, 47, 49)	-0.05	0.00	-0.42	-0.11
11.	% of total manufacturing employees (excluding SIC 11) employed in heavy industries, 1981 (SICs 14, 21, 22, 23, 24, 25)	0.02	0.00	-0.12	-0.07
12.	% of production employees employed in easy entry industries, 1981 (SICs 46, 47, 49)	0.08	0.00	0.67	0.93
13.	% of production employees employed in easy entry industries, 1981 (SICs 31, 32, 46, 47, 49)	0.14	0.00	1.09	0.87
14.	% of production employees employed in heavy industries, 1981 (SICs 11, 14, 21, 22, 23, 24, 25)	-0.29	0.07	-2.38*	-0.00
15.	% of total employment in agriculture, 1978	0.23	0.04	1.88	0.30
16.	% of total employment in manufacturing, 1978	-0.44	0.13	-3.85***	-0.22
†17.	% of total employment in production, 1978	-0.55	0.29	-5.14***	-0.25
†18.	% of total employment in services, 1978	0.57	0.32	5.48***	0.29
19.	% of total employment in agriculture, 1981	0.10	0.01	0.77	0.03
20.	% of total employment in manufacturing, 1981	-0.39	0.14	-3.28**	-0.22
21.	% of total employment in production, 1981	-0.55	0.29	-5.10***	-0.27
(c)	Agglomeration.				
22.	Total employment, 1978 as a proportion of total land area (ha.)	0.05	0.00	0.37	0.00
23.	Total production employment, 1978 as a proportion of total land area (ha.)	-0.09	0.00	-0.66	-0.25
(d)	Plant Size of Manufacturing Establishments				
†24.	% of manufacturing employment in establishments, 1979 with 11-19 employees (SP20)	0.64	0.40	6.53***	131.81
25.	% of manufacturing employment in establishments, 1979 with 11-50 employees	0.59	0.34	5.74***	48.11
†26.	% of manufacturing employment in establishments, 1979 with >500 employees (LP500)	-0.42	0.17	-3.64***	-12.13
(e)	Socio-Economic Structure/Occupational Structure				
27.	% of total employees in employment in S.E.G.s 1 & 2 (managerial + professional, 1981)	0.08	0.00	0.66	3.10

SHEFFIELD UNIVERSITY LIBRARY

		Pearson Correlation	Adjusted R^2	Significance of 'r' : t values	Slope (b)
28.	% of total employees in employment in S.E.G.s 1, 2 + 13 (managerial, professional + farmers), 1981	0.08	0.00	0.61	2.52
29.	% of total employees in S.E.G.s 9, 10 + 111 (manual), 1981	-0.15	0.01	-1.20	-1.38
†30.	% of total economically active people self-employed, 1981 (SELF)	0.64	0.40	6.53***	63.51
31.	% of total employed who are self-employed, 1981	0.60	0.35	5.82***	57.13
(f)	'Push' Factors				
32.	% unemployed, July 1979	-0.20	0.02	-1.56	-0.19
†33.	% unemployed, October 1983 (%UNEMPL)	-0.36	0.12	-3.03**	-0.44
†34.	Change in % unemployed between 1979-1983	0.40	0.15	3.43***	0.04
†35.	Stop rate per 1,000 production employees, 1981 (STOP)	0.89	0.79	15.28***	0.99
(g)	Market Demand				
36.	% change in total employment, 1975-1981	0.43	0.17	3.70***	0.20
†37.	% change in total employment, 1978-1981	0.47	0.21	4.15***	0.35
38.	% change in manufacturing + utilities employment, 1975-1981	0.38	0.13	3.20**	0.12
39.	% change in manufacturing + utilities employment, 1978-1981	0.35	0.11	2.90**	0.16
(h)	Personal Income				
†40.	% owner occupiers, 1981 (OWNER)	0.48	0.22	4.27***	0.18
†41.	% of dwellings rented from local authority, 1979 (LOCAL)	-0.63	0.39	-6.41***	-0.26
42.	Domestic average rateable value, 1979	-0.04	0.00	-0.27	-382.34
43.	Average domestic rates per hereditament, 1979 per person, 1981	0.01	0.00	0.10	223.93
44.	Average male weekly earnings as per male in employment, 1981	0.07	0.00	0.56	22.58
45.	Average female weekly earnings per female in employment, 1981	0.14	0.00	1.12	39.32
46.	GDP per head 1979 per employee in employment, 1978	0.06	0.00	0.49	4.37
(i)	Educational Qualifications				
†47.	% school leavers with no graded results, 1981 (GRADES)	-0.38	0.13	-3.24**	-0.20
48.	% pupils >16 years old staying at school, 1979	-0.10	0.00	-0.76	-0.00
49.	% school leavers with 1 or more 'A' levels, 1981	-0.10	0.00	-0.81	-0.07
50.	% male employees with higher educational qualifications, 1981	0.29	0.06	2.33*	0.39
51.	% female employees with higher educational qualifications, 1981	0.27	0.06	2.18*	0.62
(j)	Floorspace Availability				
52.	Total non-domestic floorspace, 1979 per resident, 1981	-0.38	0.13	-2.99**	-0.66
53.	Total non-domestic floorspace, 1979 per hectare	-0.08	0.00	-0.54	-0.00
54.	Total floorspace, 1979 per production employee, 1981	0.36	0.11	2.80**	0.09
55.	Total industrial floorspace, 1979 per resident, 1981	-0.52	0.26	-4.44***	-1.08
56.	Total industrial floorspace, 1979 per hectare	-0.18	0.01	-1.29	-0.01
57.	Total industrial floorspace, 1979 per production employee, 1981	-0.28	0.06	-2.11*	-0.12

Notes: * Significant at $p < 0.05$;
** Significant at $p < 0.01$;
*** Significant at $p < 0.001$.

Sources: Population Census, 1981; Central Statistics Office; Census of Production; British Business; and the Department of Employment.

better. They broadly support the presumption that certain features of counties' employment and occupational structures are positively associated with high formation rates, particularly self-employment, employment in small firms, and home-ownership. The most closely-related variable was clearly a county's 'stop' rate : the number of businesses of all kinds ceasing to be registered for VAT purposes during 1980-3 per thousand production workers. High 'birth' rates often accompanied (and perhaps stimulated) high 'death' rates. Some expected relationships were not sustained, however; for example the 'enhancing' effect of a high percentage of Managers and Professionals and 'inhibiting' effect of a high percentage of manual workers detected by Gould and Keeble (1984) at a regional scale were not found here.

