

**T.R.
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS
PhD PROGRAMME IN ECONOMICS**

PhD DISSERTATION



**THREE ESSAYS ON THE FIRM SIZE DISTRIBUTION
IN TURKEY**

**YIGIT AYDOGAN
13729010**

**THESIS ADVISOR
PROF. DR. MURAT DONDURAN**

**ISTANBUL
2019**

**T.R.
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS
PhD PROGRAMME IN ECONOMICS**

PhD DISSERTATION



**THREE ESSAYS ON THE FIRM SIZE DISTRIBUTION
IN TURKEY**

**YIGIT AYDOGAN
13729010**

**THESIS ADVISOR
PROF. DR. MURAT DONDURAN**

**ISTANBUL
2019**

T.R.
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS
PhD PROGRAMME IN ECONOMICS

PhD DISSERTATION

**THREE ESSAYS ON THE FIRM SIZE DISTRIBUTION
IN TURKEY**

YIGIT AYDOGAN
13729010

Submission Date of Thesis to the Institution: 14/01/2019

Thesis Defense Date: 18/02/2019

Thesis is Found Successful with Unanimity.

	Title Name Surname
Thesis Advisor	: Prof. Dr. Murat DONDURAN
Jury Members	: Prof. Dr. Ercan EREN
	Prof. Dr. Suut DOĞRUEL
	Prof. Dr. Cem BAŞLEVENT
	Dr. Öğr. Üy. Tolga AKSOY

Signature



ISTANBUL
FEBRUARY 2019

ABSTRACT

THREE ESSAYS ON THE FIRM SIZE DISTRIBUTION IN TURKEY

Yigit Aydogan

February, 2019

This study is a combination of three independent articles regarding the firm size distribution in Turkey for the period between 2005 and 2016. The first one looks for the validity of the Gibrat's Law, i.e. the Law of the Proportionate Effect, in Turkey. A random parameters estimation model is considered to make use of the heterogeneity of individual firms. Unlike the somehow weak results of the rejecting studies from the literature, with the present method and a complete dataset, Gibrat's Law is strongly rejected. The second article surveys the evolution of the firm size distribution in Turkey for the subjected period through detailed data decomposition and density estimations. The analysis exhibits permanent effects of the Global Financial Crisis to Turkish businesses as derailing the growth trend of oldest firms. It further presents the lack of selection and resistance against adaptation among firm groups in the economy. The final article compares the fit of Pareto and lognormal distributions to the firm size distribution. It is concluded that Turkish firm sizes are not Pareto distributed. While extremely divergence is observed in 2008 and 2011, lognormal distribution well approximates the tail for the years 2006, 2007, 2010, 2015 and 2016.

Key Words: Firm size distribution, Gibrat's Law, selection, power law, lognormal distribution.

ÖZ

TÜRKİYE'DE FİRMALARIN BÜYÜKLÜK DAĞILIMI ÜZERİNE ÜÇ DENEME

Yiğit Aydoğın
Şubat, 2019

Bu çalışma Türkiye'de 2005 ve 2016 yılları arasında bulunan firmaların büyüklük dağılımları ile ilgili üç bağımsız makaleden oluşmaktadır. İlk makalede Gibrat Yasası'nın, bir diğer adıyla Orantılı Etki Yasası'nın, Türkiye'de geçerliliği sorgulanmıştır. Firmaların heterojenliğinden faydalanabilmek için veriye rassal parametre tahmin yöntemi tercih edilmiştir. Literatürdeki kısmen zayıf ya da karma sonuçlara karşın mevcut yöntem eldeki tam veri setine uygulandığında Gibrat Yasası güçlü bir şekilde reddedilmiştir. İkinci makalede Türkiye'deki firmaların büyüklük dağılımının belirtilen dönem içerisindeki evrimi verinin farklı katmanları göz önüne alınarak ve yoğunluk grafikleri kullanılarak incelemektedir. Analizde Küresel Finansal Krizin en Türk firmalarında bıraktığı kalıcı etkiler ve büyüme trendlerinin değişimi gösterilmiştir. Ayrıca dönemsel olarak firma gruplarına karşı seleksiyonun eksikliği ve bazı grupların adaptasyon sorunlarına işaret edebilecek bulgular elde edilmiştir. Son makalede ise firma büyüklük dağılımının Pareto ya da lognormal dağılım ile yakınsanabilirliği incelenmiştir. Sonuç olarak Türkiye'de firmaların Pareto dağılmadıkları gösterilmiştir. 2008 ve 2011 yıllarında dağılımda gözlenen aşırı aykırılık yanında, firma büyüklük dağılımının 2006, 2007, 2010, 2015 ve 2016 yıllarında en azından kuyrukta lognormal dağılım ile iyi derecede yakınsanabildiği ortaya konmuştur.

Anahtar Kelimeler: Firma büyüklük dağılımı, Gibrat Yasası, seleksiyon, güç yasası, lognormal dağılım.

PREFACE

The subjected dissertation is resulted from a strong effort despite all setbacks. The majority of the gratitude belongs to my advisor Prof. Dr. Murat Donduran who has been always extremely supportive and motivating. I would like to thank Prof. Dr. Ercan Eren and Prof. Dr. Suut Dođruel, respected Jury members for nearly four years, who participated in the discussions about the work sincerely, supported during downturns and shown great guidance. All my professors' high trust in me kept this study going in problematic times. Finally, I would like to thank for the patience of my family and especially Ezgi, who shown to be not only a caring partner but also supportive colleague.

Istanbul; February, 2019

Yigit Aydogan

CONTENTS

ABSTRACT	iii
ÖZ	iv
PREFACE	v
CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
1. TESTING GIBRAT’S LAW WITH TURKISH FIRM DATA	1
1.1. Introduction.....	1
1.2. Literature Review.....	2
1.3. Data	4
1.4. Methodology	7
1.4.1. Correlation Matrix.....	7
1.4.2. Growth Plots	8
1.4.3. Mixed Effects Estimation.....	9
1.4.3.1. Random Intercept Estimation.....	9
1.4.3.2. Random Intercept and Random Slope Estimation	11
1.5. Conclusions	13
1.6. Bibliography.....	14
2. EVOLUTION OF THE FIRM SIZE DISTRIBUTION IN TURKEY	17
2.1. Theorizing the Selection in Economics	17
2.2. Data	19
2.3. Evolution of the Firm Dynamics.....	22
2.3.1. Survey of the Firm Dynamics	22
2.3.2. Density Estimations	27
2.3.2.1. Distribution of the Entrants.....	27
2.3.2.2. Distribution of the Closing Firms	28
2.4. Age and Size Distribution of Firms	30
2.4.1. Survey of the Age Groups.....	31
2.4.2. Density Estimations	35
2.5. Conclusions	36

2.6. Bibliography.....	39
3. ESTIMATIONS ON THE FIRM SIZE DISTRIBUTION IN TURKEY	41
3.1. Introduction.....	41
3.2. Literature Review.....	42
3.3. Data	45
3.4. Definitions.....	46
3.5. Fitting to the Firm Size Distribution.....	48
3.5.1. OLS Estimation of the Pareto Exponent.....	49
3.5.2. Power Law and Lognormal Distribution Fitting Using MLE.....	50
3.5.3. Comparing the Fits.....	52
3.6. Conclusions.....	57
3.7. Bibliography.....	59
APPENDIX.....	62
Appendix to Chapter 2.....	62
Appendix to Chapter 3.....	66
RESUME.....	72

LIST OF TABLES

Table 1.1: Selected Literature of Verifying the Gibrat's Law	2
Table 1.2: Selected Literature of Rejecting the Gibrat's Law	3
Table 1.3: Selected Literature Providing Mixed Evidence on Gibrat's Law	4
Table 1.4: Summary of Employment Data for Surviving Firms, 2005-2016	5
Table 1.5: Summary of Growth Data for Surviving Firms, 2005-2016.....	6
Table 1.6: Correlation Matrix of the Variables	7
Table 1.7: Random Intercept Model Output	11
Table 1.8: Random Intercept and Random Slope Model Output.....	12
Table 2.1: Summary Statistics of Turkish Firms, 2005-2016	20
Table 2.2: Shares of Selected Sections in Firm Count, 2005-2016	22
Table 2.3: Size Statistics of Entering Firms, 2005-2016	23
Table 2.4: Size Statistics of Closing Firms, 2005-2016.....	24
Table 2.5: Age Statistics of Firms by Year, 2005-2016.....	25
Table 2.6: Age Statistics of Closing Firms, 2005-2016	25
Table 2.7: Size Statistics of Age Groups, Selected Years.....	34
Table 3.1: Summary Statistics of Firm Employment Data in Turkey.....	46
Table 3.2: OLS Estimations of the Power Law Exponent	50
Table 3.3: Power Law and Lognormal Fits to the Data	51
Table 3.4: Assessments and Comparisons of the Fits	55
Table 3.5: Algorithm to Compute p-value (Clauset et al., 2009).....	56
Table A.1: Sectoral Composition Between 2005-2008.....	63
Table A.2: Sectoral Composition Between 2009-2012.....	64
Table A.3: Sectoral Composition Between 2013-2016.....	65

LIST OF FIGURES

Figure 1.1: Growth Plots for Years 2006-2016, and All Years Combined.....	9
Figure 2.1: Size Distribution of All Firms, 2005-2016.....	21
Figure 2.2: Age and Firm Dynamics, 2005-2016	26
Figure 2.3: Firm Size and Dynamics, 2005-2016	28
Figure 2.4: Size Distribution of the Entrants, 2005-2016	29
Figure 2.5: Size Distribution of the Closing Firms, 2005-2016	30
Figure 2.6: Plot of the Age Groups	31
Figure 2.7: Tracking the Age Groups, 2005-2016.....	32
Figure 2.8: Size Distribution of the Age Groups, Selected Years	36
Figure 3.1: Histogram of Averaged log-firm Sizes Between 2005-2016 and Fit.....	49
Figure 3.2: Power Law (Solid) and Lognormal (Dashed) Fits Between 2005-2010	53
Figure 3.3: Power Law (Solid) and Lognormal (Dashed) Fits Between 2011-2016	54
Figure 3.4: Bootstrap Output for the Year 2009	57
Figure A.1: Size Distribution of All Entrants, 2005, 2009, 2016	62
Figure B.1: All Histograms of the log-firm Sizes and Fits-1	66
Figure B.2: All Histograms of the log-firm Sizes and Fits-2.....	67
Figure B.3: Bootstrap Results of Powerlaw, 2012.....	68
Figure B.4: Bootstrap Results of Powerlaw, 2013.....	69
Figure B.5: Bootstrap Results of Powerlaw, 2014.....	70
Figure B.6: Bootstrap Results of Powerlaw, 2016.....	71

1. TESTING GIBRAT'S LAW WITH TURKISH FIRM DATA

1.1. Introduction

Gibrat's Law is a simple but strong assumption that firm growth is a purely random process (Gibrat, 1957). Also referred as the Law of the Proportionate Effect, or Proportionate Growth Theory, i.e. the change of a firm size is proportional to its size, this widely examined phenomenon builded upon constant elasticity cost or demand functions. The main importance of the validity of the Law is as follows. Gibrat's Law being verified suggests that firm size distribution is non-stationary resulting in the lognormal size distribution. On the other hand, a rejection of the Law proposes a stationary process of firm sizes to follow, yielding a power law -or Pareto distribution (Luttmer, 2011).

Gibrat's Law, or the Law of the Proportionate Effect is formally stated as follows:

$$x_t - x_{t-1} = \epsilon_t x_{t-1} \quad (1.1)$$

where x is a firm's size at time t , and ϵ_t is a random parameter denoting a proportionate rate. For incremental time periods, logarithm of the 1.1 yields

$$\log x_t \cong \log x_0 + \epsilon_1 + \epsilon_2 + \dots + \epsilon_t \quad (1.2)$$

While the random term assumed to be independently distributed, randomized effects dominate the value in the limit. So, the distribution of $\log x_t$ becomes well approximated by normal distribution (Sutton, 1997).

As the Occam's razor principle leads, scientists traditionally treat the simplest tool as the superior if there is no evidence against it. However, the power of rejection or validation seems to be caught in the middle lacking a common acceptance. This could be overcome with a complete dataset and conclusively distant -or close findings to the Law.

When looking for a conclusion about the Gibrat's Law for the firm growth, the approaches in literature gather around dividing already limited samples of firms into

subgroups and measuring the randomness of the growth. It could be a subgroup regarding industry, or age, it holds for some, while may not hold for others. On the other hand, Hall (1987) posits that a balanced panel data of firms is a must for this quest which is somehow overlooked in some papers. While avoiding the search for the perfect partitioning of firms, in this study, an exhaustive data based testing of the Gibrat's Law is aimed with a balanced panel of nearly a million Turkish firms for twelve years by utilizing a newly used effective innovative addition to the methodology.

1.2. Literature Review

There has been an outstanding effort to validate or reject the Gibrat's Law, during the last three decades. Available papers depict a similar structure, but diverse procedures. It is not an easy job to do a summary of the literature, but it is easier to point at some work and tabulate the majority of efforts as under three groups, namely those who find the Law of the Proportionate Effect plausible, those who reject and those who find evidence for both cases.

Table 1.1: Selected Literature of Verifying the Gibrat's Law

Study	Region	Time span	Sample size
Hart and Prais (1956)	UK	2-10	60-2,103
Simon and Bonini (1958)	UK&USA	2	500
Singh and Whittington (1975)	UK	13	2,000
Farinas and Moreno (2000)	Spain	6	1,971
Geroski et al. (2003)	UK	31	147
Audretsch et al. (2004)	Netherlands	3	1,170
Fujiwara et al. (2004)	Europe	10	260,000
Lensink et al. (2005)	Netherlands	5	1,097
Choi (2010)	USA	10	823
Leitao et al. (2010)	Portugal	7	39

In terms of data utilization, although divided into 45 European countries, Fujiwara et al. (2004) gather input from 260,000 firms for 1992-2001 can be considered as the largest set ever on the validity of the Gibrat's Law. Despite the coverage, it has a considerable bias in terms of size. It was reported that firms in the dataset has at least 100 employees, or 20 million Euros in assets, or 10 million Euros in operating revenue,

varying among the countries. They confirm the Gibrat's Law, while the representativeness of the sample issues remains unsolved. The bias in the sample could be more severe because the selection criteria are based on "at least one of the size measures". This procedure means that firms with lots of employees with very little revenue, or assets are included, while also small enterprises with high revenue or assets, too. It should put a serious unnatural disturbance in terms of efficiency, therefore, growth distribution of those firms.

