Zipf's and Gibrat's laws for migrations

Clemente, Jesús and González-Val, Rafael and Olloqui, Irene

4 January 2010

Zipf's and Gibrat's laws for migrations

Jesús Clemente

Rafael González-Val

Irene Olloqui

Departamento de Análisis Económico

Universidad de Zaragoza

Facultad de CC. Económicas y Empresariales

Gran Vía, 2, 50005 Zaragoza (Spain)

+34 976761000 ext. 4728

+34 976761996

rafaelg@unizar.es

Abstract: This paper analyzes the evolution of the size distribution of the stock of immigrants in the period 1960–2000. In particular, we are interested in testing the validity of two empirical regularities: Zipf's law, which postulates that the product between the rank and size of a population is constant; and Gibrat's law, according to which the growth rate of a variable is independent of its initial size. We use parametric and nonparametric methods and apply them to absolute (stock of immigrants) and relative (migration density, defined as the quotient between the stock of immigrants of a country and its total population) measurements. We find that both the stock of immigrants and migration density follow similar size distributions to those of cities and of countries. Contrary to what traditional migrations models predict, growth in the stock of immigrants is independent of the initial stock. Moreover, the growth of migration density shows a divergent behaviour, which could be explained by the lower birth rates of host countries and the reduction in the cost of emigration produced by the presence of a previous stock of immigrants in the country.

Keywords: *Migration distribution, Zipf´s law, Gibrat´s law.*

JEL: *J61, R11, R12.*

Acknowledgements:

The authors would like to thank the Spanish Ministerio de Educación y Ciencia (SEJ2006-04893/ECON project and AP2005-0168 grant from FPU programme), the DGA (ADETRE research group) and FEDER for their financial support.

1. Introduction

In the study of the economic landscape, the influence of the location of productive factors on economic activity is an important element. Some of these factors cannot be moved from one geographical space to another (natural resources, amenities, etc.) but others, such as physical capital, human capital or technology, can. Therefore, the analysis of the distribution of the population in space is an extremely interesting question.

Two laws have been widely considered: Zipf's and Gibrat's. The first refers to the city size distribution and the second to population growth (Eeckhout, 2004). Recently, Rose (2006) analyzed whether other phenomena associated with population size, such as the number of inhabitants of countries, also follow a characteristic distribution, and concluded that the size distributions of cities and of countries are similar.[1] In this context, we analyze the distribution of the number of immigrants by country from the perspective both of stock and of the percentage of this stock over the total population of the country, the migration density.

The recent evolution of migratory flows has led to a considerable growth in the stock of immigrants. Therefore, given that the total population is the sum of natives and immigrants, it is useful to analyze whether Zipf's and Gibrat's laws hold for both groups. In recent decades, population growth in developed countries has been mainly due to immigrant population. The growth of the stock of immigrants (M) in country i can be represented by the following function:

$$\frac{M_{it}}{M_{it-1}} = f\left(a_{+}\left(M_{it-1}\right), b_{-}\left(M_{it-1}\right)\right),$$

[1] A new critique of Rose's work is Gonzalez-Val and Sanso-Navarro (2009).

where $a_{+}(M_{it-1})$ and $b_{-}(M_{it-1})$ are, respectively, the positive and negative external effects[2] of the stock of immigrants on the growth rate. The positive external effects are associated with the so-called scale effect and the social network effect, while the negative external effect is related to the influence that the stock of immigrants has on migratory flows through wages. Therefore, the net sign of the combined effect would be a priori undetermined.

In this context, the determinants of migration are of key importance. The traditional model of Harris and Todaro (1970) predicts that migrations will disappear due to the mobility of the factors leading to convergence in expected wages between countries. However, empirical studies do not support this conclusion, as shown by Ghatak and Wheatley (1996). Authors such as Carrington et al. (1996) indicate the importance of the presence of social networks, that is, the existence of a stock of immigrants in the host country prior to the individuals' decisions to emigrate. The stock of immigrants reduces the cost of emigration, increasing the rate of migration. Additionally, authors such as Larramona and Sanso (2006) show that the differences that exist between countries do not always disappear in the long term, so the convergence achieved is limited or conditional because it does not necessarily imply the equalization of per capita income, of the capital/labour ratio or of wages. Thus, the final result of the size distribution of the immigrant population is an open question which may be important in order to explain the size of countries.

Another useful perspective for this type of analysis is that adopted by Alesina et al. (2000) and Spolaore and Wacziarg (2005). These authors find empirical

[2] We call these effects "external" because the present stock of immigrants influences the decisions of other immigrants to emigrate later. But the effects are "internal" in geographical terms, given

evidence in favour of the so-called scale effect, that is, that countries with larger populations or GDPs have a larger potential market and their incomes exhibits greater growth rates. In this context, the migratory stock contributes to increasing the market potential and this has positive effects on the productivity of the country, partly canceling out the tendency to the equalization of wages predicted by the traditional models. This perspective introduces interesting elements related to the effects (at the aggregate level) of immigration from the point of view of developed economies, which have lower birth rates.

