



3304

**The link between firm births and job creation:
Is there a Upas Tree effect?**

by

André van Stel
EIM Business and Policy Research
and
Max Planck Institute for Research into Economic Systems

David J. Storey
Center for Small and Medium Sized Enterprises
at University of Warwick

Number of Pages: 32

The *Papers on Entrepreneurship, Growth and Public Policy* are edited by the
Group Entrepreneurship, Growth and Public Policy, MPI Jena.
For editorial correspondence,
please contact: egppapers@mpiew-jena.mpg.de

ISSN 1613-8333
© by the author

Max Planck Institute for
Research into Economic Systems
Group Entrepreneurship, Growth and
Public Policy
Kahlaische Str. 10
07745 Jena, Germany
Fax: ++49-3641-686710

**The link between firm births and job creation:
Is there a Upas Tree effect?**

Adriaan J. van Stel
EIM Business and Policy Research
P.O. Box 7001, 2701 AA Zoetermeer, the Netherlands
Tel. +31 (0)79 343 02 00
Email: ast@eim.nl
and
Max Planck Institute for Research into Economic Systems, Jena, Germany

David J. Storey
Centre for Small and Medium Sized Enterprises at University of Warwick
Coventry CV4 7AL, United Kingdom
Tel. +44 (0)24 7652 3741
Email: smeds@wbs.warwick.ac.uk
and
Fellow at EIM Business and Policy Research

Version: June 2004

Abstract: This paper examines the relationship between firm births and job creation in Great Britain. We use a new data set for 60 British regions, covering the whole of Great Britain, between 1980 and 1998. The relationship between new-firm startups and employment growth has previously been examined either with no time-lag or with only a short period lag. We find, for GB as a whole, no significant relationship between startups and employment creation in the 1980s, but a negative relationship for the 'low enterprise' area of the North East of England. For the 1990s we find a significant positive relationship for GB as a whole but for Scotland, which focussed policy on startups, a negative relationship. We feel this raises questions over policies designed to raise rates of new firm formation as a strategy for employment creation, particularly in 'low enterprise' areas.

Keywords: new-firm startups, employment growth, entrepreneurship, Great Britain

JEL-classification: J23, L10, M13, R11

Acknowledgement: This paper has been written in the framework of the research programme SCALES (SCientific AnaLysis of Entrepreneurship and SMEs) which is financed by the Dutch Ministry of Economic Affairs. The authors wish to thank Anne Green and Rhys Davies of the Institute of Employment Research (IER) at Warwick University. Anne helped us to avoid the many pitfalls of UK regional data. Rhys provided us with the local wage data. Ian Dale of the Small Business Service helped us interpret the small firm data. We also particularly valued the comments of Peter Broer, Andrew Burke, Jim Love, Roy Thurik, Sander Wennekers, Gerrit de Wit, and four anonymous referees on earlier versions of the paper. Responsibility for what remains lies with us.

1. Introduction

This paper examines the relationship between new-firm startups and employment change in Great Britain. This relationship is of considerable policy importance, since national and sub-national governments in Britain have, for more than two decades, sought to raise business startup rates in order to enhance wealth- and job-creation. An example of a central government policy was the Enterprise Allowance Scheme (EAS). At its peak in 1987-88, public expenditure on EAS was virtually £200 million, subsidising more than 106,000 unemployed people to start a new business [Storey (1994)]. A second example is the Business Birth Rate Strategy initiated in Scotland in the early 1990s, which sought to raise new-firm formation rates. A third example was the Entrepreneurship Action Plan for Wales announced in 2001. Finally, in 2004, the UK government announced that “building an enterprise culture” and “encouraging a dynamic start up market” are the first two of the seven pillars of small business policy. The assumption of a strong positive relationship between increased new-firm startup rates and subsequent employment growth underpinned all such policies.

This paper tests for that underpinning. It begins by presenting the theoretical arguments for the presence of a relationship between startups and job creation, going on to provide an overview of current evidence. The central theme of the paper is that, with the exception of a recent paper by Audretsch and Fritsch (2002) for Germany, the relationship between startups and job creation has previously been examined either with no time-lag or with only a short period lag.

The current paper claims to make seven advances on prior work. The *first* is to construct and use a long-run (1980-98) data set that facilitates a valid comparison between the results for Great Britain and Germany. A *second* innovation is the explicit choice of variables. It argues that the appropriate measure of new firm formation is the sectorally adjusted number of private sector new firms, normalised by the sectorally adjusted working population. It also argues that the appropriate measure of employment change is the sectorally adjusted private sector employment. *Thirdly* it incorporates, for the first time, data on private sector wages in the locality. *Fourthly*, the paper explicitly incorporates various tests for misspecification which virtually all models pass. *Fifth*, the paper explicitly corrects for multicollinearity caused by strong intertemporal correlations between startup rates for different periods. *Sixth*, it utilises the concept of the “Upas Tree” to see whether Scotland and Wales differ from England in the relationship between startups and job creation. *Seventh*, and finally it links the findings to changes in Enterprise Policy both for the UK as a whole and for Scotland in particular.

The key results in the paper call into question the impact of policies seeking to raise new firm formation, so as to enhance employment creation, particularly in areas where new firm formation rates are low. Specifically we find that, in the 1980's when national public policy was focussed on raising new firm formation, there is no evidence that this led to increased employment creation during that decade. Furthermore, although it is non significant for the UK as a whole in the 1980's, it is significantly negative for the North East of England, an area with notably low rates of new firm formation.

In the 1990's, when UK national policy shifts away from stimulating new firm formation, a positive relationship emerges between firm formation and employment creation. Crucially, however, in Scotland which implemented a policy to stimulate new firm births in the 90's, a significant negative relationship between new firm births and employment creation appears in this decade, although our data do not extend sufficiently in time to imply that Scotland's business birth rate policy lead to lower employment.

2. The Issues

This section reviews the theoretical basis for believing a relationship exists between the extent to which a geographical area is “entrepreneurial” and the extent to which it is “economically successful”. We show there are a priori reasons for expecting a positive relationship, but that there are also reasons for expecting no relationship or, in extreme cases, a negative relationship.

There are three reasons why more “entrepreneurial” areas might generate more jobs- where jobs are a measure of “economic success”. The first is that if “entrepreneurial” is reflected in “new-firm formation” then these new firms themselves create jobs directly and so add to the stock of jobs. The second is that the new firms constitute a (real or imagined) competitive threat to existing firms, encouraging the latter to perform better [Disney, Haskel and Heden (2003)]. Finally, new firms provide a vehicle for the introduction of new ideas and innovation to an economy, which has been shown to be a key source of long-term economic growth [Romer (1986)]. Indeed Audretsch and Thurik (2001) argue that the role of new firms in technological development has been enhanced by a reduced importance of scale economies and an increasing degree of uncertainty in the world economy, creating more room for innovative entry.

The reasons for not expecting firm formation rates to be related to job creation are also three-fold. The first is that new firms directly contribute only a very small proportion of the stock of jobs in the economy [5.5% of the stock of UK employment in 1989 was in firms that had been born in the previous two years- Storey (1994)]. Secondly, innovation is very much the exception rather than the rule amongst new firms. For example, during the 1990s, twice-yearly Surveys were taken of (primarily) small firms in the West Midlands. The proportion of firms claiming to have introduced a product or service new to the marketplace in the prior twelve months varied from 4% to 17% [Price Waterhouse Coopers (1999)]. Third, the scale of job creation in new firms varies considerably from firm to firm. Storey and Strange (1992) show that 2% of all new firms created 33% of jobs in new firms, reflecting the extent of skewness in the distribution of employment. This skewness is taken to reflect differences in the human capital of founders [Frank (1988)] or their ability to learn [Jovanovic (1982)]. For these reasons job creation, even in new enterprises, may be more strongly influenced by the human capital of the founders, than by the absolute number of startups [Cooper, Woo and Dunkelberg (1989), Van Praag and Cramer (2001)].

The case for a negative relationship between new firm births and subsequent job creation derives from examining policies to stimulate new firm formation in “unenterprising” areas. Since these are frequently areas where human capital is low, the new firms tend to be in easy to enter sectors such as vehicle-repairing, window cleaning and hairdressing [Storey and Strange (1992)]. Subsidising entry means entrants temporarily have a competitive advantage over incumbents who are forced out of business. Once the subsidy is removed, the no-longer subsidised entrants may be forced out either by newly subsidised entrants or by re-entrants. The effect of this ‘churn’ is to lower customer confidence leading to lower expenditure and hence lower employment [Greene, Mole and Storey (2004)].

3. The Evidence

Prior empirical studies of the relationship between “entrepreneurship” and “economic success” have adopted different approaches, yielding different results. Three studies, albeit using very different dependent and independent variables, find a positive relationship. GEM (2000) examines the relationship across 21 countries between “total Entrepreneurial Activity” and per cent growth in GDP. They show that “Entrepreneurship is strongly associated with economic growth. Amongst nations with similar economic structures, the correlation between entrepreneurship and economic growth exceeds 0.7 and is highly significant”. Second, Johnson and Parker (1996) find “robust evidence that growth in births (and reductions in deaths) *significantly* lowers unemployment”.¹ Finally, taking the period 1981-89, Ashcroft and Love (1996), find new-firm formation to be strongly associated with net employment change in Great Britain.