Because the initial variables deliberately included some near-duplicates within the ten domains of interest, a hard core of sixteen independent variables was picked out which represented those most strongly-associated with the formation rate (Y) in each of the ten domains, at at least the 0.01 level. They are marked (†) in Table 2. Variables discarded were normally instances where the same variable applied to two different time-periods, in which case the strongest relationship was retained. Full listings of the regression equations and test statistics not included here for reasons of space are available from the authors.

Inevitably, the individual variables are by no means independent of each other. A formal backward elimininination method of multiple regression can be used to establish the joint interaction of the variables; the most powerful regression statement to emerge from this procedure was (t values in parenthesis):

$$Y = \underset{(3.933)}{2.889} + \underset{(16.74)\ ***}{0.753STOP} + \underset{(6.69)\ ***}{0.161POPCH} + \underset{(5.97)\ ***}{0.455SP20} - \underset{(-2.72)\ **}{0.047LOCAL}$$

The t values show that the most influential value in this expression was STOP, with population change, and the relative level of employment in small plants (11-20 employees) being positively associated with high new firm formation rates, and with a high proportion of Local Authority-owned property (a surrogate for personal wealth) being significantly and negatively associated with new firm formation. Multivariate classification is needed to clarify the relationships further.

A MULTIVARIATE CLASSIFICATION OF COUNTIES IN GREAT BRITAIN

The remainder of this paper applies the two-stage method of data reduction first popularised by Spence (1968). Each county has its own unique profile of scores on each of the sixteen variables identified as closely correlated with new firm formation rates. In principle, each county can be located at a single co-ordinate within 16-dimensional space. Counties can then be grouped within that 16-dimensional space, so that those 'close' to each other in multidimensional terms can be isolated : counties which have similar characteristics in terms of the new firm environmental variables. Principal Components Analysis can be used to reduce the number of dimensions involved, and then each county located within that simpler multi-dimensional space. The second stage of the analysis groups the counties into a smaller number of clusters whose members are more like each other than those in other clusters.

RESULTS OF THE PRINCIPAL COMPONENTS ANALYSIS (PCA)

A matrix of correlations of the sixteen selected 'incubator' variables across the 63 areas is presented in Table 3. Clearly, the variables are associated with each other, so that it is likely that they will share common components. An R-mode PCA was performed on these coefficients using the SPSSX program. The component model's key

Table 3

Correlation matrix

	1. POPCH	2. PROD	3. SERVICES	4. HEAVY	5. MINING	6. COMMERCE	7. SP20	8. LP500	9. SELF	10. UNEMPL	11. CHUNEMPL	12. STOP	13. OWNER	14. LOCAL	15. GRADES	16. TOTCH
1, POPCH	1.00															
2. PROD	-0.19	1.00														
3. SERVICES	0.05	-0.90***	1.00													
4. HEAVY	-0.21	0.43	-0.40	1.00												
5. MINING	-0.20	0.47	-0.47	0.22	1.00											
6. COMMERCE	-0.07	-0.47	0.73**	-0.29	-0.50*	1.00										
7. SP20	0.20	-0.61*	0.51*	-0.39	-0.17	0.25	1.00									
8. LP500	-0.32	0.36	-0.16	0.43	0.04	-0.02	-0.73**	1.00								
9. SELF	0.52*	-0.57*	0.36	-0.36	-0.27	0.03	0.64**	-0.63**	1.00							
10. %UNEMPL	-0.37	0.21	-0.25	0.48	0.30	-0.37	-0.17	0.27	-0.22	1.00						
11. CHUNEMPL	0.37	0.06	0.05	-0.17	-0.30	0.32	0.15	-0.29	0.21	-0.42	1.00					
12. STOP	0.20	-0.43	0.53*	-0.35	-0.39	0.60*	0.49	-0.25	0.40	-0.29	0.39	1.00				
13. OWNER	0.42	-0.04	0.19	-0.01	-0.19	0.29	0.07	-0.04	0.37	-0.25	0.39	0.39	1.00			
14. LOCAL	-0.51*	0.34	-0.38	0.27	0.30	-0.30	-0.30	0.25	-0.62**	0.38	-0.35	-0.48	-0.89***	1.00		
15. GRADES	-0.38	-0.01	-0.14	0.07	0.20	-0.37	0.08	-0.13	-0.19	0.36	-0.41	-0.31	-0.65**	0.57	1.00	
16. TOTCH	0.44	-0.54*	0.47	-0.47	-0.36	0.33	0.36	-0.36	0.38	-0.73**	0.18	0.36	0.13	-0.36	-0.17	1.00

* Significant at 0.05 level of significance.
** Significant at 0.01 level of significance.
*** Significant at 0.001 level of significance.

assumptions (Norusis, 1985) were satisfied, and in this and subsequent uses of PCA the sums of squares of the component loadings are all greater than one. In the resultant unrotated model, the first five components (out of a possible 16) explained 80.3% of the variance in the original 63-by-16 data set. However, as often happens with preliminary use of model, the component loadings provide confusing signals as to which variables are most closely associated with each component. Following common practice, a factor analytic varimax rotation procedure is used, so as to achieve a situation in which each of the original variables is highly correlated with only one factor rather than with many. The outcome is presented in Table 4.

In the rotated solution, the first five components also accounted for 80.3% of the initial total variance, but in comparison with the unrotated version the variance was more evenly distributed. The first component accounted for 20.4% of the total variance, the second 18.7%, the third 18.2% and so on. The final column of Table 4 records the extent to which variations in each of the original variables are captured by the five components (communality : h^2). The highest (PROD) approaches 94% communality; most of the other variables are well explained by the first five components but two in particular (HEAVY : 56% and MINING : 47%) are less successfully represented.

