Economics Bulletin

Volume 32, Issue 2

The size and growth of state populations in the United States

Kwok Tong Soo
Lancaster University

Abstract

This paper explores the population distribution across U.S. states over time. We test for Zipf's Law on the size distribution of state populations and Gibrat's Law for the growth of state populations. State populations follow a lognormal distribution more closely than they do a Zipf or Pareto distribution. State population growth is negatively related to current state population in the 19th century but not in the 20th century, and is positively related to market potential in the 20th century but not in the 19th century.

Thanks to Wendy Beekes, Steve Redding, an anonymous referee, and seminar participants at Lancaster University for helpful suggestions on an earlier draft. The author is responsible for any errors and omissions.
Citation: Kwok Tong Soo, (2012) "The size and growth of state populations in the United States", *Economics Bulletin*, Vol. 32 No. 2 pp. 1238-1249.
Contact: Kwok Tong Soo - k.soo@lancaster.ac.uk.
Submitted: September 28, 2011. **Published:** April 19, 2012.

1. Introduction

At the time of the first population census in 1790, the total population of the United States was about 3.9 million people. By the fourth census in 1820, this had more than doubled, to 9.6 million, and by the seventh census in 1850 it had more than doubled again, to 23.2 million, and yet again by 1880, to 50.2 million. After that the rate of growth slowed; the next doubling was in 1920, when population reached 106 million, followed by 1980, when population reached 227 million. This extremely rapid population growth especially in the 19th century was driven primarily by migration. At the same time, the 19th century was a time of rapid geographical expansion of the United States. Starting from the original 13 colonies on the East coast which declared independence in 1776, the national boundaries expanded westwards throughout most of the 19th century, reaching the Pacific coast when California joined the Union in 1850, although it was not until 1912 when Arizona and New Mexico joined the Union that the whole of the continental United States came into being, and in 1959 Alaska and Hawaii became states as well, completing the 50 states.

This paper explores the size distribution and growth of U.S. state populations, using census data from 1790 to 2000. The approach used is to test for the existence of Zipf's Law (1949) on the size distribution of state populations and Gibrat's Law (1931) on the growth rate of state populations. According to Zipf's Law, the size distribution follows a simple Pareto distribution with shape parameter equal to 1. According to Gabaix (1999), such a distribution can be obtained from Gibrat's Law of random growth, which implies that there are constant returns to population size, and also that there is no convergence in population size over time.

The main contributions of the present paper are as follows. First, whilst tests for Zipf's Law and Gibrat's Law have most often been performed for cities (see for example Nitsch (2005), Rosen and Resnick (1980), Soo (2005)), recent work has extended this literature to country populations (Gonzalez-Val and Sanso-Navarro (2010), Rose (2006)). The present paper, by using state-level data, strikes the middle ground between cities on the one hand, and countries on the other hand, enabling us to uncover whether both Laws hold at this level of aggregation. Using state-level data enables us to consider the entire population of a country, hence allowing us to study the entire size distribution of state populations, whereas most work using city data focuses on the upper tail of the distribution, although recent work has used data on the full set of incorporated and unincorporated places in the U.S. (Eeckhout (2004), Gonzalez-Val (2010), Michaels *et. al.* (2008)). Also, because of U.S. federalism, states are the administrative level at which many policies vary[1], and because of these different policies, states may evolve differently from each other.

The second main contribution of the paper is to more carefully consider the estimation methods for Gibrat's Law, and to present the results of alternative estimation methods. This consideration of different methods has not been done in previous work using parametric methods (Dobkins and Ioannides (2001), Black and Henderson (2003)), and enables us to assess the appropriate estimation methods for this literature[2].

---

[1] Although local governments in the U.S. also have legislative authority, the extent of this authority varies across states. On the other hand, the Tenth Amendment to the United States Constitution limited the powers of the federal government to those delegated to it by the Constitution.

