Does Population Density have Any Effect On Entrepreneurial Rates? Evidence From US Counties

Hatice Karahan
TÜBİTAK, The Scientific and Technological Research Council of Turkey, Turkey
hatice.karahan@tubitak.gov.tr

Abstract: Inspired from the theory of division of labor, I test the hypothesis that self-employment rate declines with a greater density of the population. The motivation is the existence of a huge variation of population densities and a considerable difference in self-employment rates across US states and counties. Results show that a negative and statistically significant impact of population density is present. Hence, people who are located in less crowded areas are more likely to be self-employed. Considering the relatively low availability of good-matching jobs in smaller places, self-employment can be explained as a failure of the size of the market place.

Keywords: *Self-Employment . Entrepreneurship . Population Density . Specialization*

1. Introduction

As a large country, the United States has very different places in terms of size and population density. Concentration of population in some states and counties of the US is fundamentally different from the others. On the other hand, a critical factor in understanding the macro context of entrepreneurial behavior in a particular location is the spatial patterns like population density which shapes the overall climate for entrepreneurship. An immediate question arises: How does the distribution of the US self-employment differ by the concentration of population? This brings to one's mind the possibility of a relationship between population density and self-employment rates. As well known, population density in the US thins as one moves away from Midwest to Northwest. Statistics by the US Census Bureau as of 2000 shows that the least populous states of the country are mostly located in the Northwest. Interestingly, most of these states have also higher than national average self-employment rates. Montana, North Dakota, South Dakota, Idaho, Wyoming and Nebraska are some of the states which fit this fact perfectly. Based on the US Census Data the national self-employment rate in 2000 is 13% for men and 7.7% for women. The rate for the Northwest states ranges from 15.3% to 22.2% for men and from 9.07% to 12.98% for women. This interesting deviation does not seem to be trivial and could refer to the potential relationship between concentration of population and entrepreneurial activity. So, it seems important to explore the impact of population density on people's self-employment decisions. The basic explanation behind this possible relationship comes from the traditional discussion of division of labor which was started by Adam Smith centuries ago. According to this theory, division of labor is limited by the extent of the market and therefore it is greater in cities than in small towns. This theory is also supported by a number of studies like Baumgardner (1988), Krugman (1987) and Lancaster (1975). Because the extent of the division of labor is strongly associated with local demand for services, the theory also suggests that an increase in the local population increases specialization. This explains why there are fewer opportunities for high skilled workers in towns. In larger markets, there exist more companies that seek different skill characteristics so the probability of having a good-matching job would be high. Analysis made by Kim (1998) proves this fact.

A good example to this would be the case of specialized physicians. In a sparsely populated area, there would be insufficient business for a specialized doctor but if he practices in a more densely populated area he would find enough demand for his skill and in that case specialization would be optimal. So, the low demand in smaller places for a certain skill might be thought of a reason for highly skilled people to move away from such areas toward more populous locations. For the other people though, the small size of the market would force them to accept less secure forms of employment and volatile incomes, namely to choose self-employment. This brings up the idea that self-employment could be a failure of the size of the market place. Another fact that affects the supply side of the story is that thinly populated areas have many dispersed small villages. So, these low-density areas will often have many small retail outlets and workshops whereas urban areas give rise to economies of scale, through which small-sized entrepreneurship in trade and craft comes under pressure. To date, regional entrepreneurship has been

ignored except for the contributions by Georgellis and Wall (1999) and Robson (1998) on the British case. Population density, along with some other demographic characteristics, can be an important economic determinant in predicting the level of entrepreneurship across the US states and counties. My hypothesis is that the self-employment rate declines with a greater fraction of the population. To my knowledge, there are no studies to date on the relationship between population density and entrepreneurial activity so this study will be a contribution to the literature on self-employment. This article proceeds as follows: Section 2 presents a theoretical background on specialization. Section 3 discusses the data used in this study and provides descriptive statistics. Section 4 explains empirical strategy and provides estimation results. Conclusions are presented in Section 5.

2. Theory of Specialization

Let s_{ij} be the specialization degree of a person i in location j:

$$s_{ij} = 1 - t_{ij} / k \quad \text{and} \quad 0 \leq s_{ij} < 1 \tag{1}$$

where k is the max number of tasks performed through a certain job and t_{ij} is the number of tasks person i performs in j. The main assumption in this model describes the efficiency gains to specialization such that the more the number of tasks a single person performs the more the time he has to spend on them. So we assume that

$$h'(t_{ij}) > 0 \tag{2}$$

where h represents the time allocated to the tasks. This means that the most effective outcome would be h(1) such that each worker is responsible for a single task as equation (2) implies

$$h(1) < h(2) < h(3) < \ldots < h(k) \tag{3}$$

If can be achieved, the desired outcome is complete specialization, in case there is enough demand for it in the local market. So we need to construct a framework as follows: The cost of organizing a production or a service must be minimized subject to two constraints: the production (outcome) constraint and workers` utility constraint.

