

Gravity Model and Zipf's Law: An In-Depth Study into the Nature of International Trade

Mislav Josic, Ph.D.

Teaching and Research Assistant, Faculty of Economics and Business Zagreb,
University of Zagreb, International Economics Department, Zagreb, Croatia
Email: mjosic@efzg.hr, Phone: +385992714482

Marija Nikic

Student Teaching Assistant, Faculty of Economics and Business Zagreb,
University of Zagreb, International Economics Department, Zagreb, Croatia
Email: marija.nikic.mn@gmail.com, Phone: +385996386406

Doi:10.5901/ajis.2013.v2n9p583

Abstract

In this paper trade pattern based on distances between countries was tested through gravity model approach. This research highlights some of the regularities in international trade that arise and are in concordance with the rank-size rule. An empirical study was conducted on static data used in Croatia in 2012 in order to test the validity of Zipf's law. The presence of Zipf's law was empirically tested using population data for city proper and city settlements as well as the export values of Croatia's most important trading partners. In both cases results corroborate the existence of Zipf's law in population and export values. Due to the heavy tailness of Pareto distribution minor adjustments to the model have been made. Gravity model approach shows significant impact of distance and GDP on international trade. Dummy variable common border indicates that most of the trade in Croatia is done due to preferable geographical position of the country.

Keywords: International trade, Zipf's law, rank-size rule, gravity model, geographical economics

1. Introduction

Most of the studies conducted in the field of international economics are aimed towards the analysis of causes and effects in international trade. Classical theories of international trade explained trade as a result of price difference based on different absolute or relative labour productivities among countries. Later on, Eli Heckscher and Bertil Ohlin introduced capital as second factor of production having both production possibilities and demand preferences. New trade theories assumed several assumption that were rather rigid in the aforementioned theories. This paper unpretendedly aims to recreate new trade pattern of international trade based on Zipfian or rank-size distribution. Why do cities emerge, how do they locate in countries and how do production inputs relocate within a country is still the key question in geographical economics.

Power laws originate from mathematics and construct relationship between two variables. Power law relation is merely a function of two variables expressed as a power of one variable. Due to the nature of power functions logarithmic decomposition can be made as well as its exponential counterpart from reverse action. Let's define power law function as follows:

$$y(x) = ax^k \quad (1)$$

where x and y are independent and dependent variables, a is estimated parameter and k is the power coefficient. The same (1) expression can be transformed in logarithmic form (2) as follows:

$$\log y = \log a + k \log x \quad (2)$$

Previously mentioned functions are deterministic which is a characteristic of hard sciences such as physics and mathematics, but not economics. In most cases natural laws incorporate stochastic component ε as a result of sampling/measurement error or any other uncertainty that affects the mean. In this respect power law in social sciences is often remodeled into nondeterministic form (3):

$$\log y = \log a + k \log x + \varepsilon \quad (3)$$

where \mathcal{E} represents deviation factor or error term. There are several varieties of power laws available and its prevalence is in mathematics, physics, biology, economics and many other social sciences in particular. In this paper we analyse specific form of power law that is known as Zipf's law and is of great importance for economists. First proposition regarding city-size relation was made by Felix Auerbach (1913). Fine tuning of Auerbach's findings was done by American researcher George Kingsley Zipf (1949) in a rule that states that frequency of some occurrence is inversely related to the rank of the same occurrence. The latter conclusion was formed out of analysis of word occurrence in text where one could conclude *that the frequency of the word occurrence is inversely related to its rank in frequency table*. Cobb-Douglas production function is also a power law function with two parameters in the exponent (α and $1-\alpha$).