[2] On the other hand, we do not consider non-parametric approaches, such as those used in Ioannides and Overman (2001, 2003, 2004), Gonzalez-Val *et. al.* (2009), and Gonzalez-Val and Sanso-Navarro (2010). Parametric approaches have the advantage in this application since they can be used to go beyond Gibrat's Law to consider other explanations for the growth of state populations. It does mean however that we are imposing a

Our main findings are as follows. First, we find that the distribution of state populations follows a lognormal distribution closely, especially in the 20th century. Nevertheless, we are unable to statistically reject Zipf's Law that state populations follow a Pareto distribution with shape parameter equal to 1. Second, we find evidence that more populous states have a lower population growth rate, rejecting Gibrat's Law of proportional growth, suggesting the presence of decreasing returns to population and hence convergence in state populations. However, there is evidence that Gibrat's Law holds in the 20th century. We also find that higher state population growth is associated with states with higher market potential in the 20th century, but not in the 19th century.

The next section discusses the size distribution of U.S. state populations. This is followed in Section 3 by the growth of U.S. state populations. Section 4 concludes the paper.

2. The size distribution of U.S. state populations

2.1. Data

The primary data source is U.S. census data for state populations. This census has been held every decade since 1790. We use data for all 51 states including the District of Columbia. Given the geographic expansion of the country, the number of states increases over time. In the first census in 1790 there were only 18 states; this increased over time until 1900 when all 51 states were in the census. States are sometimes included in the census even before the state joins the Union: the most notable examples of this are Alaska and Hawaii, both of which achieved statehood in 1959, but data for Alaska are available since 1880 and for Hawaii since 1900. Table 1 documents the increase in the number of states included, as well as the increase in the average state population over time.

Table 1. The number of states in each U.S. census.

Year	States	Average Population	Year	States	Average Population
1790	18	218,290	1900	51	1,494,356
1800	23	230,804	1910	51	1,808,403
1810	28	258,567	1920	51	2,078,854
1820	28	344,249	1930	51	2,415,738
1830	29	443,472	1940	51	2,591,454
1840	31	550,431	1950	51	2,967,173
1850	38	610,312	1960	51	3,516,141
1860	43	731,240	1970	51	3,984,548
1870	48	803,750	1980	51	4,442,075
1880	49	1,024,270	1990	51	4,876,664
1890	50	1,259,595	2000	51	5,518,077

2.2. Methods and Results

According to Zipf (1949), the size distribution of state populations should follow a Pareto distribution with shape parameter equal to 1. That is, letting R be the rank of a state in terms

specific functional form on the data, which may prevent us from uncovering the true relationship between the variables of interest.

of its population (with a state with a larger population having a lower rank), and P be the population of the state, the population size distribution follows