By scanning the variables which load most highly on the components, each component can be given a descriptive label. The first clearly loads highly and positively on SERVICES and COMMERCE and highly negatively on PROD; it is termed a 'Services and Commercial' component. Component two isolates 'high personal wealth and attainment' as indicated by a high positive loading on OWNER and high negative loadings on LOCAL and GRADES. Component three stresses 'smaller scale manufacturing and high self-employment', witness the high negative loading on LP500 and high positive loadings on SP20 and SELF. Component four loads highly on %UNEMPL and has

Table 4

Varimax rotated component matrix - principal components analysis

Variable	Varimax Rotated Components					Communality
	1	2	3	4	5	(h^2)
1. POPCH	-0.201	0.607	0.283	-0.494	-0.031	0.734
2. PROD	-0.706	-0.044	-0.472	0.252	0.389	0.939
3. SERVICES	0.886	0.094	0.251	-0.157	-0.197	0.921
4. HEAVY	-0.284	0.083	-0.375	0.563	-0.120	0.560
5. MINING	-0.598	-0.180	0.014	0.282	-0.048	0.472
6. COMMERCE	0.870	0.096	-0.079	-0.113	0.327	0.893
7. SP20	0.334	-0.008	0.860	-0.056	0.043	0.856
8. LP500	0.062	0.029	-0.880	0.245	-0.198	0.878
9. SELF	0.111	0.469	0.760	-0.172	-0.149	0.862
10. %UNEMPL	-0.177	-0.188	-0.010	0.829	-0.301	0.845
11. CHUNEMPL	0.065	0.318	0.138	-0.208	0.804	0.814
12. STOP	0.617	0.279	0.339	-0.032	0.371	0.712
13. OWNER	0.152	0.912	0.023	0.057	0.177	0.891
14. LOCAL	-0.251	-0.859	-0.267	0.142	-0.034	0.891
15. GRADES	-0.179	-0.739	0.244	0.149	-0.278	0.737
16. TOTCH	0.299	0.134	0.215	-0.828	-0.097	0.848
Eigenvalue	3.271	2.995	2.916	2.297	1.372	
% of Variance	20.4	18.7	18.2	14.4	8.6	
Cumulative % of Variance	20.4	39.1	57.3	71.7	80.3	

a high negative loading on TOTCH, leading to the suggested label of 'employment decline and high unemployment'. Finally component five, loading highly only on CHUNEMPL, could be termed 'high percentage increase in unemployment 1979-1982'. These labels simply state the main numerical characteristics of the components : too much causal significance should not be attached to them.

A 63-by-5 matrix of component scores can now be created which shows how each county performs on each of the five scales which the components represent. These component scores are mapped out on Figs. 3(a) to 3(e). In keeping with the Component 1 label, counties with the highest component scores on Fig. 3(a) are those generally urbanised ones with a strong commercial-services nexus : Greater London, East Sussex, South Glamorgan, Surrey, Merseyside, Avon and Essex. Fig. 3(b) highlights counties with high personal wealth and educational attainment (as measured by owner-occupation and GRADES); those which perform poorly on. Component 2 are all of the Metropolitan Counties, Greater London, and include the older industrial areas of central Scotland, the north east of England and industrial South Wales. Conversely the more rural counties of England and Wales generally score more highly. Fig. 3(c) shows the most peripheral counties with small-plant rather than large-plant manufacturing employment and high levels of self-employment, probably partly related to agriculture. Scores on Component 4, illustrated in Fig. 3(d) suggest that characteristics of high unemployment and large absolute employment decline were shared by both the older industrial counties (for example, Cleveland, West Midlands) and more peripheral and rural counties like Dyfed and Cornwall. Finally, map 3(e), dominated by the theme of percentage increase in unemployment 1979-1983 points to common characteristics along a London-Manchester axis, with only Borders Region also scoring strongly and positively on this component.

Each map thus shows how counties fare on scales constructed from variables which are taken as yardsticks of potential for new firm formation. Counties clearly have their own unique places on these five scales but intuitively there seem to be recurrent themes - urban, rural, core or periphery - which underly the individual distribution patterns. These four maps can now be synthesised.

GROUPING OF COUNTIES

Each county can now be located in five-dimensional space and then joined with the county with which it is closest in that space. The process continues until relatively homogeneous groups are formed which have the maximum between-group, and minimum within-group, variances. Ward's method of clustering (Ward, 1963) produces a sequence of fusions most conveniently displayed by means of a dendrogram (Fig. 4), which visually records the sequence of grouping starting with that of Cambridgeshire with Suffolk and culminating in the grouping of a mass of mainly Scottish sub-areas with the rest of Great Britain. Hence any number of groups can exist between 63 and 1.

As grouping proceeds the amount of detail in the original distance matrix is reduced; as each county is fused with another, so their pooled characteristics are fed into the next stage of the grouping procedure rather than their separate characteristics. The Error Sum of Squares contribution is a measure of that loss of detail. At early stages, little detail is lost but the rate of loss becomes exponential. The choice as to how many groups are optimal is subjective and dependent on the purposes of the exercise. But a reasonable break of slope can be envisaged at Stage 55 of the sequence. This gives eight groups of counties (as opposed to the original 63) for the loss of detail of 50.26% (bearing in mind that the component scores used derived from components which explained 80.3% of the original variance). The eight clusters are mapped in Fig. 5.

INTERPRETATION

Are the clusters distinct? Table 5 shows the average value for each variable within each of the clusters, and compares it with the grand mean for the 63 counties. Cases where the cluster means differ by at least half a standard deviation from the respective grand means highlight the distinguishing features of each cluster. They help description although, as before, should not be causally over-interpreted.

Cluster 1 contains thirteen members. The variables which distinguish the cluster most sharply from the average are their above-average population growth, and manufacturing dependence on large plants, correspondingly low levels of small-plant manufacturing employment, and below-average mining employment, unemployment, local authority housing and educational attainment. Fig. 5 shows the counties in this cluster as particularly in the eastern part of the South West region, in East Anglia, Cheshire, North Yorkshire and Cumbria. Cluster 2 in comparison strongly emphasises heavy industry, declining employment, high unemployment albeit increasing less fast than average, and large plants. This cluster include some English and Welsh steelmaking counties (Cleveland, Humberside, Clwyd, South and West Glamorgan), and also Merseyside and perhaps surprisingly Devon, perhaps affected by the industrial structure of the Plymouth area. Cluster 3 picks out counties with above-average increases in unemployment accompanied by high rates of business closure but nevertheless strong bases of population growth, below-average employment fall, and economies with over-endowment of services and commerce. All counties contiguous to London fell into this group, as did those fringing the West Midlands metropolitan county to south and west, with Lancashire, Dorset and the Isle of Wight completing this group of fifteen members.