2. Application of Zipf's Law in Economics

How can Zipf's law (or rule) be applied in economics today? Due to its nonlinearity (convexity) Zipf's law is very useful in convex analysis of a single variable calculus. Well known Pareto distribution has a power law tail and is very similar to Zipf's distribution¹. Pareto distribution is strictly used for continuous functions while Zipf's one is for discrete cases. In this respect they are somewhat different and much alike at the same time. One of the first applications of Zipf's law was in geographical economics. Population of the cities and their rank follow Zipf's law. The second biggest city in a country is approximately $\frac{1}{2}$ of the largest city size, and third biggest city is $\frac{1}{3}$ of the largest city size, etc. The rank-size rule applied in geographical economics emerges from Auerbach's paper from the beginning of 20th century and can be expressed in Pareto distribution (4):

$$y(x) = ax^{-k} \quad (4)$$

which is equivalent to (5):

$$\log y = \log a - k \log x \quad (5)$$

where x is population size, y is the number of cities larger than x , and a, k as constants. Zipf's work on Auerbach's idea goes further and assumes that distribution should have the following values of parameters: $k = 1$ and $a = \text{the size of the biggest city}$ in order for data to fit. Several papers have tested Zipf's law using different approaches. Let us mention few of the most promising:

Ioannides and Overman (2000) empirically examined validity of Zipf's law for metro areas in the United States on a large set of data ranging from 1900 – 1990. They used nonparametric test in order to estimate variabilities in endogenous variables. Results have showed that Zipf's law holds for a large range of city sizes.

General investigation in the area of power laws was systematically laid out by Newman (2004) showing remarkable concordance with Zipf's law. It was found in variations of the following variables: word frequency, citations, web hits, books sold, telephone calls received, earthquake magnitude, crater diameter in km, net worth in US dollars, name frequency and most notably population of cities.

A cross country investigation of Zipf's law was again tested by Kwok Tong Soo (2005) using OLS and Hill's estimator. With the use of OLS method Zipf's law was rejected for majority of countries while Hill's estimator rejected Zipf's law for minority of countries. Overall conclusion was that political economy variables better explain variations in parameters than geographical ones.

Hinloopen and van Marrewijk (2008) concluded that Balassa index of revealed comparative advantage also follows Zipf's law using estimator developed by Gabaix and Ibragimov (2006). Despite the general acceptance of Zipf's law in data there are differences between countries, sectors, population size, etc.

Recent comprehensive study in the area of Economics and Finance was conducted by Xavier Gabaix (2009) leaving several questions open for discussion, e.g. why do origins of cities, firms, international trade, CEO pay or market fluctuations follow power laws.

3. Data and Methodology

In order to test the presence of Zipf's law an appropriate population data for city proper and city agglomeration was taken from Croatian Bureau of Statistics based on census survey in 2011. Population was matched for top 100 cities or

¹ Sometimes referred as zeta distribution.

settlements ranked by size. Along with classic relationship between rank and size of the cities presented by G. K. Zipf, this paper goes further on establishing the link between export values and its rank among top 100 importing countries of Croatian products.

The following regression equations for city proper and city agglomerations were used in testing the validity of Zipf's law:

$$\log\left(\text{Rank}_i - \frac{1}{2}\right) = \beta_0 + \beta_1 \log \text{City Proper}_i + \varepsilon_i \quad (6)$$

for each i -th city proper ranging from largest to smallest (Zagreb to Zlatar), $i = 1 \dots 100$.

and

$$\log\left(\text{Rank}_i - \frac{1}{2}\right) = \beta_0 + \beta_1 \log \text{City Agglomeration}_i + \varepsilon_i \quad (7)$$

for each i -th city agglomeration ranging from largest to smallest (Zagreb to Buje), $i = 1 \dots 100$.

Experimental test of power law, more specifically Zipf's law, present in export values of Croatia's top 100 trade partners was carried out very similar:

$$\log\left(\text{Rank}_i - \frac{1}{2}\right) = \beta_0 + \beta_1 \log \text{Exports}_i + \varepsilon_i \quad (8)$$

for each i -th importing partner country of Croatia ranging from largest to smallest (Italy to New Zealand), $i = 1 \dots 100$.