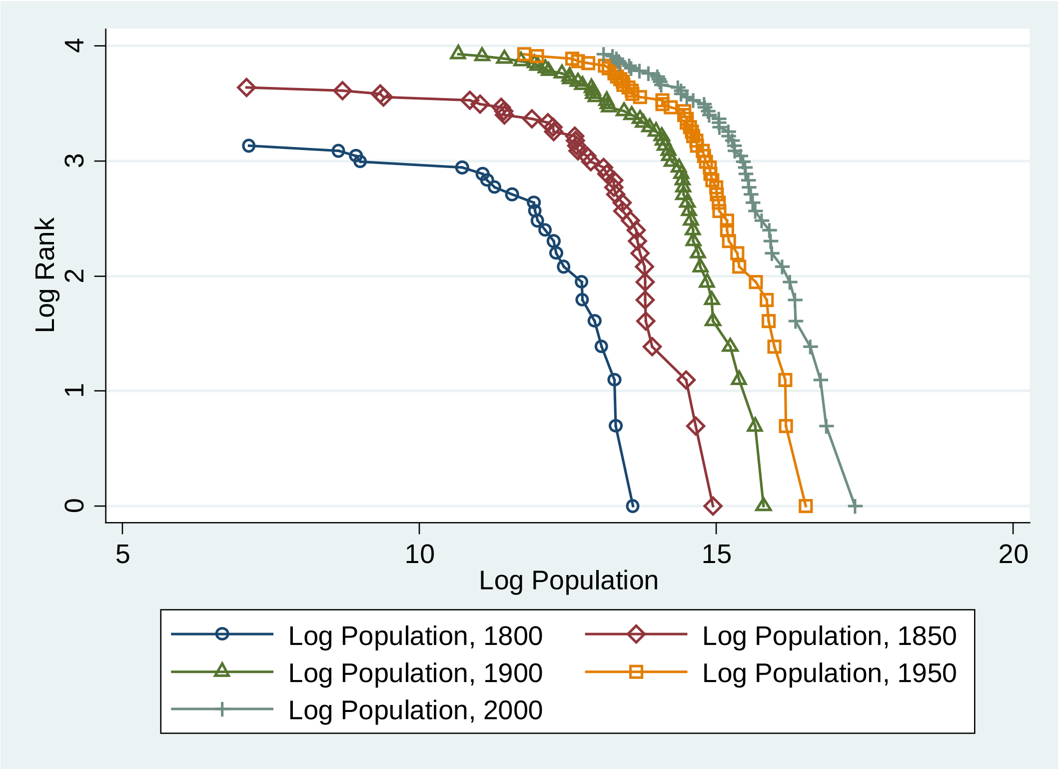
$$R = A P^{-\alpha} \quad (1)$$

Or,

$$\ln R = \ln A - \alpha \ln P \quad (2)$$

Zipf's Law states that α = 1. Equation (2) suggests a linear relationship between the natural log of a state's rank and the natural log of its population. Figure 1 plots equation (2) for 1800, 1850, 1900, 1950 and 2000. The rightward shift of the distribution over time indicates population growth. From this figure it is clear that linearity is not a good approximation for the relationship, and hence that Zipf's Law does not hold for U.S. states, although a linear relationship appears to hold quite well for the upper tail of the distribution. Drawing on a similar relationship between city sizes and city ranks, Eeckhout (2004) suggests that city sizes may be lognormally distributed.

Figure 1. Zipf plots for 1800, 1850, 1900, 1950 and 2000.



We therefore test whether the population size distribution of U.S. states is lognormally distributed. If state populations are lognormally distributed, then the standard Shapiro-Wilks and Shapiro-Francia tests for normality should fail to reject the null that the log of state populations follows a normal distribution. The results are in Table 2. Using either test, from 1900 onwards we cannot reject the null hypothesis that the distribution of state populations is lognormal at any conventional significance level. The p-value of the test increases further in the 20th century, indicating that the distribution is becoming more similar to the lognormal over time. One possible explanation for why U.S. state populations follow a lognormal distribution is that they are aggregations of lognormally distributed county populations (see Eeckhout (2004)).[3]

[3] The sum of lognormally distributed variables is not lognormally distributed, but can be reasonably approximated by one, using the Fenton (1960) method.

Table 2. Results from the Shapiro-Wilks and Shapiro-Francia tests for lognormality of the size distribution of U.S. states.

Year	Observations	Shapiro-Wilks test		Shapiro-Francia test	
		z	p-value	z	p-value
1800	23	2.41	0.0081	2.30	0.0108
1850	38	3.25	0.0006	3.08	0.0010
1900	51	1.17	0.1218	1.03	0.1517
1950	51	0.41	0.3424	0.30	0.3820
2000	51	0.16	0.4357	-0.27	0.6060

To facilitate comparisons with other results we estimate Zipf's Law, using the approach developed by Gabaix and Ibragimov (2011). They propose, instead of estimating equation (2) using OLS, to estimate the following equation instead:

$$\ln\left(R - \frac{1}{2}\right) = \ln A - \alpha \ln P \tag{3}$$

With the standard error of α being given by $(2/n)^{1/2}\alpha$. This alternative specification is used since OLS estimation of equation (2), whilst consistent, leads to a downward bias in the estimate of α in small samples. Similarly, OLS estimation of the standard errors underestimates the true standard errors. Given the curvature of the Zipf plots in Figure 2, we estimate equation (3) for different sample sizes in each year.

Table 3. Results from the Zipf regression equation (3).