The striking features of cluster 4 are severe population and job

Table 5. Characteristics of the Clusters.

Variables	Grand Mean	Standard Deviation	Coefficient of Variation	Clusters 1	2	3	4	5	6	7	8
1. POPCH	3.92	5.99	152.81	7.05*	0.84*	7.30*	-5.68**	3.14	9.00*	-2.33**	4.15
2. PROD	34.87	8.49	24.3	34.24	31.70	32.03	39.20*	46.51**	25.70**	34.80	27.40*
3. SERVICES	55.95	7.46	13.35	56.27	59.08	59.81*	55.90	46.62**	58.97	56.15	55.03*
4. HEAVY	4.45	3.45	77.53	3.52	8.18**	3.06	3.84	6.35*	3.84	3.96	2.32*
5. MINING	1.78	2.70	151.69	0.41*	1.08	0.42*	0.87	6.32**	2.85	2.53	0.32*
6. COMMERCE	19.08	2.81	14.73	18.95	19.26	21.31*	22.18**	16.54*	16.78*	18.32	16.88*
7. SP20	4.25	1.85	43.53	3.26*	3.66	4.70	4.97	3.21*	9.25***	3.29*	6.54*
8. LP500	42.40	13.24	31.23	49.86*	54.33*	38.36	40.10	43.73	10.99***	48.97	22.21**
9. SELF	10.16	3.84	37.80	10.53	9.61	11.55	7.74*	7.58*	21.15***	5.91**	11.76
10. %UNEMPL	12.64	3.14	24.84	10.29*	16.40**	11.14	13.25	13.46	10.63*	15.00*	10.90*
11. CHUNEMPL	117.61	39.06	33.21	109.16	85.81*	157.06**	143.09*	113.25	114.32	85.00*	92.61*
12. STOP	6.36	3.41	53.62	5.74	5.77	9.08*	9.28*	4.13*	8.36*	3.33*	5.11
13. OWNER	56.41	10.05	17.82	60.79	59.29	63.77*	54.58	56.10	60.90	34.57***	40.88**
14. LOCAL	28.47	9.10	31.96	23.75*	27.71	21.55*	32.85	31.97	21.23*	47.57***	35.00*
15. GRADES	14.50	7.31	50.41	9.60*	15.59	9.68*	14.88	14.76	15.87	25.82**	27.28**
16. TOTCH	-3.94	5.14	130.46	-1.42	-8.44*	-0.93*	-8.24*	-7.84*	-4.36	-6.59*	5.82**

Notes: (i) Cluster mean marked (*) deviate by more than half the standard deviation from the respective global mean.
(ii) Cluster mean marked (**) deviate by more than one standard deviation from the respective global mean.
(iii) Cluster mean marked (***) deviate by more than two standard deviations from the respective global mean.

loss, and accelerated unemployment but with a heavy bias towards commerce; it comprises Greater London, Manchester, West Yorkshire and West Midlands. In comparison cluster 5 isolates ten counties heavily involved in production industries and mining, with weak service employment, little small-plant manufacturing and self-employment and relatively low business stop rates. A solid phalanx of the East Midlands, South Yorkshire and Staffordshire is accompanied by Northumberland, Durham, Gwent and Mid Glamorgan in this group. Cluster 6 contains only Dyfed, Powys and Cornwall; production industries where they occur are of small-plants, self employment was prominent, overall unemployment was low and population increase high. Cluster 7 counties occupy central Scotland with Tyne and Wear as its only member elsewhere, characterised by falling population and employment but most strikingly by low personal wealth (via the housing variables) and self-employment, and low educational attainment. Finally, cluster 8 involves all the other, more rural Scottish sub-areas: on balance, these show exceptional employment growth, and presence of small manufacturing where it exists, and with little heavy industry, mining or commerce and poor educational attainment.

SUMMARY AND CONCLUSIONS

The mechanical process of classification and grouping has allowed each of the 63 sub-areas of Great Britain to be allocated to eight mutually-exclusive groups. Each group has been defined by the distinctive ways in which various new firm promoting or impeding factors are present within them. Many of the groupings of counties are in line with expectations: the coherence of all groups except perhaps Cluster 2 is remarkable. The separateness of Scotland (apart from Tyne and Wear's inclusion in Cluster 7) may well arise from the distinctiveness of housing tenure and educational attainment throughout Scotland. The

ultimate acid test is whether the groups bring together counties with similar new firm formation rates. Analysis of variance indeed suggests that the groups are decidedly non-random (Table 6). The clusters have a low internal variability of formation rates that could have arisen by chance only once in a hundred trials if the 63 counties were to be combined into other groups at random.

Table 6. Rates of new production registrations, 1980-83 by Cluster Type.

Cluster type	Number of counties	Mean ($\bar{x}$)	Standard deviation (σ)	Coefficient of variation
1	12	8.62	1.92	22.27
2	8	7.80	2.87	36.79
3	16	11.86	3.36	28.33
4	4	9.90	4.61	46.57
5	10	6.37	2.35	36.89
6	3	14.27	4.11	28.80
7	6	4.22	0.77	18.25
8	4	8.25	3.50	42.42

Analysis of variance (N-k) = 63-8 = 55; (k-1) = 8-1 = 7

F = 7.58 Reject Ho. Significant sub-groups at the 0.01 level of significance.

The characteristics of the groups can now be related to the mean formation rate. In relative terms the least fertile new production firm environment was Cluster 7 : the central Scottish-Tyne and Wear environment. Next came the English-South Wales coalfield environment (Cluster 5), followed by the dominantly steel closure environment of Cluster 2.