Table 1.2: Selected Literature of Rejecting the Gibrat's Law

Study	Area	Time span	Sample size
Ijiri and Simon (1964)	simulated	1,000	247
Kumar (1985)	UK	17	2,000
Evans (1987)	USA	7	21,218
Hall (1987)	USA	8	1,349
Hart and Oulton (1999)	UK	5	29,230
Hart and Oulton (2001)	UK	10	8,103
Cefis et al. (2002)	USA&Europe	12	199
Elston (2002)	Germany	5	341
Pagano and Schivardi (2003)	Europe	5	136
Harris and Trainor (2005)	UK	26	26 ind.
Oliveira and Fortunato (2006)	Portugal	12	7,653
Coad et al. (2008)	France	9	8,496
Falk (2008)	Europe	5	20,000
Hoxha (2008)	Kosovo	6	289
Petrunia (2008)	Canada	13	2,029
Lotti et al. (2009)	Italy	8	3,285
Mukhopadhyay and AmirKhalkhali (2010)	USA	8	500
Piergiovanni (2010)	Italy	11	41,859
Daunfeldt et al. (2012)	Sweden	7	18,141
Daunfeldt and Elert (2013)	Sweden	7	91,691

Daunfeldt and Elert (2013) claim employing the data of all limited liability firms in Sweden of which entered to the registers at any time during the period 1998-2004, being 288,747 firms in total. However, it is unclear from the article that how many of them are used in analyses. By looking at the regression output, one can calculate that 91,691 is the actual surviving firm count for the period. In the end, Gibrat's Law was rejected when the aggregated sample was considered but confirmed for half of the

industries. Further comments put on sectoral based validity of the Gibrat’s Law, while discussing of the term’s acceptability as a true Law.

Aslan (2008) inspected Turkish firms to find evidence on Gibrat’s Law. The data is drawn from ICI-500 (largest firms in the economy) and examined with panel unit root tests. The methodology has validity issues considering a panel with high correlation between units, i.e. firms. It was concluded on this issue of collinearity and rejected Gibrat’s Law except for ceramics and electronics industries. However, it should be noted that dealing with such a small dataset significantly lowers the confidence on the output. In this case mentioned industries have 4 and 3 firms, respectively.

Table 1.3: Selected Literature Providing Mixed Evidence on Gibrat’s Law

Study	Area	Time span	Sample size
Mansfield (1962)	USA	42	hundreds
Dunne and Hughes (1994)	UK	5-5	1,172-1,696
Lotti et al. (2001)	Italy	7	428
Calvo (2004)	Spain	12	751
Ganugi et al. (2004)	Italy	3	17,709
Kostov et al. (2005)	Ireland	7	1,290
Aslan (2008)	Turkey	20	103
Serrasqueiro et al. (2010)	Portugal	7	350
Teruel-Carrizosa (2010)	Spain	9	139,922
Park and Sydnor (2011)	USA	12	5,818
Tang (2015)	Sweden	15	2,185

Finally, Tang (2015) runs a random parameter model on a data of 2,185 Swedish energy firms for 15 years similar to the current study. It was concluded that the Law of the Proportionate Effect is rejected for the whole sample. However, since overall the rejection is by a coefficient close to the acceptance range, further analysis conducted on individual firms. It is concluded that despite the overall rejection, 70% of the Swedish energy firms could still be considered as obeying the random growth model.

1.3. Data

To study Gibrat’s Law, there needs to be an unbiased sample of firms for subjected group. Although there are various size measurements employed in the firm dynamics

literature such as sales (Cefis et al., 2002), assets (Serrasqueiro et al., 2010), revenue (Tang, 2015), output and value added (Harris and Trainor, 2005) it can be said that employment is the generic size measure of firms starting as early as Gibrat (1957). Numerous studies concerning multiple size definitions report the results being similar to each other (e.g. Axtell (2001), Tang (2015), Daunfeldt and Elert (2013)). The availability and accuracy of employment data makes it indispensable for a comprehensive study. TURKSTAT holds a large dataset regarding Turkish firms of which contributes to the GDP for a given year, called Annual Business Registers Framework between 2005-2016. There are 3 to 3.5 million firms in the dataset each year and tracking on their unique IDs, a balanced panel of 953,177 firms was gathered to test Gibrat's Law.

Table 1.4 provides descriptive statistics of the surviving firms from 2005 to 2016 in Turkish economy. Mean firm size starts from 5 employees per firm in 2005, increasing 70% to 8.49 in 2016. Increase in employees continued until 2008, then two years of shrinking occurred in 2009 and 2010. Another positive momentum was observed between 2011-2014, followed by a negative but slower decrease in size until 2016. Nevertheless, the trackable surviving firms in Turkey had enjoyed a substantial growth over the period in average terms.

Table 1.4: Summary of Employment Data for Surviving Firms, 2005-2016

Year	Obs.	Mean ^a	Std. ^b	b/a	Variance	Skewness	Kurtosis
2005	953,177	5.0	90	18	8,109	247	87,375
2006	953,177	5.1	96	19	9,293	351	181,101
2007	953,177	5.4	88	16	7,784	254	98,621
2008	953,177	6.9	186	27	34,638	217	59,767
2009	953,177	6.8	141	21	19,868	448	291,559
2010	953,177	6.5	94	14	8,759	176	45,826
2011	953,177	7.0	101	14	10,099	158	35,997
2012	953,177	7.8	102	13	10,312	146	31,735
2013	953,177	8.5	120	14	14,288	197	66,548
2014	953,177	8.9	132	15	17,497	192	66,595
2015	953,177	8.7	116	13	13,419	115	19,242
2016	953,177	8.5	115	14	13,181	121	21,674

Source: TURKSTAT

Increasing standard deviation of the mean firm size in Turkey may indicate a peak, according to the Table 1.4. It is observed that increasing mean size is accompanied by an increasing variance, except for 2007, but decreasing mean always lowers the variance for that year. One can conclude on this movement as Turkish firms benefit

from growing employee count sparsely, yet collectively suffer from negative conditions towards a more common size distribution.

Looking at the detailed statistics to build upon, it is evident that the firm size has become more symmetric in the subjected twelve years. On the way to symmetry, there are sharp disturbances in 2006, 2009 and 2013 in terms of increased skewness. Kurtosis also gets highly affected those three years making the distribution have more concentrated peaks or making the middle firm sizes closer. After so-called shocks, firms get a lower concentration around the center highlighting the previous benefit-suffer remark.

Table 1.5: Summary of Growth Data for Surviving Firms, 2005-2016

Year	Obs.	Mean ^a	Std. ^b	b/a	Variance	Skewness	Kurtosis
2006	953,177	0.01	0.3	34	0.11	1.4	37
2007	953,177	0.06	0.2	4	0.05	2.7	27
2008	953,174	0.08	0.5	6	0.20	2.2	24
2009	953,174	0.35	0.5	1	0.25	-2.0	16
2010	952,889	-0.36	0.5	-1	0.27	1.6	12
2011	952,881	0.05	0.3	6	0.10	1.8	32
2012	952,903	0.15	0.4	3	0.17	0.5	13
2013	950,451	0.03	0.4	12	0.12	0.9	25
2014	948,557	0.01	0.4	41	0.17	-0.5	22
2015	950,076	0.03	0.4	12	0.12	-1.2	25
2016	943,919	0.01	0.3	32	0.10	-0.3	26

Table 1.5 summarizes the growth rate statistics of tracked 953,177 firms over the period 2006-2016. Growth for 2005 is not available due to the obvious reason that the data starts at the year. A fraction of surviving firms does not have any registered employees for some years which causes the observation count slightly differ in time. Mean growth rate of firms unsurprisingly close to zero in non-disturbance periods, while significantly closer than pre-crisis in recent years. Also variance of the growth rates has a stable nature despite increasing overall firm sizes during the period as presented in Table 1.4, excluding 2008 to 2010. During the Global Financial Crisis, firm growth rates in Turkey changed drastically. There is an extraordinary high mean growth rate in 2009, yet another extraordinary high shrinking rate in the following year of 2010.

A positively skewed growth rate distribution means that majority of the firms grow higher than the mean growth, which is the case except years 2009, 2014, 2015, 2016. In those years, majority of the firms could not grow above economywide average in employee count. Despite the negative skewness in recent years, distribution of the growth rates occurred more symmetrically. Kurtosis of the distribution of growth rates remained relatively stable since 2013, like 2007 and 2008. Yet, an unexpected positive shock in 2009 and negative in 2010 effected more uniformly across the firms. This uniformity also backs the prior finding about the dispersion of the benefits and sufferings from turbulent years in Turkish economy.

1.4. Methodology

As mentioned before, there are plenty of models used to validate or reject Gibrat's Law of which several employed here. In this section, conclusive evidence about the validity of the Gibrat's Law in Turkey can be found building upon a wide range of applications from as simple as investigating correlation matrix and data plots, to random and mixed effects estimations.

1.4.1. Correlation Matrix

Building a correlation matrix is important before modeling relationships. Table 1.6 presents the correlations between the variables considered in this study and their lagged values. log size is log of employment indicating the size, year is a time variable to capture time-varying heteroskedasticity of firms' growth.

Table 1.6: Correlation Matrix of the Variables

	growth	log size	log size _{t-1}	growth _{t-1}	year
growth	1				
log size	0.2338	1			
log size _{t-1}	-0.1745	0.9166	1		
growth _{t-1}	-0.2872	0.1248	0.2444	1	
year	-0.0697	0.0694	0.0989	-0.0413	1

Correlations above $|0:1|$ in Table 1.6 indicates probable relationships between variables. While year can be only slightly correlated with the last year's size, all selected variables seem to be correlated to each other. On the other hand log size is highly correlated with its lagged value, with a coefficient 0.91, which needs to get confirmed to be exactly equal to unity in order to validate the Gibrat's Law.

It is promising to conduct a research to find out to what extent of these correlations gets confirmed. Namely, growth of a firm in a given period is negatively correlated with its last year's growth by a coefficient 0.28, last year's size by 0.17, positively correlated with current firm size 0.23.

1.4.2. Growth Plots

Plottings of each year's growth rates against the size distribution is an important visual benchmark. It could show a relation between size and growth distributions, so has been widely used in the literature. Figure 1.1 displays the yearly growth plots for years 2006-2016, and in the southeast panel all data plotted in a single graph.

To validate Gibrat's Law, growth rates should be scattered around without a visible shape. Figure 1.1 pictures significant tent shape for each year's growth plot against log transformed size variable. General conclusion would be that it shows a negative impact of size on growth for firms, undermining the Gibrat's Law. Although the observed tent shaped plot of growth rates against size, the lower body of the plot is very thick. This keeps hopes high for econometric analysis. On the other hand, the shape is changing year-over-year, allows further investigation.

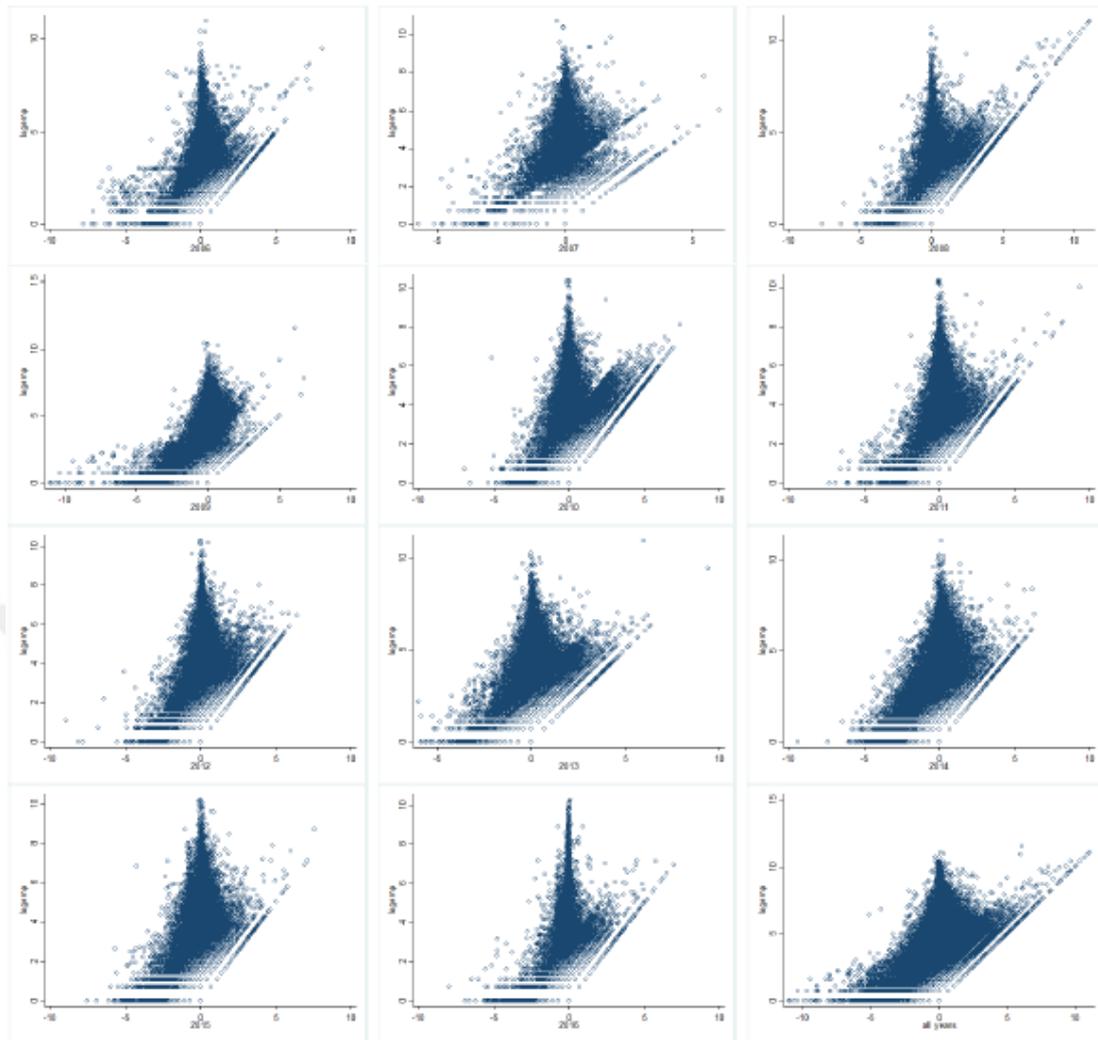


Figure 1.1: Growth Plots for Years 2006-2016, and All Years Combined

It is evident that 2007 has the least size-related growth distribution in subjected time period. Taken together with 2013 plot, they have a wide body relative to 2010 and 2016. So, there is another support for the extraction that Turkish firms suffers from bad times in similar manner, while good economic environment benefits actors differently.