Thus, it is interesting to study if there are variations in the distribution of the stock of immigrant population. This is because mobility of labour is usually associated with differences in socioeconomic characteristics such as wages and the previous stock of immigrants. From this point of view, the flows of labour tend to equalize the labour conditions of the different countries, so there would be some decrease in the stocks of immigrants and a tendency to a lower migration density.

Furthermore, migration is a phenomenon closely related to the job market because it increases the labour supply in the host country. However, labour mobility has more restrictions than capital mobility. Until the mid 20th century, most of these restrictions were imposed by transport technology. Since then, the reduction of these costs has been enormous. This decrease has recently been counterbalanced in many countries with the rise of protectionist immigration legislation. So, the cost is decreasing but the legislation has an influence in the opposite direction. The question is whether the distribution of the size of the stock of immigrants has become more uneven or whether there has been some convergence. This is the topic discussed in this work. Convergence implies the change to a less uneven size

---

that the equation shown describes the dynamics in the growth of the stock of immigrants in each distribution, with countries becoming more homogeneous in their stocks and/or their density of immigrants. This would indicate that the economic and social driving forces of migratory flows are leading the distribution to an equilibrium outcome. This situation would be related to an equalization of the positive and negative external effects of the stock of immigrants on the growth rate.

The results of our research are interesting because they provide evidence in favour of some kinds of migration models. A more uneven distribution of migrations stock means that the migration rate does not tend to zero and migration models with permanent migration flows would be more appropriate. If the result is the opposite (convergence), the traditional migration model cannot be rejected.

The rest of the article is structured as follows. In Section 2, we introduce the data. In Section 3, the results relative to Zipf's and to Gibrat's laws are presented. Finally, the last section concludes.

2. Data and descriptive analysis

The data correspond to the total stock of immigrants by country and the source is the Department of Economic and Social Affairs of the Population Division of the United Nations (2004). The Population Division maintains a data bank on international migration statistics covering most countries of the world. The data bank includes information from censuses on the number of foreign-born individuals or, in some cases, the foreign population living in a country. These data provide the basis for estimating the number of international migrants in the world at different points in time.

---

country.

Four types of data are used by the Population Division to obtain the estimate of the migrant stock. For most countries, the estimate of the migrant stock is derived from data classified by place of birth, so it represents the foreign-born population. Sometimes it is derived from data classified by citizenship, so it represents the foreign population. In some cases, there was no data on either the foreign or the foreign-born population for the country or area concerned, so the estimate is an assigned value (this situation occurs in very few countries and corresponds to countries with wars in their territories[3]). And finally, the Population Division deals with the issue of refugees (which are an important percentage of the population of some African and Asian countries) considering the number of refugees as reported by UNHCR or UNRWA. These data were added to the estimates of the international migrant stock for developing countries where they were likely not to have been included in the census data available.

The sample includes all of the 214 United Nations member countries.[4] The period considered is from 1960 to 2000, presenting information by decade. The data on the total population of the countries was taken from the same source.

A first analysis consists of describing the evolution of the stocks of immigrants and of the migration density. Panel (a) of Table 1 shows the total stock by geographical area and Panel (b) its growth in the period 1960–2000. The most important point is the rise in the number of immigrants, which increased by 130.48% during these 40 years. There was a particularly marked increase in North America and Oceania, while Europe is slightly above the mean. In Latin America

[3] We have repeated the analysis excluding the African countries, characterised by an unstable political situation, and this has no effect on the qualitative outcomes of our analysis.

[4] Including the former USSR as a single country, because the disintegration of the USSR in 1991 produced a transformation of internal migrants into international migrants generating data discontinuity. The former USSR includes 15 countries: Armenia, Azerbaijan, Belarus, Estonia,

and the Caribbean the stock has decreased. Thus, the evolution of foreign population by country is not homogeneous.

If we look at the migration density, Panel (c) of Table 1, we see that this impression is corroborated. The variable grows in North America, in Oceania and, to a lesser extent, in Europe, while it decreases in Africa and Asia. So, we can conclude that there has been a change in the behaviour of international migration.

Returning to the growth rates of the stocks of immigrants, (Panel (b) of Table 1), we can also point out that the rate never reaches its maximum in each area in the last decade (1990–2000), so it appears that total immigration has not increased notably. A clear example is Europe, where the rate grows faster in the decade 1960–1970 than in 1990–2000. This fact, together with the increased migration density, which rose 3%, indicates that the birth rate of European countries is responsible for this situation.

Moreover, from the study of the changes in the ranking of the stock of immigrants, it can be concluded that the countries that show the greatest variation in the ranking are some African and Arabic countries, while the most developed countries do not present relatively high variations .[5]

Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Republic of Moldova, Russian Federation, Tajikistan, Turkmenistan, Ukraine and Uzbekistan.