¹ Their italics.

Fritsch (1996), however, obtains more ambiguous results. In a pioneering study that can be considered as the fore-runner to this study, he examines 74 (former) West German planning regions, 1986-89. He finds "a positive statistical relationship between entry rates and employment change for manufacturing in the longer run, ...(but)... this relationship proves to be negative for the service sector as well as for all sectors together" [Fritsch (1996), p. 247]. A recent paper by Audretsch and Fritsch (2002) provides new insights for (West) Germany. Taking the same 74 planning regions, they present three key findings. First, confirming the Fritsch (1996) findings, startup rates in the 1980s are found to be unrelated to employment change. Second, in the 1990s, those regions with higher startup rates have higher employment growth. Third, and perhaps most interesting, is that regions with high startup rates in the 1980s had high employment growth in the 1990s.

In summary therefore the evidence to date generally points to a significant and positive relationship between new firm formation and measures of employment creation. There seems no prior empirical support for a negative relationship, although some non-significant relationships have been found.

4. Modelling Issues

The relationship to be modelled is of the simple form in Equation (1) below

$$\Delta EMP_t = f(BIR_{t-1}, CON) \quad (1)$$

where ΔEMP_t = change in employment,
 BIR_{t-1} = firm birth rates at start of period,
 CON = control variables.

(i) Choice of Measures

Whilst, in principle, the model is simple to estimate there are five clear problems of definition. The *first* relates to the measure of BIR to be used. Given that the units of account are geographical areas that vary in size, BIR needs to be normalised by a size measure. The denominator should both control for the different absolute sizes of the regions concerned, and represent the source from which startups or firm formations are most likely to come [Ashcroft, Love and Malloy (1991)]. The two variables normally used as denominators are the stock of existing firms, and the size of the regional workforce [Keeble, Walker and Robson (1993)]. This is called the Business Stock (BS) approach and the Labour Market (LM) approach, respectively. The BS approach assumes new firms arise from existing ones, whereas the LM approach assumes that new firms arise from (potential) workers.² The choice of measure can be highly significant. For example, for a given number of startups, regions which are equally large in terms of workforce but which are different in terms of average firm size, will have the same startup rate according to the LM approach but different startup rates according to the BS approach.³ Garofoli (1994) makes a robust case in favour of LM over BS. The latter, he argues, is misleading in areas with small numbers of (generally large) firms. Here small numbers of new firms would provide an artificially high birth rate, primarily because of the small denominator. Audretsch and Fritsch (1994) also show that, in West Germany, the statistical relationship between

² In Ashcroft and Love (1996), total population is used as denominator. However, this assumes that new firms may arise from children or elderly persons as well. This seems less plausible.

³ In Van Stel, Dielbandhoesing, Van den Heuvel and Storey (2002) the (differences between the) two approaches are illustrated in detail by means of a numerical example for actual GB data.

unemployment and startup activity crucially depends on the BS or LM methods used to measure startup rates.⁴ We favour the Garofoli arguments and in this paper present only results from the LM approach.⁵

(ii) Lags

The *second* key problem relates to the lag structure specified in Equation (1). The case for the lag is that the employment impact of new firms is not likely to be immediate. Storey (1985), for example, shows that new manufacturing firms are generally eight or nine years old by the time they reach their peak employment, at which time they are about twice the size they were at the end of Year 1. However, because of their high exit rates, total employment in a cohort of new firms is lower in Year 5 than in Year 1. This means that the maximum employment impact of a cohort depends on the scale of these two influences and is an empirical, rather than theoretical, issue.

The above discussion is framed in terms of simple arithmetic, but more complex social processes could also influence the lag. For example, new businesses started in time period t may stimulate the formation of other new firms in period $t+1$. This may be because the t period firms constitute a market for the $t+1$ firms; alternatively the success of the t firms could stimulate individuals to seek to emulate them, so the t firms become "role-models". In turn, the $t+1$ firms stimulate more firms in later time periods, with the result that employment in that economy in $t+n$ is stimulated. Theory, again however, is not helpful in specifying the value of n .

Nevertheless, the above theoretical arguments discourage the use of contemporaneous startup rate variables in the model, i.e., employment change in period t being explained by new-firm startups in period t . Although correlations might be significant, the implied causal relation from births to (immediate) employment growth is potentially misleading. Positive correlations between startup rates and growth in the same period are often due to reversed causality, i.e., regions with high growth attracting new firms.⁶ In our empirical work we will include lagged startup rates only, but the precise nature of that lag is the subject of tests.

⁴ In Audretsch and Fritsch (1994) the business stock approach is called the ecological approach.

⁵ Analyses comparing the LM and BS approach are in Van Stel and Storey (2002). In that paper we also pay extensive attention to some other empirical matters discussed later in this section, such as the sector adjustment of the startup rates and the impact of public sector employment on regression results.

⁶ Even if there is a lag in this reversed causality process, the measured correlation is often still positive, because of path dependency in the growth performance of regions.

(iii) Sectoral Comparisons

A *third* problem relates to differences in industrial structure between regions. This raises the question of whether the different sectoral structures of regions should be taken into account, since this influences both the number of startups and also employment change. Taking only the difference between services and manufacturing, startup rates are higher in service industries than in manufacturing [Audretsch and Fritsch (2002)], partly because entry barriers are lower, Minimum Efficient Scale (MES) is also likely to be lower and, for some services, demand is high. For all these reasons, regions with a high share of services in the local economy are more likely to have higher startup rates than regions with a low service share.

But this does not necessarily mean these regions are also more “entrepreneurial”, in the sense that startup rates are higher for each sector of the local economy (or most sectors of the local economy). Therefore, to correct for different sectoral structures, the Ashcroft, Love and Malloy (1991) shift-share procedure is applied to derive a measure of sector-adjusted startup activity. The sector-adjusted number of startups is defined as the number of new firms in a region that can be expected to be observed if the composition of industries was identical across all regions. Thus, the measure adjusts the raw data by imposing the same composition of industries on each region [Audretsch and Fritsch (2002)]. An identical process is used to derive a measure of sectorally adjusted employment change. Appendix 1 provides an illustration of the shift-share procedure.

Another sector issue concerns the impact of the public sector on estimated model coefficients. Ideally, analysis should be restricted to private sector enterprises and private sector employment. Unfortunately, however, both private and state-owned enterprises can be present within some SIC groups. Furthermore, SIC groups with a relatively large employment share of public sector organizations (such as universities and hospitals) may disturb estimations as changes in public sector employment may create a bias in the estimated employment effect of new-firm startups. Therefore, we eliminate SIC groups dominated by state-owned enterprises or other public sector organizations from our analysis.⁷

(iv) ‘Control’ variables

A *fourth* issue relates to the choice of control variables (CON) used in Equation 1. In addition to the sectoral composition effects, noted above, previous studies have shown urban and rural areas differ in both employment change and in new-firm formation rates. In their review of regional variations in firm birth rates, Reynolds, Storey and Westhead (1994) pointed to urban areas consistently having higher formation rates in the 1980s than non-urban areas. Employment change, however, has been more mixed, with an urban-rural shift in the 1970s and 1980s [Fothergill and Gudgin (1979)] but a

⁷ This involves SIC92 industries L, M, and N (Public administration, defence and compulsory social security; Education; and Health and social work, respectively) for post-1991 data, and SIC80 industry 9 (“other services”) for pre-1991 data; we utilise data according to different SICs before and after 1991, see Tables A2b and A3 in Appendix 2.

more mixed picture in more recent times [Green and Turok (2000)]. Account of urban/rural differences is taken by the inclusion of a population density variable, and by Standard Region dummies.⁸

Another control factor is the nature of the labour market, reflected in local wage rates. Rees and Shah (1986) assume the welfare maximising individual chooses between utility in self-employment compared with paid employment, for which wages are taken as the proxy. Hence rises in wage rates would be expected to lead to movements into wage-employment and out of self-employment, consistent with a positive effect on employment change (which in the present study is defined to include employees only). Furthermore, wage rises may also stimulate labour supply which could also lead to increased employment at the regional level. However, there is also a possible negative effect as a higher price of labour may lead to a lower demand for labour (substitution between capital and labour).⁹ These opposite effects make the sign of wage rates indeterminate from theory.

A further control factor relates to the issue of reversed causality discussed earlier. Even if we include lagged startup rates only, the employment impact of new-firm startups might be overestimated, due to positive path dependency in the economic performance of regions (i.e., the business cycle effect). We correct for this by including lagged employment growth.¹⁰

(v) Public Policy and Region-specific effects

The 1980's and 1990's saw radical changes in Enterprise Policy in the UK. Greene (2002) argues that the decade of the 1980's saw, following the election of a Conservative government in 1979, the first explicit attempt to create an enterprise culture in Britain. Policy was directed towards maximising the number of new business starts so as to achieve this 'enterprise culture' and to seek to create jobs so as to offset the high levels of unemployment. In the 1990's, however, British policy changed towards a focus on established business with "growth potential". This we refer to as the policy effect.

