County-Level Determinants of the Rate of New Firm Formation in West Virginia

by

Leticia Garcia

RESEARCH PAPER 2015

Abstract: This paper analyzes county-level variations in new enterprise formation in the State of West Virginia 1993-1998 using data from the County Business Patterns Economic Profile and the American Business Disk. Two econometric models are used. The first model considers differences among the 55 counties, while the second model focuses on differences among counties and among the 9 SIC sectors. The purpose of the models is to identify a range of key determinants for the creation of new firms. The choice of variables is based on previous theoretical and survey research. The dependent variable in both models is the rate of new firm formation. The explanatory variables are chosen using demand and supply considerations. On the demand side, changes in population and income are significant. Variables that affect the cost of the firm and market concentration are significant on the supply side. The intuition about both sets of variables may be that giving relevant information to the firms regarding expected profit might be important in designing incentive programs.

INTRODUCTION

Observing data from the State of West Virginia it is possible to find specific characteristics that make it particularly interesting to study some aspects of its economy. For example the average population annual growth rate between 1990-98 was 0.12% compared with the 1.8% in the US. A 0.2% decrease is estimated for 98-99. In 1996, 42% of the population was living in metropolitan areas while the number for the US was 79.8%. Turning to employment, 51.4% of the civilian population was employed in the same year. In the US 63.2% of the civilian population was employed, putting West Virginia in 50th place among the states.

Statistics related to income and poverty are also dramatic. In 1997, personal income per capita measured with constant dollars from 1992 was $16,281, or 49th nationally. West Virginia is similarly positioned with respect to the median household income. The percentage of population below the poverty level puts West Virginia in 5th place in the country, with 18.5% compared with 13.7% for the US in 1996.

The common impression is that those numbers should be closely related. From neoclassical macroeconomics, we expect to find that income, unemployment, migration and poverty are interdependent. High unemployment decreases the level of income, increasing poverty and migration. The question that immediately arises then is to determine which variables cause the high rate of unemployment.

Within this context, I observe in more detail some economic aspects of the State. The composition of the firms in a West Virginia county may be related to the high unemployment rate. More specifically, the rate of birth, growth and death of firms in the last seven years may affect substantially the unemployment rate and it is also affected by migration and poverty. So there may be a feedback effect between demand and supply side variables that are making the situation worse as in a downward spiral. Combining these supply-side and demand-side variables, are spatial patterns important in explaining the birth, growth and rate of firms in West Virginia?

The purpose of this paper is to study the influence of regional issues in the pattern of formation of new firms in the State of West Virginia using county-level data for the period 93-98. I will present some descriptive statistics that allow us to characterize some regional differences in economic structure. Furthermore, I will use two econometric models to measure the sensitivity of the dependent variable –birth rate of firms- to demand-side and supply-side influences specific to the county. The first model considers only differences among counties. The second model also incorporates differences among sectors to which the new firms belong.

The rest of the paper is divided into four sections: I. Literature Review, II. The models and the data, III. The rate of new firm formation, IV. Results, and V. Conclusions.

I. LITERATURE REVIEW

Keeble and Walker (1994) analyze county-level spatial variations in new enterprise formation, growth in numbers of small businesses, and business failures in the UK 1980-90. They use multivariate econometric models for different periods and sectors to identify a range of key determinants chosen from previous studies. The authors present the results of a series of spatial modeling analysis, which attempt to identify the most important influences of regional patterns in new firm formation, growth in numbers of small business and firm dissolution or death rates.

The method of investigation is via econometric analysis of VAT business statistics, the most widely used measure of UK business formation and dissolution in terms of a set of independent variables deliberately selected on the basis of previous theoretical and empirical work as probable determinants of such spatial variations. These include both demand-side and supply-side influences on entrepreneurship and small business growth, and while focusing primarily on key possible economic labor market

and demographic influences, also incorporate measures of policies, both national and local, which might have influenced local variations.