[5] For this purpose, we used the Spearman coefficients for different periods and the relationship between the initial value of the variable and its final relative position. The results are available upon request.

3. Zipf's law and Gibrat's law

In this section, we present an analysis of two laws traditionally associated with the populations of cities and, recently, with the size of nations. These two laws can be studied from the point of view both of volume and of migration density.

3.1. Zipf's law

In this section, we examine the size distribution of stock and of migration density in order to see if there has been convergence or divergence between the different countries of the world. To do this, we use Pareto's distribution (1896), also known as a power law, that can be expressed as:

$$R(M) = aM^{-b}, \qquad (1)$$

where M denotes the stock of immigrants, and a and b are parameters. This expression has been used extensively in urban economics to study the size distribution of cities (see, for example, Eeckhout, 2004, and Ioannides and Overman, 2003) or the size distribution of countries (Rose, 2006; Gonzalez-Val and Sanso-Navarro, 2009), and a theoretical justification can be found in Eeckhout (2004) and Duranton (2007).

A particular case of Pareto's distribution is Zipf's law (1949), which appears when $b = 1$, and means that, ordered from larger to smaller, the stock of immigrants of the first country is twice the stock of the second, three times the stock of the third, and so on. Another empirical regularity related to Zipf's law and Pareto's distribution is Gibrat's law (1931), which postulates that the growth rate of the variable is independent of its initial size. However, Eeckhout (2004) and Duranton (2007) demonstrate that there is a possibility that only the upper tail fits this distribution and that, when the total sample is considered, the distribution which fits best is the lognormal. In this work, we test these empirical regularities for the stock of immigrants.

It is also interesting to test whether Pareto's parameter is more or less than one and the evolution of this coefficient in time. The higher the coefficient, the more homogeneous are the stocks of immigrants. A growing evolution would mean a process of convergence in the immigrant stock and a decreasing evolution would mean a process of divergence.

The expression (1) of Pareto's distribution is usually estimated in its logarithmic version:

$$\ln R = \ln a - b \cdot \ln M \ . \qquad (2)$$

Different sample sizes have been used, considering the 50, 100 and 150 countries with the biggest stock of immigrants, and the estimation has also been carried out with the total of all countries.[6] Table 2 presents the results of the OLS estimation.[7]

The estimation of the parameter is significantly different from one except when we consider only the 50 largest stocks of immigrants. It is very close to one in the upper tail of the distribution, obtaining a good adjustment level, while, as the size of the sample increases, the estimated value of b and the degree of adjustment decrease. The value of the coefficient increases over time, so some convergence is detected, especially when we consider the 100 countries with the biggest stock. Table 3 presents some measures of concentration, but the Gini coefficient indicates that the distribution is very uneven and remains almost constant throughout the period examined.

---

[6] We also consider the possibility of differentiating immigrants by sex. The estimations show that differentiated behaviours do not exist.

However, the Hill (maximum likelihood) estimator could be more efficient than the OLS estimator in the upper tail, as shown by Gabaix and Ioannides (2004), so Table 4 presents the results of the Hill estimator[8] in the upper tail (top 50 and top 100 countries). Although the estimated coefficients are slightly different from the OLS estimates, for the top 50 they are not significantly different from one at the 5% significance level.

The presence of a decreasing Zipf coefficient with the sample size, as shown in Eeckhout (2004), may be because distribution is lognormal and so, on selecting the countries with the highest, stock only the upper tail is considered, which is a good approximation to a Pareto distribution. In fact, for our migration data the Pareto distribution provides a better fit for the upper tail than the lognormal. Figure 1 shows the Rank-Size plot corresponding to the year 2000 (the graphs for the other years are very similar), and we can see that, although the lognormal has a good fit for most of the distribution, for the upper tail, the Pareto fitted by maximum likelihood is closer to the data. In order to test whether the distribution is lognormal throughout the sample, in Figure 2, we present the adaptive kernels which represent the estimated empirical density functions for three representative decades. We observe an approximation to lognormal distribution.[9]

Finally, we analyze the distribution of migration density. If we consider the earlier results referring to the stock of immigrants, as well as those obtained by Rose (2006) for the total population, it is to be expected that the difference between the

---

[7] The residues resulting from this regression usually present problems of heteroskedasticity so, to analyze the significance of the parameters, the typical corrected deviation proposed by Gabaix and Ioannides (2004) is used: $\text{GI s.e.} = \hat{b} \cdot (2/N)^{1/2}$, where N is the sample size.

[8] Calculated using the PARETOFIT stata module, developed by Jenkins and Van Kerm (2007).