Scottish rural (Cluster 8) and the growing rural Wessex/East Anglia/Cheshire/imageable northern England group (Cluster 1) occupy a middle position. Moving towards the highest rates we then have the major English conurbations (Cluster 4), the ring of suburban counties around London, Birmingham and Manchester (Cluster 3) and finally the western peripheries with the highest rates (Cluster 6). Obviously the picture of absolute numbers of new firms would be different. But given the great variety of sizes of counties, some form of standardisation of the data is obligatory, and that chosen here can be defended as a realistic measure of relative business dynamism (within the limitations of the data used, of course).

From a policy-relevance point of view the results point towards: (1) the suggestion that particular policy packages may be spatially targeted depending on the "incubator" characteristics highlighted here; (2) that the groups form a logical areal sampling frame for case-based new firm research, subject to awareness of the ecological fallacy. From an academic standpoint it would be interesting to apply this methodology to recently released complete year-by-year 1980-1987 VAT registration and deregistration data to see how far the relationships explored here change over time, and secondly to explore performance from county to county for other than production industries. By so doing, the extent to which formation of new businesses in the economy at large followed similar lines could be established, though several practical and conceptual problems would need to be resolved before that aim could be achieved.

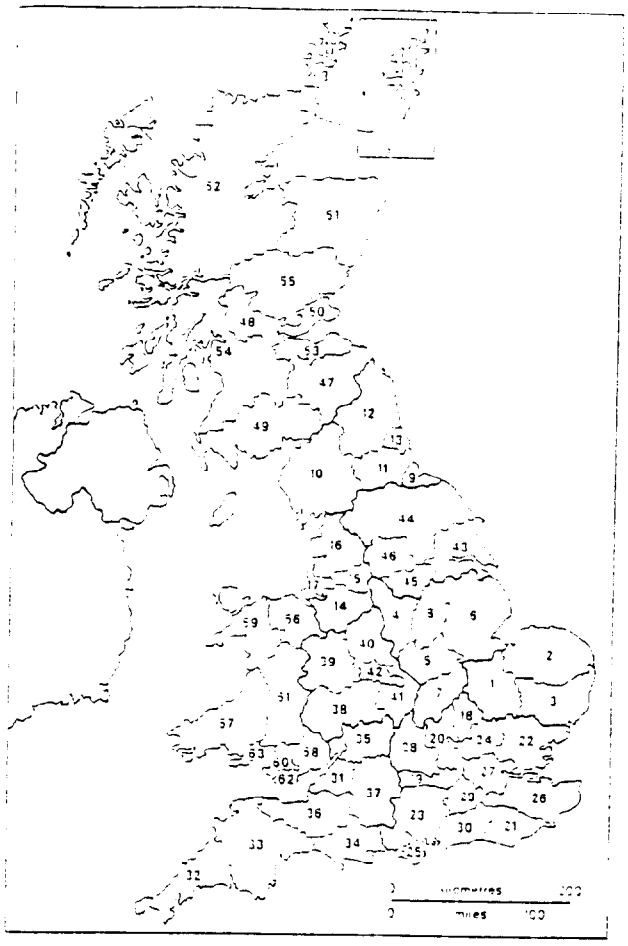
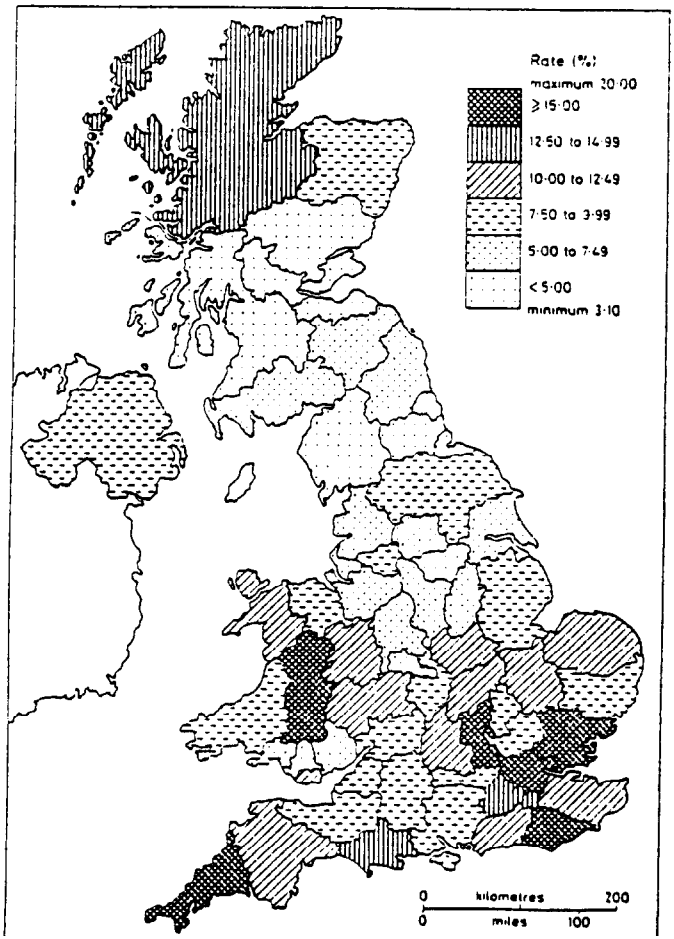


Fig. 1: Subdivisions



Key to Subdivisions

1. Cumbria
2. Norfolk
3. Suffolk
4. Derbyshire
5. Leicestershire
6. Lincolnshire
7. Northamptonshire
8. Nottinghamshire
9. Cleveland
10. Cumbria
11. Durham
12. Northumberland
13. Tyne and Wear
14. Cheshire
15. Greater Manchester
16. Lancashire
17. Merseyside
18. Bedfordshire
19. Berkshire
20. Buckinghamshire
21. East Sussex
22. Essex
23. Hampshire
24. Hertfordshire
25. Isle of Wight
26. Kent
27. Greater London
28. Oxfordshire
29. Surrey
30. West Sussex
31. Avon
32. Cornwall
33. Devon
34. Dorset
35. Gloucestershire
36. Somerset
37. Wiltshire
38. Hereford and Worcester
39. Shropshire
40. Staffordshire
41. Warwickshire
42. West Midlands
43. Humberside
44. North Yorkshire
45. South Yorkshire
46. West Yorkshire
47. Borders
48. Central
49. Dumfries and Galloway
50. Fife
51. Grampian
52. Highland plus Western Isles, Orkney and Shetland
53. Lothian
54. Strathclyde
55. Tayside
56. Clwyd
57. Dyfed
58. Gwent
59. Gwynedd
60. Mid Glamorgan
61. Powys
62. South Glamorgan
63. West Glamorgan