A review of previous theoretical and empirical research (Keeble, 1991) resulted in identification of the 31 separate independent variables listed in the table for investigation as possible influences (this table is included in the appendix). Keeble and Walker estimate 6 main models. Two of them use the rate of variation of firm's registration as the dependent variable. The first model presents the rate of new firm formation per 1000 in the labor force in 1981, while the second uses the rates of new formation per 100 businesses at the end of 1979. The explanatory variables are chosen from the list as those that can mainly explain the pattern of new firm creation.

A similar procedure is done for the growth rate of small business and for the firm deregistration or death. The Cliff-Ord test for spatial autocorrelation is applied satisfactorily to the models and f-test values yield high levels of explained variance.

Kohler (1997) analyses firm location patterns considering the existence of a *third nature* that potentially shapes the regional structure of an economy. He recognizes that *first nature* –geographic production conditions and transportation costs - and *second nature* –other agglomeration advantages- affect location patterns. Without objecting to the relevance of these factors, he shows that emerging new industries are rarely restricted by those considerations in their location choices. He models regional amenities as an important new firm location variable, showing that when "the amenity function is nonlinear, these firms bifurcate into companies that follow the workers completely to the most attractive places, and firms who choose intermediate locations at the edge or periphery of high-amenity areas". Three distinct patterns arise from the presence of regional amenities: a) Concentration of firms where the level of amenities is rapidly changing, b) a gradual change of industrial composition at intermediate locations and c) a concentration of specific industries at the periphery. In all the cases the presence of human capital in a region showed a strong effect in the decisions of the firms that tend to follow the workers towards amenable locations.

Shukla and Waddell (1991) study new firm location decisions in Dallas. They include as explanatory variables "a set of structural, land supply and agglomeration-

interaction attributes". They incorporate in the model differences among industries. Almost all hypothesized effects are found to impact location in an expected manner, though with important similarities and contrasts across industries. They employ broad spatial and industrial aggregates, one-digit SIC groups.

Fritsch (1992) presents a similar study for Germany. He takes into account the sectors to which the firms belong and concludes that the effect of regional variables in firms' decisions depends strongly on the industry.

Hansen (1990) compares the effect of differences in sectors with regional variables in the rate of small firms' creation in Mediterranean France. Using data for 22 regions, firms with 22 to 500 employees and 9 industry classes, he finds that small firm creation depends strongly on regional variables and not on the sector which the firm belongs.

Gudgin and Fothergill (1985) use some of the variables that Keeble lists to study the spatial effects in East Anglia firms. They include in the model the size of the firm, using for size the number of employees. They find that geographical variables have more influence in small firms than in big firms.

II. THE MODEL AND THE DATA

The estimated models are multivariate linear regressions where the dependent variable is the rate of new firm formation in each county in the State of West Virginia. Some of the explanatory variables were selected from the table in the appendix.

The demand-side variables are annual change in the median household income –INC-, annual change in total population –POP-, population density –DENS- and a dummy variable -D1- that takes the value of 1 if in the county more than 50% of the population is below the poverty level. INC, POP and DENS were selected assuming the hypothesis that they are the key measures of growing local market demand and hence stimuli for firm

creation. D1 works in the opposite direction; that is, it will be a disincentive for opening new firms.

In the supply side, I incorporate the variables that can affect the decision of firm creation. In general these are variables that affect the firm's operating costs. Considering the labor market, I include economically active population > 16 –ACPOP- and change in unemployment rate –UNEM-. A measure for available human capital is also included because it can affect the decision of the firm (Kohler, 1997). This variable is the percentage of population that is high school graduated or higher –EDUCPOP-.

Also from the supply side, I include three variables for measuring market concentration. This issue may have a strong impact in the firms' profits expectations and hence in the decision of firm creation. Firms' earnings –FE-, sales in county i sector j as a percentage of total sales in sector j –MARSHA-, number of employees in county i sector j as a percentage of total number of workers in sector j –SIZE-.

Another variable that may affect strongly profit expectations is taxes –TAX-. Then, this is also incorporated using the change in the total taxes from the local government general revenue.