Therefore the problem to be solved is as follows:

$$\text{Minimize C} = n_j \sum_{t=1}^{k} \{ w_0 \; h(t_{ij}) \} \tag{4}$$

subject to

$$F(.) >= n_j \tag{5}$$

and

$$U_i = n_j \sum_{t=1}^{T_j} \{ w_0 \; h(t_{ij}) \} \tag{6}$$

where n_j is a measure of total demand for a specific job in area j, w_0 is the hourly wage and F(.) is the production function . This is a discrete optimization problem so one cannot solve it using the classical derivatives approach. However, what we can immediately see from this minimization problem is that to minimize cost, the t_{ij} chosen should be as small as possible. Then, with the help of the constraints, we can easily conclude that the total demand in an area is negatively associated with the number of tasks each worker does, i.e.

$$\Delta t_{ij} / \Delta n_j < 0 \tag{7}$$
which implies that
$$\Delta s_{ij} / \Delta n_j > 0 \tag{8}$$

Using (1), Equation (8) shows that the specialization degree of a person i in location j should be greater if the local demand is more. Because local demand is an increasing function of local population, this result implies that specialization is more optimal in densely populated areas.

3. Data and Descriptive Statistics

For the empirical analysis, I use the Integrated Public Use Microdata Series (IPUMS) for the year 2000. IPUMS consists of thirty-seven high-precision samples of the American population drawn from fifteen federal censuses and from the American Community Surveys of 2000-2003. The data series includes information on a broad range of population characteristics, including labor-force participation, occupational structure, education, ethnicity, and household composition. It also provides information on geographic variables like state, city, county and public use microdata area (PUMA). In this study, I use "state" and "PUMA" together as the geographic variables in my population density calculations. In the 2000 State sample, PUMAs generally follow the boundaries of county groups, single counties, or census-defined "places"; if these areas exceed 200,000 residents, they are divided into as many PUMAs of 100,000+ residents as possible. PUMAs do not cross state lines. The target sample is composed of individuals who are in the labor force and employed as of 2000.[1] The number of self-employed men and women in the sample turns out to be 360,873 and 186,707, respectively.[2] This corresponds to a 13% of male and a 7.7% of female self-employment rate.

Table 1 presents the male and female self-employment rates across states in year 2000, ordered by population density. As can be seen from the table, the states with the least concentration of population are mostly located at the Northwest part of the country. Nebraska, Idaho, South Dakota, North Dakota, Montana and Wyoming are some of the states with extremely low population densities. This table also helps us see the huge deviation across geographical areas of the US in terms of crowdedness.

Table 1: Population Density Ranking and Self-Employment Rates of the US States as of 2000

Rank	State	Population Density (per square km)	Male Self-Employment Rate	Female Self-Employment Rate
1	New Jersey	438	12.49	5.94
2	Rhode Island	387.35	12.17	6.44
3	Massachusetts	312.68	12.83	7.25
4	Connecticut	271.40	13.62	7.20
5	Maryland	209.23	10.88	7.21
6	New York	155.18	13.01	6.97
7	Delaware	154.87	10.89	7.39
8	Florida	114.43	14.95	8.41
9	Ohio	107.05	10.62	5.92
10	Pennsylvania	105.80	12.03	6.01
11	Illinois	86.27	11.19	6.82
12	California	83.85	13.99	9.75
13	Hawaii	72.83	12.81	8.85
14	Virginia	69.03	10.65	6.66
15	Michigan	67.55	11.65	7.26
16	Indiana	65.46	10.09	6.69
17	N. Carolina	63.80	12.88	7.03
18	Georgia	54.59	12.88	6.89
19	Tennessee	53.29	12.99	6.95
20	New Hampshire	53.20	15.97	8.49
21	S. Carolina	51.45	11.97	6.08
22	Louisiana	39.61	12.85	6.54
23	Kentucky	39.28	12.55	6.49
24	Wisconsin	38.13	13.05	7.22
25	Washington	34.20	12.68	9.67
26	Alabama	33.84	12.34	6.18
27	Missouri	31.36	12.94	7.64

---

[1] I put an age restriction on the data and keep only the persons between 25 and 64 years old in the sample. Farmers are dropped out of the data as that type of self-employment is not of interest.