[9] The Kolgomorov-Smirnov test shows that the null hypothesis of normality is not rejected for any decade, for either the stock of immigrants and migration density, providing evidence against Zipf's law when considering all the countries.

logarithms of the two variables follows a lognormal distribution. The results of the OLS estimation of the Pareto exponent are shown in Table 2, while Table 4 displays the results for the upper tail using the Hill estimator.

Again, if we take the whole sample, the distribution is uneven, although not as much as in the case of the stock of immigrants. We reject Zipf's law, and the concentration indexes (see Table 3) indicate a slight tendency to divergence between countries. The graphic representation of the adaptive kernel of migration density, Figure 2, also shows an evolution towards a lognormal distribution, starting from a very leptokurtic distribution in 1960. The centre of distribution has lost importance compared to the tails, which indicates that growth has not been convergent.

Therefore, we can conclude that both the stock of immigrants and migration density follow similar distributions to those found in Eeckhout (2004) and Rose (2006) for the size of US cities and of countries, respectively, confirming the presence of an empirical stylized fact when the spatial distribution of the population is considered, even though the determinants of migratory flows between countries are different than those within the same country. Moreover, we can affirm that no sign of convergence appears in the stock of immigrants and in migration density. This may be considered as evidence in favour of the theoretical models which find long term migration rates different to zero, recognizing the presence of factors which compensate for the effects of traditional factors such as income or wages. However, the question of convergence must be analyzed in the framework of Gibrat's law, which we do in the next section.

3.2. Gibrat's law

The previous section has shown that there is a certain stability in the distribution of the immigrant population size and in migration density, although a small tendency towards divergence was observed in the latter. However, for a dynamic analysis, we have to use growth rates. We are interested in verifying whether Gibrat's law holds or not.[10] We will use the methodology followed by Ioannides and Overman (2003) and Eeckhout (2004). It consists of taking the following specification:

$$g_i = m(S_i) + \varepsilon_i , \qquad (3)$$

where g_i is the normalized growth rate (subtracting the mean and dividing by the standard deviation) and S_i is the logarithm of the stock of immigrants and, instead of making suppositions about the functional relationship of m and supposing a linear relationship, $\hat{m}(s)$ is estimated as a local average around point s and is smoothed using a kernel, which is a symmetrical, weighted and continuous function around s.

In order to analyze the period 1960–2000, the Nadaraya-Watson method is used, exactly as it appears in Härdle (1990), based on the following expression:[11]

$$\hat{m}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - S_i) g_i}{n^{-1}\sum_{i=1}^{n} K_h(s - S_i)} , \qquad (4)$$

where K_h denotes the dependence of the kernel K (in this case an Epanechnikov kernel) on the bandwidth h (0.5). Starting from this calculated mean, $\hat{m}(s)$, the

[10] Gibrat (1931) observed that the size distribution (measured by sales or the number of employees) of firms tends to be lognormal, and his explanation was that the growth process of firms can be multiplicative and independent of the size of the firm.

variance of the growth rate, g_i is also estimated, again applying the Nadaraya-Watson estimator starting from:

$$\hat{\sigma}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s-S_i)(g_i - \hat{m}(s))^2}{n^{-1}\sum_{i=1}^{n} K_h(s-S_i)}. \qquad (5)$$

The estimator is very sensitive, both in mean and in variance, to atypical values. So, we eliminate 5% of the lowest observations of the distribution, both for the stock of immigrants and for migration density, as these observations are characterized by a very high dispersion both in mean and in variance.[12] For the case of stock, four more very atypical values are eliminated.[13]

If growth were independent of size, the estimated kernel would be a straight line on the zero value. Values different from zero imply deviations with respect to the mean. Variance would also be a straight line if it does not depend on the size of the variable analyzed.

Figure 3 presents the nonparametric estimates of the mean growth and of the variance of that growth, for both the stock of immigrants and migration density. Bootstrapped 95% confidence bands are also displayed.[14] For the calculation, all the available observations covering the entire sample period have been taken into account (more than 800 observations). Some conclusions can be highlighted. Regarding the stock of immigrants, the estimated kernel of the mean is close to zero. However, the trend is slightly decreasing; the bigger the size, the smaller the

---

[11] The calculation was done with the KERNREG2 stata module, developed by Cox et al. (1999). This program is based on the algorithm described by Härdle (1990) in Chapter 5.

[12] The majority of these 43 excluded observations of the stock of immigrants correspond to African countries and islands which constitute independent states. In the case of migration density, Asian countries also appear (China and Vietnam, for example), which, due to their high populations, have low migration density.

[13] These four observations correspond to the United Arab Emirates (1960–1970), Djibouti (1970–1980), Mozambique (1970–1980) and Somalia (1970–1980).

[14] They have been calculated using 500 random samples with replacement.

growth rate.[15] The null hypothesis of this mean being equal to zero can be rejected at a 5% significance level only for some values in the upper tail so, except for these values of the distribution, Gibrat's law holds in the period examined. Thus, we find evidence of slight convergence in the stock of immigrants because the biggest countries have had less mean growth.