Finally, a transportation variable is incorporated as a measure of the accessibility of the cities and towns in the county -TRANS. This variable can be considered as affecting both supply and demand sides. Specifically, for transportation, the variable is the change in the number of exits from 4-lane highways in the county between 1992 and 1999.

Two models are estimated using some of the explanatory variables listed above. The first model uses cross-section data for the 55 counties in the state. The variables are used as averages from 93 to 98. The explanatory variables are incorporated with a lag, meaning they are observed in year 0 for making the decision in year 1. I did not incorporate all the explanatory variables because some of them are only relevant if differences among sectors are considered.

The second model incorporates the differences among sectors, using the two-digit SIC classification. The hypothesis here is that there is a fixed effect that produces different intercepts in the linear regression regarding the sectors. Then, the model is

estimated using all the explanatory variables in the right-hand side plus 9 dummy variables, one for each sector. This estimation is compared with a restricted model that assumes there are no differences in the intercept. For testing the restriction, an *F* statistic is calculated.

I use data from the Economic Census for the years 1993 to 1997 and from the American Business Disk 1998 for the rate of new firm's formation. For the explanatory variables the data were collected from the Census and from the 1988 and 1994 County and City Data Books. The highway access variable was collected from the McNally's Road Atlas 1992 and 1999. The rates of new firm's formation are calculated using the number of years that the firm has been operating based on the data in the AB disk.

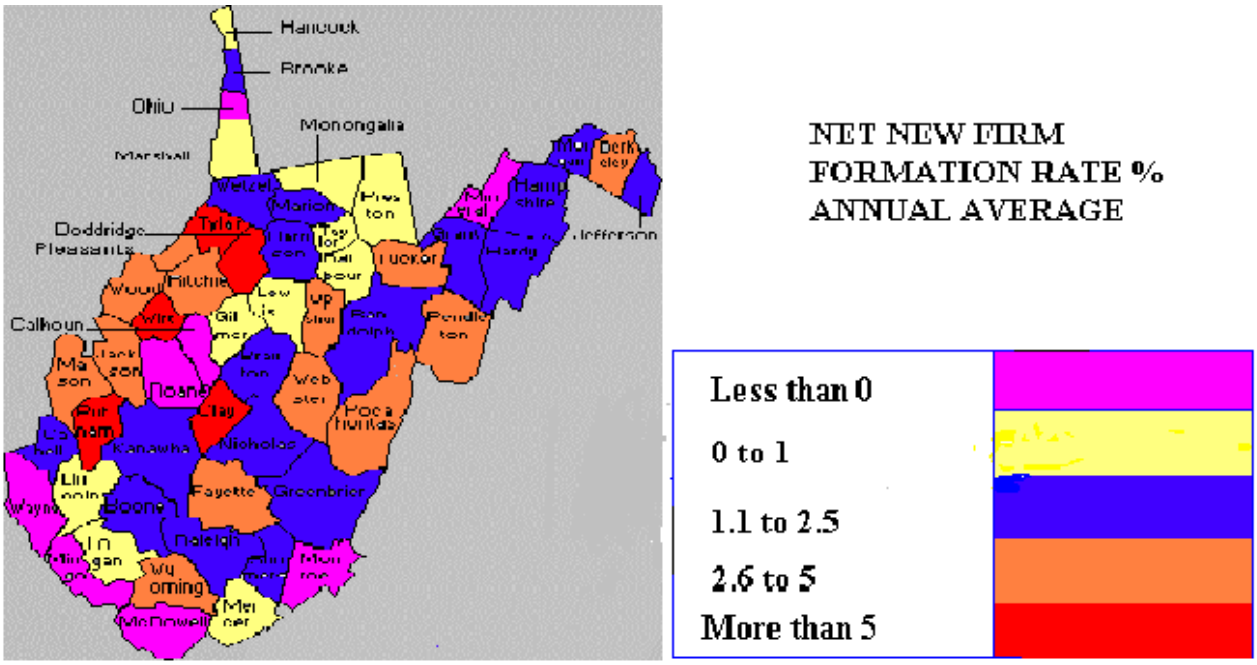
In the first model the dependent variable is calculated by county, as an annual average between 1993 and 1998. The right-hand side variables are measured as average percent variations between 1990 and 1998. In the second model the dependent variable is the new firm creation rate by county and SIC sector annual average 93-98.