In addition we also argue for the presence of region-specific effects reflecting the major cultural differences, within Great Britain, in attitudes towards enterprise and self-employment. We call this the Upas Tree effect. The term was originally used by Checkland (1976) to describe economic change in the city of Glasgow, and was derived from a description of the Upas Tree that was native to Java. According to legend, the Upas Tree was able to destroy other growths for a radius of 15 miles, and Checkland viewed it as analagous to the destructive effect that the heavy engineering sector had

⁸ According to Audretsch and Fritsch (2002, p. 120), who also use population density as a control in their regressions for Germany, "Population density here represents all kinds of regional influences such as availability of qualified labour, house prices, local demand and the level of knowledge spillovers".

⁹ For a selection of European countries, Van Stel (1999) estimates the real wage elasticity (the response of labour demand on an exogenous rise in real wages at constant output level and price of capital) to lie between -0.2 and -0.4 in the period 1970-1994.

¹⁰ The concept of using lagged dependent variables to correct for reversed causality is known in the econometric literature as Granger-causality. The Granger (1969) approach to the question of whether x causes y is to see how much of the current y can be explained by past values of y and then to see whether adding lagged values of x can improve the explanation. y is said to be Granger-caused by x if x helps in the prediction of y , or equivalently if the coefficients on the lagged x 's are statistically significant (Audretsch, Carree and Thurik, 2001).

upon the growth of other industries in Glasgow for much of the twentieth century.¹¹ We use it to characterise Scotland and Wales, both of which appear to have a long-standing antipathy to “entrepreneurship”, but also North East England [McDonald and Coffield (1992); Greene, Mole and Storey (2004)].

However, the policy and the region-specific effects interact with one another. This is because, whilst Britain as a whole, in the 1990’s, was shifting its policy away from a focus on business start-ups, Scotland explicitly chose the opposite policy. It established a “business birth rate” strategy [Fraser of Allander (2001)] the focus of which was to raise new firm formation in that country. Account therefore has to be taken of these very different policy environments in Britain in 1980’s and 1990’s and of the differences between Scotland and the rest of Britain in the 1990’s.

We investigate the impact of new firms on employment change separately for the 1980s and the 1990s to see whether effects differ between these two decades. We also incorporate slope dummies for Standard Regions to see whether effects for certain regions deviate from the overall effect for Great Britain.

5. Variables and Data Sources

The data used is at the spatial aggregation level of NUTS3 regions in Great Britain. This is county level in England and Wales, and local authority region level in Scotland. In this partitioning, Great Britain comprises 60 regions, each disaggregated by six sectors. This facilitates correction for sectoral differences between regions, i.e., to apply the shift-share procedure described earlier. Different regional and sectoral classifications in the original data files meant some linking operations were performed to ensure uniformity for the whole period 1980-98. These linking operations and the exact classification schemes employed are reported in Appendix 2. The agricultural sector is excluded, as this sector is fundamentally different from the rest of the economy, having, during this period, exceptionally low startup and death rates.

Variable definitions and their sources are now provided:

Sector adjusted (lagged) employment change. This is the change in regional employment, expressed in percentages (excluding agriculture). For each region, sectoral employment growth rates are weighted by employment per sector for *Great Britain as a whole*. Data on employment are taken from the Census of Employment and the Annual Employment Survey and are supplied by Nomis. Employment figures include both full-time and part-time employees, and exclude self-employed workers and unpaid family workers. Employment is measured in September of each year.

¹¹ To our knowledge Lloyd and Mason (1984) were the first to use Checkland’s analogy in this context.

Sector adjusted startup rate. This is the sectoral startup rate, weighted by employment per sector for Great Britain as a whole. Using this weighting implies an identical sector structure for each region. Regional employment, rather than regional workforce, is used as the denominator for the LM approach, because of greater data reliability. Startups in the agricultural sector are again excluded. Startups are measured as VAT registrations and these data are supplied by Small Business Service. The consistency and general availability of this data source make it the most generally useful source of data on firm formation for the UK as a whole [Ashcroft, Love and Malloy (1991)]. Startup rates are expressed as the number of startups per thousand workers (LM approach).

Population density. Data on both population and area of the regions are obtained from the Office for National Statistics. The variable is expressed in thousands of inhabitants per square kilometre.

Wage growth. This variable measures changes in regional wage rates. We use data from the New Earnings Survey Panel Data-set (NESPD), which is operated by the Office for National Statistics. The estimates of regional wage rates refer to average hourly earnings excluding overtime payments. The samples from which the mean wages are calculated relate to full time employees whose pay was unaffected by absence during the survey week (which falls in April of each year) and exclude those employed in agriculture, forestry and fishing.

6. Research design and regression diagnostics

In this section we describe our research design and several regression diagnostics that we will use in order to test validity of regression results. From Table A2b in Appendix 2 we can see that we cannot utilise employment data for all the years in our sample period 1980-98. We also had to make several harmonizations to correct for changes in regional and sectoral classifications over time. Both these aspects hamper the use of panel data regression techniques. Instead, we estimate each regression cross-sectionally, i.e., using 60 observations (one for each region). Because of missing (employment) data, the region Orkney/Shetland/Western Isles had to be dropped, generating a total of 59 observations. The models are estimated using OLS.

We want to estimate separate models for the 1980s and the 1990s to see whether the relationship between firm births and job creation has changed over time. Given the (limited) availability of employment data and our preference to measure employment change in a period after the period in which we measure startups (to obtain the correct direction of causality), Table A2b demonstrates that the most appropriate periods are 1984-91 and 1991-98 for employment change in the 1980s and 1990s.¹² In this way we make optimal use of the full time length of our data set. We measure startup rates in a period directly preceding the employment change periods. We use four-year averages in order to correct for outlier years. This results in startup rates for the periods 1980-83 and 1987-90. In this way the 1980s and 1990s estimations are completely 'symmetric' (seven year pe-

¹² A further advantage of this choice is that the computation of employment change is not hampered by the change in sectoral classification in 1991, see Table A2b.

riod for employment change; four-year averaged startup rate in the preceding period), which enables comparison of the results for the two periods.

Regarding control variables, population density is measured three years before employment growth. For wage growth and lagged employment growth some small inconsistencies concerning the 'symmetry' of the 1980s and 1990s models cannot be avoided, due to the unavailability of data prior to 1981. We measure wage growth over a period of four years, and we want the wage growth period to precede the employment growth period, if possible. Using a four year length this inevitably results in the period 1981-85. For the 1990s regression we experimented with the lag which resulted in inclusion of wage growth for the period 1985-89 (based on statistical fit). Regarding the lagged dependent variable we want to use the same length as the dependent variable (seven years). For the 1990s regression this results in the period 1984-91. For the 1980s regression however, the first available year in our data set is 1981 and therefore lagged growth is measured over the period 1981-84.

In all instances, four regression diagnostics are presented. These are first, the Jarque-Bera test on normality of the disturbances; second, the Lagrange Multiplier test on heteroscedasticity; third, the Ramsey RESET test on general misspecification of the model. To facilitate direct evaluation of these tests p-values are shown. For all three tests the null hypothesis corresponds to "correct estimates", i.e., normality at the Jarque-Bera test, no heteroscedasticity at the Lagrange Multiplier test and no sign of misspecification at the Ramsey RESET test.

Finally, the fact that the data relate to spatial variations raises the potential problem of spatial autocorrelation, an issue "which has been widely ignored in the econometric literature, including most previous work on spatial variations in new firm formation" [Keeble, Walker and Robson (1993), p. 34]. Following Keeble, Walker and Robson (1993), account is taken of this by including Standard Region intercept dummies in the equations.¹³ To see whether spatial autocorrelation is actually present in our regressions, we report the Durbin-Watson statistic. We test for positive spatial autocorrelation, implying that the null hypothesis of no spatial autocorrelation is accepted (not rejected) if the DW test statistic is greater than a certain upper bound for critical values, which depends on the number of observations and regressors.¹⁴

When using the DW statistic we acknowledge this statistic is designed for times series analysis with observations being arranged along a 'time line'. In other words, it is designed for one-dimensional concepts such as time. Space, however, is a two-dimensional concept. Following Ashcroft and Love (1996), we present the data to the estimation by county within each Standard Region. It follows that many adjacent observations are from contiguous counties. Therefore we argue that the DW statistic

¹³ For this purpose the county Greater London is added to the South East region. This is because there is only one county within the London region in our data set.

¹⁴ We test for *positive* autocorrelation as neighbouring regions may be expected to benefit from each other (spillover effects).

should measure the degree of spatial dependence reasonably well. But even so, we recognise there remains some arbitrariness in the ordering of neighbouring regions.¹⁵

7. Results

To test whether startup activity has a different impact on employment growth in different time periods several models are estimated. Recalling that a key objective is to test for short or long-run relationships this section begins by examining the relationship between startups, 1980-83, on employment change 1984-91; then it examines startups in the period 1987-90 on employment change 1991-98. This provides an initial assessment of whether the short-term impact of startups differed between the 1980s and the 1990s. Next, we look at possible region-specific deviations in the effect of startups on employment growth. In the third subsection we investigate whether estimation results are affected by the periods in which startup rates are measured in terms of recession or boom periods. The fourth subsection investigates long-run effects. We also pay attention to the interpretation of the magnitude of the estimated effects. Finally, we compare our results with other studies.