For migration density, we observe two clearly differentiated behaviours: the countries with a lower rate have grown more slowly than those that began with a higher rate. Therefore, we find divergence. And it is also observed that variance is independent of size, except for some upper tail distribution values, so evidence against Gibrat's law in migration density is not found.

If we interpret the two variables together, we can say that, although the growth of the stock of immigrants does not appear to have been especially important when establishing the size distribution, migration density is. This is possibly because the host countries have lower birth rates than the origin countries so that, while the stock of immigrants grows at the same or a similar rate, migration density increases.

This could be important for several reasons. The first is that less population growth in the host countries possibly generates a scarcity in the labour supply, which creates wage differences that encourage migration. As long as these differences are maintained, migration will continue. On the other hand, the stock of immigrants reduces immigration costs and, as long as the importance of this fact is greater than the rate at which wages converge, migration will continue. Finally, in a context where a scale effect of population on economic growth is

[15] Although clusters of countries are detected that differ from the trend, both in mean and in variance, we cannot reject that the mean growth is equal to zero or that the variance is equal to one.

important, this result will be useful for understanding the existence of persistent differences in wages. Furthermore, it may help us to understand why migratory patterns are maintained and to give us new perspectives on the mechanisms by which migration affects the economic growth and welfare of the host countries.

4. Conclusions

This paper studies the evolution of the worldwide distribution of the stock of immigrants, focusing on two well-known empirical regularities in urban economics, Zipf's law and Gibrat's law. The analysis shows that both the stock of immigrants and migration density, defined as the percentage of immigrants over the total population of the country, follow similar size distributions to those of cities and of countries, although population movements at an international level are more difficult than at a sub-national level. We use parametric and nonparametric methods and obtain the following results.

First, for the stock of immigrants, the estimated Pareto exponent is very close to one in the upper tail of the distribution while, as the size of the sample increases, the estimated values decrease. Moreover, the Gini coefficient indicates that the distribution is very uneven and remains almost constant throughout the period examined. Also, the estimated kernels show that the distribution that fits best is lognormal, while the upper tail is represented by a Pareto distribution, a statistical regularity already shown in Eeckhout (2004) for the case of North American cities. We show that growth is independent of size, although we found a weak convergence in the size distribution of the stock of immigrants.

We have repeated the analysis for migration density. In this case, if we take the whole sample, the distribution is uneven, although not as much as in the case of the stock of immigrants, and the concentration indexes indicate a slight tendency to divergence between countries. This coincides with the results offered by the estimated kernels, which show a loss of kurtosis in the centre of the distribution. Although the estimated kernel of the growth of migration density is close to zero, we observe two clearly differentiated behaviours: the countries with a lower rate have grown more slowly than those that began with a higher rate. Consequently, in the period examined, we observe a divergent behaviour.

These results support theoretical models with a non-null equilibrium migration rate. Therefore, we find empirical evidence in favour of models that consider the importance of factors such as capital stock, the social cost of migration, the skill composition of the native labour force, migration policy, market potential… That is to say, the conclusion is that the models which use the wage gap as the most important element in the migration process could be extended in an important way by other determinants.

Finally, the most important conclusion of this paper is that knowledge of the migration process is not independent of the spatial distribution of the population and this is, obviously, an important point that researchers in migration and in economic geography must take into account.

References

Alesina, A., E. Spolaore and R. Wacziarg, (2000). Economic Integration and Political Disintegration. American Economic Review 90: 1276–1296.

Carrington, W. J., E. Detragiache and F. Vishwanath, (1996). Migration with Endogenous Moving Cost. American Economic Review, 86, 909–930.

Cox, N, I. Salgado-Ugarte, M. Shimizu and T. Toru Taniuchi, (1999). KERNREG2: Stata module to compute kernel regression (Nadaraya-Watson). Available online at http://fmwww.bc.edu/repec/bocode/k/kernreg2.ado

Duranton, G., (2007). Urban Evolutions: The Fast, the Slow, and the Still. American Economic Review 97(1), 197–221.

Eeckhout, J., (2004). Gibrat's Law for (All) Cities. American Economic Review, vol. 94(5), 1429–1451.