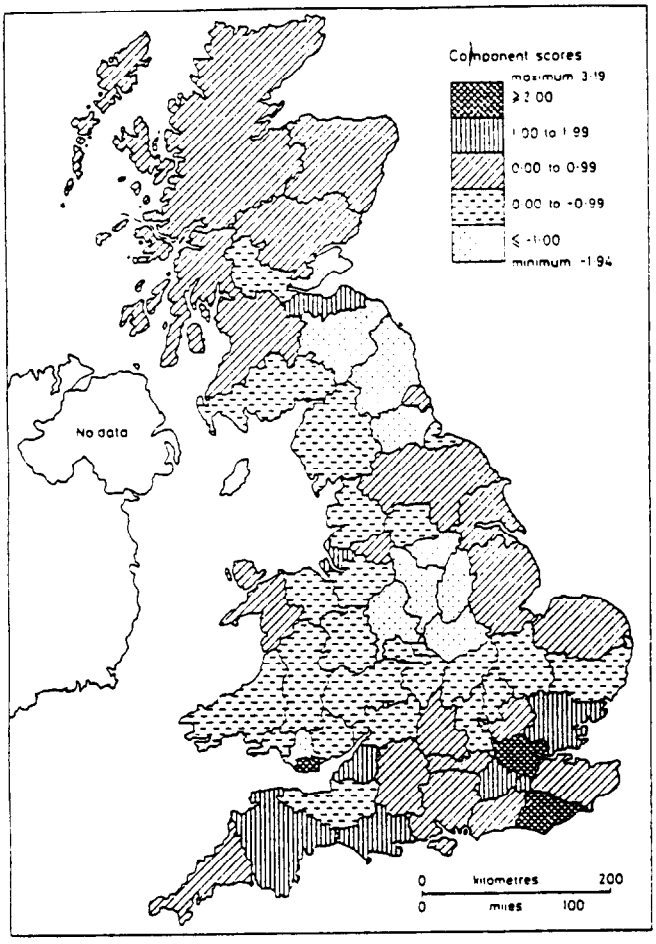


Fig. 3(a): Service and commercial component

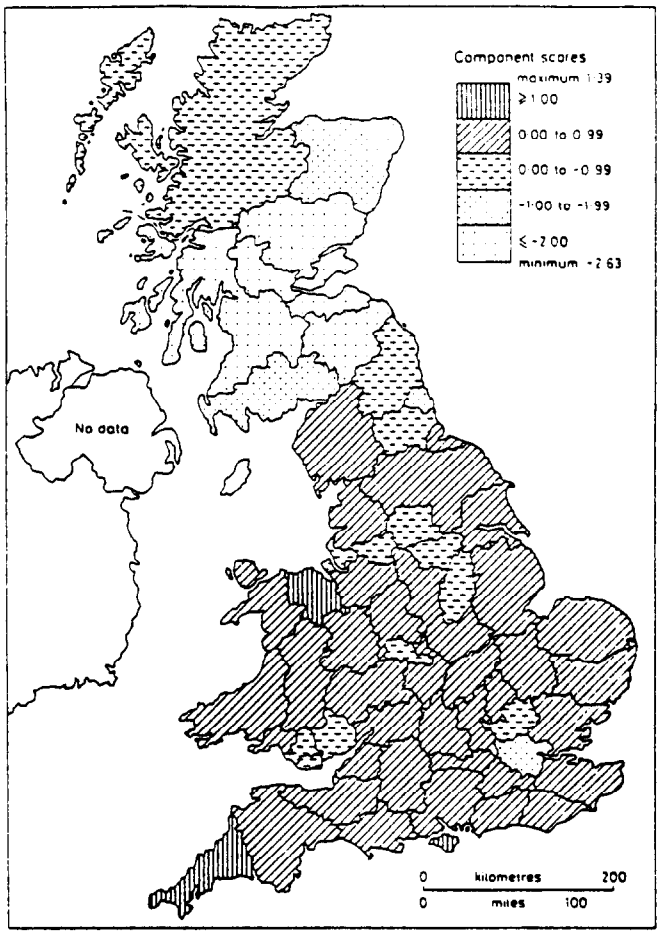


Fig. 3(b): High personal wealth, educational attainment component

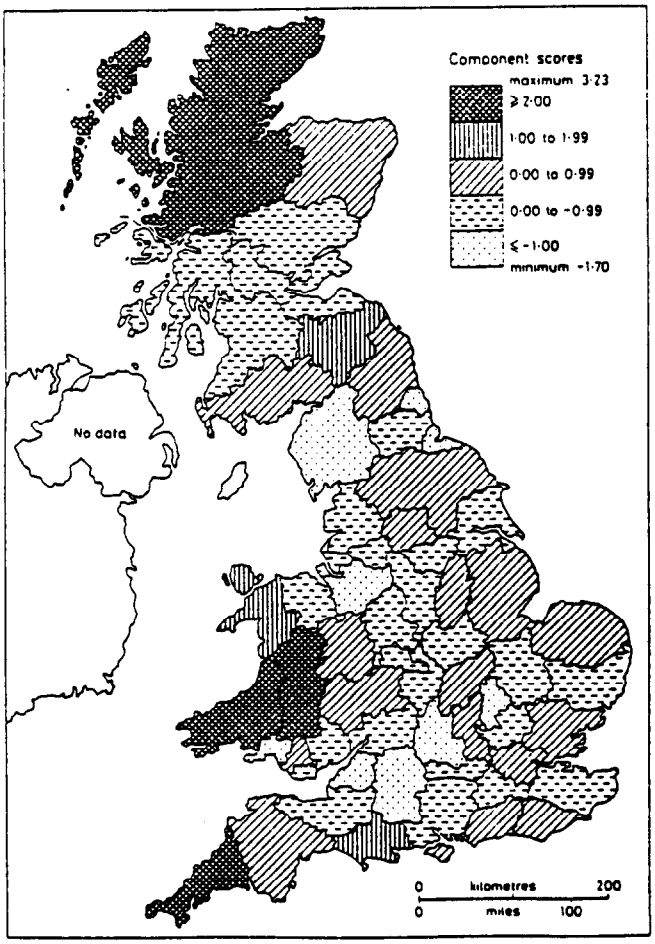


Fig. 3(c): Small plant, self-employment

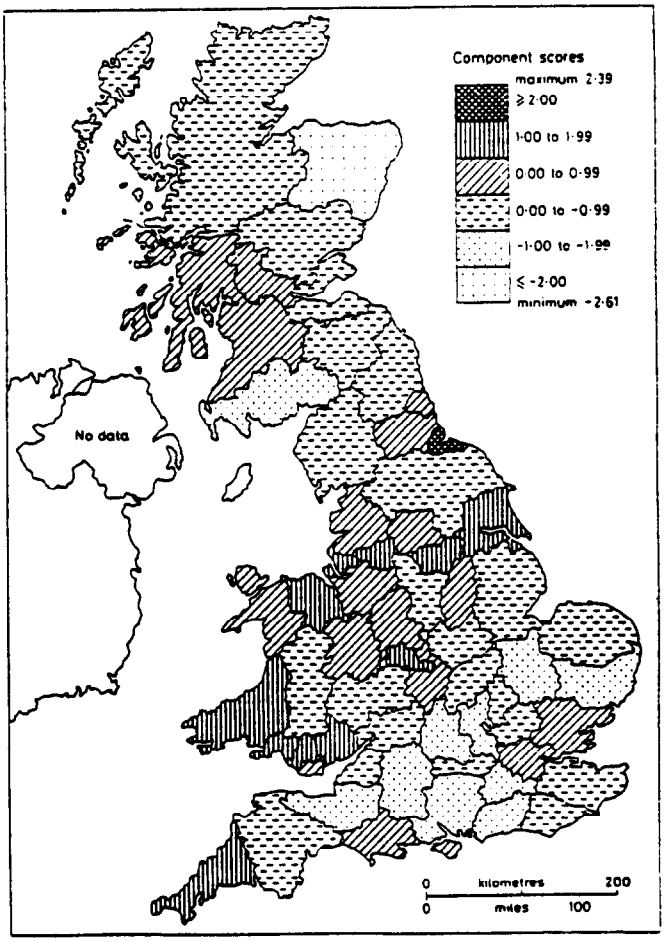


Fig. 3(d): High unemployment, employment

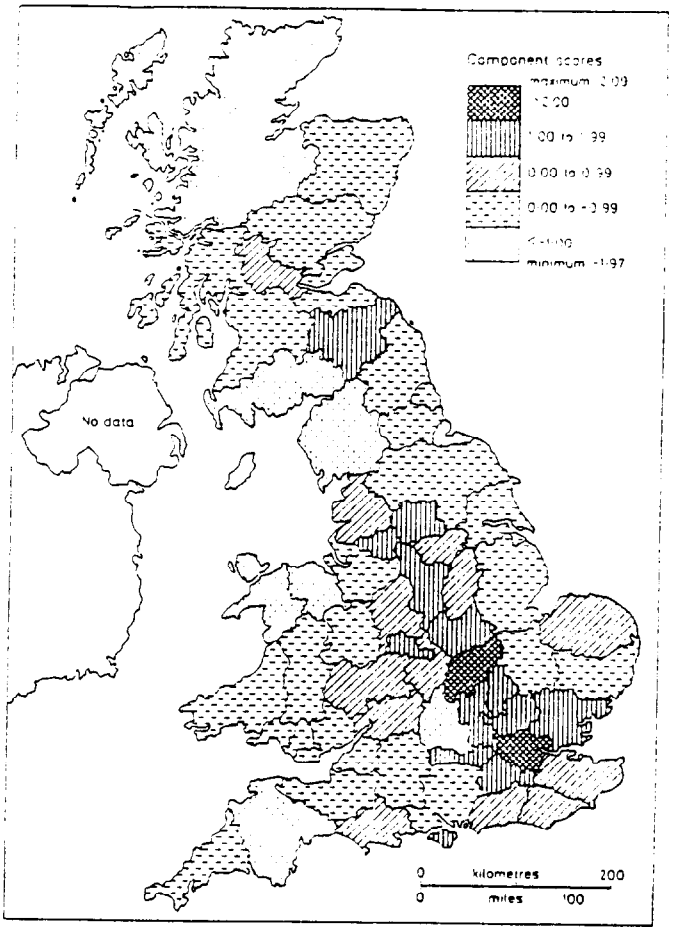


Fig. 3(e): High % increase in unemployment component

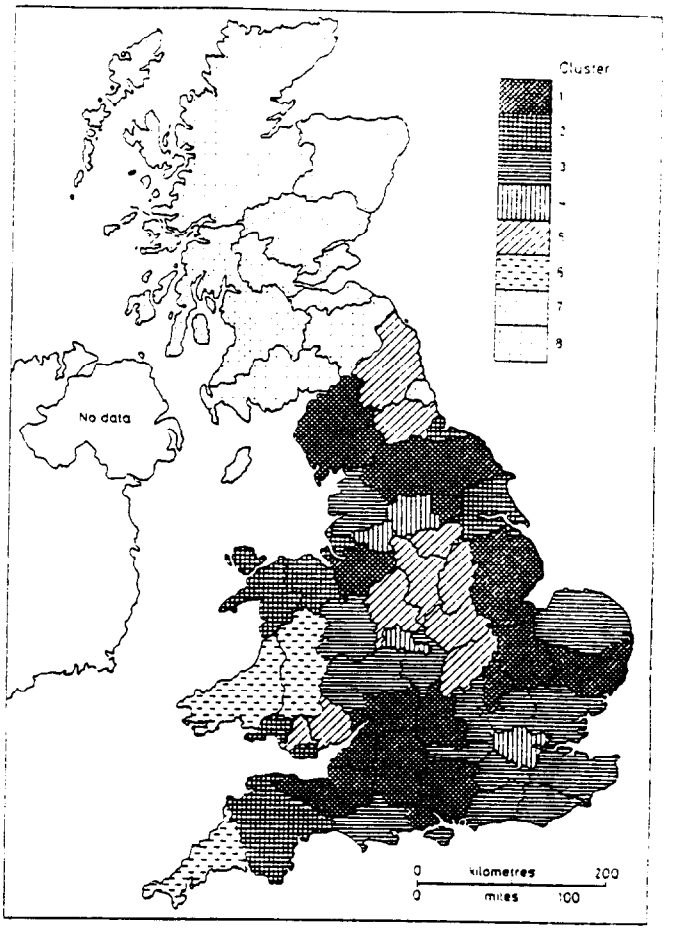


Fig. 5: Eight cluster ecological incubator environments

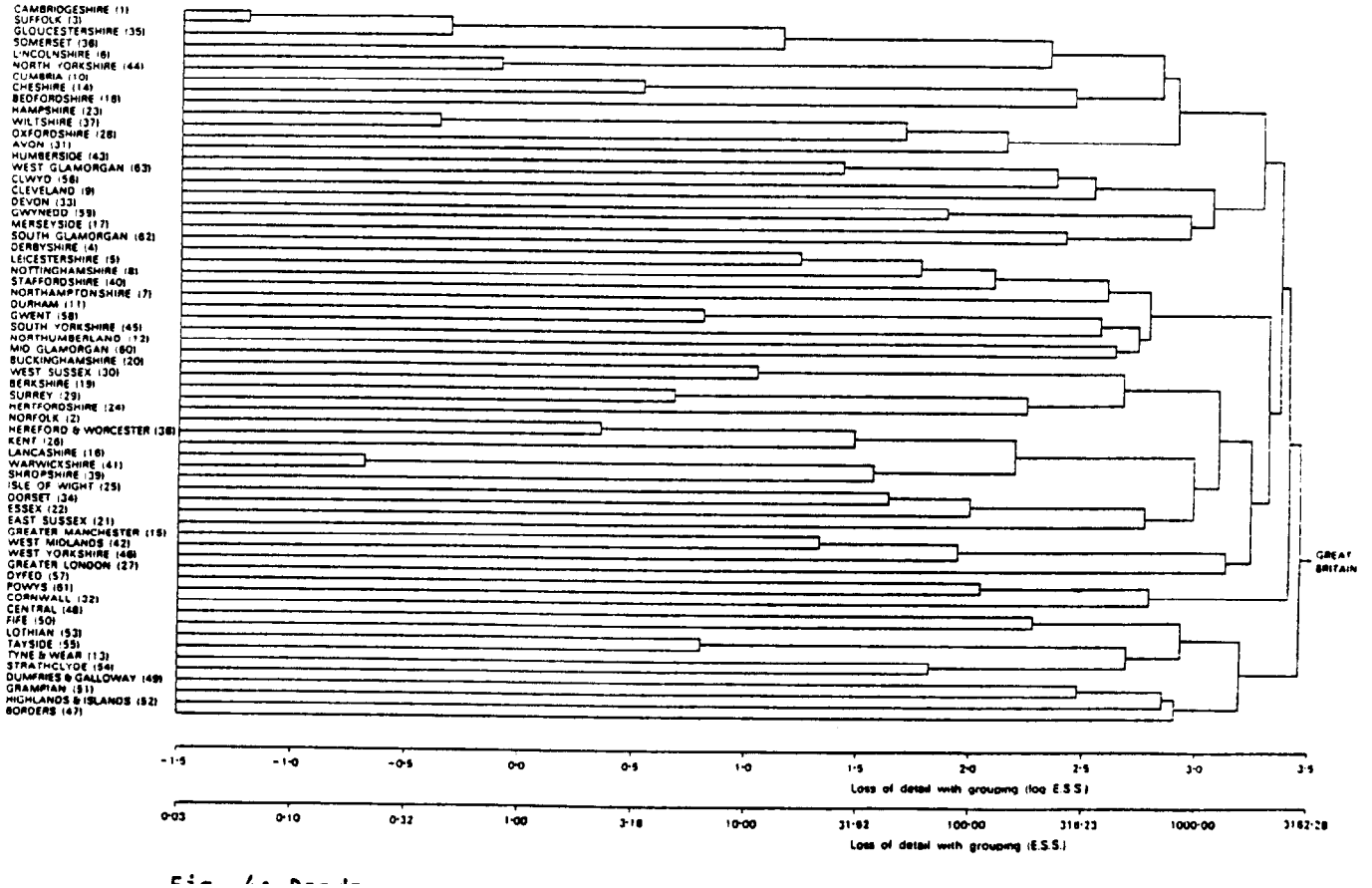


Fig. 4: Dendr

REFERENCES