III. THE RATE OF NEW FIRM FORMATION

In this part, I include the description of the dependent variable used in the models. The new firm formation rate is computed at a county level as an annual average between 1993 and 1998. The new firm formation rate is calculated as a growth rate using the difference between the number of firms operating in a specific year and the number of firms operating the year before. One limitation of this rate is that it is not adjusted for the number of firms in the county. So, if the county had 77 firms (which is the case in 1997 for the county with the smallest number of firms) in a specific year, the opening of one more firm represents much more in percentage terms than if the county had 5975 firms (the county with the highest number of firms in 1997).

FIGURE 1

REGIONAL VARIATIONS IN NET NEW FIRM FORMATION RATE IN WEST VIRGNIA 1993-1998[a]



[a] The net new firm formation rate is calculated as a growth rate using the number of firms that enter the market in a specific year and the number of firms that enter the market the year before. A negative rate means that the dynamic of new firm formation from one year to the next one has decreased in the last 6 years.

Differences among sectors in the rate of new firm creation are computed as an annual average between 1993 and 1998.

Table 1

NET NEW FIRM FORMATION RATE BY COUNTY: SECTOR DIFFERENCES

	SIC SECTOR	RATE
7	AGRICULTURAL SERVICES, FORESTRY AND FISHING	5.41%
10	MINING	-1.6%
15	CONSTRUCTION	6.45%
20	MANUFACTURING	2.86%
40	TRANSPORTATION AND PUBLIC UTILITIES	5.08%
50	WHOLESALE TRADE	3.12%
52	RETAIL TADE	0.28%
50	FINANCE, INSURANCE AND REAL ESTATE	9.23%
70	SERVICES	3.09%

These numbers show significant differences. The new firm formation rate by county goes from –16% to 11% for the period considered. At the state level the rate is 1.4%. This can be compared with the number of businesses, in absolute terms, that existed in the county in 1997, which ranged from 77 to 5975. The mean is 756.4 and the standard deviation is 953. The number of firms in West Virginia in 1997 was 41,625. The rates of net new firm creation by SIC sector shows that mining decreased 1.6% while Finance, Insurance and Real Estate (F.I.R.E.) increased 9.23%.

IV. RESULTS

In table 2 the results for the first model are presented.

Table 2

MODEL 1

COUNTY-LEVEL VARIABLES AND THE NEW FIRM FORMATION RATE

Independent variable	**Coefficient**	***t*-statistics**
CONSTANT	-3.899	-0.872
INC	13.412**	5.763
POP	2.699**	3.71
DENS	4.022	0.102
D1	0.799	0.624
ACPOP	0.421**	2.663
UNEM	-0.606*	-2.106
EDUCPOP	5.395**	4.162
FE	4.821**	2.966
TAX	-0.662*	-2.066
TRANS	7.256**	3.865

R^2 adjusted = 0.655

F-statistic = 42.84

***** Significant at 5%

****** Significant at 1%

From the demand side variables only INC and POP were significant. Both have positive signs which is consistent with theory. More and wealthier people mean more demand and hence are an incentive for opening more firms.

On the supply side ACPOP, UNEM, EDUCPOP, FE and TAX are significant[1]. The positive signs for ACPOP, EDUCPOP and FE coefficients were as expected. An increase in the supply of labor (ACPOP) reduces wages increasing incentives for new firm creation. The growth rate of more educated people is also an incentive for opening firms. According to Kohler (1997), firms tend to follow workers based in human capital. But this is only true for firms that are human capital intensive. For firms that are not human capital intensive, this variable must work in the opposite direction, that is, more educated employees represent higher wages and hence a disincentive for opening new firms. Another possibility to explain the positive relation between human capital and new firm creation is that more human capital increases labor productivity and so creates incentives. An alternative specification was used regarding the EDUCPOP variables. It consisted of respecifying education as two variables: percentage of population with only high school and percentage that are college graduates. The overall significance of the model did not change significantly but the *t*-value for each coefficient decreased.

