Births-Deaths Nexus of Firms: Estimating VAR with Panel Data

Aki Kangasharju
Antti Moisio

ABSTRACT. This paper investigates dynamic interrelationships between the births and deaths of firms at the level of the 88 Finnish subregions during 1989–1993. The births-deaths nexus of firms is analysed using vector autoregressions (VAR) and an instrumental variable estimator. The main findings are: first, that the firm births equation has a two-year lag structure, which means that firm deaths cause firm births, and firm births cause subsequent births following two-year lags; and second, that according to the firm deaths equation, which has a one-year lag structure, firm births do not cause firm deaths, but firm deaths cause subsequent deaths. A further finding is that the use of the ordinary least-squares estimator produces inconsistent, and clearly different results from those produced by the instrumental variable estimator.

1. Introduction

In recent years the births and deaths of firms have been widely studied in a spatial context (see e.g. Reynolds, Storey and Westhead, 1994a; and Spilling, 1996). In particular, regional determinants of firm births, such as unemployment and output, as well as other factors on the supply and demand side, have received much research attention (Audretch and Fritsch, 1994). A general finding is that a growth in local demand and the dominance of the small-firm sector are the primary conditions associated with firm formation at regional levels (Kangasharju, 1997). Most studies have been conducted in a static environment, applying the OLS method (see e.g. Reynolds, Storey and Westhead, 1994b). There are also studies that have analysed the regional determinants of firm deaths (see Moyes and Westhead, 1990; Keeble and Walker, 1994; and Westhead and Birley, 1994); or the joint emergence of births and deaths, i.e. studies of business turbulence such as that by Tervo and Niittykangas (1994).

Attention is now turning to the interdependencies of births and deaths; and in some cases, investigations have been carried out even in dynamic settings (see e.g. Rosenbaum and Lamort, 1992 and Love, 1993). For example, Carree and Thurik (1993) used a one-period lag structure when investigating the determinants of births and deaths. Johnson and Parker (1994) allowed for the existence of a longer lag structure, and found that firm births would have a lag structure of as many as six years, and deaths a lag structure of two years. Johnson and Parker (1996) reported that when they included variables proxying regional characteristics in the analysis, they ended up with a two-period lag structure. They concluded (1994 and 1996) that births cause deaths and deaths cause births at a regional level, although their 1994 study concluded that the latter connection was not demonstrated to any conventional level of significance. They investigated the interdependencies using OLS estimator.

The present paper continues this line of enquiry, and begins by investigating the interdependence between the births and deaths of firms in a dynamic context using the method pioneered by Holtz-Eakin, Newey and Rosen (1988) where vector autoregressions (VAR) are estimated with panel data. This method includes an instrumental variable estimator, the use of which avoids the problem caused by the correlation between the lagged dependent variable and the lagged error term. The analysis is carried out across the 88 Finnish subregions, which reflect real trading and commuting areas, and covers the years 1989–1993. The results obtained in the present paper are quite the opposite of those found by Johnson and Parker. Unlike them, we find that

Final version accepted on September 4, 1997

Centre for Economic Research
School of Business and Economics
University of Jyväskylä, Helsinki
Finland

Small Business Economics **11**: 303–313, 1998.

deaths (Granger) cause births, whereas births do not cause deaths. Further, obtained lag lengths and the signs of the parameters differ between the studies.

Second, in an attempt to elucidate the probable reasons for the different results, our paper briefly describes the differences of detail between the method used by Johnson and Parker (1994 and 1996) and that used in our study and shows that, the different estimators produce clearly different results, and thus, are the probable reasons for the conflicting results between the studies.

This paper is organised as follows. Section 2 provides a concise treatment of the theoretical framework, and section 3 discusses the relevant econometrics. Section 4 describes the data, and section 5 presents the results. Section 6 shows the effects of the OLS estimator on the results, and section 7 concludes the paper.

2. Theoretical framework

A dynamic interrelationship between the entries and exits of firms implies the existence of two opposing forces, termed the "multiplier" and the "competition" effects (Johnson and Parker, 1994). The multiplier effect occurs when births give rise to yet more births, and slow down the incidence of deaths; or when deaths give rise to further deaths and slow down the rate of new births. The competition effect occurs when births give rise to deaths and decrease the rate of births; or when deaths increase the rate of new births and slow down the rate at which deaths occur.

The multiplier effect may be generated, for example, by a new firm which stimulates output in a region through forward and backward linkages. A new firm may need to subcontract from, or it may produce suitable intermediates for, other firms. On the other hand, the bankruptcy of a crucial firm may cause the deaths of firms that had forward or backward connection with the closed-down firm. Furthermore, a new firm may serve as a demonstration of a good business idea for other potential founders of firms. This demonstration effect may be particularly applicable to the small business context. This effect is also empirically supported by the finding that the presence of small firms is one of the regional characteristics, which most clearly encourages firm formation (Kangasharju, 1997). Finally, the multiplier effect may work through changes in the income level of a region. A new firm may increase the income level of employees in a region, which leads to further foundations of firms through increased consumption. Similarly, the closing down of a firm may decrease incomes and thus cause the future deaths of firms.