1.4.3. Mixed Effects Estimation

1.4.3.1. Random Intercept Estimation

There are several main pillars of model selection with panel data. Considering current exhaustingly large dataset with not much of a time length, the fixed and random effect models can be candidates. The practicality of these models dealing with heterogeneity

is a bonus. When specifying the fixed effects model, all time-invariant heterogeneity between entities are accounted for in order to single out the variance within unit. By this notion it is impossible to detect timeinvariant answers because the model incorporates all time-independent characteristics of the entities (Kohler and Kreuter, 2012, p.335). So, fixed effects modeling is ruled out for testing the Gibrat's Law over a time period. Then mixed effects estimation is taken into consideration. It is largely put forward that Random Effects (and Mixed Effects) modeling is to be preferred due to its high flexibility and saves the researcher from the blockage on extracted information caused from Fixed Effects due to the extended controlling variables (Bell and Jones, 2015). This method should also allow to have the heterogeneity of growth distributions over time as seen in Figure 1.1

If the unobserved individual heterogeneity is uncorrelated with the elements in the model, random effects could be used. If we assume that unobserved firm specific heterogeneity is uncorrelated with the parameters in the model, which are the last period's size and time vector, we may estimate the random effects with Maximum Likelihood. General form of the equation is:

$$Y_{it} = \beta_1 X_{it} + \alpha + \gamma_i + \epsilon_{it} \quad (1.3)$$

where α is the time-invariant intercept, $\gamma_i \sim N(0, \sigma_\gamma)$ is error between entities, and $\epsilon_{it} \sim N(0, \sigma_\epsilon)$ is error within an entity over time (Greene, 2008, p. 183).

To incorporate into current framework, one can use:

$$\log size_{it} = \beta_0 + \beta_1 \log size_{it-1} + \beta_2 time_t + u_{it} \quad (1.4)$$

where,

$$u_{it} = \gamma_{1i} + \epsilon_{it} \quad (1.5)$$

Here, $time_t$ is a vector of fixed effects capturing time-dependent heterogeneity in firm growth, $\gamma_{1i} \sim N(0, \sigma_1)$ is time-invariant random effects parameter which accounts for the variation not captured by β_0 , and $\epsilon_{it} \sim N(0, \sigma_\epsilon)$.

In this model, $\log size_{it-1}$ and $time_t$ are assumed to be exogenous, bringing on $E(\gamma_{1i} | \log size_{it-1}, time_t) = 0$ and $E(\epsilon_{it} | \log size_{it-1}, time_t, \gamma_{1i}) = 0$. The model also can be stated as:

$$\log size_{it} = (\beta_0 + \gamma_{1i}) + \beta_1 \log size_{it-1} + \beta_2 time_t + \epsilon_{it} \quad (1.6)$$

or

$$\log size_{it} = \beta_{0i} + \beta_1 \log size_{it-1} + \beta_2 time_t + \epsilon_{it} \quad (1.7)$$

where $\beta_{0i} = \beta_0 + \gamma_{1i}$ incorporates the random intercept and leaves β_1 as the unbiased estimator of the effect of the firm size over firm growth. Gibrat's Law is to be confirmed if $\beta_1 = 1$, rejected otherwise.

Table 1.7: Random Intercept Model Output

Coef.	MLE	Std. Err.
logsize_{t-1}	0.9313**	0.0001
Year	-0.0038**	0.0001
Cons	7.7632**	0.0823
sd(_cons)	0.0001**	0
sd(residual)	0.4153**	0.0001
Wald p-val	[0.0001]	
LR p-val	[0.99]	

**p<0.001

Table 1.7 shows highly significant coefficients for all the parameters, leaving the Gibrat's law rejected with a β_1 estimation smaller than 1. But the random effect parameter γ_{1i} has a very small value. To improve variance capturing, an additional term of random slope might be added to the equation in the following section, allowing to benefit further from the unprecedented magnitude of the dataset.

1.4.3.2. Random Intercept and Random Slope Estimation

After incorporating the random effect on the unobserved individual heterogeneity, one can also include a random effect regarding the observed heterogeneity as a slope, namely firm size. There would be an interesting feature of the developed model, if the estimated coefficient of $\log size_{it}$ turns out close to 1: the Gibrat's Law could be tested for individual firms similar to Tang (2015) with Swedish energy firms. In any case, the new model would become:

$$\log size_{it} = \beta_0 + \beta_1 \log size_{it-1} + \beta_2 time_t + v_{it} \quad (1.8)$$

where,

$$v_{it} = \gamma_{1i} + \gamma_{2i} \log size_{it-1} + \epsilon_{it} \quad (1.9)$$

where $\gamma_{1i} \sim N(0, \sigma_1)$ is again the random part missed in β_0 , $\gamma_{2i} \sim N(0, \sigma_2)$ is the random slope on the last period's firm size, and $\epsilon_{it} \sim N(0, \sigma_\epsilon)$ is the error of the model. Rearranging yields:

$$\log size_{it} = \beta_0 + \beta_1 \log size_{it-1} + \beta_2 time_t + \gamma_{1i} + \gamma_{2i} \log size_{it-1} + \epsilon_{it} \quad (1.10)$$

Assuming the terms $\log size_{it-1}$ and $time_t$ exogenous, $E(\gamma_{1i} | \log size_{it-1}, time_t) = 0$, $E(\gamma_{2i} | \log size_{it-1}, time_t) = 0$ and $E(\epsilon_{it} | \log size_{it-1}, time_t, \gamma_{1i}, \gamma_{2i}) = 0$. Thus a more general form can be specified as:

$$\log size_{it} = (\beta_0 + \gamma_{1i}) + (\beta_1 + \gamma_{2i}) \log size_{it-1} + \beta_2 time_t + \epsilon_{it} \quad (1.11)$$

This model was estimated using Maximum Likelihood method. Table 1.8 presents the estimated parameters of the model incorporating both random intercept and random slope.

Table 1.8: Random Intercept and Random Slope Model Output

Coef.	MLE	Std. Err.
log size_{t-1}	0.6093**	0.0004
year	0.0001*	0
cons	-0.0697	0.0781
sd(log size_{t-1})	0.2711**	0.0004
sd(_cons)	0.0868**	0.0005
sd(residual)	0.3786**	0.0001
Wald p-val	[0.0000]	
LR p-val	[0.0001]	

**p<0.001, *p<0.01

According to the estimates on Table 1.8, slope of the last period's employment is far lower from 1. There are only a handful of firms of which Gibrat's Law may hold individually, if any. Adding the firm specific random slope one can assert that Gibrat's Law is strongly rejected in Turkey.

1.5. Conclusions

To the best of our knowledge, the most comprehensive dataset in the literature to test the validity of Gibrat's Law is employed in the present study. The inspection of correlation matrix and growth plots has shown early warnings about strong relation between growth and size of firms in Turkey. Growth distribution over the years yielded an unexpected claim for turbulent times. It could be further examined that Turkish firms gets affected from economic downturn collectively, while benefits from booms distributed sparsely.

Without random effect parameter on the size from previous year, econometric analysis provide evidence on Gibrat's Law as almost holds with a coefficient close to one. However, when all surviving firms' data incorporated to the equation with firm specific variance component, coefficient significantly drops. Since it is a general benchmark for firm dynamics, this result puts importance on reconsidering the validity with complete data for developed countries.

In this study, it is shown that when a complete dataset considered, Gibrat's Law is rejected by a certain figure. In other words, Turkish firms do not grow in proportion to their size, i.e.: growth is not independent of size. Economic environment in Turkey during the period of 2005-2016 seems to support smaller firms to grow faster, with a considerably high random effect parameter. Further research opportunities provided with a focus on heterogeneity of macroeconomic effects on individual firms offer prospect, as seen above, shall be fruitful in terms of combining the macro and micro theory of the firm.

1.6. Bibliography

- Aslan, Alper. 2008. Testing gibrat's law: Empirical evidence from panel unit root tests of turkish firms. **RePeC**. [01.12.2018]
<http://mpira.ub.uni-muenchen.de/10594/>.
- Audretsch, David, Klomp, Ludwig, Santarelli, Enrico. and Thurik, Roy. 2004. Gibrat's law: Are the services different? **Review of Industrial Organization**, v. 24. i. 3: 301–324.
- Axtell, Robert. 2001. Zipf distribution of u.s. firm sizes. **Science**. v. 293. i. 5536: 1818–1820.
- Bell, Andrew, Jones, Kelvyn. 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. **Political Science Research and Methods**. v. 3 i. 1: 133–153.
- Calvo, Jose. 2004. Testing gibrats law across regions. evidence from spain. The 44th ERSA Conference, 25 August 2005, Porto.
- Cefis, Elena, Ciccarelli, Matteo, Orsenigo, Luigi. 2002. From gibrat's legacy to gibrat's fallacy. A Bayesian Approach to Study the Growth of Firms, **Working papers at the Hyman Minsky Department of Economic Studies**. v. 19.
- Choi, Paul. 2010. The us property and liability insurance industry: Firm growth, size, and age. **Risk Management and Insurance Review**, v. 13. i. 2: 207– 224.
- Coad, Alex. 2008. Firm growth and scaling of growth rate variance in multiplant firms. **Economics Bulletin**. v. 12. i. 9: 1–15.
- Daunfeldt, Sven-Olov, Elert, Niklas. 2013. When is gibrat's law a law? **Small Business Economics**. V. 41. i. 1: 133–147.
- _____, Lang, Åsa. 2012. Does gibrat's law hold for retailing? evidence from sweden. **Journal of Retailing and Consumer Services**. v. 19. i. 5: 464–469.
- Dunne, Paul, Hughes, Alan. 1994. Age, size, growth and survival: Uk companies in the 1980s. **The Journal of Industrial Economics**, v. 42. i. 2: 115–140.
- Elston, Julie. 2002. An examination of the relationship between firm size, growth and liquidity in the Neuer Markt. **Volkswirtschaftliches Forschungszentrum der Deutschen Bundesbank**. Berlin
- Evans, David. 1987. Tests of alternative theories of firm growth. **Journal of political economy**, v. 95. i. 4: 657–674.
- Falk, Martin. 2008. Testing Gibrat's Law for European Multinational Enterprises. **FIW Research Reports**. Vienna.
- Farinas, Jose, Moreno, Laurdes. 2000. Firms' growth, size and age: A nonparametric approach. **Review of Industrial organization**, v. 17. i. 3: 249–265.
- Fujiwara, Yoshi, Di Guilmi, Corrado, Aoyama, Hideaki, Gallegati, Mauro, Souma, Wataru. 2004. Do pareto–zipf and gibrat laws hold true? an analysis with european firms. **Physica A: Statistical Mechanics and its Applications**. v. 335. i. 1: 197– 216.

- Ganugi, Piero, Grossi, Luigi, Crosato, Lisa. 2004. Firm size distributions and stochastic growth models: a comparison between ict and mechanical italian companies. **Statistical Methods and Applications**. v. 12. i. 3: 391–414.
- Geroski, Paul, Lazarova, Stepana, Urga, Giovanni, Walters, Teng. 2003. Are differences in firm size transitory or permanent? **Journal of Applied Econometrics**. v. 18. i. 1: 47–59.
- Gibrat, Robert. 1957. **On Economic Inequalities**. trans. E. Henderson. International Economic Papers, The Macmillan Company.
- Greene, William. 2008. **Econometric Analysis**. Pearson Education Inc. Prentice Hall.
- Hall, Bronwyn. 1987. The relationship between firm size and firm growth in the us manufacturing sector. **The Journal of Industrial Economics**. v. 35: 583–606.
- Harris, Richard, Trainor, Mary. 2005. Plant-level analysis using the ard: another look at gibrat's law. **Scottish Journal of Political Economy**. v. 52. i. 3: 492–518.
- Hart, Peter, Oulton, Nicholas. 1999. Gibrat, galton and job generation. **International Journal of the Economics of Business**. v. 6. i. 2: 149–164.
- _____, _____. 2001. Galtonian regression, company age and job generation 1986–95. **Scottish Journal of Political Economy**. v. 48. i. 1: 82–98.
- _____ and Prais, Sigbert. 1956. The analysis of business concentration: a statistical approach. **Journal of the Royal Statistical Society**. v. 119. i. 2: 150–191.
- Hoxha, Durim. 2008. Determinants of growth-an empirical analysis of new firms and fast growing firms in kosova. **Journal of Labor Economics**. v. 23: 81–114.
- Ijiri, Yuri, Simon, Herbert. 1964. Business firm growth and size. **The American Economic Review**. v. 54 i. 2: 77–89.
- Kohler, Ulrich, Kreuter, Frauke. 2012. **Data analysis using Stata**. Stata press.
- Kostov, Philip, Patton, Myles, Moss, Joan, McErlean, Seamus. 2005. Does Gibrat's Law Hold Amongst Dairy Farmers in Northern Ireland? **European Association of Agricultural Economists**. Ireland.
- Kumar, Madhavan. 1985. Growth, acquisition activity and firm size: evidence from the united kingdom. **The Journal of Industrial Economics**. v. 33. i. 3: 327–338.
- Leitao, Joao, Serrasqueiro, Zelia, Nunes, Paulo. 2010. Testing gibrat's law for listed portuguese companies: A quantile approach. **International Research Journal of Finance and Economics**. v. 37. i. 1: 147–158.
- Lensink, Robert, Van Steen, Paul, Sterken, Elmer. 2005. Uncertainty and growth of the firm. **Small Business Economics**. v. 24 i. 4: 381–391.
- Lotti, Francesca, Santarelli, Enrico, Vivarelli, Marco. 2009. Defending gibrat's law as a long-run regularity. **Small Business Economics**. v. 32. i. 1: 31–44.
- _____, _____, Vivarelli Marco. 2001. The relationship between size and growth: the case of italian newborn firms. **Applied Economics Letters**. v. 8. i. 7: 451–454.
- Luttmer, Erzo. 2011. On the mechanics of firm growth. **The Review of Economic Studies**. v. 78. i. 3: 1042–1068.