Gabaix, X., and Y. M. Ioannides, (2004). The evolution of city size distributions. Handbook of urban and regional economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, 2341–2378.
Ghatak, L., and P. Wheatley, (1996). Migration Theories and Evidence: an Assessment, Journal of Economic Surveys, 10, 159–198.
Gibrat, R., (1931). Les inégalités économiques. Paris: Librairie du Recueil Sirey.
Gonzalez-Val, R., and M. Sanso-Navarro, (2009). Gibrat's Law for Countries. Journal of Population Economics, forthcoming. DOI 10.1007s00148-009-0246-7.
Härdle, W., (1990). Applied nonparametric regression. Econometric Society Monographs. Cambridge, New York and Melbourne: Cambridge University Press.
Harris, J. R., and M. P. Todaro, (1970). Migration, Unemployment and Development: A Two-Sector Analysis. American Economic Review, 60, 126–142.
Ioannides, Y. M., and H. G. Overman, (2003). Zipf's law for cities: an empirical examination. Regional Science and Urban Economics 33, 127–137.
Jenkins, S. P., and P. Van Kerm, (2007). PARETOFIT: Stata module to fit a Type 1 Pareto distribution. Available online at http://ideas.repec.org/c/boc/bocode/s456832.html.
Larramona, G., and M. Sanso, (2006). Migration dynamics, growth and convergence. Journal of Economic Dynamic and Control, 30(11), 2261–2279.
Pareto, V., (1896). Cours d'Economie Politique. Geneva: Droz.
Rose, A. K., (2006). Cities and countries. Journal of Money, Credit, and Banking, 38(8), 2225–2246.
Spolaore, E., and R. Wacziarg, (2005) Borders and Growth. Journal of Economic Growth 10: 331–386.
United Nations, (2004). Trends in total migrant stock: The 2003 revision. Database maintained by The Population Division of the Department of Economic and Social Affairs.
Zipf, G., (1949). Human Behaviour and the Principle of Least Effort. Cambridge, MA: Addison-Wesley.

Tables

Table 1. Descriptive analysis

Panel (a): Total Stock

Area	Stock of Immigrants				
	1960	1970	1980	1990	2000
ASIA	29,280,680	28,103,771	32,312,541	41,754,291	43,761,383
EUROPE	14,015,392	18,705,244	22,163,201	26,346,258	32,803,182
NORTHERN AMERICA	12,512,766	12,985,541	18,086,918	27,596,538	40,844,405
AFRICA	8,977,075	9,862,987	14,075,826	16,221,255	16,277,486
LATIN AMERICA AND THE CARIBBEAN	6,038,976	5,749,585	6,138,943	7,013,584	5,943,680
OCEANIA	2,134,122	3,027,537	3,754,597	4,750,591	5,834,976
TOTAL	75,900,698	81,527,177	99,783,096	154,005,048	174,933,814

Panel (b): Growth of the Stock of Immigrants

Area	Growth of the Stock of Immigrants (%)				
	1960–1970	1970–1980	1980–1990	1990–2000	1960–2000
ASIA	-4.02	14.98	29.22	4.81	49.45
EUROPE	33.46	18.49	18.87	24.51	134.05
NORTHERN AMERICA	3.78	39.29	52.58	48.01	226.42
AFRICA	9.87	42.71	15.24	0.35	81.32
LATIN AMERICA AND THE CARIBBEAN	-4.79	6.77	14.25	-15.25	-1.58
OCEANIA	41.86	24.01	26.53	22.83	173.41
TOTAL	7.41	22.39	54.34	13.59	130.48

Panel (c): Migration Density

Area	Migration Density (%)				
	1960	1970	1980	1990	2000
ASIA	1.76	1.34	1.25	1.35	1.21
EUROPE	3.30	4.08	4.59	5.28	6.42
NORTHERN AMERICA	6.13	5.60	7.06	9.73	12.93
AFRICA	3.24	2.76	3.00	2.61	2.05
LATIN AMERICA AND THE CARIBBEAN	2.77	2.02	1.70	1.59	1.14
OCEANIA	13.43	15.57	16.45	17.80	18.80
TOTAL	2.51	2.21	2.25	2.93	2.88

Table 2. Pareto coefficients by decade estimated by OLS

	Stock of Immigrants						Migration Density					
	Top 50			Top 100			Top 50			Top 100		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	0.966	0.193	0.981	0.641	0.091	0.915	1.327	0.265	0.916	1.118	0.158	0.944
1970	0.939	0.188	0.973	0.685	0.097	0.935	1.363	0.273	0.909	1.071	0.152	0.931
1980	1.035	0.207	0.983	0.719	0.102	0.931	1.449	0.290	0.877	1.047	0.148	0.898
1990	0.925	0.185	0.982	0.726	0.103	0.947	1.654	0.331	0.904	1.106	0.156	0.893
2000	0.939	0.188	0.981	0.743	0.105	0.952	1.713	0.343	0.890	1.071	0.151	0.874
	Top 150			All (214)			Top 150			All (214)		
Year	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2	b < 0	(GI s.e.)	R^2
1960	0.523	0.060	0.909	0.333	0.032	0.790	0.941	0.109	0.935	0.558	0.054	0.775
1970	0.551	0.063	0.916	0.348	0.033	0.787	0.874	0.101	0.915	0.535	0.052	0.773
1980	0.572	0.066	0.910	0.354	0.034	0.780	0.843	0.097	0.895	0.541	0.052	0.794
1990	0.580	0.067	0.921	0.352	0.034	0.780	0.838	0.097	0.872	0.521	0.050	0.776
2000	0.569	0.066	0.912	0.343	0.033	0.770	0.804	0.093	0.862	0.499	0.048	0.771

All coefficients are significantly different from zero at the 0.05 level.
(GI s.e.) Gabaix-Ioannides (2004) corrected standard error.