The FE coefficient is positive while the TAX coefficient is negative, which is consistent with theory; the greater the earnings and the lower the taxes, the more firms will enter the market.

UNEM has a negative coefficient. It goes against intuition from the supply point of view because of market labor considerations. However, it may be correct from the demand side: more unemployment reduces the purchasing power of the families and hence the demand.

The transportation variable is significant at the 1% level and positive. Meaning that more accessibility to highways increases the rate of new firm formation. This variable was included as a change rate because most of the other variables were used as percentages too. Including TRANS as a change rate produced a better overall result compared with the use of the variable in levels, increasing the fit of the model and the significance level of some of the coefficients[2].

---

[1] To test for collinearity between ACPOP and UNEM, I regressed these two variables and use the regression results to calculate the variance inflation factor ($1/1-R^2$). The value of this factor allows rejecting the hypothesis of high collinearity between the variables.

[2] Other specifications were used, such as the transportation variable in levels and excluding the mentioned variable, the results of those regressions are included in the appendix.

The results for the second model are presented in table 3

Table 3

MODEL 2

COUNTY-LEVEL VARIABLES AND THE NEW FIRM FORMATION RATE

2-DIGIT SIC GROUPS DIFFERENCES

Independent variable	Coefficient	t-statistics
CONSTANT	-9.9143**	-2.68501
INC	7.4547**	6.79682
POP	0.9867*	1.83047
ACPOP	1.2716**	12.35077
EDUCPOP	2.4068*	2.261
FE	3.223**	4.17748
MARSHA	-0.2255**	-4.1461
SIZE	1.5928	1.04851
TAX	-0.1584**	-4.6207
TRANS	4.3582**	2.8147
D1 (Agricultural services, forestry and fishing)	4.3153	0.85528
D2 (Mining)	8.1895	1.49599
D3 (Construction)	14.7213**	3.11076
D4 (Manufacturing)	9.0552	1.33214
D5 (Transportation and public utilities)	14.1383**	6.87479
D6 (Wholesale trade)	10.461**	2.8147
D7 (Retail trade)	7.4536*	2.10845
D8 (Finance, Insurance and real state)	26.1261**	3.24071
D9 (Services)	19.393**	4.1255

R^2 adjusted = 0.533

F-statistic = 66.422

***** Significant at 5%

****** Significant at 1%

The goodness of fit for this model was evaluated with a test of structural change (Green, 1990, page 218). I used as a restricted model the one without dummies. The complete model is model 2 (I do not include the results for the restricted model). The F statistic calculated for the structural change test was 4.162. The critical value is 1.888 for 5% significance. So it is possible to reject the null hypothesis in favor of model 2.

Hence I found support for the hypothesis of spatial differences in the rate of new firm formation. Even though the differences are only captured here through the intercept and not in the explanatory variables. From the demand side variables income and population are significant with the expected signs.

Supply variables ACPOP, EDUCPOP and FE are significant and they have the expected positive signs. MARSHA and TAXES are also significant and the show a negative relationship. Assuming that the simple way I used to calculate the proxy for market concentration is correct, the more concentrated is the sector the less new firms are going to enter the market.

The TRANS variable is also significant for explaining the rate of new firm formation when considering differences among sectors. The positive sign shows a direct effect of the change of the number of exits to highways and formation of new firms.[3]

Finally, some of the dummy variables are significant showing positive shifts in the regression line. Construction, Transportation and public utilities, wholesale trade, Retail Trade, F.I.R.E. and Services are significant and positive.

[3] Here alternative specifications were used too. The results of the other regressions are included in the appendix.

V. CONCLUSIONS

The models presented here have some limitations, as usual; there are no perfect models to explore reality. However, it is possible to draw some conclusions and also to learn more about the specific economic situation of the State.

The strongest limitation that I found in the model is that in measuring the effect of county-level variables in the rate of new firm creation with supply and demand side variables it is assumed that the activities of the firms –producing and selling- are done in the boundaries of the county. This assumption is not realistic especially in the current time when the mobility among regions is easier than ever.