The well-known concept of Schumpeterian creative destruction is an example of the competition effect. According to this model, new, more innovative and efficient new firms take over the market from older and less efficient ones, and thus firm births cause subsequent deaths. The well documented "push effect" of a recession serves as an example of deaths increasing the subsequent rate of firm births. A firm's death may cause new firm foundation, if a former employee of a closed-down firm decides to start a business of his or her own. Similarly, in the oligopolistic sense a rise in business deaths may increase the subsequent rate of births, if the closing-down of firms has the effect of lowering entry barriers in the market. A firm's death may also decrease the incidence of deaths, if it decreases competition in the market.

According to Johnson and Parker (1994), a third effect, the "Marshall effect", must also occur because the birth of a firm must necessarily be followed at some stage by its own death. There is clear evidence that most new firms do not last more than a few years (see e.g. Storey, 1988; Bennet, 1989; and Littunen, Storhammar and Nenonen, 1997). Many studies talk about the "death valley" of the first few years, which many new firms do not survive (see e.g. Gibb, 1990). In other words, the Marshall effect can be expected to occur particularly during the early years.

Table I summarises the three effects, demonstrating their expected signs assuming a one-year lag. In reality, effects may also manifest themselves over longer and varying time periods. For example, firm foundation takes time, and hence the effects leading to it may take a longer period to work through. Similarly, an incumbent does not usually close a firm down immediately difficulties begin, but instead will often struggle for a long time. Below, we follow Johnson and Parker (1996) and assume that the multiplier, competition and Marshall effects mainly occur following two-year

TABLE I
Expected sign of the multiplier, competition and Marshall effects

Partial derivative	Multiplier	Competition	Marshall
$\partial Bt/\partial B_{t-1}$	+	–	n.a.
$\partial Bt/\partial D_{t-1}$	–	+	n.a.
$\partial Dt/\partial D_{t-1}$	+	–	n.a.
$\partial Dt/\partial B_{t-1}$	–	+	+

Note: n.a. denotes not applicable.
Source: Johnson and Parker (1994).

lags. The analysis will show that this initial lag-length turns out to be appropriate.[1]

The empirical analysis given below can provide evidence only of the net effect of these different impacts. Thus, in the first four cases an empirical analysis can answer the question as to which one of the two effects, multiplier or competition, has been the stronger. The last row presents the most difficult case, because here all three effects may be at work.

3. Econometrics

To empirically investigate the three effects discussed above we estimate dynamic panel data regressions, which take the possibility of the presence of a fixed individual effect into account (see Holtz-Eakin, Newey and Rosen, 1989):

$$y_{it} = \alpha_0 + \sum_{l=1}^{m} \alpha_l y_{it-l} + \sum_{l=1}^{m} \delta_l x_{it-l} + f_i + u_t, \quad i = 1, \ldots, N \quad \text{and} \quad t = m+1, \ldots, T, \tag{1}$$

where the symbol y denotes firm births when the symbol x is firm deaths, and vice versa. The symbols α and δ are parameters, m is a lag length long enough to ensure that u_t is a white-noise error, and f_i is a possible individual effect. The individual effect summarises the influence of unobserved variables, which have a persistent effect on the dependent variable. For example, the number of firm births in a subregion might be affected by a higher-than-elsewhere tendency on the part of its population to found firms, or by a stronger-than-elsewhere political will on the part of local government to encourage the founding of firms.

Because of the possibility that each subregion will have an individual effect, we cannot simply leave term f_i out and stack all the time series-cross section observations together and use them to estimate equation (1) using OLS. Such omission of the individual effect will result in inconsistent estimates, if other right-hand side variables are correlated with it. In static panel data analyses the usual way of removing the individual effect is to use a so-called "within estimator", in which all variables are transformed by subtracting the values from the time-means. In a dynamic context such a procedure leads to inconsistent estimates, however (Nickell, 1981). Instead, we take the time difference of equation (1) to eliminate f_i and then, to avoid the correlation between the first lag of the dependent variable and that of the error term, employ an instrumental variable estimator:[2]

$$\begin{aligned} y_{it} - y_{it-1} = &\sum_{l=1}^{m} \alpha_l (y_{it-l} - y_{it-l-1}) \\ &+ \sum_{l=1}^{m} \delta_l (x_{it-l} - x_{it-l-1}) \\ &+ (u_{it} - u_{it-1}), \end{aligned} \tag{2}$$