In all three log-linear regression equations (6)-(8) rank values were adjusted using Gabaix-Ibragimov estimator due to the fact that most of the data were grouped on the right side of the Pareto distribution. Empirical results in the next section should confirm model's *goodness of fit* (high R^2 value) and estimated β_1 close to -1.

Gravity empirics was tested using 2012 export values (in EUR) for Croatia's trading partner countries, again first 100 countries. Data was collected from Croatian Bureau of Statistics based on values for foreign trade in goods in 2012. Partner GDP values were downloaded from Eurostat and World Bank sources for year 2012. Aerial distance was calculated as a shortest distance between two capital cities of partner countries engaged in bilateral trade.

Proposed statical gravity model is as follows:

$$\log \text{Exports}_{ijt} = \beta_0 + \beta_1 \log \text{GDP}_{jt} + \beta_2 \log \text{Distance}_{ij} + \beta_3 \text{common_border}_{ij} + \varepsilon_{ij} \quad (9)$$

where

$i = 1$ - reference country (Croatia),

$j = 1 \dots 100$ - partner country (Italy to New Zealand),

$t = 2012$ - base year,

Exports_{ijt} - exports from Croatia to j -th partner country in year t (in EUR),

GDP_{jt} - j -th partner country GDP in year t (in EUR),

Distance_{ij} - distance from Croatia's capital Zagreb to j -th country capital city,

$\text{common_border}_{ij}$ - 0 if countries do not share common border, otherwise 1

ε_{ij} - noise/error term

If practice (empirics) complies with theory then expected signs of regression parameters β_i should be: $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$ along with high R^2 value and statistical significance of all independent variables.

4. Empirical Results

With the methodology already mentioned in previous section the presence of Zipf's law is tested for city proper and city agglomeration values based on 2011 data. Graphical relations are shown at Figure 1 and Figure 2. Experimental data analysis provided for export values in Croatia's goods trade yields somewhat different results, but still within the theoretical bounds, which can be carefully observed at Figure 3.

Figure 1: Zipf's law for city proper in Croatia in 2011

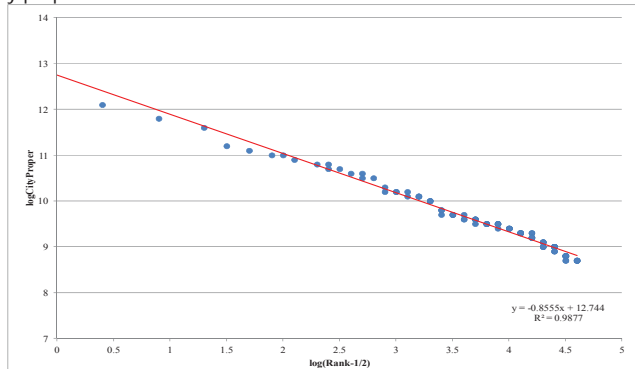
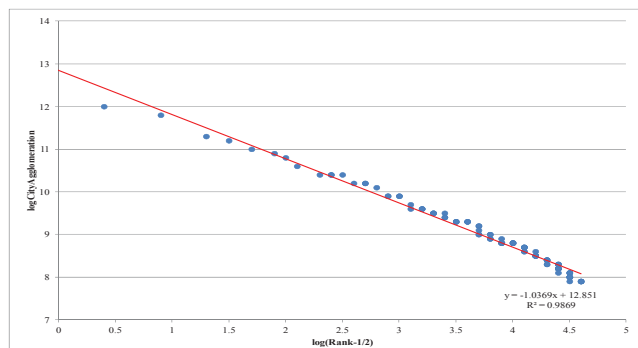