- Mansfield, Edwin. 1962. Entry, gibrat's law, innovation, and the growth of firms. *The American economic review*. v. 52. i. 5: 1023–1051.
- Mukhopadhyay, Arun, AmirKhalkhali, Sal. 2010. Profitability performance and firm size-growth relationship. *Journal of Business & Economics Research*. v. 8. i. 9: 121-126.
- Oliveira, Blandina, Fortunato, Adelino. 2006. Testing gibrat's law: empirical evidence from a panel of portuguese manufacturing firms. *International Journal of the Economics of Business*. v. 13. i. 1: 65–81.
- Pagano, Patrizio, Schivardi, Fabiano. 2003. Firm size distribution and growth. *Scandinavian Journal of Economics*. v. 105. i. 2: 255–274.
- Park, Kwangmin, Sydnor, Sandra. 2011. International and domestic growth rate patterns across firm size. *International Journal of Tourism Sciences*. v. 11. i. 3: 91– 107.
- Petrunia, Robert. 2008. Does gibrat's law hold? evidence from canadian retail and manufacturing firms. *Small Business Economics*. v. 30. i. 2: 201–214.
- Piergiovanni, Robert. 2010. Gibrat's law in the “third italy”: Firm growth in the veneto region. *Growth and Change*. v. 41. i. 1: 28–58.
- Serrasqueiro, Zelia, Nunes, Paulo, Leitão, Joao, Armada, Manuel. 2010. Are there non-linearities between sme growth and its determinants? a quantile approach. *Industrial and Corporate Change*. v. 19. i. 4: 1071–1108.
- Simon, Herbert, Bonini, Charles. 1958. The size distribution of business firms. *The American economic review*. v. 48. i. 4: 607–617.
- Singh, Ajit, Whittington, Geoffrey. 1975. The size and growth of firms. *The Review of Economic Studies*. v. 42. i. 1: 15–26.
- Sutton, John. 1997. Gibrat's legacy. *Journal of economic literature*. v. 35. i. 1: 40–59.
- Tang, Aili. 2015. Does gibrat's law hold for swedish energy firms? *Empirical Economics*. v. 49. i. 2: 659–674.
- Teruel-Carrizosa, Manuel. 2010. Gibrat's law and the learning process. *Small Business Economics*. v. 34. i. 4: 355–373.

2. EVOLUTION OF THE FIRM SIZE DISTRIBUTION IN TURKEY

2.1. Theorizing the Selection in Economics

Evolutionary studies on firm theory does not reach back in time. Among the pioneers, Jovanovic (1982) shows that firms learn about their individual abilities and evolve towards an efficient size. This evolution is said not from a scarcity of capital but a self discovery of firms about their talents. The findings come with the claim that smaller firms exit the industry, while efficient ones grow and survive. Supporting the claim of learning through the business process, Evans and Leighton (1989) further asserts that the probability of being an entrepreneur does not depend on age or experience on a job, and in seven years, almost 50% of undertakers switch back to paid work.

Later work by Audretsch and Mahmood (1994) and Audretsch and Mata (1995) propose an efficient size for firms to achieve in order to survive, while those who could not fail and exit. Combining results with claims from Lucas (1978) they all find it plausible that selection occurs, even if one might not already observe. Firm sizes agreed to be skew towards smaller ones but this is, as claimed, not due to stagnant survivors of smaller firms but continuing the entry of such size groups.

Cabral and Mata (2003) builds their model of industry evolution on the financial constraints that a firm face due to social status, or contacts, of the entrepreneur. In order to overcome a venture, either the owner needs to possess some wealth, or achieve equivalent financing opportunities. Besides, selection is assumed to play a little role.

On the other hand, Akcigit et al. (2016) assert that selection lacks in nondeveloped economies. It was observed that many small firms nevertheless survive in such economies which possess a problem in terms of theory, comparing the US and Indian firm data. They find that in developing economies people run businesses not always seeking growth and prosperity, but for subsistence. Therefore, they lack incentives to incorporate managerial talent from outside. This choice of not growing is said to be also originated from the difficulties of delegating managers which firm needs.

Whether static or dynamic, ideas on firm theory can be built and backed with econometric explanations. But when those ideas taken to comprehensive data exemplified by this study, one would see the complicatedness of the evolution of firm sizes might need simpler pillars of action sequences. Complex outcomes do not necessarily originate from equally complex procedures. Agent-Based Computational (ABC) modeling of a virtual economy consists of daily transactions depend not on the extreme-or even bounded rationality of entities, but the realistic approaches taken against actual problems bounded with the individual reality. As exemplified in Delli Gatti et al. (2011) not only the business entities, but the whole economic environment could be formed in 8 basic steps. These steps start with firm's decision on output and labor, vacancies get posted, job applications get done and contracts occur. When finances of firms lack resources to produce, they apply to credit market and after evaluation gets credit or not. If fails, workers get laid off, planned output drops. After production, goods market operates with viable distance constraints on consumers. After sales, there might be unsold goods remained for the firms and higher than desired prices for the consumers. Profiting firms pay their debts, invest in R&D, distribute dividends. With possible unpaid loans, banks get non-performing loans. With varying profits, firms might add to or lower their net worth. Firms and banks with negative net worth at the end of the period go bankrupt, new firms enter to the economy and a cycle is completed. While incorporating firm dynamics and industry evolution, this model also accounts for the inflation, business cycles and other macro phenomenon yielding the term bottom-up macroeconomics.

The extremely straightforward but flexible environment of ABC modeling relies on strong relations to the data, offering a prospect future in economic theory. Building up an ABC model does not demand extensively mathematical proofs, or vague equation systems, but simulated computerized environments for agents to interact. The validity of such models are dependent on accurate and detailed observations which goes beyond sorting and regressing limited samples. All of the actions depend on simple, practical and far from fully-rational decision making procedures. No matter how plausible to the mind- or to the math, each economic environment has its special dynamics which should be taken into considerations prior to modeling. If the aim is to capture the spirit of businesses, building concrete models must depend on not only a snapshot of an industry but an evolutionary mapping capturing the actual history of

majority of the subjected firms. In this study, Turkish business entities between 2005 and 2016 are exhaustively surveyed with a data oriented perspective. The unprecedented exhaustive dataset and sifting out highly useful characteristics of agents while providing benchmarks for computational modeling of Turkish economy are the key contributions of this work. The hope on the span of the data and diversity of the inferences lay a cornerstone for the next step of desired modeling.

In this paper, Turkish firms are examined with an evolutionary perspective. The raw data of nearly 3.5 million firms per year from 2005 to 2016 is handled by two dynamic and multi layered points of view in the following two sections. The first angle consists of entering and exiting firms to and from Turkish economy during the subjected period. Additionally, age statistics of the firms went out of business is incorporated to the analysis. For the second part, the firms in each period get divided into 6 age groups, i.e.: 0-1, 2-4, 5-9, 10-19, 20-29, 30 or more years in business. Both type of analysis are enriched by visualization, density estimations are plotted in a dynamic manner. In the last part, concluding remarks and inferences from the evolution of Turkish firms is made together with references to the contemporary literature.

2.2. Data

Regarding the Turkish firm statistics, TURKSTAT holds a very comprehensive dataset called “Annual Business Registers Framework” which consists all registered firms contributing to the GDP of Turkey. Table 2.1 possesses some statistics on approximately 3.5 million firms per year for the period from 2005 to 2016.

Table 2.1: Summary Statistics of Turkish Firms, 2005-2016

Year	Median	Mean	3rd-Q	Max	#Firms	#Employees
2005	1	3.1	1	40,393	3,455,700	10,678,113
2006	1	3.1	1	59,146	3,205,948	9,842,260
2007	1	3.4	1	43,555	3,175,493	10,796,676
2008	1	4.3	1	64,000	3,490,786	15,010,380
2009	2	4.1	2	101,386	3,227,238	13,070,314
2010	1	3.8	1	30,724	3,003,116	11,291,716
2011	1	4.1	2	83,700	3,422,163	14,065,090
2012	1	4.5	2	31,760	3,474,992	15,776,464
2013	1	5.2	3	70,679	3,529,541	18,459,499
2014	1	5.4	3	139,576	3,525,431	19,002,073
2015	1	5.4	3	27,852	3,584,832	19,322,244
2016	1	5.4	3	29,309	3,652,521	19,650,563

Source: TURKSTAT – Micro Data Collections

As seen in Table 2.1 mean firm size in Turkey significantly increased since 2005, by 74% from 3.09 to 5.38 employees. This increase has not been robust to shocks, however. The hit of Global Financial Crisis in 2009 created a 3-year setback in terms of mean size. More interestingly, median firm size reached to 2 in the crisis year meaning a more symmetrical firm size distribution than pre-crisis years. During 2011-2014, firm sizes increase steadily in average but stagnate afterwards. The third quantile in firm size distribution reached 3 from 1 in a decade, noting a general growth in most of the firms in the economy.

Figure 2.1 depicts firm size distribution in Turkey for the period 2005-2016. It is evident that the overall distribution remained relatively stable, except for the year 2009. There is a critical slope around the firms with 50 or more employees, where the density of the larger firms drop instantly. These statistics and visualization below emphasize that Turkish firm size distribution possesses an increasingly skew distribution as expected, with medians stay further from the mean as the mean rises. According to the literature, a versatile economic environment and efficient firms should carry the lower or at least medium parts of the distribution up. Lack of this situation needs to be diagnosed in order to infer a way out.

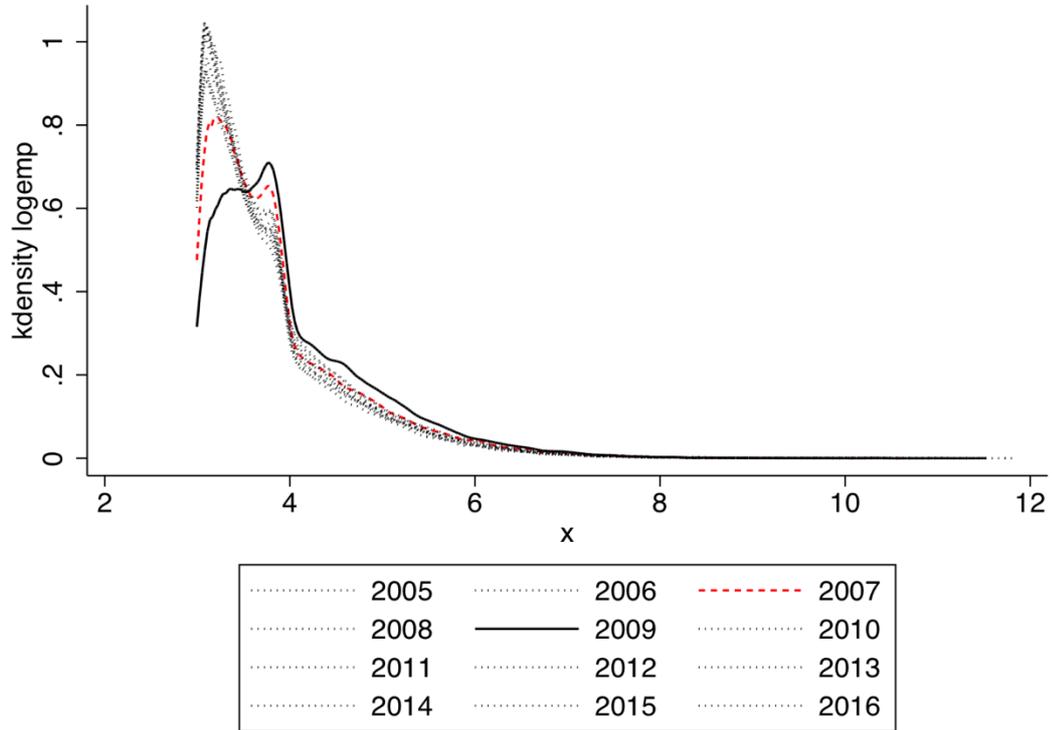


Figure 2.1: Size Distribution of All Firms, 2005-2016

Regarding the sectoral composition of the subjected data set, Table 2.2 summarizes the selected sections. TURKSTAT reports the Business Registers complying NACE Rev. 1.1 for the period of 2005-2008, and NACE Rev. 2 from 2009 to 2016. Detailed tables are to be presented in the Appendix A.

Three sections of Business Registers data are selected for demonstration purposes, which are Manufacturing as the back-bone of modern economies, Construction as a highly debated boosting element of 2000s Turkey, and Trade as the largest part . It is seen from Table 2.2 that Turkish private sector composition get negatively hit by the Global Financial Crisis in terms of Manufacturing firms, and Construction and Trade sections increased in percentage. Yet, Manufacturing firms regain their share after the crisis years. Interestingly enough, Construction firms increases in share throughout the period, except 2015, which could be subjected in a further study with a focus on sectoral evolution.

Table 2.2: Shares of Selected Sections in Firm Count, 2005-2016

Year	Manufacturing	Construction	Trade
2005	12.30%	4.20%	39.50%
2006	12.70%	4.60%	39.60%
2007	12.70%	4.80%	39.60%
2008	13.00%	5.00%	39.70%
2009	11.90%	5.30%	40.20%
2010	11.60%	5.30%	40.50%
2011	12.00%	6.10%	39.50%
2012	12.20%	6.10%	38.60%
2013	13.00%	7.20%	37.60%
2014	13.00%	7.60%	37.10%
2015	13.20%	6.50%	37.50%
2016	12.90%	7.70%	36.80%

2.3. Evolution of the Firm Dynamics

The primary forces behind the evolution of the industry are closing firms due to selection and new enterprises emerging with hopes to survive. Entry and exit conditions of firms might possess important information regarding the core functions of the economy. In the economic literature, both entering and closing firms are considered to be small with respect to the overall distribution. A disturbance on this status defines an important characteristic about the subjected economy.