(i) Startups and employment change in the 1980s and the 1990s: short-term effects

Table 1 presents the regression results for the 1980s and the 1990s. Startup rates are related to subsequent employment growth, while controlling for population density, wage growth, lagged employment growth, and regional dummies. All control variables are measured prior to the period of the dependent variable.

The final rows show all diagnostic tests are passed (p-values are well above 0.05), except for the RESET test in the 1990s, possibly indicating a missing variable. As regards spatial autocorrelation, the null hypothesis of no autocorrelation is accepted (not rejected) as the DW test statistic exceeds the upper bound critical value (which is about 2, in our case).¹⁶

The impact of startup activity on subsequent regional employment change is different for the 1980s and the 1990s. In the 1980s startups and employment change are unrelated while in the 1990s startups have a significantly positive impact on employment growth. The bigger employment impact of 1987-90 births compared to 1980-83 births *might* reflect that the importance of new and small firms in the process of innovation and economic growth has increased in the last two decades of the 20th century. In this interpretation Great Britain would have moved from a more “managed” type of economy toward a more “entrepreneurial” type of economy [Audretsch and Thurik (2001)]. However, perhaps a more testable explanation is that the increased employment impact reflects “Enterprise

¹⁵ For an alternative approach see the methods covered by Anselin (1988).

¹⁶ The Durbin-Watson test should be interpreted with caution in the presence of a lagged dependent variable in the model (Stewart, 1991, p. 168). However, as the DW test statistic is clearly greater than the upper bound critical value, we think it is safe to assume that our estimates do not suffer from first-order spatial autocorrelation. Furthermore, when the Standard Region intercept dummies are removed, the DW test statistic falls to 2.10 for the 1980s regression and to 1.64 for the 1990s regression. The latter value falls within the inconclusive region, indicating that the regional dummies are indeed helpful in correcting for spatial autocorrelation.

Policy” changes, with public policy switching from being quantity-oriented in the 1980s towards being more quality-oriented in the 1990s [Greene (2002)].

As for the control variables, we see a significant negative impact of population density (in the 1980s), and a positive impact of wage growth (in the 1990s) and lagged employment growth. The latter effect points at positive path dependency. Regions that perform relatively well in a certain period, still perform relatively well in the next period.

Table 1: Determinants of regional employment growth (%), short-term equations 1980s and 1990s

	Employment growth 1984-91	Employment growth 1991-98
Constant	5.5 (0.5)	-21.3 (2.8)
Average startup rate, 1980-83 (left column) 1987-90 (right column)	-0.25 (0.3)	1.11 (2.3)
Population density, 1981 (left column) 1988 (right column)	-4.6 (3.2)	-0.36 (0.3)
Wage growth, 1981-85 (left column) 1985-89 (right column)	0.28 (0.8)	0.53 (2.6)
Lagged employment growth, 1981-84 (left column) 1984-91 (right column)	0.46 (2.1)	0.20 (1.8)
Adjusted R ²	0.281	0.696
JB test: [p-value]	[0.517]	[0.820]
LM het. test: [p-value]	[0.630]	[0.264]
RESET test: [p-value]	[0.743]	[0.015]
DW test	2.23	2.18

Note: Intercept dummies for Standard Regions not reported. Employment growth rates and startup rates are sector adjusted. Employment growth is measured exclusive of the non-private sector. Absolute t-values in parentheses.

(ii) Regional specific effects

In this subsection we investigate whether certain regions deviate in the employment effect of new-firm startups. For this purpose we compute slope dummies (startup rate multiplied by regional dummy) for the (ten) Standard Regions. Given the specifications in table 1 (hence, including all *intercept* dummies), we include, one at a time, a *slope* dummy for each Standard Region. Those slope dummies which are significant at 10% level when included separately, are included in table 2. For the 1980s this is the North East region, and for the 1990s Scotland and Wales. The effects for the other regions are not significantly different from the overall effect. The improved value for the RESET test for the 1990s regression (compared to table 1) implies that the slope dummies for Scotland and Wales contribute to the validity of the model.

For the 1980s, the overall effect of the startup rate is nil on employment in Great Britain. However, for North East England new-firm startups contribute *negatively* to employment growth in that region in the 1980s. For the 1990s the overall startup rate effect is positive (and stronger than in Table 1), but for Wales the effect is nil, and for Scotland the effect is *negative*.¹⁷

Whilst the significant negative sign for North-east England in the 1980s may seem a surprise the plausible explanation comes from a reading of McDonald and Coffield (1992). They paint a picture of unemployed young people in North-east England with very modest human capital being press-ganged by public agendas into starting their own enterprise and ending up more disadvantaged than before they started. The effect is to erode confidence leading to declining economic performance.

Matters change in the 1990s. In the right hand side of Table 2 it can be seen that it is Wales and particularly Scotland that exhibit the significantly negative signs. The common thread is that North-east England, Scotland and Wales all have rates of new firm foundation that are well below the GB average. Second, for two regions, the North-east and Scotland, the negative deviation coincides with a policy thrust to raise new firm formation. This raises the question whether such policies are productive in 'un-enterprising' areas.

It will be recalled that after October 1993 Scotland implements an active policy to raise business birth rates (BBRS) [Fraser of Allander Institute (2001)]. Although the periods studied in the current paper do not entirely coincide with the period during which the BBRS is active (from 1994 onwards), the negative value for the Scotland dummy indicates that the BBRS actually might have had a negative effect on job creation in Scotland.

The results from Table 2 call into question the impact of policies seeking to raise new-firm formation, for two reasons. *First*, in the 1980's, when UK policy was to stimulate starts, there is no effect on

¹⁷ The slope dummies refer to the deviation from the overall effect. For instance, the significant parameter estimate for Wales means that the effect for Wales deviates significantly from England. It does not mean that it deviates significantly from zero. Indeed, the effect for Wales is -0.3 which is not significant (t-value -0.3). The effect for Scotland is -2.7 which is significantly different from zero at 10% level (t-value -1.8).

employment in the UK as a whole, and even a negative effect for the North East. *Second*, in the 1990s there is a significantly positive overall effect after the UK policy changed towards more emphasis on established businesses with the potential to grow. However, for Scotland, which has a business birth rate strategy in the 1990s, the effect is negative.

Table 2: Examining region-specific deviations in employment impact of startups

	Employment growth 1984-91	Employment growth 1991-98
Constant	66.7 (2.2)	-28.9 (3.9)
Startup rate, <i>overall effect</i>	-0.18 (0.2)	1.88 (3.8)
Startup rate, <i>slope dummy North East</i>	-10.5 (2.2)	
Startup rate, <i>slope dummy Wales</i>		-2.2 (2.0)
Startup rate, <i>slope dummy Scotland</i>		-4.6 (2.9)
Population density,	-5.1 (3.6)	0.39 (0.3)
Wage growth,	0.26 (0.8)	0.55 (2.9)
Lagged employment growth,	0.54 (2.5)	0.25 (2.5)
Adjusted R ²	0.335	0.747
JB test: [p-value]	[0.493]	[0.030]
LM het. test: [p-value]	[0.843]	[0.534]
RESET test: [p-value]	[0.225]	[0.576]
DW test	2.30	2.22

Note: Intercept dummies for Standard Regions not reported. Employment growth rates and startup rates are sector adjusted. Employment growth is measured exclusive of the non-private sector. Except for startup rate slope dummies, variable specifications are as in Table 1. Absolute t-values in parentheses.

(iii) Recession births versus boom births

In the previous sections we argued that the different short-term impacts of startups in the early and late 1980s may have been caused by “Enterprise Policy” changes. An alternative explanation is that the 1980-83 startups may be a different *type* of startups, compared with the 1987-90 startups. The obvious difference is that, while 1980-83 were recession years, 1987-90 was a “boom” period. During recessions, a higher proportion of startups may be from individuals with lower human capital, who find employment in the employee labour market more difficult [Cressy (1996)]. These startups may be less likely to generate jobs. On the other hand, during a period of economic prosperity, it

may be the more “entrepreneurial” type of person who starts a business. This type of startup may be more likely to generate jobs in the short and the long-run. So, while recession births may be the result of “push”-factors being at work (possibly creating fewer jobs), boom births may be more “pull-factor” in nature (possibly creating more jobs).

To test this we examine in Table 3 the relationship between firm births in the 1990s recession and short-term employment change. Using the same control variables as those reported in Table 2, we estimate a regression in which employment change in the period 1993-98 is explained by the average startup rate over the period 1990-93. To facilitate comparison, the results from the right column of Table 2 are reported again in Table 3. The results are similar: we find a significant positive impact, implying that the lack of a relationship in the 1980s is not because of the choice of recession-years.¹⁸ Instead, it seems to be the case that (new) firms in the late 1980s and early 1990s contribute more to employment change than firms started in the early 1980s irrespective of macro-economic conditions.