$$i = 1, \ldots, N \quad \text{and} \quad t = (m+2), \ldots, T.$$

This procedure makes the standard assumption that f_i is time-invariant. Parameters in Equation (2) are identifiable only in the presence of a sufficient number of instrumental variables, which also change for each estimation period. Equation (2) has $2m$ right-hand side variables for each time period, the parameters of which can be identified only when there are at least this number of instrumental variables. To be specific about the identification, and to show why the instrumental variables change for each period, let us take a look at the orthogonality conditions:

$$E(y_{is}u_{it}) = E(x_{is}u_{it}) = E(f_i u_{it}) = 0, \quad s < t \tag{3}$$

The orthogonality conditions imply that the disturbance term $(u_{it} - u_{it-1})$ will be uncorrelated with y_{it-s} and x_{it-s} for $s \geq 2$. Thus, the instrumental variables change, because the $2(t-2)$ variables $[y_{it-2}, \ldots, y_{i1}, x_{it-2}, \ldots, x_{i1}]$ are available as instruments for each time period, t. In order to have at least as many instrumental variables as right-hand side variables, it must be that $2(t-2) \geq 2m$, i.e. estimable periods t must be equal or greater than $m + 2$ (i.e. $t \geq m + 2$).

Holtz-Eakin, Newey and Rosen (1988) suggest the following estimation procedure for the above situation. First, one has to make an assumption about lag length, m. Then, one has to think of (2) as a system of ($T - m - 1$) equations with constraints across the equations, viz. αs and δs are the same in each equation (This assumption will be relaxed below). An efficient estimator consists of three steps:

1) Estimate the equations for each time period using 2SLS;
2) Estimate the joint covariance of the error terms, using the residuals and the matrix of instruments;
3) Estimate all the parameters simultaneously using the GLS of the stacked equations.[3]

The present paper uses this procedure to test three linear constraints. First, we test whether the parameters are constant. A standard assumption in studies using panel data is that all parameters, including the individual effect, are constant. The present study tests the constancy of the parameters, and so our first and the most general model therefore allows parameters α and δ initially to be time-variant. Individual effects are still assumed to be time invariant because of data availability (we do not have instruments enough for this specification). Thus, our first null hypothesis is that the parameters can be restricted to be constant. Second, we also try to restrict the model by shortening the lag length, m, conditional upon the outcome of parameter constancy. In other words, we use nested testing,[4] and test the second null hypothesis in the case of constant parameters if we cannot reject the first null hypothesis. If we do reject the first null hypothesis, then we test the second null hypothesis in the case of nonconstant parameters. The second null hypothesis is that, compared with the situation where the lag equals m, lag length can be restricted to $m - 1$. Last but not least, we try to further restrict the model and test whether or not x (Granger)[5] causes y conditional upon the outcome of both parameter constancy and lag length. Here our hypothesis is that all δ's are jointly zero (see equation 2).

The test statistic used is denoted as $L = SSE_R - SSE$, where SSE_R is the sum of the squared residuals of the more restricted model and SSE is the sum of the squared residuals of the less restricted model (Holtz-Eakin, Newey and Rosen, 1988). More and less restricted models refer to the method of nested testing. L is χ^2-distributed as N grows, and it has degrees of freedom equal to the degrees of freedom of SSE_R minus those of SSE (this equals the number of restrictions in the more restricted model). Degrees of freedom for SSE and SSE_R are calculated by subtracting all the estimated parameters from the instrumental variables. In practice, the test procedure is such that we compare the received sums of the squared residuals obtained using the more restricted equation with SSE obtained using the more general (unrestricted) equation. If the rise in SSE (i.e. L) exceeds the critical χ^2 value, we reject the null hypothesis, whereas we accept it where the rise is smaller than the critical value.

Holtz-Eakin, Newey and Rosen (1988) stress the fact that, when testing hypotheses, the same covariance matrix must be used. Therefore we use the covariance matrix of the most general model in all estimations. In addition, Holtz-Eakin, Newey and Rosen (1989) point out that inferences about causality will be incorrect if the lag distribution or parameter constancy is wrongly chosen. To avoid these type II errors, we evaluate tests of significance for lag length and parameter constancy at the 0.10 level rather than the conventionally used 0.05 level.

4. Data

The sources of the data used consist of issues of Statistics Finland.[6] At regional level, data on firm births and deaths are available from 1989. Until the year 1994 a business enterprise was considered as a firm on the basis of the amount of turnover tax paid, or if it was registered as an employer. Because data are not available for the whole of 1994, only the five years from 1989 to 1993 are available for regional analysis. Since 1995 Statistics Finland has started to compile a new time series on the basis of value-added tax (VAT). Because the identification of parameters implies that $t \geq m + 2$, we have two periods to estimate, 1992 and 1993.