2.3.1. Survey of the Firm Dynamics

The strength of an economy comes from its resilience to shocks, and the key factor to overcome destructive time periods is the capability of keeping the influx of new venturers to the economic environment. This nature of the economy can be observed through the statistics of the entrants. Table 2.3 provides the business units that got enlisted by the registers during the period 2005-2016.

Table 2.3: Size Statistics of Entering Firms, 2005-2016

Year	Obs.	Mean^a	Std.^b	b/a	Variance	Skewness
2005	565,364	2.2	49.2	22	2421	595
2006	465,124	1.8	23.8	13	567	378
2007	427,475	3.9	82.9	21	6880	138
2008	273,048	2	8.5	4	72	91
2009	246,856	2.1	14	7	197	167
2010	97,494	1	0.2	0	0	133
2011	343,529	2.2	64.5	29	4159	392
2012	117,459	4	43.5	11	1890	245
2013	372,920	2.8	86.3	31	7442	559
2014	445,292	2.5	14.6	6	213	45
2015	411,093	2.9	18.5	6	343	48
2016	461,834	2.4	17.7	7	314	83

New firm formations play crucial role in overall economic prosperity. There is a significant higher level of innovation comes from the entering firms, rather than the older business entities, favoring also the higher size (Huergo and Jaumandreu, 2004). Although the starting and ending values are close, Table 2.3 depicts a significant variance of the new establishments mean size over the years. Mean size of the new firms turned out to be especially low in 2006, 2008 and 2010. On the other hand, mean size of entrants in 2007 and 2012 are close to 4, former being higher than the overall mean of the economy. This is a peculiarity because the small size of the entrants is an important common pre-acceptance carried by almost all firm dynamics research.

Total number of newly formed businesses also varies during the period. While entrant count stayed moderately low in 2008 and 2009, it was shockingly stayed below a hundred thousand in 2010. It is surprising due to the notion that more firms were supposed to be added after a crisis year to seize the opportunity. The year 2010 formed a deadlock of Turkish businesses by such a low number of firms with a mean of 1, compressed with an extremely low variance. As a final remark, year 2012 possesses a very low number of entrants just above 2009, yet their mean is nearly equal to the overall distribution of firms.

Table 2.4 possesses information on the firms which got out of business during the period 2005-2016¹. In theory, closing firms should rather be small in size and this is

¹ Data of 2006 is of limited scope due a collection problem

mostly verified by the Turkish firm data. Again in 2010, the smallest of the closing firm sizes occur with a small variance. A higher mean size of closing firms than the mean size of entrants depicts an unfavorable picture, further lowering the overall mean of firm sizes. Although the total number of entrants and closing firms differ a lot, this statistic lays the general climate about the economic outlook.

Table 2.4: Size Statistics of Closing Firms, 2005-2016

Year	Obs.	Mean^a	Std.^b	b/a	Variance	Skewness
2005	250,533	1.7	23.4	14	547	94
2006	705	49.1	346.1	7	119763	18
2007	74,146	2.7	45.4	17	2065	78
2008	50,060	13.5	125.3	9	15693	56
2009	183,545	2.6	39.5	16	1558	367
2010	214,664	1.4	6.6	5	44	94
2011	270,825	1.6	7.5	5	56	113
2012	242,045	1.9	28.5	15	810	418
2013	270,345	2.4	137.8	57	18997	499
2014	245,010	2.1	14.4	7	207	86
2015	222,231	2.3	16.4	7	269	80
2016	251,483	2.8	25.2	9	635	67

In this notion, Turkish economy could not make up for the spots opened by closing firms in terms of relative size in 2010 and 2016. In 2013, closing firms' distribution varied significantly in terms of size, indicating a widespread hazard problem also for larger firms.

Table 2.5: Age Statistics of Firms by Year, 2005-2016

Year	Obs.	Mean ^a	Std. ^b	b/a	Variance	Skewness
2005	3,040,531	4.8	5.7	1.2	31.9	2.5
2006	2,890,595	5.1	5.8	1.1	33.5	2.3
2007	2,928,276	5.3	5.9	1.1	34.7	2.3
2008	3,182,797	6	7.1	1.2	50.5	4.1
2009	3,182,797	7	7.1	1	50.5	4.1
2010	2,968,474	8.7	8.7	1	76.3	5.5
2011	3,344,796	8.3	9.1	1.1	82.6	4.9
2012	3,173,469	10.6	8	0.8	64.6	1.2
2013	3,529,165	7.9	7.4	0.9	54.5	1.6
2014	3,525,430	7.6	8.2	1.1	67.1	3.2
2015	3,311,021	7.7	9.3	1.2	86.6	-189.5
2016	3,652,521	7.8	8.3	1.1	68.6	2.4

Firms in a stable economic environment enjoy a long life cycle, depending on their individual capabilities. In terms of growth, older firms are shown to be more stable (Evans, 1987b) and more robust to shocks (Evans, 1987a; Adelino et al., 2017). According to Table 2.5, Turkish firms were young in 2005 with a mean of 4.8 years, showing a steady aging to a mean of 10.6 in 2012. Post-crisis period shows a sudden increase in age, backing the theory of resilience of older firms to shocks. Between 2013-2016, combining with Table 2.3, influx of new firms kept the mean age lower. The year 2012 marks an important symmetrical age distribution, reflecting the sweeping of the crisis on younger firms².

Table 2.6: Age Statistics of Closing Firms, 2005-2016

Year	Obs.	Mean ^a	Std. ^b	b/a	Variance	Skewness
2005	233,925	3.8	5.3	1.4	28.3	2.9
2006	625	5.5	6.8	1.2	45.7	2.7
2007	70,878	4.9	5.6	1.1	31.4	2.8
2008	49,907	7.3	8.2	1.1	67.9	4.1
2009	183,487	6.3	6.2	1	38.2	3.8
2010	214,396	6.8	8.1	1.2	65.4	3.3
2011	270,562	6.5	8.3	1.3	68.2	3.5
2012	241,499	9.3	7.8	0.8	61.4	1.4
2013	270,341	6.1	7	1.1	48.5	2
2014	245,010	5	7.1	1.4	49.7	14.3
2015	222,231	4.7	6.6	1.4	43.9	2.6
2016	251,483	4.9	6.6	1.4	43.6	2.4

² Skewness and Kurtosis calculations gets effected even with a few outliers' data disturbance, as in Table 2.5 for 2015.

Despite the nourishing nature of economic environment in any country, firms get out of business all the time. Distribution of these firms ages signals different drawbacks. It is seen in Table 2.6 that mean age of firms who went out of business during the period 2005-2016 changes significantly. Average age of closing firms raised from 3.81 to 4.88 in twelve years, but it reached a shocking maximum of 9.27 in 2012. The period of 2009-2013 constitutes a massive shutting down of grown up businesses with respect to overall age distribution in Turkey.

As noted before, closing firms should be picked by selection. Combining Tables 2.4 and 2.6, one can see that such firms were long lived despite their small sizes. This supports the claim of Akcigit et al. (2016) regarding the business ownership for subsistence, hence lacking incentives to grow. Yet these firms could had been enjoying the easing access to credit during the 2000s (Ayyagari et al., 2016), and deterred the inevitable end, while efficient firms with easy financing grows to the future (Beck and Demirguc-Kunt, 2006).

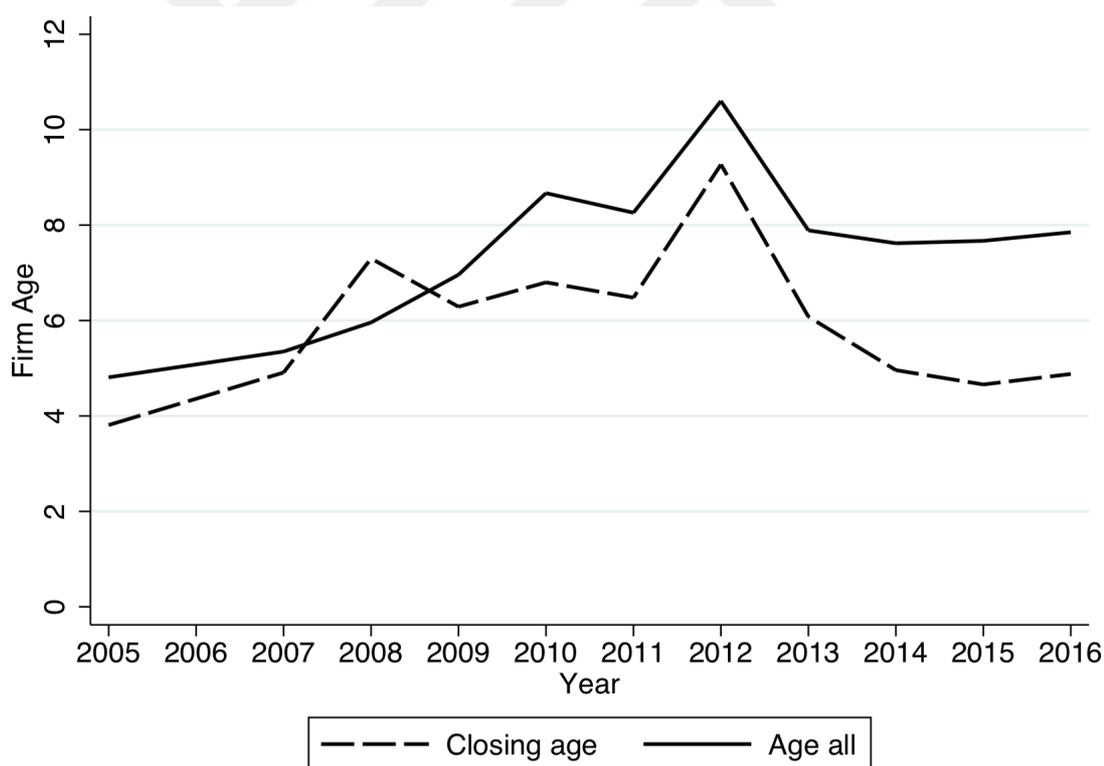


Figure 2.2: Age and Firm Dynamics, 2005-2016

Figure 2.2 graphs the closing firms mean age together with the all firm mean over the period 2005-2016. It is surprising to infer that until recently, selection did not work in

Turkish economy building on the fact that closing firms are increasingly older between 2005 and 2012. After 2013, closing firms' age decouples from mean age of all firms.

Figure 2.3 depicts mean sizes of entering and closing firms within the period of 2005 and 2016. The extreme peak in mean closing firm size of 2008 portrays the severity of the Global Financial Crisis reflections in Turkey. During 2008-2010, not only the larger firms gone out of business, but also entrants mean size was significantly decreased. Entrants' mean size increase sharply during 2011-2012 period signaling a post-crisis boom. A more interesting note on the graph is that in the last year, i.e. 2016, mean size of closing firms are higher than the entrants' which was only the case during the Crisis period. It warns about another possible turmoil in Turkish economy.

2.3.2. Density Estimations

2.3.2.1. Distribution of the Entrants

Numerical assessments of the previous part deserve visualization for an extra emphasis. Figure 2.4 depicts the size distribution of firms who enter to the business environment in Turkey between 2005 and 2016³. The lines of consecutive three years are plotted into the same panel to capture year-over-year variance. Figure 2.4a shows a disturbance on lower medium to medium size firms. There are significantly more firms entered to the economy in 2005 and 2006, while smaller firms dominate the 2007 cohort.

To assess the effect of the crisis perception, Figure 2.4b plots 2008, 2009 and 2010 together. It can be seen that the majority of the entrants did not exceed 20 employees in 2008. Yet the larger firms start to be formed in 2009 and 2010. Comparing to other panels and keeping in mind that mean size of entrants equals to 1 in 2010 from Table 2.3, it can be seen that the weight of the medium-size firms decreased for this period. This might be caused by the discouraging nature of the unfavorable economic environment, letting subsistence-seeking people in but keeping moderately endowed risk averse entrepreneurs. The fading effect of the crisis years can also be seen in Figure 2.4c. The peak in 2011 at the density of entering smaller firms gets a significant

³ To make the distribution plots visually tractable, firms with size 10 and above are considered. Complete plotting is exemplified in the Appendix.

correcting shift in 2012, and almost staying the same in 2013. Figure 2.4d, on the other hand,



Figure 2.3: Firm Size and Dynamics, 2005-2016

depicts another shock towards smaller entrants in 2014, with a more gradual smoothing over 2015 to 2016. But recalling from Table 2.3, the cluster of smaller entrants in 2014 is much less skewed and sparsely distributed around the middle size firms. This fact lowers the noteworthiness of the visual disturbance observed, at least compared to Figure 2.4c.

2.3.2.2. Distribution of the Closing Firms

Investigating the distributional properties of closing firms has two major outcomes. First, it helps distinguishing the main reasons that drive the firms out of business during the subjected time period, being a contemporary shock or a structural problem. Second, it reveals the direction of which selection levies on the economic environment, hence allows intervening if the evolution results undesirable outcomes. The lines of consecutive three years are plotted into the same panel to capture year-over-year variance in Figure 2.5, same as before.