¹⁸ The estimated effect for the recession period is even stronger, although not significantly. As regards the dummy variables, the deviations of Wales and Scotland seem to be smaller compared to table 2 (t-values – 1.3). However, the isolated effects are nil for both Wales (effect –0.8; t-value –0.4) and Scotland (effect – 0.7; t-value –0.3), while the effect for the English regions is significantly positive. This implies that Wales and Scotland still lag behind in the employment effect of new firms started in the period 1990-93.

Table 3: Examining the impact of recession or boom period

	Employment growth 1991-98	Employment growth 1993-98
Constant	-28.9 (3.9)	-26.8 (4.0)
Startup rate, <i>overall effect</i>	1.88	2.39
1987-90 (left column)	(3.8)	(4.1)
1990-93 (right column)		
Startup rate, <i>slope dummy Wales</i>	-2.2 (2.0)	-3.2 (1.3)
Startup rate, <i>slope dummy Scotland</i>	-4.6 (2.9)	-3.1 (1.4)
Population density,	0.39 (0.3)	-0.06 (0.06)
Wage growth,	0.55 (2.9)	0.52 (2.5)
Lagged employment growth,	0.25 (2.5)	0.15 (1.4)
Adjusted R ²	0.747	0.737
JB test: [p-value]	[0.030]	[0.473]
LM het. test: [p-value]	[0.534]	[0.289]
RESET test: [p-value]	[0.576]	[0.534]
DW test	2.22	2.51

Note: Intercept dummies for Standard Regions not reported. Employment growth rates and startup rates are sector adjusted. Employment growth is measured exclusive of the non-private sector. Except for startup rate slope dummies, variable specifications are as in Table 1. Absolute t-values in parentheses.

(iv) A long-term effect?

In this subsection we test for long-run effects. Given our data set we can only test for long-run effects for employment growth in the 1990s, as we have no startup data prior to 1980. The easiest way to test for long-run effects is to run separate regressions which include different lags of the startup rate. Using the same control variables as in Table 2, the coefficients of startup rate in separate regressions explaining employment change 1991-98, are 1.88 for for 1987-90 startups, 2.25 for 1984-87 startups, and 2.44 for 1980-83 startups. So, the impact increases with the lag, seemingly indicating that the long-run effect exceeds the short-run effect.

However, we must be cautious in comparing these coefficients. To avoid multicollinearity we estimated the impact of the startup rates from different periods in separate regressions. A disadvantage of this approach is that, because of the strong intertemporal correlation between startup rates (correlations of up to 0.9), the estimated startup rate coefficient may pick up some of the effect of startup activity from other periods. This means comparing coefficients of the long-term and short-term equations is complex.

A better way of establishing the individual impacts of startup rate variables from different periods draws upon the distributed lag literature [Stewart (1991)]. By including startup rates from different periods in one regression, but imposing restrictions on the individual parameters, an accurate approximation of the shape of the lag response can be obtained. In the Almon method, parameter restrictions are imposed in such a way that the coefficients of the lagged variables are a polynomial function of the lag length. In this way the startup rate coefficients are reparameterized in a “smooth” way.

We apply the Almon method for a quadratic polynomial function (i.e., a polynomial of second degree). This choice corresponds to imposing one parameter restriction.¹⁹ The results are shown in Table 4, with further details presented in Appendix 3.

¹⁹ This can be seen as follows. In the unrestricted regression three startup rate variables are included in the model, while in the first unrestricted regression column, only two variables are included (COMBI1 and COMBI2 in Table 4). In the second unrestricted regression column, only one startup rate variable is included (COMBI3), and this corresponds to two parameter restrictions. The startup rate coefficients in the restricted regressions are linear combinations of the combinatory variable coefficients. See equation (A3) in Appendix 2.

Table 4: Examining the lag structure

	Employment growth 1991-98		
	Unrestricted regression	Restricted regression (one restr.)	Restricted regression (two restr.)
COMBI1 = $X_{-1}+2X_{-2}+3X_{-3}$		1.29 (0.9)	
COMBI2 = $X_{-1}+4X_{-2}+9X_{-3}$		-0.39 (0.6)	
COMBI3 = $-2(X_{-1}+X_{-2})$			-0.53 (3.8)
Startup rate 1987-90 (X_{-1})	1.2 (0.8)	0.89 (1.1)	1.06 (3.8)
Startup rate 1984-87 (X_{-2})	0.48 (0.2)	1.00 (2.6)	1.06 (3.8)
Startup rate 1980-83 (X_{-3})	0.47 (0.3)	0.32 (0.2)	
Adjusted R ²	0.737	0.743	0.748
JB test: [p-value]	[0.062]	[0.087]	[0.069]
LM het. test: [p-value]	[0.554]	[0.578]	[0.581]
RESET test: [p-value]	[0.657]	[0.661]	[0.642]
DW test	2.21	2.21	2.22
Validity Almon restrictions:			
F-test statistic		0.062	0.053
Critical value (5% level)		4.1	3.2

Note: Except for startup rates, model specifications are as in right column Table 2. Intercept dummies for Standard Regions, startup rate slope dummies for Scotland and Wales, and coefficients of population density, wage growth, and lagged employment growth are not reported. Absolute t-values in parentheses. Null hypothesis for JB test, LM het. test, RESET test, and DW test is "correct model specification". Null hypothesis for F-test is "valid restrictions". Critical values for F-tests are according to F(1;41) and F(2;41) distributions.

In Table 4, regression results using unrestricted regression (i.e., free estimation) and restricted regressions (i.e., using the Almon method) are presented. For the *unrestricted regression* we see that t-values of the separate startup rates are low. This is due to multicollinearity. In the first *restricted regression* column a corrected lag pattern is presented. We see that the impact of the startup rate 1984-87 is strongest. The impact of 1980-83 startups, however, is zero: the t-value is extremely

low.²⁰ This pattern suggests that the lag is approximately 4 to 7 years. The validity of imposing the Almon restriction is formally confirmed by the F-test on parameter restrictions.

However, as t-values for both combinatory variables are low, we suspect that multicollinearity may still influence results in the middle column to some extent. Therefore, we test an additional restriction. As both parameter estimate and t-value of 1980-83 startup rate are low, we impose the effect of 1980-83 startups to be zero. This extra restriction, which can be written as $\beta_3 = 0$, also implies that the employment impacts of 1987-90 startups and 1984-87 startups are equal.²¹ In the last column we see that both the unrestricted and the restricted parameter estimates are significant. Also, the F-test on valid restrictions is not rejected. We therefore conclude that the employment impact of 1980-83 startups is zero and that the employment impacts of 1987-90 startups and 1984-87 startups are equal and significantly positive.

Using the estimation results from the last column in Table 4, the employment impact of the startup rate can be written as a function of the lag length of the startup rate as $\beta_{i^*} = 1.58(i^*/3) - 0.53(i^*/3)^2$, where i^* is the lag length in years.²² The employment impact of startup rates is maximised after 4.5 years and extinguished after 9 years, counting backwards from 1991.²³ So, according to this formula, startups from 1986-87 contribute most to employment growth 1991-98, whereas new-firm startups founded in 1983 or earlier do not contribute to employment growth beyond 1991.

The different results for the unrestricted and restricted regression clearly demonstrate the necessity to take account of intertemporal correlations between the different lags of the startup rate.

(v) Magnitude of the effects

We now examine the magnitude of the effects. The coefficients from “separate regressions” overestimate the employment effect as these coefficients partly reflect the impact of new-firm startups from different periods, as was shown above. To establish the correct average impact of one new-firm startup, we use the coefficients from the last column of Table 4. The estimated parameter of the sector adjusted startup rate 1987-90 is 1.06. But this requires interpretation. The dependent variable

²⁰ Recall that in the restricted regression columns in Table 4, the coefficients of the startup rate variables 1987-90, 1984-87, and 1980-83 are linear combinations of the coefficients of the combinatory variables COMBI1, COMBI2, and COMBI3. In other words, the bold-printed coefficients are *restricted* parameter estimates.

²¹ This is clear when the restriction $\beta_3 = 0$ is substituted in equation (A3) in Appendix 2: this results in $\beta_1 = \beta_2 = -2\gamma_2$. Again, we refer to Appendix 2 for further details.

²² The lag length in years is denoted as i^* . One unit in i corresponds to a period of three years, i.e., $i = i^*/3$. Again, details are in Appendix 2.

²³ Fritsch and Mueller (2004) apply the Almon method for West-German regions in the period 1983-2002. Using a polynomial of third degree, they find an optimal lag of eight years.

equals $100(Empl_{1998} - Empl_{1991})/Empl_{1991}$, where *Empl* stands for employment. The independent variable equals $1000 \sum_{i=1987}^{1990} NFF / (4 Empl_{1987})$, where *NFF* stands for new-firm formation.

Due to data limitations we use four times 1987-employment, instead of the sum of employment over the years 1987-1990. For simplicity we assume that employment in 1987 equals employment in 1991, so the impact of one new-firm startup on absolute employment change is $(1.06 \times (1000/4))/100 = 2.7$. So, *ceteris paribus*, one new firm started in the period 1987-90 on average created 2.7 net new jobs in the period 1991-98.²⁴ The employment impact of 1984-87 is also 2.7 jobs per startup. Note that these jobs are additional to the jobs created by the 1987-90 startups.