Year	1800	1850	1900	1950	2000
N = 10	1.874	1.927	2.146	2.056	2.023
	(1.043)	(1.076)	(1.194)	(1.148)	(1.131)
N = 15	1.444	1.933	2.154	1.829	1.804
	(0.842)	(1.322)	(1.467)	(1.241)	(1.221)
N = 20	0.705	1.76	2.152	1.784	1.746
	(1.323)	(1.366)	(1.693)	(1.390)	(1.351)
N = 25	0.421	1.465	2.029	1.717	1.682
	(4.664)***	(1.122)	(1.793)*	(1.476)	(1.434)
N = 30		1.075	1.722	1.65	1.57
		(0.270)	(1.624)	(1.526)	(1.406)
N = 35		0.674	1.286	1.425	1.406
		(2.023)*	(0.930)	(1.248)	(1.208)
N = 40		0.421	1.069	1.134	1.192
		(5.995)***	(0.289)	(0.528)	(0.720)
N = 45			0.897	1.001	1.013
			(0.545)	(0.005)	(0.061)
N = 51			0.675	0.771	0.843
			(2.431)***	(1.500)	(0.940)

Notes: The table shows the estimated values of α from equation (3). Figures in parentheses are the t-statistics relative to the null hypothesis that α = 1. * significant at 10%; ** significant at 5%; *** significant at 1%. The maximum number of states in 1800 was 23 and in 1850 was 38. In 1900, 1950 and 2000 the maximum number of states was 51. See Table 1 for details.

The results are reported in Table 3. In every year the Zipf coefficient decreases as the sample size increases to include less-populous states, reflecting the concavity indicated in Figure 1. In parallel to the discussion on Zipf's Law for cities, from 1900 onwards the Zipf coefficient is fairly stable at the upper tail of the distribution (when $N \leq 25$); in these years the average Zipf coefficient at the upper tail is between 1.7 and 2.0, larger than the results usually obtained for cities and countries where it is about 1.1 (see for example Nitsch (2005), Rose (2006), Soo (2005)). In the earlier part of the sample, the Zipf coefficient decreases much more quickly as the sample size increases. However, it is only when all states are included in the sample that the Zipf coefficient is ever significantly different from 1 in a statistical sense, and even then only for 1900 and before. The least-populous states in the earlier periods are states that most recently joined the Union. These states subsequently experienced rapid population growth and caught up with other states, reducing the curvature of the Zipf plots.

3. The growth of U.S. state populations

3.1. Methods

Gabaix (1999) shows that Zipf's Law is an outcome of Gibrat's Law, which in this context would state that the population of a state and its growth rate are independent. Following Black and Henderson (2003), Gibrat's Law may be tested by estimating the following equation:

$$ln(P_{i,t}) = \beta + \theta_i + \gamma_t + \delta\, ln(P_{i,t-1}) + \varepsilon_{i,t} \qquad (4)$$

Where θ_i is a set of state fixed effects and γ_t is a set of year dummies. If Gibrat's Law holds, then $\delta = 1$, and the error term $\varepsilon_{i,t}$ is i.i.d. Hence, equation (4) may be estimated using OLS. Equation (4) may be augmented with additional explanatory variables that may explain the growth of state populations:

$$ln(P_{i,t}) = \beta + \theta_i + \gamma_t + \delta\, ln(P_{i,t-1}) + \eta\, \boldsymbol{Z_{i,t}} + \varepsilon_{i,t} \qquad (5)$$

Where $\mathbf{Z_{i,t}}$ is a vector containing both time varying and time invariant variables. If Gibrat's Law does not hold, then the inclusion of lagged state population in equations (4) and (5) means that conventional OLS, fixed- and random-effects models are all biased (see Baltagi (2005) for details). Since our dataset has more states than time periods, the appropriate estimation method is a dynamic panel data model.

Our dataset includes a relatively large number of both cross-sectional (51 states) and time-series units (21 census periods after dropping the first census from the sample; see the results section for more discussion). Therefore, to estimate equation (5), we use the method originally proposed by Kiviet (1995), which involves correcting the bias of the fixed effects models. Judson and Owen (1999) show that for samples with relatively large time and cross sectional dimensions, fixed effects estimation with Kiviet's correction outperforms OLS, uncorrected fixed effects, the Arellano-Bond (1991) GMM estimator and the Anderson-Hsiao (1982) instrumental variables estimator, in terms of the root mean square error, for every sample size which they tested. At the time of Judson and Owen (1999), Kiviet's correction was not available for unbalanced panels. This was rectified by Bruno (2005a, 2005b), who shows that Kiviet's correction outperforms alternative estimation methods in unbalanced panels in terms of bias and root mean square error. We estimate the model using Bruno's (2005a, 2005b) corrected fixed effects, with the correction initiated by the Anderson-Hsiao (1982) estimator, and the bias correction is O(1/T) (see Bruno (2005a) for details). Nevertheless, one key limitation of the Kiviet-Bruno correction is that all the other explanatory variables must be strictly exogenous. In the robustness section we compare the results of using this method, with alternative estimation methods.