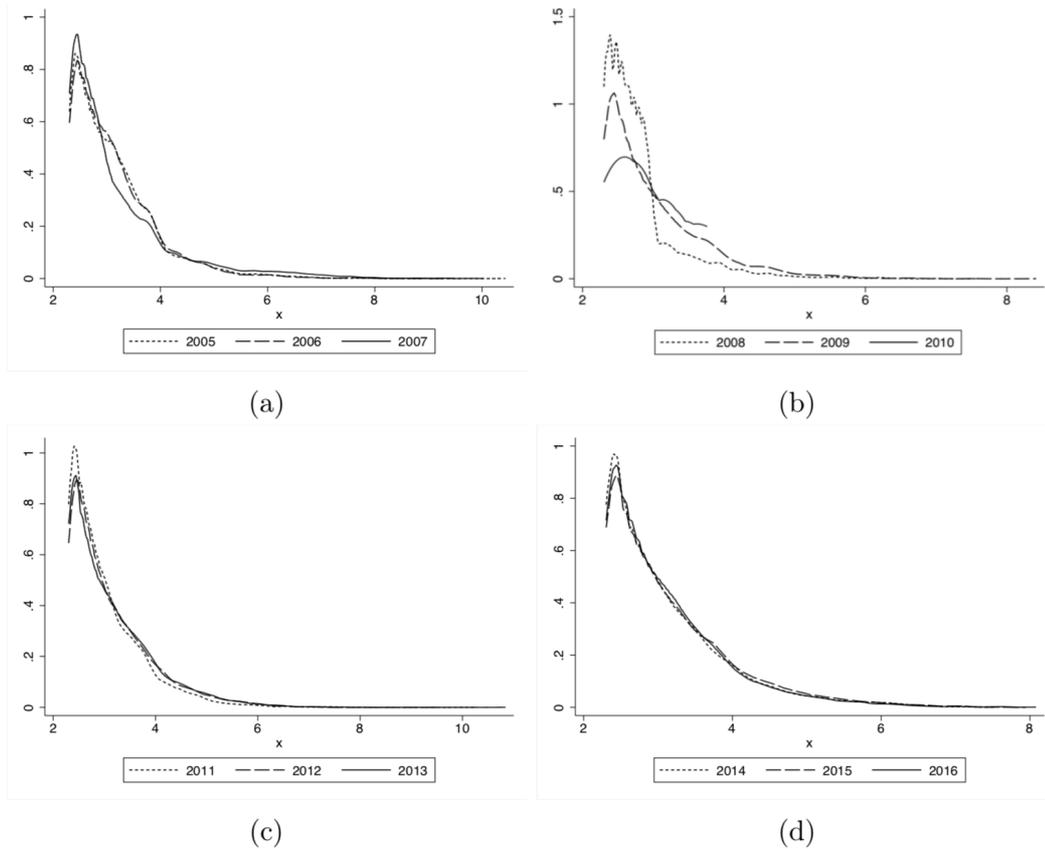


Figure 2.4: Size Distribution of the Entrants, 2005-2016

Figure 2.5a shows a smoothing distribution of closing firms from 2005 to 2007. The concentration of closing sizes is more sparsely distributed than all other years. This observation supports the idea that small and inefficient firms could deter their exit with easing financing opportunities until the crisis. Previously fallen density of smaller firms starts to dominate the distribution from 2008 onwards, as Figure 2.5b depicts, reaching a high in 2009. Another inference from these two lines would be that moderately sized firms who survived through 2008, succeeded to sparing themselves also from troubles in 2009. Adapting protection measures from crisis might worked for some of the firms for some time. Figure 2.5c reverses the order: domination of small firms in closing distribution fades from 2011 to 2013.

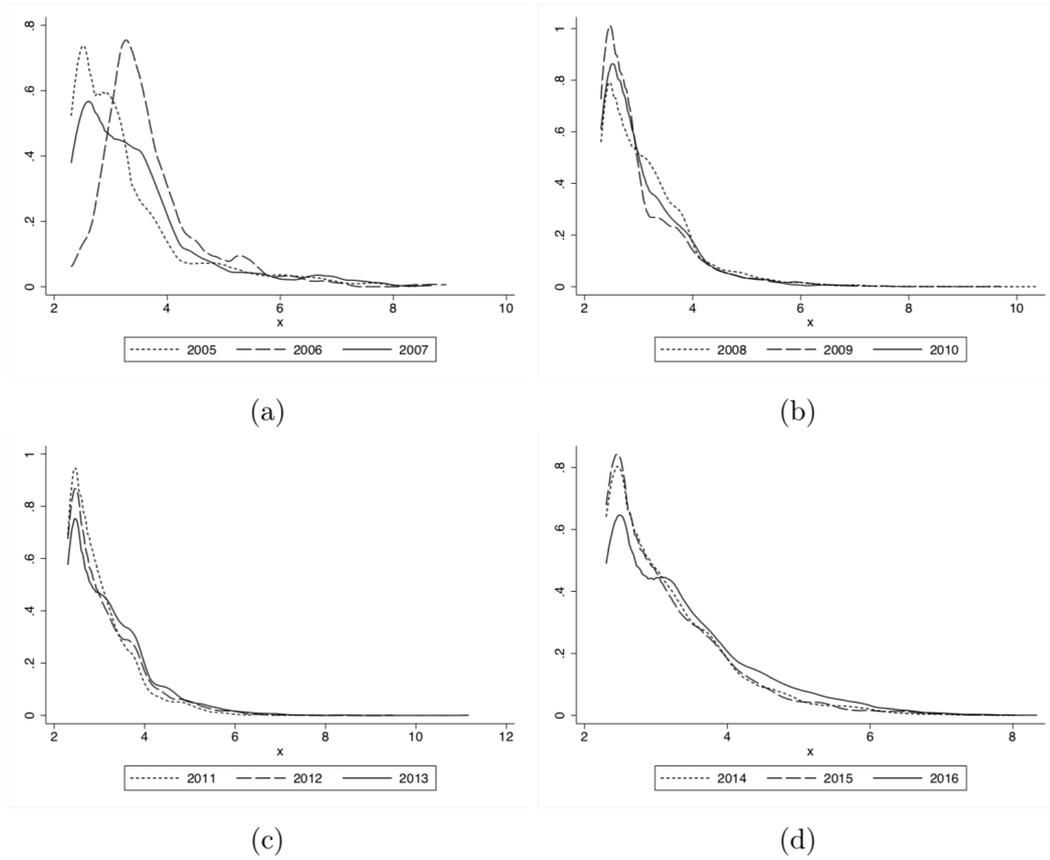


Figure 2.5: Size Distribution of the Closing Firms, 2005-2016

The similarity of the years 2008 and 2013 are striking. When Figure 2.5d examined, things about the closing firm size distribution converges back to a similar shape with 2005, but this time with less skewness and more density on the smaller firms as theory predicts.

2.4. Age and Size Distribution of Firms

Firms of different ages could possess diverse characteristics, both as individuals and groups. The age grouping of Cabral and Mata (2003) is adopted for Turkish firms in the following section for the available data. There are 6 age groups of firms identified. Firms with 0-1, 2-4, 5-9, 10-19, 20-29 and 30 or more years of lifetime are considered to form a cohort. Apart from the yearly changes, the overall evolution of the age groups is inspected. So, from the period of 2005 to 2016, selections of 5 years are to be tabulated in the Section 2.4.1 and 2.4.2, namely 2005, 2007-2009, and 2016.

2.4.1. Survey of the Age Groups

Handling the data is crucial in order to extract unique information. Figures 2.6 and 2.7 depicts the same data with reverted plots. In Figure 2.6, evolution of each age group can be observed. Line patterns follow the age order from dots to a solid line. In 2005, an interesting disorder of older firms being smaller is observed. This irregularity gets corrected during the Global Financial Crisis and age-size order is kept in line afterwards, backing the claim on the lack of selection in pre-crisis period. The oldest age group, namely 30 and above, is boosted in 2008. The Crisis could be said to hit the small and inefficient old firms severely, and dropping out of this group lets the remaining firms portray a larger mean in terms of size.

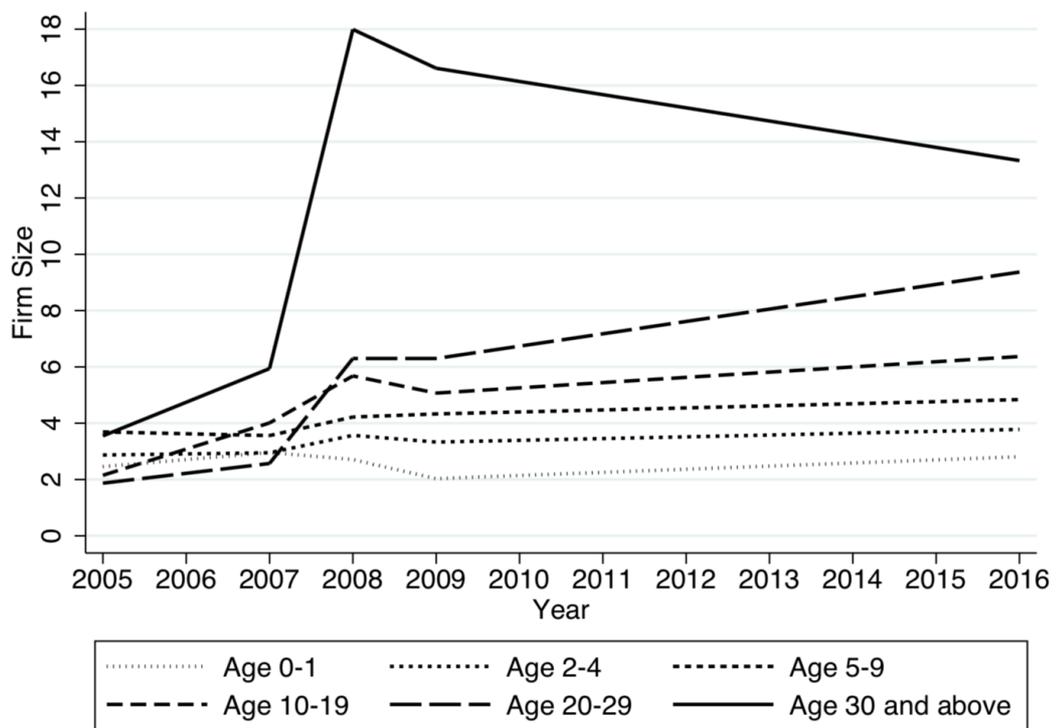


Figure 2.6: Plot of the Age Groups

In Figure 2.7, distribution of age groups are presented. It is seen that before the Crisis, Turkish firms were not ordered according to their age in terms of size: older did not mean larger. Especially the age groups of 10 to 19 and 20 to 29 in 2005, and 20 to 29 in 2007 stands out as extraordinarily small in mean size with respect to the younger groups. A shock to the economy can be represented with a kinked age group-size plot of 2008 and 2009, indicating the oldest age group grow in size and create a gap with

the successor group. This kinked shape smooths out in 2016 towards a more natural outlook.

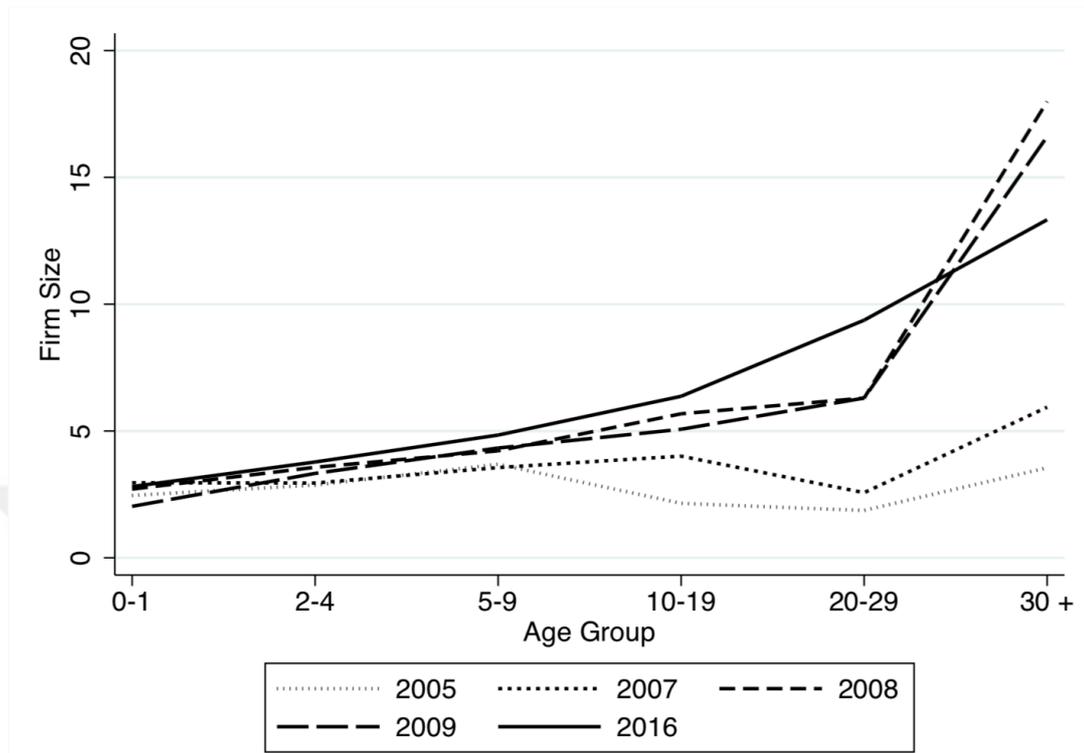


Figure 2.7: Tracking the Age Groups, 2005-2016

Table 2.7 portrays Turkish firms regarding the age groups nominated above. In 2005, the first thing to notice is that the youngest age group is the largest one, followed by 5-9 group. Both of these placements, youngs being the most dominant group and 5-9 exceeding the 2-4, are specific to 2005. 20-29 age group has the lowest mean size with 1.87 employees per firm not only in 2005, but in the whole period. Moreover, 5-9 age group has a mean size larger than 30 or more age firms. When we date these firms' origins, it clears the mystery yet demanding a conclusive research. The smallest sized 20-29 group were founded between 1976-1985, first 5 years including an embargo, domestic turbulence, coup d'etat, and second 5 years under military partly in control. Firms emerged in this period could be thought of a mostly subsistence oriented, or at least coded in a stability seeking nature. The 5-9 age group dates to 1996-2000 period. This group of firms has entered the Turkish economy in a turbulent inflationary and high government spending period with high interest rates, yet survived through 2001 crisis. So, it should not be a surprise for these firms being tough and versatile,

achieving a larger size than their companions. As for the 30 or more age group, founded before 1975, they originate from a planned economic environment with a low portfolio of skilled workers. Firms of those era were nurtured by the government protection with import substitution industrialization perspective, yet a handful of them grow and get corporate while most businesses in the country simply aim subsistence and daily trade.