(vi) Comparing these results with those from other studies

Our findings for Great Britain show similarities to those of Audretsch and Fritsch (2002) for German regions. They also find no short-term effect on employment of startups in the early to mid 1980s, but they do find a short-term employment effect of the early 1990s startups.²⁵ The common finding, for both Britain and Germany, is that the short-term effect of new-firm startups is higher in the 1990s than in the 1980s.

Our results for the 1980s, however, differ from those of Ashcroft and Love (1996) for virtually the same British counties. As noted earlier, they find a strong positive effect of new firms started in the period 1980-88 on net employment change in the period 1981-89. They employ a model in which both employment change and new-firm formation are explained with only a one year lag, allowing for interdependencies between these two variables. The employment effect in their study is certainly stronger than our short-term result for the 1980s.

One possible explanation for the differences may again be the different lag structures employed in the two models. In their model Ashcroft and Love relate new-firm formation 1980-88 to net employment change 1981-89, whereas in this paper the lags are of a minimum of three years (taking the mid year of our startup rate variables as reference year).²⁶ Given the findings of this paper that the relationship strengthens over time, we believe our results to be more robust.

²⁴ It is important to realize that these 2.7 jobs do not necessarily have to be created in the new firms themselves. It is also possible that (part of) these jobs are created in incumbent firms, but that this is induced by competitive pressure from the new entries. In other words, the 2.7 jobs is the total net effect; we cannot distinguish between direct and indirect employment effects.

²⁵ Audretsch and Fritsch however do not control for region-specific effects (by means of regional dummies), or wage growth.

²⁶ Note that a lag of three years in the present paper is not comparable with the one year lag used by Ashcroft and Love. In their method, the one year lag is counted backward from the *end* year of the employment change period, whereas we count back from the *start* year of the employment change period. So the lags in the present paper are considerably larger than the difference between 3 and 1 year suggests. In fact, in Ashcroft and Love, the years in which employment change and startup activity are measured display an 80% overlap, possibly resulting in the reversed causality problems described earlier. In the present paper we deliberately choose non-overlapping periods.

7. Discussion and Implications

In contrast with the expectations of the policy makers at the time, this paper finds no evidence that changes in GB new firm formation rates in the 1980-83 period explained changes in employment 1984–91. Indeed for the “unenterprising” and high unemployment area of the North East of England, raising rates of new firm formation is associated with employment reduction. It is only later in the decade that increased rates of new firm formation nationally appear to lead to job creation. Nevertheless the 1980’s was a decade in which national policy focussed on raising new firm formation as a key strategy for creating jobs and lowering unemployment.

That policy, however, began to be reviewed in the early 1990’s and, by 1993, had been radically switched. Instead of a focus on startups, British policy, with the exception of Scotland, was directed towards, established, rather than new firms, and the job creation impact of new firms in that decade, nationally, was positive and significant.

Scotland, however, adopted the reverse strategy. It sought, explicitly, to raise new firm formation rate as a mechanism to promote job creation. We show that, in GB, new firm formation in the 1987-90 period was significantly positively associated with employment growth in the 1991-8 period. In Scotland however, increases in new firm formation lead to falling employment. The results for Scotland therefore provide no support for policies which seek to raise new firm formation as a mechanism for stimulating job creation, particularly in areas deemed to be “lacking in enterprise.” Our interpretation is that the ineffectiveness of the Business Birth Rate Strategy in Scotland [Fraser of Allander (2001)] was probably a blessing in disguise since a rise in new firm formation would, on these grounds, have lead to falling employment.

Overall, our findings are important for public policy makers for several reasons. First, the considerably bigger short-term (and possibly long-term) employment impact of 1990s births, compared with early 1980s births, is likely to reflect “Enterprise Policy” changes. As Greene (2002) argues, the 1980s in Britain was a decade in which the key objective was to maximise the number of business startups. In contrast, the 1990s saw a shift towards policies to improve the “quality” of the SME sector as a whole. Given that major policy shift it is unsurprising -although reassuring- to observe bigger employment impacts in the 1990s, than in the previous decade.

Nevertheless this paper makes it clear that increases in birth rates can lead to additional job creation in the short and medium term. Much less clear is whether a public policy-induced increase in birth rates is a consistently cost-effective way of enhancing employment in the medium term. Our interpretation of our findings is that it is not for two reasons. The first is that the only area, in the 1990s, with a clear (public) policy to promote new-firm births was Scotland. Yet it was Scotland, (along with Wales), where the job creation impact of a new startup was significantly lower than elsewhere.²⁷ The fact that the business birth rate policy is implemented in 1993 – after the 1990-93 period when we

²⁷ In 2002 Scottish Enterprise announced the effective abolition of its Business Birth Rate Strategy, replacing it with a greater focus on SMEs with potential for growth. However, in 2001, an Entrepreneurship Action Plan for Wales was announced with a £300 million budget, one key element of which was to raise birth rates of firms in Wales to the UK average by 2006 [National Assembly for Wales (2001)].

measure birth rates – means that the definitive link cannot be made . However, given the evidence presented in the paper the likelihood that the business birth rate policy was effective in terms of job creation seems very small.

The second key finding is that startups had a much greater impact on job creation in the 1990s than in the 1980s, even though raising the startup rate was the key policy objective in the 1980s. Our interpretation is that “birth rate policies” lead to individuals with limited human capital -who are often unemployed- being encouraged to start in business. Such individuals are likely to be very transitory business owners and very unlikely to start and develop businesses with employees [Storey and Strange (1992); Greene, Mole and Storey (2004)]. This suggests that, if the objective is to enhance employment, implementing old- fashioned “birth rate” policies is difficult to justify from this research.

Unfortunately current UK policy documents appear to signal a return to such policies. HM Treasury (2002) refers consistently to an “enterprise gap”, and in its Foreward says “...and across the UK, start-up rates in the best performing areas are ten times those of the worst, contributing to an enterprise gap in our inner cities of 88,000 companies, £5 billion in turnover and tens of thousands of jobs... we cannot close that overnight”. In 2004 the focus on startups and enterprise culture is apparent in the government’s seven pillars of policy [Small Business Service, 2004]. The clear implication is that it is current policy to seek to close the gap by raising new firm formation, particularly in “unenterprising” areas. The lessons from this paper are that public policies to raise new firm formation, particularly in “unenterprising” areas are likely to be unproductive at best and counter-productive at worst.

References

- Anselin, L. (1988), *"Spatial Econometrics: Methods and Models"*, Dordrecht, NL: Kluwer Academic Publishers.
- Ashcroft, B. and Love, J.H. (1996), "Firm Births and Employment Change in the British Counties: 1981-89", *Papers in Regional Science* **75**, 483-500.
- Ashcroft, B., Love, J.H. and Malloy, E. (1991), "New Firm Formation in the British Counties with Special Reference to Scotland", *Regional Studies* **25**, 395-409.
- Audretsch, D.B. and Fritsch, M (1994), "On the Measurement of Entry Rates", *Empirica* **21**, 105-113.
- Audretsch, D.B. and Thurik, A.R. (2001), "What is New about the New Economy: Sources of Growth in the Managed and Entrepreneurial Economies", *Industrial and Corporate Change* **10**, 267-315.
- Audretsch, D.B., Carree, M.A. and Thurik, A.R. (2001), *Does Entrepreneurship Reduce Unemployment?*, Tinbergen Institute Discussion Paper TI 2001-074/3, Erasmus University Rotterdam.
- Audretsch, D.B. and Fritsch, M. (2002), "Growth Regimes over Time and Space", *Regional Studies* **36**, 113-124.
- Checkland, S.G. (1976), *"The Upas Tree: Glasgow 1875-1975"*, Glasgow, UK: Glasgow University Press.
- Cooper, A., Woo, C. and Dunkelberg, W (1989), "Entrepreneurship and the Initial Size of Firms", *Journal of Business Venturing* **4**, 317-332.
- Cressy, R. (1996), "Are Business Start Ups Debt Rationed?", *Economic Journal* **106**, 1253-1270.
- Disney, R., Haskel J., and Heden Y. (2003), "Restructuring and Productivity growth in UK Manufacturing", *Economic Journal* **113**, 666-694.
- Fothergill, S. and Gudgin, G. (1979), *"Unequal Growth"*, London: Heineman.
- Frank, M.Z. (1988), "An Inter-temporal Model of Industrial Exit", *Quarterly Journal of Economics* **103**, 333-344.
- Fraser of Allander Institute (2001), *"Promoting Business Start-ups: A New Strategic Formula; Stage 1: Progress Review; Final Report"*, Glasgow, UK: The Fraser of Allander Institute for Research on the Scottish Economy, University of Strathclyde.
- Fritsch, M. (1996), "Turbulence and Growth in West Germany: A Comparison of Evidence from Regions and Industries", *Review of Industrial Organisation* **11**, 231-251.
- Fritsch, M., and P. Mueller (2004), "Employment Effects of New Firms", *Regional Studies*, this issue.
- Garofoli, G. (1994), "New Firm Formation and Regional Development: The case of Italy", *Regional Studies* **28**, 381-393.
- GEM (2000), *"Global Entrepreneurship Monitor: 2000 Executive Report"*, Kansas City, MO: Kauffman Center for Entrepreneurial Leadership.
- Granger, C.W.J. (1969), Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* **37**, 424-438.
- Green, A.E. and Turok, I. (2000), "Employability, Adaptability and Flexibility. Changing Labour Market prospects", *Regional Studies* **34**, 599-600.
- Greene, F.J. (2002), "An Investigation into Enterprise Support for Young People", *International Small Business Journal* **20**, 315-336.
- Greene, F.J., Mole, K.F. and Storey, D.J. (2004), "Does more mean worse? Three decades of enterprise policy in the Tees Valley", *Urban Studies* **41**, 1207-1228.