A problem which arises with the data used is that a change in the ownership or legal status of a firm means a new firm is recorded in the register. This problem will be kept in mind when interpreting the results obtained. Another problem was that a change in the system of turnover tax affected the number of firm registrations in 1991. Because of this change, about 4,500 additional firms were subjected to turnover tax. The effects of this change on the number of firm formations in 1991 have been nullified in the data used.

The data contained no information about the economic and demographic characteristics of the subregions. Such variables typically play an important role in regression analyses of firm births and deaths. However, to the extent that economic and demographic characteristics can be regarded as individual effects, this absence of information will cause no problems. The statistical procedure discussed above eliminates these effects via differencing. The caveat is, however, that the differencing may not eliminate all effects. This means that the results obtained may not be entirely conclusive; but they are, at least, suggestive of the causal relations between firm births and deaths at the subregional level.

By the year 1991 a severe recession had overtaken Finland. Therefore, we included a dummy variable for the year 1992 in every equation. This dummy will capture any underlying trend in the data between the years 1992 and 1993, as well as important macroeconomic influences that affected all subregions similarly in 1992 when compared to 1993. The analysis presented in the next section uses differences of logarithmic variables of firm births and deaths. Therefore the variables can be regarded as growth rates; and they need not be proportioned, for instance by workforce or by the number of existing businesses as is the practice in static single-equation studies.

Table II provides descriptive statistics for the whole country and its subregions. The total number of firms increased in 1989 and 1990 as the number of firm births exceeded that of firm deaths. In 1991 and 1992, the severe recession caused the number of firm deaths to exceed that of births. In 1993 the Finnish economy was already showing some signs of recovery since more firms were born than died. The following econometric analysis is carried out using data from the 88 Finnish subregions. These represent relevant-sized units for the present investigation, because the vast majority of new firms operate in markets formed by common trading and commuting areas, and are most likely to co-operate with firms located in those areas. In addition, location and agglomeration benefits (accruing from the presence of firms in the same and interlinked industries), which contribute to firm foundation, are most likely to occur within a common trading and commuting area. The subregions have been drawn up by the Finnish Government on the basis of small economic areas. They are sub-divisions of the 19 regions, which constitute the basic units of Finnish and European regional policy in Finland. In 1993 the average population of the subregions was 57,500 inhabitants (range 2,407 to 1.038,851), the average number of establishments per thousand inhabitants was 27 (range 22 to 40), and the average per capita gross subregional product was FIM 69,366 (range FIM 37,932 to FIM 122,371).

TABLE II
Descriptive statistics

Number of firm births and deaths in the whole country					
	1989	1990	1991	1992	1993
Births	22,229	19,582	18,000	18,852	19,600
Deaths	15,127	17,481	20,250	20,400	18,900

Annual average growth rate of firm births and deaths across subregions, %				
	1989–1990	1990–1991	1991–1992	1992–1993
Births	–14.1 (0.175)	–7.2 (0.150)	–1.3 (0.184)	2.6 (0.187)
Deaths	8.9 (0.178)	22.9 (0.171)	–11.5 (0.229)	–5.2 (0.183)

Note: Standard deviations are in parentheses.

At this subregional level, the number of firm births declined by 14 percent on average from 1989 to 1990, and the decline continued until 1992 (Table II). In 1993 an average of 2.6 percent more firms were born in the subregions than had been the case in 1992. The number of firm deaths rose both in 1989 and 1990, but began to decline thereafter, as shown by the figures: in 1992 an average of 11.5 percent fewer firms died than in 1991; compared to 5.2 percent fewer in 1993 than in 1992.

5. Results

Table III shows the results for the *firm births* equations. First, we started with m = 2 and assumed time variant parameters and constant fixed effects. This yielded a SSE value of 4.31 (line i). At the 0.10 level, the χ^2 value with 2 degrees of freedom[7] is 4.61. Therefore we accepted constant fixed effects and m = 2 as the starting point.

Second, we tested to see whether, in addition to the constant fixed effects, all the parameters were constant. Specifically, we ascertained whether the parameters for 1992 were similar to those for 1993. The SSE value for this specification is 6.23 (line ii), and the value for L, 1.92 (6.23 – 4.31). With 4 degrees of freedom our hypothesis of constancy is confirmed, as the critical χ^2 value is 7.77. In this restricted model there are 6 degrees of freedom, which were subtracted from those obtained in the unrestricted model. (Otherwise stated, there are 4 degrees of freedom in the restricted model, because we restricted the four lag parameters for 1992 to their 1993 values.)

Third, we tried to ascertain whether or not we could drop the second lag parameters. The L value for this specification is 13.93 (20.16 – 6.23). We subtracted the SSE statistic in line (ii) (and not those in line (i)) from the received one, as we accepted the hypothesis of parameter constancy (nested testing). With 2 degrees of freedom the hypothesis of m = 1 had to be rejected. Now, in this case the number of restrictions is, i.e. degrees of freedom are, 2; because the second lag parameters for both variables were restricted to zero.