Models of economic geography (see for example Fujita *et. al.* (1999)) suggest that market potential plays an important role in determining the attractiveness of locations to workers and firms. This may influence state population growth. Therefore, in addition to the lagged dependent variable, we include in equation (5) market potential as a measure of economic geography, defined as:

$$MP_{j,t} = \sum_{i \neq j} \left(\frac{P_{i,t}}{D_{i,t}}\right) \quad (6)$$

That is, the market potential of a state j in year t is the sum of the population of all other states weighted by the inverse of the distance from the state, where the distance is the great circle distance between state capitals[4]. Market potential captures the potential market and competition facing a state. Small states on the East Coast such as Connecticut, Delaware, District of Columbia, Maryland and Rhode Island have the largest market potential, for two reasons: first, they are in a densely populated region, and second, being small states, they are closer to their neighbours than larger states. At the other end of the spectrum, perhaps unsurprisingly Alaska and Hawaii have the lowest market potential, but California has a low market potential as well, since despite being the most populous state, it borders large and less densely populated states. [5]

3.2. Results

Table 4. State population growth and Gibrat's Law. Dependent variable: log state population, $\ln(P_{i,t})$.

	Full sample: 1800-2000		19th century: 1800-1910		20th century: 1920-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(P_{i,t-1})$	0.760**	0.751**	0.674**	0.701**	1.029**	0.952**
	(0.017)	(0.021)	(0.029)	(0.038)	(0.030)	(0.040)
ln(Market potential)		0.038		-0.361		0.497**
		(0.092)		(0.239)		(0.146)
N	826	826	367	367	408	408
F-Test $\delta = 1$ p-value	0.00	0.00	0.00	0.00	0.32	0.23
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation method is the Kiviet-Bruno corrected fixed effects estimator. All regressions include unreported constants. * significant at 5%; ** significant at 1%. Figures in parentheses are standard errors obtained through 100 bootstrap replications. The test for $\delta = 1$ is the test for Gibrat's Law. Bias corrections are initialised by the Anderson-Hsiao (1982) estimator, and the bias correction is O(1/T). See Bruno (2005a) for details.

Table 4 presents the results of regressions (4) and (5), estimated using the Kiviet-Bruno corrected fixed effects estimator for unbalanced panels. The first year a state enters the census is dropped from the sample, as states often experienced very rapid initial population growth from very low bases. All results reported include unreported time dummies. Standard errors are estimated using bootstrap methods with 100 replications. Column (1) which is for the full sample shows that the coefficient on the lagged dependent variable is 0.76, which is significantly less than 1. This indicates that Gibrat's Law of proportional growth in state populations does not hold. States with a higher population grow more slowly than states with a lower population, suggesting decreasing returns to population and convergence of state populations. Column (2) of Table 4 includes the log of market potential as an additional

[4] We use great circle distances since over the long time period under investigation, new transport links were developed, but we do not have the data to capture these developments.

[5] Because of the use of fixed effects methods, we are unable to include any time-invariant variables in the model, such as geographical measures like land area, which would be captured by the state fixed effects.

explanatory variable. This has no significant impact on state population growth[6], and does not change the coefficient on the lagged dependent variable.

There are essentially two phases in the geographical development of the United States: the first phase, which occurred in the 19th century, involved the expansion of the country to new states. This expansion was concluded with the acceptance of Arizona and New Mexico into the Union in 1912. The second phase after 1912 involved no geographical expansion, but population movement across states is not uncommon. Since some population data for states is available even before formal admission into the Union, we divided the sample into two time periods, the first from 1800 to 1910, and the second from 1920 to 2000, to investigate whether state population growth behaves differently in the two phases of development.[7]