HM Treasury and Small Business Service (2002), *Enterprise Britain: A modern approach to meeting the enterprise challenge*, London: HM Treasury.

Johnson, P.S. and Parker, S. (1996), "Spatial Variations in the Determinants and Effects of Firm Births and Deaths", *Regional Studies* **30**, 679-688.

Jovanovic, B. (1982), "Selection and Evolution in Industry", *Econometrica* **50**, 649-670.

Keeble, D., Walker, S. and Robson, M. (1993), *New Firm Formation and Small Business Growth in the United Kingdom: Spatial and Temporal Variations and Determinants*, Research Series, No. 15, Sheffield, UK: Employment Department.

Lloyd, P.E. and Mason, C.M. (1984), "Spatial Variations in New Firm Formation in the United Kingdom: Comparative Evidence from Merseyside, Greater Manchester and South Hampshire", *Regional Studies* **18**, 207-220.

McDonald, R. and Coffield, F. (1992), *Risky Business*, London: Falmer Press.

National Assembly for Wales (2001), *National Economic Development Strategy: in Search of Economic Growth, Social Justice and Sustainable and Balanced Development*, Cardiff, UK: National Assembly for Wales.

Price Waterhouse Coopers (1999), *The West Midlands Business Survey*, Birmingham, UK: Price Waterhouse Coopers.

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

Reynolds, P., Storey, D.J. and Westhead, P. (1994), "Cross-National Comparisons of the Variation in New Firm Formation Rates", *Regional Studies* **28**, 443-456.

Romer, P.M. (1986), "Increasing Returns and Long Run Growth", *Journal of Political Economy* **94**, 1002-1037.

Small Business Service (2000), *Business start-ups and closures: VAT registrations and de-registrations 1980-99*, SBS Statistical Bulletin, Sheffield, UK: Small Business Service.

Small Business Service (2004), *A government action plan for small business*, London: Department of Trade and Industry.

Stewart, J. (1991), *Econometrics*, New York: Philip Allan.

Storey, D.J. and Strange, A. (1992), *Entrepreneurship in Cleveland 1979-1989: A Study of the Effects of the Enterprise Culture*, Research Series, No. 3, Sheffield, UK: Employment Department.

Storey, D.J. (1985), "Manufacturing Employment Change in Northern England, 1965-78: the Role of Small Businesses", in: D.J. Storey (ed), *Small Firms in Regional Economic Development*, Cambridge, UK: Cambridge University Press.

Storey, D.J. (1994), *Understanding the Small Business Sector*, London: Routledge.

Van Praag, M. and Cramer, J. (2001), "The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion", *Economica* **149**, 45-62.

Van Stel, A.J. (1999), *The determinants of employment in Europe, the USA and Japan*, Research Report 9803, Zoetermeer, NL: EIM.
[this report can be downloaded at www.eim.net]

Van Stel, A.J., and Storey, D.J. (2002), *The relationship between firm births and job creation*, Tinbergen Institute Discussion Paper TI 2002-052/3, Erasmus University Rotterdam.
[this report can be downloaded at www.tinbergen.nl]

Van Stel, A.J., Dielbandhoesing, S., Van den Heuvel, W. and Storey, D.J. (2002), *Startup activity and employment growth in regions*, Research Report 0108, Zoetermeer, NL: EIM.
[this report can be downloaded at www.eim.net]

APPENDIX 1: Illustration shift-share procedure

In this appendix we illustrate the shift-share procedure, used in this paper to compute sector-adjusted startup rates, by means of a numerical example. Basically, the shift-share procedure imposes the same sector structure (that of the country as a whole) on each region. Sector structure in this paper is measured in terms of employment (labour market approach). As an illustration, we show the calculation of both the unadjusted and the sector-adjusted average startup rate 1980-83 for Derbyshire. This county is chosen because of the relatively large difference between the two types of startup rate for this region.

Table A1: Illustration shift-share procedure; the case of Derbyshire, 1980-83

sector	average startup rate 1980-83, by sector ¹	employment share 1981, Derbyshire	employment share 1981, Great Britain
production	1.519	0.470	0.323
construction	23.294	0.049	0.052
trade & catering	24.478	0.145	0.196
transport & communication	7.559	0.054	0.067
other services	3.848	0.282	0.363
average unadjusted startup rate 1980-83, Derbyshire		6.890	
average sector-adjusted startup rate 1980-83, Derbyshire			8.392

¹ Number of VAT registrations per 1000 workers.

In the second column of Table A1 the startup rate (averaged over the years 1980-83) is reported for each sector of economy (except for agriculture which is excluded from the current study). The startup rate of a region is a weighted average of the sectoral startup rates. The difference between the unadjusted startup rate and the sector-adjusted startup rate of a region is the weighting scheme applied to the sectoral startup rates. For the unadjusted startup rate, the employment shares of the sectors of the region under consideration (in this case Derbyshire) are taken as weights. This results in a value of 6.890. For the sector adjusted startup rate of Derbyshire, we take the sectoral employment shares of Great Britain as a whole as weights. This results in a value of 8.392.

The difference between the sector-adjusted and the unadjusted startup rate is equal to 21.8%, which is quite large.²⁸ Closer inspection of Table A1 reveals that the differences between the employment shares of Derbyshire and Great Britain as a whole, of especially the sectors production and trade & catering, account for the large difference between the unadjusted and the sector-adjusted startup rates. The relatively large share of production in Derbyshire (a sector with a low startup rate) and the relatively small share of trade & catering (a sector with a high startup rate) account for the lower value for the unadjusted startup rate, compared to the sector-adjusted startup rate.

²⁸ In our data set the difference between the sector-adjusted and the unadjusted startup rate ranges from – 20.9% to +21.8%.

APPENDIX 2: Data sources

The startup rate and employment change variables used are all constructed from a data base which contains four basic variables: startups, closures, number of enterprises, and employment. This data-base was constructed by EIM. These four variables are available at the sectoral (1-digit) and regional (NUTS3) aggregation level for the period 1980-99. Except for employment, each of these four variables is available on a yearly basis according to uniform regional and sectoral classifications, for the whole period 1980-99. Achieving this uniformity is not straightforward, since the crude data were delivered according to different regional and sectoral classifications. In this appendix the exact regional and sectoral aggregation levels, at which the four variables are available in the EIM-data set, are presented. Furthermore, the data sources and some characteristics of the variables are described.

Basic data

In Tables A2a and A2b, we give an overview of the different classifications (regional and sectoral), according to which the four variables are available in the basic data files. Also, the exact years for which the variables are available (for employment there are some missing years), are tabulated.

Table A2a: Available years and classification schemes in basic data files: startups, closures and number of enterprises ^a

Period	Available years	Regional classification	Sectoral classification
1980-1993	All	pre-LGR ^b	VTC ^c
1994-1999	All	post-LGR	SIC92

^a The figures of these variables are supplied by Small Business Service.

^b LGR = local government reorganisation 1995-98.

^c VTC = VAT Trade Classification. This is effectively SIC68.

Table A2b: Available years and classification schemes in basic data files: employment ^a

Period	Available years	Regional classification	Sectoral classification
1980-1991	1981; '84; '87; '89; '91	pre-LGR ^b	SIC80
1991-1999	1991; '93; '95-'98	pre-LGR	SIC92

^a The figures of this variable are supplied by Nomis.

^b LGR = local government reorganisation 1995-98.

Startups, closures and number of enterprises: source and description

The figures on startups, closures, and number of enterprises are supplied by Small Business Service. This organisation publishes yearly figures on VAT registrations, VAT deregistrations, and the stock of VAT registered enterprises, based on data from the Inter-Departmental Business Register (IDBR; this register is administered by the Office for National Statistics). See Small Business Service (2000). The VAT-registrations and VAT-deregistrations represent the number of enterprises registering and de-registering for VAT each year. Because there is a turnover threshold for VAT (£52,000 in 2000, for example), the very smallest one person businesses are excluded from the figures. The stock of VAT registered enterprises represents the number of enterprises registered for VAT at the start of the year.

Employment: source and description

The figures on employment are taken from the Census of Employment (until 1993) and the Annual Employment Survey (from 1995 onwards) and are supplied by Nomis. The employment figures only relate to employees. Self-employed workers and unpaid family workers are thus excluded from the data. The employment figures include both full-time and part-time employees, and relate to the situation in September of each year.