Finally, since the previous hypothesis was rejected, we returned to specification (ii) and tested our causality hypothesis against it. The null hypothesis was that variables for business closures could be dropped from the equation. The L value for this specification is as much as 21.28 (27.51 – 6.23). This specification, too, has 2 degrees of freedom, as we restricted both the lag parameters of the death variable to zero. Even if we accepted the less strict 0.05 level of significance, we had also to reject the hypothesis that deaths would not cause entries. Thus, we concluded that explicable both in their own terms, and as a result of firm deaths (assuming a two-year lag structure).

Table IV shows the parameter estimates of the specifications. On the one hand, in specification (i) the parameters seem to differ somewhat from each other. However, as the test above showed, the differences are not large enough to disallow the hypothesis of constancy. The inclusion of the time dummies, which control for common macroeconomic shocks, is one factor contributing to this acceptance of the hypothesis. On the other hand, the parameters for 1993 seem to be rather similar in all specifications. Specification (ii) turned out to be the most parsimonious.

TABLE III
Firm births equations: IV estimator

Specification	SSE	DFs for the specification	L	DFs for L	χ^2
i) Stationary fixed effects, m = 2	4.31	2	–	–	4.61
ii) All parameters stationary (given i)	6.23	6	1.92	4	7.77
iii) m = 1 (given ii)	20.16	8	13.93	2	4.61
iv) Exclude deaths (given ii)	27.51	8	21.28	2	5.99

Note: For lines i through iii, χ^2 is evaluated at the 0.10 level of significance and for line iv, at the 0.05 level. SSE is the sum of squared residuals for the specification and DFs are degrees of freedom.

TABLE IV
Parameter estimates in the firm births equations

	(i) Coef. (*t*-values)	(ii) Coef. (*t*-values)	(iii) Coef. (*t*-values)	(iv) Coef. (*t*-values)
Constant	0.115 (0.4)	0.099 (1.3)	0.039 (0.8)	–0.119 (–1.9)
1992				
Dummy	2.160 (0.5)	–0.046 (–0.7)	–0.037 (–0.7)	0.102 (1.9)
Δbirths$_{t-1}$	0.428 (0.5)			
Δdeaths$_{t-1}$	0.677 (0.5)			
Δbirths$_{t-2}$	0.131 (0.2)			
Δdeaths$_{t-2}$	0.047 (1.2)			
1993				
Δbirths$_{t-1}$	–0.301 (–1.0)	–0.293 (–1.1)	–0.212 (–1.4)	–0.593 (–2.2)
Δdeaths$_{t-1}$	–0.558 (–4.3)	–0.537 (–4.5)	–0.303 (–3.5)	
Δbirths$_{t-2}$	–0.139 (–0.6)	–0.163 (–0.9)		–0.374 (–2.2)
Δdeaths$_{t-2}$	–0.390 (–2.8)	–0.426 (–3.5)		

In this specification, the coefficients of the death variables are negative; which suggests that positive rates of growth of firm deaths around the turn of the decade reduced the rate of growth of firm births in 1992 and 1993. With respect to the conceptual model outlined above, this result implies that the multiplier effect prevailed over the competition effect. This suggests that the increased death rate of firms decreased incomes and consumption, and hence firm formation in the subregions. Because the data does not allow us to distinguish between real firm death and a change in legal status or ownership, this result may also mean for example that increased firm takeovers discouraged subsequent firm births, and made firm owners to protect themselves against takeovers.

In specification (ii) in Table IV, the coefficients of the firm birth variables are also negative, which suggests that the negative growth rates of firm births around the turn of the decade increased the birth rate of firms in 1992 and 1993. With respect to the conceptual model, this implies that the competition effect was stronger than the multiplier effect in this case. The negative growth rates of firm births reduced competition, or at least eased Schumpeterian creative destruction, as a smaller number of enthusiastic firms entered the market. This contributed to subsequent firm births in the subregions.

The numerical values and the level of significance of the coefficients in this specification imply that the effect of lagged deaths on subsequent births was stronger than that of lagged births on subsequent births. Therefore these results imply that the net effect of firm births and deaths during the turn of the decade on firm births in 1992 and 1993 was negative. Statistically insignificant variables even raise the question of whether firm births have autoregressive features at all. We examined this question, and found that lagged birth variables cannot be excluded from the firm birth equation.