Among the findings, I can conclude that the results show a response in the rate of new firm creation to regional patterns, such as growth in income and population from the demand side; economically active population, degree of education, firms earnings and market concentration form the supply side.

One especially interesting variable is taxes, which has a negative relationship with the rate of new firm formation. From here may be derived a policy implication looking for program of incentives in the creation of new firms.

In a regional model it is important to include a transportation variable regarding the situation of the county in terms of accessibility to highways because this may affect the incentives for opening new firms. Changes in the accessibility of the county to highways showed its importance for explaining the rate of new firm formation. This fact also needs to be considered in the design of development policies.

Another interesting point is the marked difference among the sectors that was found with the second model showing that the influence of regional patterns in the dependent variable is strongly related with the sector to which the firm belongs.

Among the limitations of the models, I found that for the second model there are variables that are county specific and sector specific but there are other variables that are

only county specific. Future modeling efforts might take this feature into account or to find a more disaggregated data set.

Also related to the level of aggregation, another limitation of the model is that I only consider 2 digit industrial groups for the SIC industry classification. The use of 3 and 4 digit industrial groups might improve the quality of the estimation but also will make it more difficult to find explanatory variables at this level. Another way to improve the model is to consider changes in slope using dummies for each of the sectors or to estimate a panel of data using feasible generalized least squares to model random effects across counties and sectors.

Studying the regional determinants of the new firm creation may provide some information for policy recommendations, especially because in some counties it seems to be urgent to establish policies to improve their economic situation. Those figures that were presented at the beginning of the paper about population growth, unemployment and poverty show significant differences among counties and also the new firm creation rate. The results show that policy decisions need to take into account spatial differences.

A final limitation of the model is that the rate of new firm creation does not consider the size of the firm. A firm with twenty employees has the same weight as one with 1000 employees. This issue needs to be incorporated because the economic impact of a new firm entering the market in terms of employment and poverty reduction is strongly related with the size of the firm.

APPENDIX

Explanatory variables

Average annual % change in total population
Average annual net civilian migration per thousand total population
Average annual GDP per capita
Average annual % change in GDP
Average annual % change in GDP per capita
Average annual household disposable income per capita
Average annual % change in HDI per capita
Average annual % change in HDI
Population density
Average house purchase price
% dwellings owner –occupied or privately rented
% usually resident economically active population > 16 in non-manual socioeconomic groups
% usually resident economically active population > 16 in managerial and professional socioeconomic groups (SIC codes 1-4)
% usually resident economically active population > 16 in managerial and professional socioeconomic groups including farmers-employers and managers (SIC codes 1-4 and 13)
% school leavers with graded results
% employees in employment with higher education qualification
% turnover in legal units of less than 500,000
% employees in manufacturing units of 11-20 employees
% of employees in industries with low barriers to entry
Average annual % change in rate of unemployment
Average annual % change in volume of unemployment
Average annual unemployment rate

Change in percentage point unemployment rate
% Labor and Nationalist councilors on local councils
Dummy variable : value of 1 if 50% of more of county comprises an assisted area; 0 if less than 50%
Number of enterprise agencies with direct access to loan and / or grant funds per million civilian labor force
Total rates per head collected by local authorities in county
Expenditure per head on all services by local authorities in county
Lagged new registrations relative to labor force or business stock as appropriate.

MODEL 1

COUNTY-LEVEL VARIABLES AND THE NEW FIRM FORMATION RATE

Alternative specifications

Independent variable	**Coefficient**[a]	***t*-statistics**[a]	**Coefficient**[b]	***t*-statistics**[b]
CONSTANT	-4.333	-0.960	2.455	0.452
INC	12.305**	5.620	6.366	2.192*
POP	2.545*	3.706	1.421	1.365[+]
DENS	4.480	0.180	3.210	0.032
D1	0.814	0.660	-0.122	-0.056
ACPOP	0.397**	2.634	0.222	1.165
UNEM	-0.617*	-2.172	0.023	0.056
EDUCPOP	5.665**	4.230	3.856	3.219**
FE	3.940**	3.270	2.232	2.873**
TAX	-0.538*	-1.883	0.025	0.538
TRANS			2.329	1.458[+]
R^2 adjusted	0.523		0.145	
***F*-statistic**	36.330		26.120	

[a] In this specification the transportation variable is excluded.