[2] To form the self-employed group, I take advantage of the "class of worker" variable which indicates whether a respondent worked for his or her own enterprise or for someone else as an employee and I create a dummy variable called "self-employed" accordingly.

28	Texas	30.75	13.10	7.77
29	W.Virginia	29.00	9.71	5.98
30	Vermont	25.41	19.93	10.89
31	Minnesota	23.86	15.20	8.37
32	Mississippi	23.42	13.34	5.65
33	Iowa	20.22	13.86	8.08
34	Arkansas	19.82	16.03	7.63
35	Oklahoma	19.40	14.84	8.49
36	Arizona	17.43	12.65	8.88
37	Colorado	16.01	15.54	10.99
38	Maine	15.95	18.32	10.13
39	Oregon	13.76	15.30	11.37
40	Kansas	12.69	13.90	8.24
41	Utah	10.50	13.11	9.73
42	Nebraska	8.60	17.13	9.07
43	Nevada	7.03	9.11	6.76
44	Idaho	6.04	17.06	10.99
45	New Mexico	5.79	14.03	10.60
46	S. Dakota	3.84	19.60	9.00
47	N. Dakota	3.59	19.92	9.60
48	Montana	2.39	22.20	12.98
49	Wyoming	1.96	14.18	10.68
50	Alaska	0.42	11.51	9.37

Source: Calculated from US Census Bureau Data 2000 and Integrated Public Microdata Series 2000.

Apparently, the category I focus on, which is the least densely populated states, has a higher number of entrepreneurs. Male self-employment rates in the Northwest states like Montana (22.2%), North Dakota (19.92%), South Dakota (19.6%), Nebraska (17.1%) and Idaho (17.1%) are quite higher than the national rate. On the Northeast, Maine (18.3%) and Vermont (19.93%) make examples to the low density areas with a high fraction of entrepreneurship. Statistics show a similar pattern for women, too. Female self-employment has relatively big fractions in Montana (12.98%), Oregon (11.37%), Idaho (10.99%), Vermont (10.89%), Wyoming (10.68%) and Maine (10.13%). Therefore, it is important to understand if low population density is associated with a higher fraction of self-employment at any geographic location. In the 2000 data, IPUMS introduces a new variable, PUMALAND, which is basically the area of each PUMA in square meters. This variable is a major element in my study as I need it in the computation of population density. Using the "personal weights" in the data I am also able to aggregate data on population in each PUMA. Then I calculate the population density in each PUMA, dividing the population by the area of that PUMA. Table 2 presents the means of the "population density" variable, comparing the self-employed and the wage-employed individuals.[3] It is interesting that for both men and women, the mean of population density is lower for the self-employed. T-test results indicate that this difference is significant. This implies that people who run their own business are more likely to be located in a less populous location. This supports my hypothesis that areas with lower concentration of population can have greater fractions of self-employment.

Table 2: Sample Means of the "Population Density" per square meter

	Self-Employed	Wage-Employed
MALE	0.0035	0.0043
FEMALE	0.0037	0.0044

[3]The means of the other variables used in the study are not presented here. However, I would like to mention that the numbers obtained confirm the evidence found in some previous studies that self-employed people are more likely to be married, older and more educated. (Rees and Shah, 1985 and Wong, 1986)They also repeat the fact that black self-employment rate is quite low. (Meyer, 1990) Self-employed individuals also seem to have more children than the salary-employed. Within the male sample, probability of having children is higher for the self-employed group. One reason behind this could be that the increasing family size might encourage parents to start their business with the hope of earning more to support their family.

4. Results

In order to estimate the effect of population density on the self-employment status of an individual, I take advantage of the Probit model. My model estimates the role of population density and some demographic variables in the dependent variable which is the decision to be self-employed or wage-employed. The variable of interest, namely, population density, d, is calculated for each single PUMA. Regressions are separately run for men and women. The equation used is as follows

$$Y_1 = X\beta_1 + \alpha_1 d + \varepsilon_1 \qquad\qquad (9)$$

where X is a vector of demographic variables which includes age, age squared, age tripled, education, race, marital status, child ownership and number of children. The coefficient of interest in this estimation is α_1, which will determine if concentration of population significantly affects employment status in any direction. If α_1 turns out to be negative I can conclude that people who live in smaller places are more likely to be self-employed. Table 3 reports the results from the probit analysis of self-employment status indicating that for both sexes, a negative and statistically significant effect of population density is present. This confirms my expectation that α_1 is negative, implying that people who are located at less crowded areas are more likely to be self-employed. Inclusion of state fixed effects to check any possible impacts of taxes does not change the significance of the findings. Therefore, I can conclude that density at a particular geographic location is an important factor in affecting individuals' employment decisions. Considering the relatively low availability of good-matching jobs in smaller towns, self-employment can be explained as a failure of the size of the market place.