Columns (3) and (4) of Table 4 report the results for the 1800-1910 sample (henceforth the 19th century), while columns (5) and (6) report the results for the 1920-2000 sample (henceforth the 20th century). The coefficient on the lagged dependent variable is significantly smaller in the 19th century than in the 20th century. Also, the coefficient is not significantly different from 1 in the 20th century. That is, we cannot reject Gibrat's Law of proportional growth in the 20th century, but we can do so in the 19th century. In addition, market potential is positively related to population growth in the 20th century, but not in the 19th century. The results of columns (3) to (6) of Table 4 suggest that, in the 19th century, there were decreasing returns to population, possibly due to the expansion into new territories which encouraged the growth of new settlements, so that newly settled, less-populous states experienced faster population growth. On the other hand, in the 20th century, there are constant returns to population, and increasing attractiveness of proximity to other economic agents. We do not have the data to investigate this further, but it is possible that this is caused by improvements in transportation technology and infrastructure during this period (see for example Michaels (2008), Duranton and Turner (2010)).

3.3. Robustness

In this section we compare the results from the previous subsection using the Kiviet-Bruno corrected fixed effects estimator, with other methods of estimating the model. The methods we compare are: (1) OLS; (2) fixed effects; (3) the Anderson-Hsiao (1982) first-differences model with the differenced lagged dependent variable instrumented by the second lag of the dependent variable in levels; (4) the Arellano-Bond (1991) Difference GMM model in first-differences in which additional lags of the dependent variable are used as instruments; (5) the Blundell-Bond (1998) System GMM model in which the equation in levels is included in the estimated system to exploit the additional moment conditions; and (6) the Kelejian-Prucha (1999) GMM estimator of a spatial autoregressive model with spatial-autoregressive disturbances (SARAR)[8]. For a more detailed discussion of methods (1) to (5) above, see Roodman (2009a).

In both the Arellano-Bond (1991) and Blundell-Bond (1998) GMM models, to avoid the problem of too many instruments (see Roodman (2009b)), we collapse the set of instruments, and use only two lags of the endogenous variables as instruments. In addition, we estimate

[6] Equations (4) and (5) can be rearranged as equations with population growth rates (differences in log population) on the left-hand-side, without changing the rest of the equations.
[7] Estimating equations (4) and (5) excluding Hawaii and Alaska does not significantly change the results.
[8] See also Kapoor *et al* (2007) and Drukker *et al* (2011). The SARAR model was estimated using the GS2SLSXT module in STATA, written by Shehata (2011).

both the Arellano-Bond and Blundell-Bond models in their asymptotically efficient, two-step form, and correct for the standard errors using the Windmeijer (2005) correction. As noted above, OLS and fixed effects methods are biased in the presence of the lagged dependent variable. On the other hand, the properties of the Anderson-Hsiao (1982) and GMM estimators hold when N is large, so they can be biased in panels with a small number of cross-sectional units.

Table 5. State population growth and Gibrat's Law: Estimating equation (5) using alternative methods. Dependent variable: log state population, $\ln(P_{i,t})$.

Estimation method	(1) OLS	(2) Fixed effects	(3) Anderson-Hsiao	(4) Arellano-Bond	(5) Blundell-Bond	(6) Kelejian-Prucha
$\ln(P_{i,t-1})$	0.939**	0.915**	0.699**	0.395**	0.563**	0.938**
	(0.014)	(0.018)	(0.083)	(0.058)	(0.084)	(0.017)
ln(Market potential)	-0.135**	-0.113**	0.467	1.469**	-0.028	0.582**
	(0.032)	(0.040)	(0.361)	(0.336)	(0.324)	(0.084)
Spatial lag						-0.275**
						(0.089)
N	844	844	844	844	844	459
F-Test $\delta = 1$ p-value	0.00	0.00	0.00	0.00	0.00	0.00
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Time period	1800-2000	1800-2000	1800-2000	1800-2000	1800-2000	1920-2000
Instruments				24	26	

Notes: All regressions include unreported constants. * significant at 5%; ** significant at 1%. Standard errors clustered by state reported in parentheses. The test for $\delta = 1$ is the test for Gibrat's Law. See the text and Roodman (2009a) for details on the estimation methods used.