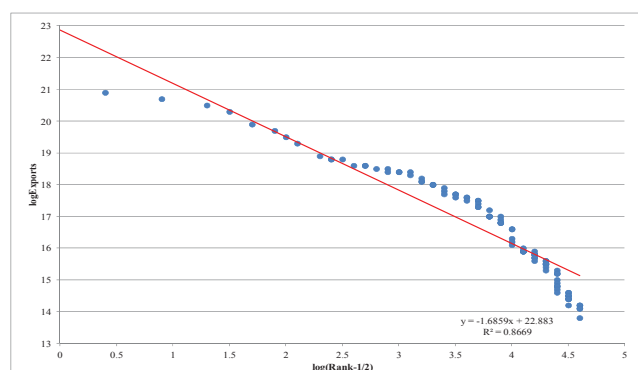
Figure 2: Zipf's law for city agglomeration in Croatia in 2011



As shown at Figure 1 and Figure 2 observed data scatter neatly along the line representing log-linear fit corresponding with Zipf's findings. Both R^2 values (0.9877 and 0.9869) are very high implying the correct use of an economic model. The slopes of the regression lines in both cases are negative with coefficients close to -1. Observed data for city agglomeration fit more better Zipf's model with $\beta_1 = -1.0369$ as opposed to city proper data where $\beta_1 = -0.8555$.

Croatia's goods export values and the rank of the importing countries also follow Zipf's law but not perfectly as shown in two previous cases.

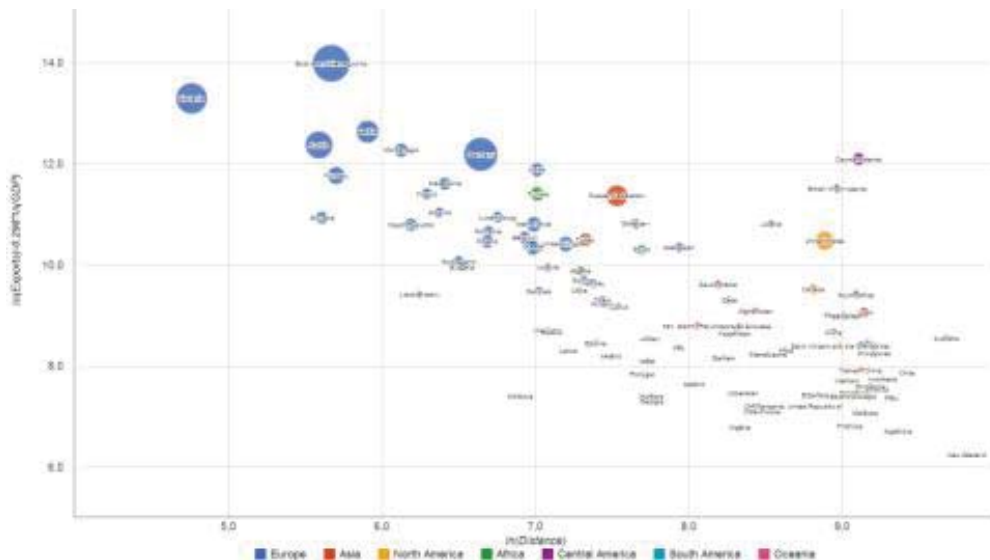
Figure 3: Zipf's law for Croatia's export partners in 2012



Coefficient of determination is still very high in this case also ($R^2 = 0.8669$) while the regression line is more steeply (negatively) inclined with $\beta_1 = -1.6859$. Overall conclusion is that population of city proper, city agglomeration and export values to partner countries follow Zipf's law.

As a complement to Zipf's law gravity model assumes several important facts that should be accounted for future investigation of trade patterns and trade flows. Most of the international trade among countries is affected by location and geographical position of a country as a result of differences in transport costs. Due to Croatia's Mediterranean orientation and its close connection with the European Union most of the trade is created with countries in near proximity, most notably Italy, Bosnia and Herzegovina, Germany, Slovenia and Austria while the rest of the non-Europe (extra) trade is created with Russian Federation and United States. Trade distribution over continents along with the size of a countries measured by GDP can be closely examined at Figure 4.