There is a serious drop in numbers of 0-1 age group in 2007, and the majority in groups transfers to 2-4 age firms. The abnormal small mean size of 20-29 age group continues. Total count of oldest firms shows a significant decrease, with pointing to that small and inefficient ones starting to drop out finally.

In 2008, 0-1 age group gets hit further and become the third most populated group in the economy with both aging of the members and decreasing number of entrants as shown in Table 2.3. The age groups finally depict a continuously increasing mean size starting from 2008. As successful and large firms from lower ages join, and small and inefficient drop out, 30 or more age group shows a significant increase in mean size reaching almost 18.

From 2008 to 2009, a noticeable decrease in all age groups but 5-9 occurs. This decoupling of the 5-9 age firms is not surprising, because of the survivors of 2001 crisis among them as mentioned before. The most significant loss of mean size is of 0-1 firms as expected with 33%. Entrants during the crisis years are the most vulnerable ones who had to face the storm in such inexperienced times, yet the evidence from previous crisis entrants offer prospect to those who survive.

Table 2.7: Size Statistics of Age Groups, Selected Years

Size Group	Obs.	Mean ^a	Std. ^b	b/a	Skewness	Kurtosis
2005						
0-1	1,047,848	2.46	45.58	19	476	320,680
2-4	748,162	2.87	36.33	13	137	30,900
5-9	843,414	3.69	76.59	21	296	128,036
10-19	307,051	2.15	29.26	14	148	33,380
20-29	63,545	1.87	33.87	18	137	22,459
30≤	38,750	3.55	63.16	18	68	6,307
2007						
0-1	829,174	2.96	60.6	20	183	48,966
2-4	880,009	2.95	29.71	10	180	61,307
5-9	683,077	3.56	40.4	11	143	33,683
10-19	435,246	4.01	72.72	18	498	295,678
20-29	77,488	2.57	27.22	11	75	7,743
30≤	26,269	5.94	163.53	28	79	7,434
2008						
0-1	610,664	2.71	87.25	32	593	380,645
2-4	854,831	3.57	40.57	11	484	337,251
5-9	791,442	4.22	36.15	9	106	21,015
10-19	542,133	5.68	151.73	27	569	371,029
20-29	104,544	6.3	101.95	16	92	11,081
30≤	32,378	17.99	448.02	25	53	3,238
2009						
0-1	519,904	2.03	11.46	6	165	51,708
2-4	940,573	3.33	78.11	23	596	406,808
5-9	818,374	4.33	35.85	8	107	21,616
10-19	747,208	5.07	129.86	26	658	501,622
20-29	120,333	6.3	100.97	16	89	10,406
30≤	36,456	16.61	422.85	25	56	3,625
2016						
0-1	860,644	2.81	18.98	7	71	8,510
2-4	832,449	3.78	27.98	7	131	33,646
5-9	796,720	4.84	58.3	12	245	93,967
10-19	806,976	6.37	70.04	11	111	20,615
20-29	276,710	9.37	123.52	13	112	19,963
30≤	79,022	13.33	306.3	23	57	3,872

At the bottom part of Table 2.7 size statistics of the age groups observed in 2016 can be found. Compared to their 2009 situation, 0-1 age firms are seen to recover to a healthier mean size. The second interesting fact is, however, 30 or older firms

continued to shrink while their total count more than doubles as they were expected to grow further following their pre-crisis trend. Not to mention that previously noticed stagnant 1976-1985 born firms now entered to the 30 or more age group. At the same time, 20-29 age group achieved a superior growth in the post-crisis period with almost 50%. These cohort are born in 1987-1996, which is a quite characteristic era with liberal economic policies, starting with ended military administration, liberating the international capital movements and ending with joining to the customs union of the European Union. These firms survived and gained experience with 1994 and 2001 crises of Turkey, justifying their outstanding performance. On the other hand, the first four age groups are almost has equal members in 2016. This uniformity informs a loss of dynamism because expanding older age groups should depict a favorable economic environment hence new firms seek entrance. On the contrary to seemingly consecutive high number of entrants during 2014-2016 as seen in Table 2.3, these figures are even behind 2005 values. The thrust and power coming from new firms are missing at the end of the subjected period, making the age distribution of Turkish firms closer to uniform.

2.4.2. Density Estimations

Figure 2.8 plots the size distribution of the age groups outlined in Table 2.7. It should be possible to track the evolution of specific groups by looking at their density estimations. The years are denoted in the vertical line labels.

Alongside with the previous literature, Figure 2.8 provides evidence that size distribution of firms by age groups depict a more symmetrical plot as the age increases. However, crisis years of 2008 and 2009 differs significantly among similar age distributions of other years. These two years' distribution plots get a more uniform shape in the center, indicating outlier firms with extremely smaller and larger sizes than the mean suffers from selection, or went into restructuring. Moreover, in these two years, middle age groups have two peaks with a taller one on the right. Which means that small firms are wiped out and marginally larger ones survived from those distributions, as noted before. This indicates there is a possible safe size for a firm to be attained in order to survive the crisis.

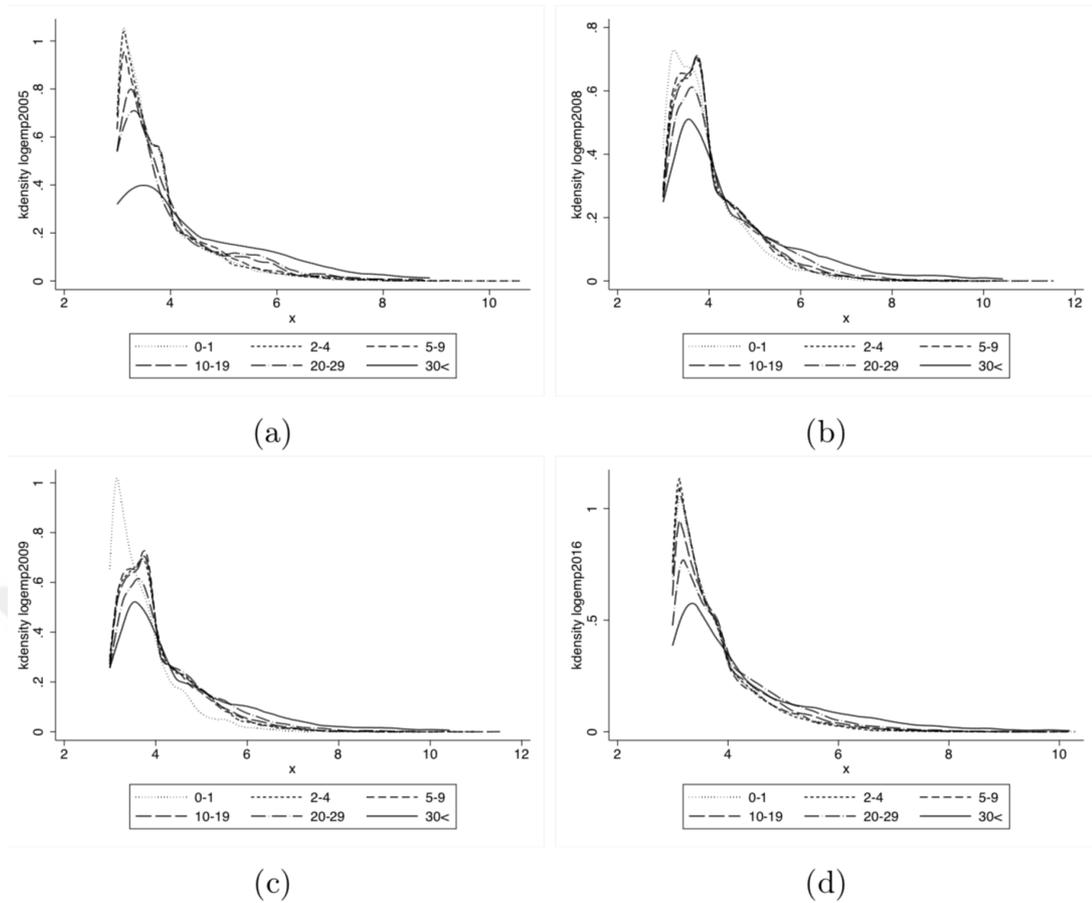


Figure 2.8: Size Distribution of the Age Groups, Selected Years

Figure 2.8c shows how extremely skew is the size distribution of youngest firms in 2009 compared to others. This age group shows a wider medium sized distribution in 2008, but mainly keeps the skew shape. There is almost no big firms (with size above 100) entering to the economy in 2009, unlike 2008, meaning a great absence of large scale investments. On the other hand, the oldest group in the distribution evolves toward a steeper peaked shape, indicating an increase in small sized firms' weight in the distribution. This was also noted earlier as 30 or older firms in 2016 possess a decreasing mean size in the post-crisis era.

2.5. Conclusions

It is shown that despite a thrust from entering firms until the Global Financial Crisis in 2008, Turkish private sector get seriously hit and could not recover until 2013. When the post-crisis data is examined, despite the expansion among older firms, there are

simply not enough new firms entering to the Turkish economy relative to its size. In 2016, age distribution of firms is almost uniform for the groups with a history of less than 20 years. Loss of dynamism and risk appetite is evident in this period. Combining with the shrinking mean size of oldest age group, there is a not-so-obvious inference: risk aversion and economic distrust keep productive people in paychecks while these lack of competition and availability of affordable professionals extend the lifetime of old and inefficient firms. These otherwise-entrepreneur productive employees could even bring the opportunity of lowering the excess size to the existing firms without self-improvement which may further explain the oldest firms' shrinking sizes. To keep in line with the scope of this paper, further comments related to policy problems and supporting evidence from macroeconomic variables are waived.

As asserted by Akcigit et al. (2016) it is evident from Table 2.7 that as Turkish firms get older, they do not necessarily grow larger. In fact, mean size of 30 years and older firms decreased significantly in the last decade. Further striking fact of this assertion is about the previous trend of the same statistic. The oldest Turkish firms, aged 30 and more, extensively grew until the crisis in 2008. This may indicate a shift of vision, or shift in the managerial ability as the paper explains. In the light of the proposed model, Global Financial Crisis in 2008 might have derailed Turkish private sector. Building further on the findings by Akcigit et al. (2016), it could be possible to claim that Turkish economic environment resembled a non-developed country until 2013 in terms of selection not functioning properly. During 2012 and 2013, many small sized but older firms get out of business. Combining with the previous assertion of oldest age groups, the shock of the crisis that disturbed the economic environment have not get stabilized until clearing the inefficient firms in 2013. In 2016 data of the oldest firms, it is seen that their mean size shrinks further which again turns the table towards the lack of selection in Turkey. Clearly, employing mainstream theories to foresee the crisis among Turkish private sector would fail. Aim to capture these special aspects puts more importance on the new and more flexible approaches like ABC models where incorporating heterogeneity of units is much simple and valued.

2010 was a year lacking firm formation of small to medium size. It was a year with mean entrants is 1 with a little total variance. This might be strikingly pointing to the fact that jobs lost in 2008 (and 2009) crisis could turn into new firms with sole employees to seek subsistence, demanding theoretical background. Furthermore,

Turkish firms laying their employees off during the crisis could have been created a serious competition against them with laid-off workers taking chances with entrepreneurship using their experience. It concurs with the evidence from age data of the closing firms after crisis years as cited above. This is an uncharted territory in economic theory, but there is evidence of the reverse. Reisenbichler and Morgan (2012) explains how the German firms kept their employees and turned a crisis to a “miracle”.

Despite the low and close mean firm size among entrants after 2011, plots in Section 2.3.2 made clear of the recovery in the last years hand in hand with the statistics. This finding emphasizes the complementarity of the methods carried out in this study.

Thinking a business entity apart from its entrepreneur can be misleading. This drawback is inevitable under modeling with the rationality assumption, since the business owners are not obliged to be rational if there is no punishment through selection. Especially in non-developed economies where delegating managerial issues are shown to be uneasy, responses to certain shocks and decisionmaking processes resemble the vision of the founder of the firm. It is humbly claimed in this study that it can be misleading to model these processes without accounting for the heterogeneity caused by originating environment. Firms born during a certain atmosphere in a country where managerial hiring is found troublesome might not renew their way of doing things, apart from innovation, even if the world surrounding them changes drastically. They might even avoid selection in big numbers. So models should also incorporate cohort characteristics as a meso source of variation, on top of the individual micro characteristics and macro level policy interactions.

2.6. Bibliography

- Adelino, Manuel, Ma, Song, Robinson, David. 2017. Firm age, investment opportunities, and job creation. **The Journal of Finance**. v. 72. i. 3: 999–1038.
- Akcigit, Ufuk, Alp, Harun, Peters, Michael. 2016. Lack of selection and limits to delegation: firm dynamics in developing countries. **National Bureau of Economic Research**. No: 21905.
- Audretsch, David Mahmood, Talat. 1994. Firm selection and industry evolution: the post-entry performance of new firms. **Journal of Evolutionary Economics**. v. 4. i. 3: 243–260.
- _____, Mata, Jose. 1995. The post-entry performance of firms: Introduction. **International Journal of Industrial Organization**. v. 13. i. 4: 413–419.
- Ayyagari, Meghana, Juarros, Pedro, Martinez Peria, Maria, Singh, Sandeep. 2016. **Access to finance and job growth: firm-level evidence across developing countries**. The World Bank.
- Beck, Thorsten, Demirguc-Kunt, Asli. 2006. Small and medium-size enterprises: Access to finance as a growth constraint. **Journal of Banking & finance**. v. 30. i. 11: 2931–2943.
- Cabral, Luis, Mata, Jose. 2003. On the evolution of the firm size distribution: Facts and theory. **American economic review**. v. 93. i. 4: 1075–1090.
- Delli Gatti, Dominico, Desiderio, Saul, Gaffeo, Eduardo, Cirillo, Pasquale, Gallegati, Mauro. 2011. **Macroeconomics from the Bottom-up**. Springer Science & Business Media.
- Evans, David. 1987a. The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. **The journal of industrial economics**. v. 35. i. 4: 567–581.
- _____. 1987b. Tests of alternative theories of firm growth. **Journal of political economy**. v. 95. i. 4: 657–674.
- _____. and Leighton, Linda. 1989. Some empirical aspects of entrepreneurship. **The American Economic Review**. v. 79. i. 3: 519–535.
- Gibrat, Robert. 1957. **On Economic Inequalities**. trans. E. Henderson. International Economic Papers. The Macmillan Company.
- Huergo, Elena, Jaumandreu, Jordi. 2004. How does probability of innovation change with firm age? **Small Business Economics**. v. 22. i. 3: 193–207.
- Jovanovic, Bojan. 1982. Selection and the evolution of industry. **Econometrica**. v. 50. i. 3: 649–670.