[b] This specification uses the transportation variable in levels instead of the change rate. The transportation variable is the number of exits from 4-lane highways in the county.

[+] Significant at 10%

* Significant at 5%

** Significant at 1%

MODEL 2

COUNTY-LEVEL VARIABLES AND THE NEW FIRM FORMATION RATE

2-DIGIT SIC GROUPS DIFFERENCES

Alternative specifications

-Independent variable	Coefficient[a]	t-statistics[a]	Coefficient[b]	t-statistics[b]
CONSTANT	-9.613**	-2.677	-11.489**	-4.365
INC	6.177**	5.714	5.247**	5.116
POP	0.297	1.319	-0.985	-0.150
ACPOP	0.556**	10.629	-0.710*	-1.670
EDUCPOP	1.588*	1.700	0.383	0.581
FE	2.330**	3.396	1.169**	2.497
MARSHA	-0.395**	-3.970	-1.718	-0.826
SIZE	0.848	0.627	-0.401	-0.631
TAX	-0.456**	-4.390	-1.783	-2.300
TRANS			2.263[+]	1.334
D1 (Agricultural services, forestry and fishing)	3.323	0.456	2.222	0.824
D2 (Mining)	6.845	1.023	5.955	0.184
D3 (Construction)	12.783**	2.452	12.249[+]	1.430
D4 (Manufacturing)	7.632	0.878	6.789	0.347
D5 (Transportation and public utilities)	12.253**	5.783	11.688**	5.194
D6 (Wholesale trade)	8.910**	2.190	8.144	1.134
D7 (Retail trade)	6.176	1.565	5.246	0.428
D8 (Finance, Insurance and real state)	23.151**	2.567	23.240[+]	1.560
D9 (Services)	17.030**	3.350	16.751**	2.445
R^2 adjusted	0.456		0.153	
***F*-statistic**	61.003		45.390	

[a] In this specification the transportation variable is excluded.

[b] This specification uses the transportation variable in levels instead of the change rate. The transportation variable is the number of exits from 4-lane highways in the county.

[+] Significant at 10%

* Significant at 5%

** Significant at 1%

REFERENCES

Fritsch M., 1992, Regional Differences in New Firm Formation: Evidence from West Germany, *Regional Studies* 26, 233-41.

Gould A. and D. Keeble, 1984, New Firms and Industrialization in East Anglia, *Regional Studies* 18, 189-201.

Greene, W., 1990, Econometric Analysis. MacMillan Publishing Company.

Gudgin G. and S. Fothergill, 1984, Geographical Variation in the Rate of Formation of New Manufacturing Firms, *Regional Studies* 18, 203-6.

Hansen, N., 1990, Innovative Regional Milieux, Small Firms, and Regional Development: Evidence from Mediterranean France, *Annals of Regional Science* 24, 107-23.

Keeble D., J.R. Bryson and P. Wood P., 1991,Small Firms, Business Services Growth and Regional Development in the United Kingdom: Some Empirical Findings. *Regional Studies* 25, 439-57.

Keeble, D. and S. Walker, 1994, New Firms, Small Firms and Dead Firms: Spatial Patterns and Determinants in the United Kingdom, *Regional Studies* 28, 411-27.

Kohler, H., 1997, The Effect of Hedonic Migration Decisions and Region-Specific Amenities on Industrial Location: Could Silicon Valley Be in South Dakota? *Journal of Regional Science* 37, 379-94.

McNally, R, 1992 and 1999, Road Atlas de Luxe.

Shukla, V. and P. Waddell, 1991, Firm Location and Land Use in Discrete Urban Space: A Study of the Spatial Structure of Dallas-Fort Worth. *Regional Science and Urban Economics* 21, 225-53.