Figure 4: Croatia's exports and distance over continents



Regression analysis on expression (9) corroborates theoretical assumptions regarding the impact of geographical distance of countries and transport costs on bilateral trade flows. Estimated regression equation based on expression (9) yields the following results:

Table 1: Regression output in gravity model for Croatia, 2012 data

Independent variable	β_i coefficient	t-statistic	p-value
Constant	16.43754***	9.97834	0
LogGDP	0.297908***	5.329575	0
LogDistance	-0.921989***	-7.857278	0
LogCommon_border	1.637159***	2.942002	0.0041
R^2	0.586568		
F-statistic	45.40086		
Prob(F-statistics)	0		

*** - variable significant at 1% level

** - variable significant at 5% level

* - variable significant at 10% level

From Table 1 it is clear that all independent variables are statistically significant at 1%, 5% and 10% level using t-statistic. F-statistic shows that at least one of the variables in group test is statistically significant as well. The signs of β_i

coefficients are consistent with economic theory, i.e. classical gravity model approach. Coefficient of determination is not too high thus reducing the risk of multicollinearity in OLS model, but still explains a great portion of variability in dependent variable.

5. Conclusions

International economics today still doesn't have one *Theory of everything*, unified theory that should explain causes and effects of international trade. Trade effects are pretty much known to us, but the causes of trade still remain a mystery. Geographical economics as a part of international economics analyses micro and macro aspects of trade explaining why do cities emerge, and most importantly where do they locate. Local trade creates micro trade patterns that also reflect in macro trade patterns, the one that exist between countries. In this paper we tested Zipf's law, one of the power laws based on Pareto distribution which puts in relation rank and size of the cities and settlements in a country. Positioning of cities greatly affects local economy due to re(a)location of inputs and outputs thus leading to new optimal equilibrium. Empirical results from this paper suggest that the size of city proper and city agglomeration in Croatia strictly follows Zipf's law based on very strong log-linear relationship in variables. Surprisingly, Croatia's export values also follow Zipf's law with less explained variance in dependent variable. In addition to Zipf's law, gravity model was constructed based on classical model by Jan Tinbergen. Estimated parameters of log-linear regression output are consistent with implied (theoretical) values and strongly support the thesis of trade based on economic size measured by GDP and distance between countries. Common border as a qualitative regressor variable also positively affects international trade giving preference to adjacent countries.

References

- Auerbach, F. (1913) Das Gesetz der Bevölkerungskonzentration, Petermanns Geographische Mitteilungen 59: 74-76.
- Croatian Bureau of Statistics: Census of Population, Households and Dwellings 2011. Available at: [http://www.dzs.hr/Hrv_Eng/publication/2011/SI-1441.pdf]
- Croatian Bureau of Statistics: Foreign Trade in Goods of the Republic of Croatia, January – December 2012. Available at: [http://www.dzs.hr/Hrv_Eng/publication/2012/04-02-01_12_2012.htm]
- Gabaix, X. (2009) Power Laws in Economics and Finance, Annual Review of Economics, Annual Reviews, Vol. 1(1), 255-294.
- Hinloopen J. and van Marrewijk C. (2008) Comparative advantage, the rank-size rule, and Zipf's law. Working Paper, Tinbergen Institute.
- Ioannides, Y. M. and Overman H. G. (2000) Zipf's law for cities: An empirical examination, Regional Science and Urban Economics, Vol 33, (2), 127-137.
- Newman, M. E. J. (2004) Power laws, Pareto distributions and Zipf's law, Contemporary Physics, Vol. 46, (5), 323-351.
- Soo, Kwok Tong (2005) Zipf's law for cities: a cross country investigation, Regional Science and Urban Economics, Vol. 35 (3), 239-263.
- World Bank database: World Development Indicators (WDI) 2012. Available at [<http://data.worldbank.org/indicator>].
- Zipf, G. K. (1949) Human Behaviour and the Principle of Least Effort, Reading, MA, Addison-Wesley.