Coombes M G, Green A E and Owen D W, "Substantive issues in the definition of 'localities'", Regional Studies 22(4), 1988, pp 303-318

Cross M New firm formation and regional development, Gower, 1981 p 197

Ganguly P, "Business starts and stops : UK county analysis 1980-83" British Business 18 January 1985 pp 106-110

Gould A and Keeble D "New firms and rural industrialization in East Anglia", Chapter 3, pp 43-71 in Storey D (Ed.) Small firms in regional economic development, Cambridge University Press, 1985

Lloyd P E and Mason C M "Spatial variations in new firm formation in the UK : comparative evidence" Regional Studies 18(3) 1984, pp 207-220

Norusis M J Advanced Statistics Guide SPSSX, McGraw Hill, Chicago, 1985

O'Farrell, P N and Crouchley R "An industrial and spatial analysis of New Firm Formation in Ireland", Regional Studies 18(3) 1984, pp 221-236

Spence N A, "A multifactor uniform regionalization of British counties on the basis of employment data for 1961", Regional Studies 2(1), 1968, pp 87-104

Storey D J and Jones A M "New firm formation: a labour market approach to industrial entry" Scot. Jnl of Political Economy 34(1), 1987 pp 37-51

Ward J H Jr. "Hierarchical grouping to optimise an objective function", J. American Statistical Association 58, 1963, pp 236-244

Westhead P, New manufacturing firm formation in the context of the economy of Wales, Ph.D. thesis, University of Wales, 1988