Table 5 presents the results for equation (5). Columns (1) to (5) report the results for the full sample; this corresponds to column (2) of Table 4. All standard errors reported are clustered by state. The OLS and fixed effects estimates of the coefficient on lagged population are larger than the Kiviet-Bruno estimates; this indicates the endogeneity bias of these methods. On the other hand, the GMM estimates (columns (4) and (5)) are smaller than the Kiviet-Bruno estimates, which again suggests that these methods suffer from bias due to the relatively small number of cross-sectional units. The Anderson-Hsiao estimate is fairly close to the Kiviet-Bruno estimate. For market potential, the OLS and fixed effects results are significantly negatively related to population growth, whereas with Arellano-Bond, it is positively related to population growth.

Finally, column (6) of Table 5 reports the results of the Kelejian-Prucha SARAR model, for the 1920-2000 sample, corresponding to column (6) of Table 4. This sample was chosen since the GS2SLSXT module in STATA requires a balanced panel. The Kelejian-Prucha model yields results for both lagged population and market potential which are very similar to the Kiviet-Bruno estimates. This suggests that the Kelejian-Prucha model may be an alternative to the Kiviet-Bruno estimator. In addition, the negative coefficient on the spatial lag term suggests that states which are close to other rapidly growing states may experience slower population growth, maybe because these neighbouring states are attracting migrants from the state. Overall, the results of Table 5 provide evidence to support our use of the Kiviet-Bruno corrected fixed effects method, as the results differ across methods, and the Monte Carlo evidence from Bruno (2005a, 2005b) indicates that the Kiviet-Bruno estimator outperforms the other estimators. The Kelejian-Prucha SARAR model appears to provide an alternative to the Kiviet-Bruno estimator.

4. Conclusions

In this paper we explore the size distribution and growth of U.S. state populations over a long time period, using census data from 1790 to 2000. Our main findings are as follows. First, we find that the size distribution of state populations follows a lognormal distribution, especially since 1900. Nevertheless, we are unable to reject Zipf's Law (that state population follows a Pareto distribution with shape parameter equal to 1) despite the concavity of the Zipf plots.

Our second main finding is that more populous states tend to have slower population growth rates than less populous states, and this result holds controlling for market potential. That is, we do not find evidence for Gibrat's Law of proportional growth in state populations. However, when we divide the sample into 19th century and 20th century samples, we find evidence that Gibrat's Law holds in the 20th century but not in the 19th century. This suggests the presence of decreasing returns to population in the 19th century, but not in the 20th century, and also that there is no evidence of convergence in state populations in the 20th century. Also, market potential is positively related to state population growth in the 20th century, but not in the 19th century.

Our results represent new findings in this literature. Most of the studies in this area have focussed on the size distribution and growth of cities. More recent studies such as Rose (2006) have suggested that similar patterns can be found for countries as well. This paper presents evidence at an intermediate level between cities and countries, and besides showing where the results follow Zipf's Law and Gibrat's Law and where they do not, we also extend the analysis to consider the changing patterns over time, and the econometric issues that arise when analysing dynamic panel data models of this type.

References

Anderson, T. W. and Hsiao, C. (1982) "Formulation and estimation of dynamic models using panel data", *Journal of Econometrics*, 18(1), 47-82.

Arellano, M. & Bond, S. (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *The Review of Economic Studies*, 58(2), 277-297.

Baltagi, B. H. (2005) *Econometric Analysis of Panel Data*. 3rd Edition, New Jersey: Wiley.

Black, D. & Henderson, J. V. (2003) "Urban evolution in the USA", *Journal of Economic Geography*, 3(4), 343-372.

Blundell, R. & Bond, S. (1998) "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87(1), 115-143.

Bruno, G. S. F. (2005a) "Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals", *Stata Journal*, 5(4), 473-500.

Bruno, G. S. F. (2005b) "Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models", *Economics Letters*, 87(3), 361-366.

Dobkins, L. H. & Ioannides, Y. M. (2001) "Spatial interactions among U.S. cities: 1900-1990", *Regional Science and Urban Economics*, 31(6), 701-731.

Drukker, D. M., Prucha, I. R. & Raciborski, R. (2011) "Maximum-likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances", mimeo, University of Maryland.

Duranton, G. & Turner, M. A. (2010) "Urban growth and transportation", mimeo, University of Toronto.

Eeckhout, J. (2004) "Gibrat's Law for (All) cities", *American Economic Review*, 94(5), 1429-1451.