Table 3: Probit Analysis of Self-Employment Decision

Variable	Male	Male	Female	Female
Population Density	-1.775 (0.366)	-1.435 (0.351)	-1.189 (0.388)	-1.312 (0.378)
Age	0.049 (0.031)	0.051 (0.031)	0.593 (0.041)	0.599 (0.041)
College	0.764 (0.011)	0.773 (0.011)	0.374 (0.014)	0.386 (0.014)
Grad/Prof School	-0.093 (0.016)	-0.084 (0.016)	0.075 (0.018)	0.083 (0.018)
Being Married	0.081 (0.004)	0.085 (0.003)	0.140 (0.005)	0.149 (0.004)
Child Owner	-0.042 (0.004)	-0.042 (0.004)	-0.081 (0.006)	-0.073 (0.005)
Number of Children	0.036 (0.002)	0.036 (0.002)	0.057 (0.002)	0.055 (0.003)
Asian	-0.021 (0.011)	-0.046 (0.012)	0.033 (0.011)	0.019 (0.012)
Black	-0.389 (0.008)	-0.393 (0.008)	-0.351 (0.007)	-0.337 (0.007)
State Fixed Effects	NO	YES	NO	YES

Notes. In terms of race, the comparison group is white people. The other control variables used include age squared, age tripled, other levels of education, American Indian race and other race. Numbers in parantheses show standard errors.

Being older or married has a significantly positive impact on self-employment status. The positive effect of age is most probably related to more job experience as well as greater wealth accumulated over the years which are fundamental factors to start a new business. As for the marriage, one can claim that it contributes to self-employment in terms of monetary and psychological support from the spouse. Results also suggest that individuals with college education have a greater tendency to have their own firm. This finding emphasizes the positive contribution of human capital investment in establishing and managing a business. Whereas having a graduate or professional degree affects entrepreneurial status of men negatively, it interestingly has a significantly positive effect on women.Moreover, it is found out that black women and men are less likely to attempt to run their business. Although having children seems to have a negative association with choosing self-employment, findings show that the higher the the number of children a person has, the more likely s/he prefers to be self-employed. People might think of wage-

employment as a safe way of getting a stable income especially when they have children. However, as new members are added to the family, parents might consider self-employment a potential lucrative source to bear the extra burden. To have an idea about the size of the coefficients I also checked the elasticities after probit using marginal effects which are listed in Table 4. The elasticity on the variable of interest is 1.4% for men and 1.1% for women. Although this might sound like a small marginal effect at the first place, considering the huge variation in moving from one of the most densely populated places of the US to a sparsely populated one, the elasticity found is not actually that small.

Table 4: Marginal Effects

Variable	Male	Female
Population Density	-0.0141	-0.0111
College	0.0384	0.0138
Grad/Prof School	-0.0024	0.0012
Being Married	0.0909	0.1628
Child Owner	-0.0358	-0.0833
Number of Children	0.0582	0.1088
Asian	-0.0015	0.0027
Black	-0.0602	-0.0823

In order to see the impact of the size of the market place on self-employment rates of the groups with more specialized occupations I also run the regressions on physicians and lawyers separately. Table 5 presents the coefficients from these regressions. The probit estimates yield significantly negative and larger impacts. This is an indication that the size of the location has a greater negative impact on the self-employment status of people with higher skills and specialization. Such people would probably prefer working in more densely populated areas where higher market demand and more good-matching job opportunities satisfy their specialization. This finding goes hand in hand with the theory of specialization. An interesting finding is that the female doctors seem to be even more adversely affected by population density in deciding on their self-employment status. In an attempt to contribute to this exploration, another regression is run on the whole sample excluding doctors, lawyers and accountants. Findings are listed in Table 6. Whereas men with other occupations still have a significantly negative coefficient, women do not appear to be affected significantly.