Table V provides the results for the *business deaths* equations. As with the firm birth equations, we started with a general specification and hypothesised that $m = 2$ and that the fixed effects are constant. As the SSE value is 0.75 and the critical value at the 0.10 level is 4.61, we inferred that this was an acceptable starting point (line i). A further restriction, i.e. the hypothesis that all parameters are constant, yielded an L value of 1.05 (1.80 – 0.75). With 4 degrees of freedom this hypothesis could not be rejected, as the critical value is 7.77 (line ii). Therefore, we went on to adopt the specification of constant parameters, and tested to see whether or not we could drop the second lag parameters. Here, our hypothesis of $m = 1$ was accepted, as the value of L was 2.45 (4.25 – 1.80) and there were 2 degrees of freedom (line iii). This result raised the question of further restricting the lag parameters. Fourthly, when we imposed the further restriction that $m = 0$, the obtained L was 6.62 (10.87 – 4.25) and the null hypothesis was rejected, as there were 2 degrees of freedom.

TABLE V
Firm deaths equations: IV estimator

Specification		SSE	DFs for the specification	L	DFs for L	χ^2
i)	Stationary fixed effects, $m = 2$	0.75	2	–		4.61
ii)	All parameters stationary (given i)	1.80	6	1.05	4	7.77
iii)	$m = 1$ (given ii)	4.25	8	2.45	2	4.61
iv)	$m = 0$ (given iii)	10.87	10	6.62	2	4.61
v)	Exclude births $m = 1$ (given iii)	4.86	9	0.61	1	3.84

Note: For lines i through iv, χ^2 is evaluated at the 0.10 level of significance and for line v through vi, at the 0.05 level. SSE is the sum of squared residuals for the specification and DFs are degrees of freedom.

Finally, we tested so ascertain whether firm births cause subsequent deaths under the conditions where the parameters are constant and $m = 1$. The L obtained was 0.61 (4.86 – 4.25). With one degree of freedom we could not reject the hypothesis that business births do not cause subsequent deaths (the critical χ^2 value is 3.84 at the 0.05 level of significance). Thus, we conclude that firm deaths explain themselves with a one-year lag structure only.

Table VI shows the parameter estimates for the firm deaths equations. Just as with the firm births equations, here too the parameters in the first specification seem to differ from each other. However, as the test showed, the differences are not large enough to lead to the rejection of the hypothesis of parameter constancy. Among the firm death equations, the specification (v) turned out to be the most parsimonious.

In this specification, the coefficient of the death variable is negative, which tends to suggest that the positive rates of firm deaths around the turn of the decade lowered the rate of deaths in 1992 and 1993. With respect to the conceptual model, this implies that the competition effect prevailed over the multiplier effect and that the positive rates of firm deaths eased competition and thus lowered the rate of deaths in 1992 and 1993. Once again, the data does not allow us to distinguish between real firm death and a change in legal status or ownership. Therefore increased firm takeovers and changes in legal status of firms may also have contributed to a decrease in the subsequent firm death rate. Because we found that firm births do not affect the subsequent rate of firm deaths, we infer that in this case the net impact of all the multiplier, competition and Marshall effects is zero. Just as in the case of the firm birth equation,

TABLE VI
Parameter estimates in the firm deaths equations

	(i) Coef. (t-values)	(ii) Coef. (t-values)	(iii) Coef. (t-values)	(v) Coef. (t-values)
Constant	0.532 (0.7)	0.045 (0.3)	–0.039 (–0.4)	–0.044 (–0.4)
1992				
Dummy	6.974 (0.7)	–0.080 (–0.6)	–0.036 (–0.3)	–0.033 (–0.3)
Δbirths$_{t-1}$	0.754 (0.4)			
Δdeaths$_{t-1}$	2.529 (0.8)			
Δbirths$_{t-2}$	0.571 (0.4)			
Δdeaths$_{t-2}$	–0.046 (–1.3)			
1993				
Δbirths$_{t-1}$	0.113 (0.3)	0.219 (0.7)	0.099 (0.8)	
Δdeaths$_{t-1}$	–0.331 (–2.2)	–0.362 (–2.5)	–0.179 (–2.3)	–0.187 (–2.5)
Δbirths$_{t-2}$	0.026 (0.1)	0.106 (0.5)		
Δdeaths$_{t-2}$	–0.192 (–1.3)	–0.220 (–1.6)		

the net effect of firm births and deaths on future deaths is negative.

6. Effects of the OLS estimator on the results

It appears that our main results for both firm birth and death equations are quite contrary to those obtained by Johnson and Parker (1994 and 1996), who used a method resembling that used in the present study. We found that deaths cause births, but births do not cause deaths; whereas Johnson and Parker found the opposite to be true at the 0.05 level of significance. They found that deaths do not cause births, but births in fact cause deaths. There also appear to be differences in the signs of the parameters and lag-lengths between the studies.

Apart from the differing time periods and regions under investigation, the probable reasons for the contradictory results are twofold. The primary reason is probably that Johnson and Parker used the OLS estimator, which produces inconsistent results in the current context, because the lagged dependent variable is correlated with the lagged error term. We avoided this problem by employing the instrumental variable estimator, which is also heteroskedasticity-consistent. Other reasons are that, unlike Johnson and Parker, we did not restrict the parameters to initially being constant, and we included time dummies to allow for common economic shocks.