Regional aggregation level and classification schemes

The regional aggregation level employed in our data set is the British NUTS3 level. This is county level in England and Wales, and local authority region level in Scotland. We thus have data at the level of the 64 regions which are listed in Table 2 of Ashcroft, Love and Malloy (1991, p. 397). In the period 1995-98, a local government reorganisation took place in Great Britain. The five tier NUTS level classification was reviewed, and the so-called unitary authorities (UAs) were introduced. As a result, geographical boundaries of some regions have changed. This implies that we have to adjust the data from before and after the reorganisation so that they become comparable (see Table A2a). For the English regions, this is easy, since the data in the basic file are given in terms of both the new and the old regions ("former counties"). But for Wales and Scotland no variables for the period 1994-99 are given in terms of the old classification. Closer inspection of the boundaries of the unitary authorities reveals that the Scottish regions can remain unchanged but that some Welsh regions have to be aggregated into larger regions, due to overlapping "new" and "old" areas. In particular, the "old" counties Gwynedd, Clwyd, and Powys are combined into one region (which might be labeled North/Mid Wales), and the "old" counties Mid Glamorgan, South Glamorgan, and Gwent are also combined (South/East Wales). This implies that the total number of Welsh regions reduces from eight to four (Dyfed and West Glamorgan remain unchanged), and the total number of British regions in our data set from 64 to 60. These 60 regions comprise 46 English counties, 4 Welsh regions, and 10 Scottish local authority regions. In the latter group of regions, the Orkney, Shetland and Western Isles are combined into one region. The 60 regions cover the whole of Great Britain.

Sectoral aggregation level and classification schemes

At the regional aggregation level described above, the four variables are all available at the sectoral 1-digit level. However, from Tables A2a and A2b, we see that three different sectoral classifications circulate: SIC68, SIC80, and SIC92. These classifications are all different, see Table A3.

Table A3: Three Standard Industrial Classifications: 1-digit level labels ^a

SIC68	SIC80	SIC92
agriculture, forestry and fishing	0 agriculture, forestry and fishing	AB agriculture; forestry and fishing
production	1 energy/water supply industries	CE mining and quarrying; electricity, gas and water supply
construction	2 extraction/manufacture: minerals/metals	D manufacturing
motor trades	3 metal goods/vehicle industries, etc	F construction
wholesale	4 other manufacturing industries	G wholesale, retail and repairs
retail	5 construction	H hotels and restaurants
catering	6 distribution, hotels/catering; repairs	I transport, storage and communication
transport and communication	7 transport/communication	J financial intermediation
finance and professional services	8 banking, finance, insurance, leasing, etc	K real estate, renting and business activities
business and other personal services	9 other services	LO public administration; other community, social and personal services
		MN education; health and social work

^a In this table, similarities in covered parts of the economy across columns are coincidental.

As was the case for the regions, some sectors have to be combined to make sectors comparable across different SICs. This results in the six-sector classification in Table A4. In this table, corresponding parts of economic activity across SICs are in the same rows. *By and large*, there are no overlapping sectors in this six-sector classification. As mentioned earlier, we do not use the data for agriculture, forestry and fishing in our analyses.

Table A4: Relation SIC68-SIC80-SIC92 classifications (1-digit level)

SIC68-sectors	SIC80-sectors (codes)	SIC92-sectors (codes)
agriculture, forestry and fishing	0	AB
production	1, 2, 3, 4	CDE
construction	5	F
trade and catering ^a	6	GH
transport and communication	7	I
other services ^b	8, 9	JKLMNO

^a This is an aggregate of four SIC68 sectors: motor trades; wholesale; retail; catering.

^b This is an aggregate of two SIC68 sectors: finance and professional services; business and other personal services.

To summarize, the EIM-data set for Great Britain contains the four variables startups, closures, number of enterprises and employment. Apart from some missing years for employment, these variables are available on a yearly basis for the whole period 1980-1999, at relatively disaggregated sectoral and spatial aggregation levels (6 sectors, 60 regions), and according to uniform sectoral and regional classifications.

APPENDIX 3: The Almon method²⁹

The Almon method is a reparameterization method that corrects for correlation between different time lags of an exogenous variable (distributed lags). Correlation between exogenous variables in a regression model is not desirable as it causes multicollinearity. This problem is often prevalent in the context of distributed lags. When the distributed lag variables are highly correlated, it is difficult to estimate individual response coefficients accurately and regular t-tests on the significance of individual parameter estimates are unreliable. The Almon method assumes that there is some “smoothness” in the lag distribution. By imposing a specific structure in the lag distribution, the multicollinearity problems inherent to free estimation can be solved. In particular, the Almon method suggests approximating the lag structure by a polynomial function. This is explained below.

Suppose we have a model of the form represented by equation (A1).

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_s X_{t-s} + \delta Z + u_t \quad (\text{A1})$$

where the X variables are the distributed lags, with maximum lag length s , and Z is a vector of other exogenous variables (either lagged or unlagged). It is clear that in our model the distributed lag variables correspond to the startup rate variables from the various periods.

Due to high correlation between the X variables with different lags, free estimation of (A1) suffers from multicollinearity. In the Almon method a “smooth” lag distribution is obtained by imposing restrictions on the parameter vector β . In particular, the Almon method suggests approximating the graph of β_i against the lag length i by a continuous function of the form

$$\beta_i = \gamma_0 + \gamma_1 i + \gamma_2 i^2 + \dots + \gamma_r i^r; r \leq s \quad (\text{A2})$$

where r is the degree of the polynomial (A2) and s is the maximum lag length.

Imposing a structure like (A2) on the estimated parameters is implemented by estimating a restricted model. The restricted model is obtained by writing explicit expressions for (A2), and rearranging the distributed lag variables, as we will show below for our employment growth model. *First*, we establish the time periods that correspond to the lags 0, 1, ..., s . A straightforward application of our model suggests that lag 0 corresponds to the period 1991-1998, while the lags 1, 2, and 3 correspond to the periods 1987-1990, 1984-1987, and 1980-1983, respectively. So s equals 3. Taking the mid years of these periods, i.e., 1988, 1985, and 1982, we see that in terms of equation (A2), the values $i=1, 2,$ and 3 correspond to time lags of 3, 6, and 9 years, respectively, measured from 1991 backwards. In other words, one unit of i corresponds to a lag length of three years. *Second*, we have not included a startup rate with lag 0 in our model, so $\beta_0=0$. This restriction reflects our argument that startup rates do not have an immediate (i.e., contemporaneous) effect on growth and inclusion of an unlagged startup rate in the model leads to problems of reversed causality. *Third*, we choose $r=2$, i.e., a quadratic polynomial form.³⁰ Writing out (A2) with $r=2$, $s=3$, and $\beta_0=0$ results in

$$\beta_0 = \gamma_0 = 0; \quad \beta_1 = \gamma_1 + \gamma_2; \quad \beta_2 = 2\gamma_1 + 4\gamma_2; \quad \beta_3 = 3\gamma_1 + 9\gamma_2. \quad (\text{A3})$$

Substituting (A3) in (A1) and rearranging terms results in

$$Y_t = \alpha + \gamma_1 (X_{t-1} + 2X_{t-2} + 3X_{t-3}) + \gamma_2 (X_{t-1} + 4X_{t-2} + 9X_{t-3}) + \delta Z + u_t \quad (\text{A4})$$

²⁹ This appendix is based on Stewart (1991, pp. 180-182).

³⁰ We consider a first degree polynomial (i.e., a straight line) too restrictive.

Equation (A4) can be estimated using OLS. The (restricted) parameters of the startup rate variables are obtained by substituting the estimates of γ_1 and γ_2 back into equation (A3). The corresponding standard errors are obtained using the ANALYZ command in TSP 4.5.

To test the validity of the parameter restrictions imposed by the Almon method a standard F-test of the form

$$F = [(S_R - S)/(s - r)]/[S/(n - k)] \quad (A5)$$

can be applied, where S_R and S are the restricted and unrestricted residual sum of squares, respectively, r is the degree of the polynomial (A2), s is the maximum lag length in equation (A1), n is the number of observations, and k is the number of regressors in the unrestricted model. Under the null hypothesis of valid restrictions, the test statistic under (A5) has an F distribution with $s - r$ and $n - k$ degrees of freedom.

In our first application, the number of restrictions $s - r$ equals $3 - 2 = 1$, while the expression $n - k$ equals $59 - 18 = 41$. The critical value of the $F(1;41)$ distribution at 5% level is 4.1. From Table 4 we see that the value of the test statistic equals 0.062, so the null hypothesis of valid restrictions is not rejected.

In our second application, where we put the employment impact of 1980-83 startups on employment growth 1991-98 equal to zero, the number of restrictions equals two. The extra restriction can be written as $\beta_3 = 0$. Substitution in equation (A3) results in $\gamma_1 = -3\gamma_2$; $\beta_1 = \beta_2 = -2\gamma_2$. So, the extra restriction also implies that the employment impacts of lags 1 and 2 (startups 1987-90 and 1984-87) are equal. Another implication is that the optimum lag is 1.5 (or 4.5 years). In this case the F-test statistic has an $F(2;41)$ distribution (critical value 3.2). The test statistic equals 0.053. So, the restriction $\beta_3 = 0$ is valid.