Fenton, L. F. (1960) "The sum of lognormal probability distributions in scatter transmission systems", *IRE Transactions on Communications Systems*, 8(1), 57-67.

Fujita, M., Krugman, P. R. & Venables, A. J. (1999) *The Spatial Economy*. Cambridge, MA, MIT Press.

Gabaix, X. (1999) "Zipf's Law for cities: An explanation", *Quarterly Journal of Economics*, 114(3), 739-767.

Gabaix, X. & Ibragimov, R. (2011) "Rank – ½: A simple way to improve the OLS estimation of tail exponents", *Journal of Business and Economic Statistics*, 29(1), 24-39.

Gibrat, R. (1931) *Les ine´galite´s e´conomiques; applications: aux ine´galite´s des richesses, a` la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel.* Paris: Librairie du Recueil Sirey.

Gonzalez-Val, R. (2010) "The evolution of US city size distribution from a long-term perspective (1900-2000)", *Journal of Regional Science*, 50(5), 952-972.

Gonzalez-Val, R., Lanaspa, L. & Sanz, F. (2009) "Gibrat's Law for cities revisited", MPRA Paper 9733.

Gonzalez-Val, R. & Sanso-Navarro, M. (2010) "Gibrat's Law for countries", *Journal of Population Economics*, 23(4), 1371-1389.

Ioannides, Y. M. & Overman, H. G. (2001) "Cross-sectional evolution of the U.S. city size distribution", *Journal of Urban Economics*, 49(3), 543-566.

Ioannides, Y. M. & Overman, H. G. (2003) "Zipf's Law for cities: An empirical investigation", *Regional Science and Urban Economics*, 33(1), 127-137.

Ioannides, Y. M. & Overman, H. G. (2004) "Spatial evolution of the U.S. urban system", *Journal of Economic Geography*, 4(2), 131-156.

Judson, R. A. & Owen, A. L. (1999) “Estimating dynamic panel data models: A guide for macroeconomists”, *Economics Letters*, 65(1), 9-15.

Kapoor, M., Kelejian, H. H. & Prucha, I. R. (2007) “Panel data models with spatially correlated error components”, *Journal of Econometrics*, 140(1), 97-130.

Kelejian, H. H. & Prucha, I. R. (1999) “A generalized moments estimator for the autoregressive parameter in a spatial model”, *International Economic Review*, 40(2), 509-533.

Kiviet, J. F. (1995) “On bias, inconsistency, and efficiency of various estimators in dynamic panel data models”, *Journal of Econometrics*, 68(1), 53-78.

Michaels, G. (2008) “The effect of trade on the demand for skill – Evidence from the interstate highway system”, *Review of Economics and Statistics*, 90(4), 683-701.

Michaels, G., Rauch, F. & Redding, S. J. (2008) “Urbanisation and structural transformation”, CEPR Discussion Paper 7016.

Nitsch, V. (2005) “Zipf zipped”, *Journal of Urban Economics*, 57(1), 86-100.

Roodman, D. (2009a) “How to do xtabond2: An introduction to ‘difference’ and ‘system’ GMM in Stata”, *Stata Journal*, 9(1), 86-136.

Roodman, D. (2009b) “A note on the theme of too many instruments”, *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.

Rose, A. K. (2006) “Cities and countries”, *Journal of Money, Credit and Banking*, 38(8), 2225-2245.

Rosen, K. T. & Resnick, M. (1980) “The size distribution of cities, An examination of the Pareto Law and primacy”, *Journal of Urban Economics*, 8(2), 165-186.

Shehata, E. A. E. (2011) “GS2SLSXT: Stata module to estimate GS2SLS Generalized spatial panel autoregressive two-stage least squares regression”, http://ideas.repec.org/c/boc/bocode/s457386.html

Soo, K. T. (2005) “Zipf’s Law for cities: A cross-country investigation”, *Regional Science and Urban Economics*, 35(3), 239-263.

Windmeijer, F. (2005) “A finite sample correction for the variance of linear efficient two-step GMM estimators”, *Journal of Econometrics*, 126(1), 25-51.

Zipf, G. K. (1949) *Human behaviour and the principle of least effort*, Cambridge, MA, Addison-Wesley Press.