In order to illustrate the effects of the OLS estimator on the results, we replicated the above analysis by applying that estimator. Tables VII and VIII show the results. As far as firm birth equations are concerned, the application of the OLS estimator suggests that neither firm births nor firm deaths would determine the subsequent births of firms. This result is quite the opposite of that obtained by the instrumental variable estimator above. There, both variables turned out to influence firm births with a two-year lag length. In the case of the firm death equations, the results are somewhat closer to those obtained above. A similarity between the results is that the application of the OLS estimator also suggests that firm births do not cause future deaths. A difference is that the lag length is two years in this instance, whereas it was one year when the instrumental estimator was used. All this suggests that the choice of the

TABLE VII
Firm births equations: OLS estimator

Specification	SSE	DFs for the specification	*L*	DFs for *L*	χ^2
i) Stationary fixed effects, $m = 2$	4.77	9	–	–	14.7
ii) All parameters stationary (given i)	5.25	5	0.48	4	7.77
iii) $m = 1$ (given ii)	7.85	3	2.60	2	4.61
iv) $m = 0$ (given iii)	11.85	1	4.00	2	4.61

Note: χ^2 is evaluated at the 0.10 level of significance. SSE is the sum of squared residuals for the specification and DFs are degrees of freedom.

TABLE VIII
Firm deaths equations: OLS estimator

Specification	SSE	DFs for the specification	*L*	DFs for *L*	χ^2
i) Stationary fixed effects, $m = 2$	3.50	9	–	–	14.7
ii) All parameters stationary (given i)	3.62	5	0.12	4	7.77
iii) $m = 1$ (given ii)	14.3	3	10.6	2	4.61
iv) Exclude births (given ii)	3.88	3	0.26	2	5.99

Note: For lines i through iii, χ^2 is evaluated at the 0.10 level of significance and for line iv, at the 0.05 level. SSE is the sum of squared residuals for the specification and DFs are degrees of freedom.

estimator in the dynamic panel data context is crucial and clearly affects the results obtained.

7. Conclusion

This paper investigated the dynamic interrelationship between firm births and deaths at the subregional level between 1989 and 1993. The main findings were that, first, the incidence of firm births does not determine the incidence of later deaths, whereas firm deaths do determine future births; second, that previous growth rates of firm births and deaths affect future firm births with a two-year lag structure, and third, that firm deaths are determined by their own one-year lags only. Furthermore, we found that the use of the instrumental variable estimator is crucial in this dynamic panel data context, as the use of the OLS estimator appeared to bias the results.

The results obtained imply that the growth rate of firm deaths has more profound effects than the growth rate of firm births on the subsequent growth rates of births and deaths. In the firm births equation, the parameters of the firm death variables were statistically significant, unlike those of the firm birth variables. In the firm deaths equation, firm births were even found not to affect subsequent deaths at all, whereas firm deaths were found to affect future deaths.

From the point of view of regional policy this result sheds light on the issue of whether to help new firm formation, or to help existing firms to survive. According to Armstrong and Taylor (1993) this dilemma, together with that of whether to help firms in manufacturing or services, and that of whether to provide uniform help or to direct it to enterprise nodes, is currently of particular significance in creating an effective regional small firm policy. The result of the present study implies that there is a rationale in concentrating regional policy measures on existing firms, i.e. on preventing firm deaths, rather than to direct attention to new firm foundation; because firm deaths have more profound effects than firm births on subsequent business dynamics at the subregional level. Of course, this conclusion ignores the possible differences in costs of policy measures directed either to incumbents or entrants. This result only suggests that policy measures directed towards incumbents are backed up by those inherent processes at work in the subregional level.

These results are conditional upon the fact that the first time difference eliminates the major economic and demographic characteristics of the subregions as factors affecting the results. Because the caveat exists that these characteristics cannot, in fact be entirely compensated for by our treatment, the results obtained are perhaps more suggestive than conclusive. This being the case, further empirical investigation of the dynamic interrelationships between births and deaths of firms at regional levels is called for.

Acknowledgements

This study was part of a research programme funded by the Academy of Finland. We thank Professor Douglas Holtz-Eakin for supplying a special program and assisting in its use, and Dr Eva Johansson and Dr. Matz Dahlberg for helpful advice. The very valuable comments of the two anonymous referees are also gratefully acknowledged. We would also like to thank Esa Storhammar for helping us in collecting the data.

Notes

[1] This presumption can also be justified on grounds of data availability, for the reasons described below.

[2] Holtz-Eakin, Newey and Rosen (1989, p. 419) call this correlation between the lagged dependent variable and the lagged error term as the simultaneity problem, which is well known from the conventional literature on time series analysis (see also Chamberlain, 1983).