Table 3. Concentration measures

		Stock of Immigrants			Migration Density		
Year	Sample Size	Herfindahl	Normalized Herfindahl	Gini Coefficient	Herfindahl	Normalized Herfindahl	Gini Coefficient
2000	All (214)	0.079	0.075	0.846	0.015	0.010	0.667
	Top 150	0.080	0.073	0.787	0.016	0.009	0.555
	Top 100	0.082	0.073	0.712	0.018	0.008	0.440
	Top 50	0.097	0.078	0.610	0.025	0.005	0.275
1990	All (214)	0.074	0.070	0.845	0.015	0.010	0.659
	Top 150	0.074	0.068	0.785	0.016	0.009	0.547
	Top 100	0.077	0.068	0.712	0.018	0.008	0.440
	Top 50	0.091	0.072	0.608	0.026	0.006	0.290
1980	All (214)	0.045	0.041	0.822	0.016	0.011	0.669
	Top 150	0.045	0.039	0.753	0.017	0.010	0.565
	Top 100	0.047	0.038	0.667	0.019	0.009	0.466
	Top 50	0.056	0.036	0.520	0.027	0.007	0.319
1970	All (214)	0.046	0.042	0.836	0.017	0.013	0.668
	Top 150	0.046	0.040	0.772	0.018	0.011	0.569
	Top 100	0.048	0.038	0.691	0.021	0.011	0.481
	Top 50	0.056	0.036	0.550	0.029	0.009	0.350
1960	All (214)	0.051	0.047	0.845	0.017	0.012	0.654
	Top 150	0.052	0.045	0.784	0.018	0.011	0.554
	Top 100	0.053	0.043	0.703	0.021	0.011	0.474
	Top 50	0.061	0.041	0.554	0.030	0.011	0.364

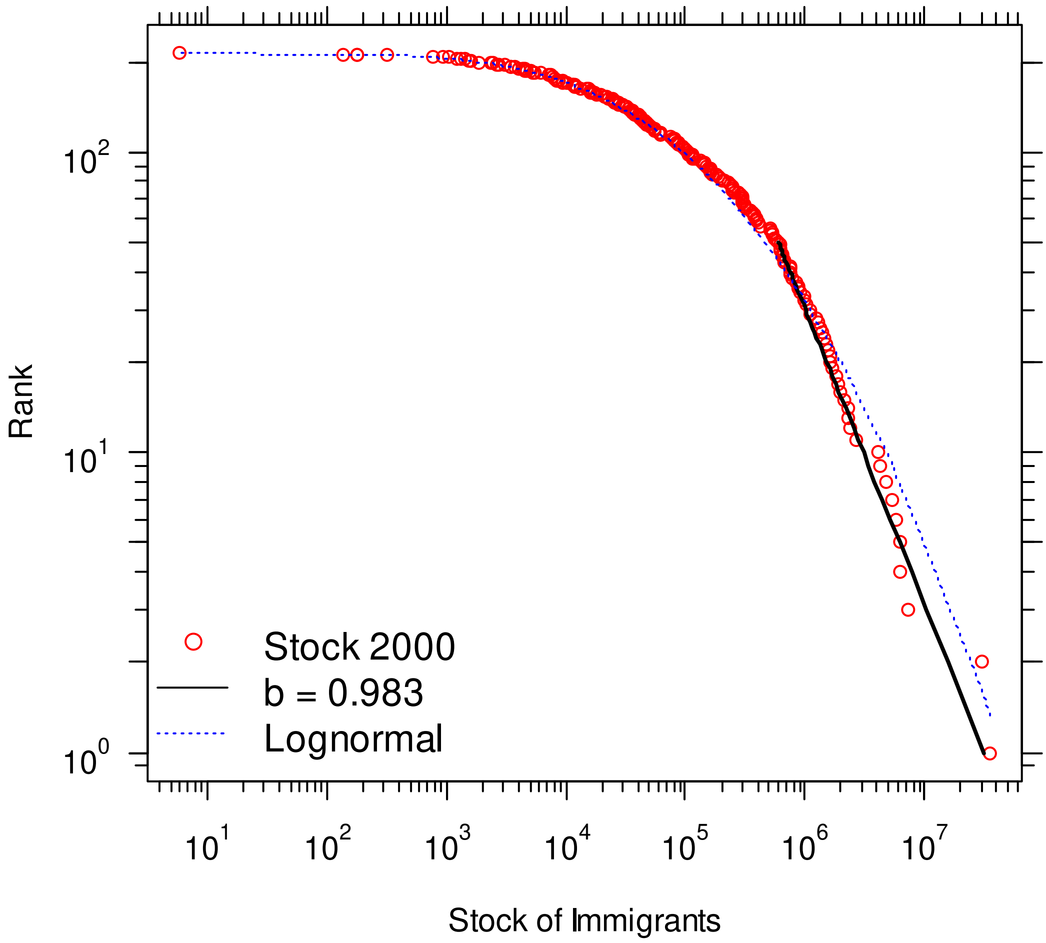
Table 4. Pareto coefficients in the upper tail obtained using Hill's estimator