Table 5: Probit Analysis of Self-Employment Decision on Physicians / Surgeons and Lawyers

Variable	Male Physicians	Female Physicians	Male Lawyers	Female Lawyers
Population	-8.781	-12.40	-6.574	-5.948
Density	(2.021)	(1.886)	(0.822)	(1.549)
Age	1.366	1.859	0.345	0.968
	(0.574)	(0.745)	(0.348)	(0.617)
College	-0.264	-0.285	1.235	-0.479
	(0.498)	(0.609)	(0.575)	(0.561)
Grad/Prof School	-0.462	-0.316	1.147	-0.557
	(0.503)	(0.608)	(0.577)	(0.570)
Being Married	0.004	0.035	0.018	0.026
	(0.033)	(0.045)	(0.026)	(0.034)
Child Owner	-0.066	-0.146	0.021	0.028
	(0.036)	(0.066)	(0.032)	(0.058)
Number of Children	0.080	0.061	0.023	0.028
	(0.012)	(0.028)	(0.014)	(0.029)
Asian	-0.026	-0.059	0.047	-0.005
	(0.029)	(0.046)	(0.061)	(0.084)
Black	-0.097	-0.036	-0.077	-0.051
	(0.064)	(0.073)	(0.057)	(0.63)

Note. Numbers in parantheses show standard errors.

Table 6: Probit Analysis of Self-Employment Decision excluding Physicians, Surgeons, Lawyers and Accountants

Variable	Male	Female
Population Density	-1.227	-0.160
	(0.369)	(0.395)
Age	0.121	0.644
	(0.325)	(0.042)
College	0.603	0.222
	(0.013)	(0.016)
Grad/Prof School	-0.204	0.040
	(0.017)	(0.187)
Being Married	0.088	0.150
	(0.032)	(0.004)
Child Owner	-0.045	-0.074
	(0.004)	(0.005)
Number of Children	0.034	0.053
	(0.002)	(0.003)
Asian	-0.032	-0.006
	(0.012)	(0.012)
Black	-0.398	-0.340
	(0.008)	(0.070)

Note. Numbers in parantheses show standard errors.

5. Conclusion

It is a well known fact that distribution of population density across the US states is very large. Concentration of population is much lower in some regions like the Northwest. Interestingly, in most of these areas, self-employment rates are higher than the national average. Considering that the spatial features in a particular region are important factors in examining the macro context of employment status, I want to explore the relationship between population density and self-employment choice. My purpose is to see if there is a significant correlation between these two variables that could prove that the case of self-employment in the Northwest states and other less populous places is not a coincidence. My intuition that there might be a negative association between population density and entrepreneurial level in an area has its roots from the theory of division of labor. Based on this theory, more densely populated places provide a greater number of opportunities for good-matching jobs, especially for people with specialized skills. Therefore, in towns, division of labor is limited due to insufficient demand. This leads me to the idea that self-employment could be a failure of the size of the market place. Another fact is that whereas small-sized entrepreneurship is a necessity in rural areas due to high number of small villages, it is not well supported in cities where economies of scale dominate the market. To my knowledge, there is no work done in analyzing the role of population density in the US self-employment rates. To fill this gap, I take advantage of the Integrated Public Use Microdata Series (IPUMS) for the year 2000 and examine the impact of population density in a PUMA on the probability of being self-employed. Regression results significantly show that individuals who live in less populous areas are more likely to be self-employed. The effect is even larger on doctors and lawyers. Hence, the findings support the hypothesis that the self-employment rate declines with a greater fraction of the population. Empirical findings also indicate that being older, married or highly educated increases the probability of being self-employed. Having more children contributes to self-employment rates, too. In addition, black men and women are found to choose wage-employment over self-employment.

References

Baumgardner, J. R. (1988). The Division of Labor, Local Markets, and Worker Organization. *The Journal of Political Economy*, 96(3), 509-527.

Georgellis, Y. & Wall, H. J. (1999). What Makes a Region Entrepreneurial? Evidence from Britain. *The Annals of Regional Science*, 34(3), 385-403.

Kim, S. (1989). Labor Specialization and the Extent of the Market. *The Journal of Political Economy*, 97(3), 692-705.

Krugman, P. R. (1987). The Narrow Moving Band, the Dutch Disease, and the Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies. *Journal of Development Economics,* 27, 41-55.

Lancaster, K. (1975). Socially Optimal Product Differentiation. *American Economic Review*, 65, 580-585

Meyer, B. (1990). Why Are There So Few Black Entrepreneurs? NBER Working Paper No. 3537

Rees, H. & Shah, A. (1996). An Empirical Analysis of Self-Employment in the U.K. *Journal of Applied Economics,* 1*,* 95-108

Robson, M. T. (1998). Self-Employment in the UK Regions. *Applied Economics*, 30(3), 313-323.

Wong, Y. (1986). Entrepreneurship, Marriage, and Earnings. *The Review of Economics and Statistics*, 68(4), 693-699.