[3] The use of the GLS corrects the possible heteroskedasticity of the error terms.

[4] See Holtz-Eakin,Newey and Rosen (1988).

[5] See Granger (1969).

[6] Population Statistics: Justice; and Information Services: Business Register.

[7] Degrees of freedom were calculated by subtracting all the estimated parameters from all the instruments used in the estimation. For 1993, we had three years of valid instruments for each variable. Adding the constant, we had seven instruments. For 1992, this number was reduced to five. Because we had two years to estimate together, there were 12 instrumental variables in the estimations. For both years, we estimated 5 parameters (including the constant). Thus the degrees of freedom were 12 – 10 = 2.

References

Armstrong, H. and J. Taylor, 1993, *Regional Economics and Policy*, Cornwall: Harvester, Wheatsheaf, Padstow.

Audretsch, D. B. and M. Fritsch, 1994, 'The Geography of Firm Births in Germany', *Regional Studies* **28**(4), 359–365.

Bennet, R., 1989, *Small Business Survival*, London: Pitman.

Carree, M. and R. Thurik, 1993, Entry and Exit in Retailing: *Incentives, Barriers, Displacement and Replacement*, Centre for Advanced Small Business Economics, Erasmus University Rotterdam, and Department of Fundamental Research, Research Institute for Small and Medium-Sized Businesses (mimeo).

Chamberlain, G., 1983, 'Panel Data', in Z. Griliches and M. Intrilligator (eds.), *The Handbook of Econometrics*, vol. 2, Amsterdam: North-Holland Publishing Co.

Gibb, A. A., 1990, *Small Business in the U.K., State of Development, Expectations and Policy*, Durham University Small Business School, Occasional Papers No. 9094.

Granger, C. W. J., 1969, 'Investigating Causal Relations by Econometric Models and Cross-Spectral Methods', *Econometrica* **56**, 424–438.

Holtz-Eakin, D., W. Newey and H. S. Rosen, 1988, 'Estimating Vector Autoregressions With Panel Data', *Econometrica* **56**, 1371–1395.

Holtz-Eakin, D., W. Newey and H. S. Rosen, 1989, 'The Revenues-Expenditures Nexus: Evidence from Local Government Data', *International Economic Review* **30**, 415–429.

Johnson, P. and S. Parker, 1994, 'The Interrelationship Between Births and Deaths', *Small Business Economics* **6**, 283–290.

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

Kangasharju, A., 1997, 'Regional Variation in Firm Formation: Cross-Sectional and Panel Data Evidence From Finland', submitted to *Papers in Regional Science*.

Keeble, D. and S. Walker, 1994, 'New Firms, Small Firms and Dead Firms: Spatial Patterns and Determinants in the U.K.', *Regional Studies* **28**(4), 411–428.

Littunen, H., E. Storhammar and T. Nenonen, 1997, 'The Success of New Firms and Local Environment', Submitted to *Entrepreneurship and Regional Development*.

Love, J., 1993, 'Spatial Entry and Exit: Recursive and Simultaneous Estimations', Strathclyde Papers in Economics 93/7, *University of Strathclyde*.

Moyes, A. and P. Westhead, 1990, 'Environments for New Firm Formation in Great Britain', *Regional Studies* **24**, 123–136.

Nickell, S., 1981, 'Biases in Dynamic Models with Fixed Effects', *Econometrica* **49**, 1417–1426.

Reynolds, P., D. J. Storey and P. Westhead (eds.), 1994a, 'Special Issue: Regional Variations in new Firm Formation', *Regional Studies* **28**.

Reynolds, P., D. J. Storey and P. Westhead, 1994b, 'Cross-national Comparisons of the Variation in New Firm Formation Rates', *Regional Studies* **28**, 443–456.

Rosenbaum, D. I., F. Lamort, 1992, 'Entry, Barriers, Exit and Sunk Costs: An Analysis', *Applied Economics* **24**, 297–304.

Spilling, O., 1996, 'Regional Variation of New Firm Formation: The Norwegian Case', *Entrepreneurship & Regional Development* **8**, 217–243.

Storey, D., 1988, 'The Role of Small and Medium Sized Enterprises in European Job Creation: Key Issues for Policy and Research', in M. Giaoutze, P. Nijkamp and D. J. Storey (eds.), *Small and Medium Size Enterprises and Regional Development*, London: Routledge.

Tervo, H. and H. Niittykangas, 1994, 'Developments in the Growth, Location and Turbulence of Business Services in Finland', Papers in Regional Science: *The Journal of the RSAI* **73**, 73–95.

Westhead, P. and S. Birley, 1994, 'Environment for Business Deregistrations in the U.K., 1987–1990', *Entrepreneurship & Regional Development* **6**, 29–62.