Stock of Immigrants				
	Top 50		Top 100	
Year	Hill	(Hill s.e.)	Hill	(Hill s.e.)
1960	0.830	0.117	0.489	0.049
1970	0.880	0.124	0.548	0.055
1980	0.775	0.110	0.569	0.057
1990	0.891	0.126	0.586	0.059
2000	0.983	0.139	0.554	0.055
Migration Density				
	Top 50		Top 100	
Year	Hill	(Hill s.e.)	Hill	(Hill s.e.)
1960	1.133	0.160	0.961	0.096
1970	1.086	0.154	0.908	0.091
1980	1.188	0.168	0.838	0.084
1990	1.159	0.166	0.852	0.085
2000	1.208	0.171	0.825	0.083

All coefficients are significantly different from zero at the 0.05 level.

Figures

Figure 1. Rank-Size plot (log scale) for the stock of immigrants (year 2000)



Note: Top 50 data are fitted by a power-law whose exponent is estimated as b = 0.983 (±0.272) by using Hill's estimator. Also shown is the fit by log-normal distribution for the entire range based on the maximum likelihood estimation.

Figure 2. Adaptive kernels (ln scale)

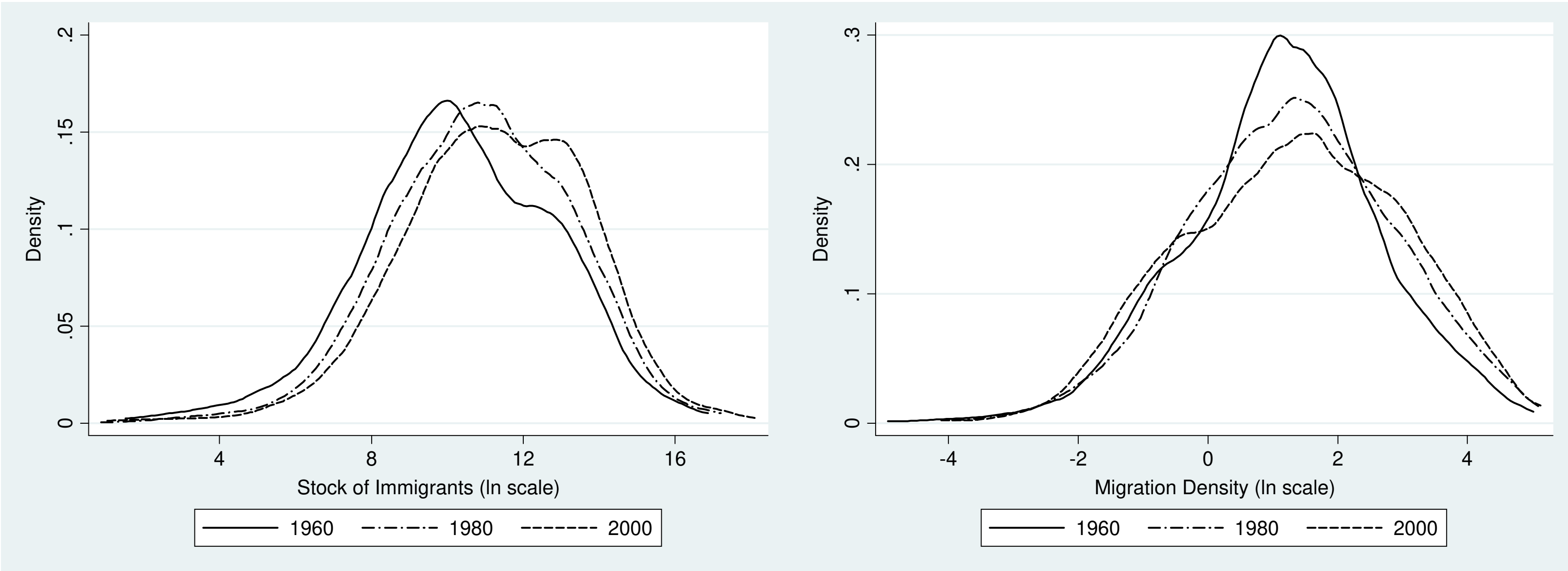


Figure 3. Kernel estimates, 1960–2000 (bandwidth 0.5)