Lucas, Robert. 1978. On the size distribution of business firms. **The Bell Journal of Economics**. v. 9. i. 2: 508–523.

Reisenbichler, Alexander, Morgan, Kimberly. 2012. From “sick man” to “miracle” explaining the robustness of the german labor market during and after the financial crisis 2008-09. **Politics and Society**. v. 40. i. 4: 549–579.



3. ESTIMATIONS ON THE FIRM SIZE DISTRIBUTION IN TURKEY

3.1. Introduction

An extraordinary assertion by Simon and Bonini (1958, p. 616) is that if there was a need to predict behaviors of a system of adaptive organisms which tends to stray away from equilibrium by some forces that are stronger than the ones forcing equilibrium, one shall replace the static equilibrium with a remedy, i.e. size distribution of firms. Building on this notion, firm size distribution has been studied mostly in perspectives of firm behavior, market concentration, and obtaining stylized facts about its self portray. The last objective constitutes the main drive for the current paper.

Quite a few serious and overwhelming mathematical modeling of firm dynamics are well presented in the literature, leading to some certain size distribution of firms. Gibrat's Law (1957), or the Law of the Proportionate Effect, were shown to be leading to a lognormal distribution of firm sizes. With some deviations added like Kesten's (1973), firm sizes should follow a power law, or Pareto distribution. Both processes assume that the firm growth is independent of current size, and the latter assumes further a minimum size for firms to obtain power law instead of lognormal (Gabaix, 2009, p. 261). These well founded competing regularities might provide the desired output when assumptions hold, yet economic theory undergoes a serious encounter with the recently blooming complexity science where reliance on concrete assumptions decay.

The main innovation in economics might come from not altering the theory but altering the links to the empirics. Micro foundation of macroeconomics now progressively evolves towards the macro foundations of micro. Aggregate regularities can be obtained by simple and multi level interactions among agents. By micro, meso and macro level interactions, some regularities might emerge which are impossible to be predicted by looking at the individual (Delli Gatti et al., 2011, p. 85). In doing so, experimenting with economics in simulated environments need correct ties to the

observed facts. Therefore, it puts importance on conclusive remarks regarding the distribution of firm sizes, supported with computational evidence.

It is beyond the scope, and out of the reach of this paper to undermine the validity of firm dynamics models leading to size distributions, yet the empirical evidence is critical in all terms. In this study, firm size distribution in Turkey is tried to be approximated by power law and lognormal distributions. The major contribution to the literature is both on data and methodological perspective. To the best of our knowledge, combining such a comprehensive firm level data with implementations of the direct comparison of proposed distributions is unprecedented among neither developed nor developing countries.

3.2. Literature Review

There is a long line of research with statistical distributions in economics, since the pioneering work of Pareto (1897) related to income distribution. Pareto found a linear relationship between the logarithm of people with an income higher than x and $\log x$ (Schumpeter, 1949, p. 155). It is recently stated that it is not an easy job to do a clear cut literature review on statistical size distributions in economics, while providing an exceptional one, because there are neither common educational resources nor direct links between seemingly related publications on close matters (Kleiber and Kotz, 2003, p. 1).

The size distribution of firms has long been considered as lognormal, under the assumption that growth of a firm is independent of its size. Using French firm level data, Gibrat (1957) puts forward his famous the Law of Proportionate Effect and asserts that firm sizes should be lognormally distributed. Kalecki (1945) confirms the lognormality and extends the Gibrat Distribution using limited samples of US factory employment and income distribution of the UK. Hart and Prais (1956), Hart (1962) and Hart and Oulton (1997) also find the representation of the firm size distribution by lognormal useful.

In his seminal book, George Kingsley Zipf (1949) put forward that word frequencies follow a similar, i.e. linear, relationship as their rank due to a principle called the Least Effort. The rank-size plot of the observed variable which results from human effort shall be regressed with a coefficient equal to negative unity (Zipf, 1949, p. 25). Later

on came the varying empirical and theoretical work on diverse fields e.g. measuring city and firm size distributions obeying the so called “Zipf’s Law” of Zipf distribution naming as a special case of Pareto distribution with a coefficient of minus one.

As a path breaking insight, Simon and Bonini (1958) finds that Pareto distribution fits the firm sizes better than lognormal, with a Pareto exponent of 1.11 for British and 1.23 for US firms. They also label the fitting distribution as being a Yule type, formally. Steindl (1965) also puts forward that firm sizes follow a Pareto distribution. Another influencing paper by Barabási and Albert (1999) brought the scale-free property of the power law under the spotlight using data from actor popularities, internet connectivity and power grid networks, making the subject extremely famous among scientists from all disciplines. Power law or Pareto distribution has been abundantly encountered among the diverse fields since the 2000s. There is a line of research about the competing ideas of lognormal and Pareto type representation of the size distribution of cities. Being representative examples, Gangopadhyay and Basu (2009), Peng (2010), Chen (2012a,b), Bee et al. (2013), Luckstead and Devadoss (2014b,a, 2017), Fazio and Modica (2015), Luckstead et al. (2017) pursues similar goals of evaluating the assumed Pareto or lognormal distribution of the city sizes with diverse methods for the US, China and India. Urzúa (2011) criticizes the OLS estimation procedures applied to validate the Zipf’s Law among works related to city size. Veneri (2016) utilizes a large dataset of OECD countries and validates the Zipf’s Law among different scales i.e. country, continent, OECD levels. Despite the lack of mentioning, this finding could be interpreted an important aspect of the data which relates to complexity theory and for city size distribution being a scale-free property. On the other hand, Newman (2005), Fragkias and Seto (2009), Ioannides and Skouras (2013) and Chen (2016) takes on an evolutionary approach to analyzing the city size distributions. Finally, Arshad et al. (2018) combines this literature and directs a fruitful future research agenda embracing the idea that cities being complex adaptive systems. Apart from city sizes, power laws have been found in studies on bacteria (Almaas et al., 2004), animal behavior (Niwa, 2003), world wide web (Adamic and Huberman, 2000) and many more fields which are not being mentioned to avoid divagation.

Okuyama et al. (1999) considers income distribution studies of Pareto (1897) and Gini (1922) as a benchmark for firm income distribution analysis. Utilizing two data sets covering 85,375 Japanese, an undisclosed number of Italian and approximately 10,000

US firms' annual incomes, OLS estimations were employed. They confirm Zipf's law for Japan and Italy, but not for the USA. and they also conclude by emphasizing that competitive structure of the economy might lead to such distributions of firm sizes.

Criticizing the dependence on incomplete data to validate- or reject the power law distribution of firm sizes, Axtell (2001) utilizes a large data set of the 5.5 million US taxpaying firms. OLS estimations on logarithmically binned data of both employment and revenue of the firms at hand were conducted. Zipf's law was persistently confirmed and it was concluded for the Law that it should be valid for all economies with a complete data set under spotlight. Artige and Reginster (2018) follows the steps of Axtell (2001) with complete data set of Belgian firms. Their results do not conform with his finding of Zipf's law but favors Generalized Pareto Distribution with cut-off. Cabral and Mata (2003, p. 1083) claim that the firm size distribution fails to be labeled as lognormal but evolves towards a more symmetric shape when the data set is complete. Bee et al. (2017) study French firms in order to determine whether Pareto or lognormal distributions explain the data. Two methodologies have been carried out, i.e. Uniformly Most Powerful Unbiased test of Del Castillo and Puig (1999), the Maximum Entropy estimation and the test of Gabaix and Ibragimov (2011). They find that firms that operate in more than one market and produce multiple products are Zipf distributed, whereas Pareto distribution does not have a good fit on the overall firm size data.

After some years of Jovanovic (1982) provides an evolutionary selection model of firm dynamics, Luttmer (2007) further builds up a clear framework implying Zipf distributed firm sizes as follows. According to Luttmer (2007, p. 1103) either the entry costs for new firms should be high or copying successful ones were too difficult, or both. Later on Luttmer (2011) confirms Gibrat's law and attaches the observed size distribution to market diversifying and innovation by firms. On the other hand, Hopenhayn (1992) presents a stationary equilibrium model and relates the skewed size distribution to skewed productivity distribution of firms, claiming that the stochastic part of the firm dynamics is limited to determining the time of exit. Toda (2018) successfully proves that tail behavior of power law in wealth distributions occur through random discounting of the agents.

In a recent paper, Clauset et al. (2009) criticizes the power law fitting attempts using least squares estimation and offer a wide scope to apply and assess the validity of the fits to the data. They successfully implement a thorough determination process for a power law distribution, considering maximum likelihood estimations, likelihood ratios and Kolmogorov Smirnov (KS) statistics. They skip the firm sizes among the 24 data set of theirs to test the power law validation procedures, opening a partial spot for this paper. Apart from the validation procedures, one can easily see the fragmented and disconnected literature on firm size distributions, providing another important goal of the current work. Finally, Stumpf and Porter (2012, p. 666) stresses out the advice that it is more important to acknowledge that power laws are not rare and they should motivate further research with possible simplifying insights rather than attributing some sort of hazy magical universality on the subject matter.

3.3. Data

TURKSTAT holds a micro dataset of Turkish firms under two different sets. “Annual Business Registers Framework” dataset consists limited type of information with almost full coverage of firms (more than 3 million per year), fortunately registered employees are also reported. Table 3.1 presents the summary statistics of all firm sizes who contribute to the GDP of Turkey.

Table 3.1: Summary Statistics of Firm Employment Data in Turkey

Year	Min	Median	Mean	3rd Q	Max	n
2005	1	1	3.09	1	40,393	3,455,700
2006	1	1	3.07	1	59,146	3,205,948
2007	1	1	3.4	1	43,555	3,175,493
2008	1	1	4.3	1	64,000	3,490,786
2009	1	2	4.05	2	101,386	3,227,238
2010	1	1	3.76	1	30,724	3,003,116
2011	1	1	4.11	2	83,700	3,422,163
2012	1	1	4.54	2	31,760	3,474,992
2013	1	1	5.23	3	70,679	3,529,541
2014	1	1	5.39	3	139,576	3,525,431
2015	1	1	5.39	3	27,852	3,584,832
2016	1	1	5.38	3	29,309	3,652,521

Source: TURKSTAT - Micro Data Collection

Turkish firms can be seen as rather small in average employee count. Mean being higher than the median means that the firm size distribution in Turkey is positively skewed. There are some interesting notes to be made. Mean firm size increased almost 40% from 2005 to 2008 then fall back as the Global Financial Crisis passes by. In 2009, possibly as a side effect of the crisis, median firm size reached 2 making the size distribution of firms more symmetric. Third quartile of firms increased from 1 to 3 in 12 years, while median stays stable, implying a general employment increase in a sizeable number of firms. Nevertheless, utilization of a dataset with such span for a non-industrialized economy could be seen as an important contribution to the firm size distribution analysis literature.

3.4. Definitions

Also referred in the literature as Pareto, Yule or Yule-Simon distribution, a power law or Pareto's law is defined in its original form (Pareto, 1897) as:

$$N = Ax^{-\alpha} \quad (3.1)$$

where N is the number of people with income greater than x, and α is the scaling/shape parameter, or also referred as Pareto coefficient/constant or simply exponent (Johnson et al., 1994). Zipf's Law is a special case of the Pareto distribution with $\alpha=1$.

A more relevant definition of the cumulative distribution function of firm sizes above a minimum size is:

$$\Pr[S \geq s_i] = \left(\frac{s_0}{s_i}\right)^\alpha, s_i \geq s_0, \alpha > 0 \quad (3.2)$$

where s_0 is the minimum size for firms (Axtell, 2001, p. 1818).

Despite the criticisms of the loss of information, logarithmic binning of the data for least squares estimation of the parameters is one way to approach the firm sizes. Implied (log transformed) linear form of the CDF is:

$$\log p(x) = c + \alpha \log x \quad (3.3)$$

where it can be simply estimated by building up the histogram of the firm size data and fitting with the least squares estimation (Clauset et al., 2009, p. 690).

As a more demanding alternative, if the minimum size for a power law distribution to be estimated is already known, one can use maximum likelihood estimation to fit such parametric model. But explicit estimation of α is not possible when the data is discrete, and if $x_{\min} \geq 6$, then the discrete data ML-estimator was shown to be well approximated by the following expression:

$$\hat{\alpha} \simeq 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min} - \frac{1}{2}} \right]^{-1} \quad (3.4)$$

which only differs from the continuous case with $-1/2$ in the denominator (Clauset et al., 2009, p. 696). Importance of the determination of the x_{\min} has been noted earlier. The power law exponent could only be calculated for a known/supposedly power law distributed variable, so there is a need to determine the tail length of the fit beforehand. In this study, Kolmogorov-Smirnov (KS) statistic is used to determine the x_{\min} , formulated as follows:

$$D = \max_{x \geq x_{\min}} |S(x) - P(x)| \quad (3.5)$$

where $S(x)$ is the CDF of the data and $P(x)$ is the CDF of the proposed power law model. KS statistic measures the maximum distance between the proposed distribution and the observed empirical distribution. It would be beneficiary to set x_{\min} where the KS statistic is minimized, which pins the longest tail that one could fit a power law by estimating the power law exponent (Clauset et al., 2009, p. 672).

Next, considered as the historical competitor against the Pareto distribution for the firm sizes, Lognormal distribution is defined for a positively defined random variable X if $\log X = Y$ conforms with the normal distribution. The cumulative density function for a lognormally distributed X is:

$$f(X|\mu, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)} X} \exp\left[-\frac{(\ln(X) - \mu)^2}{2\sigma^2}\right] \quad (3.6)$$

where μ and σ correspond to the parameters of Y , being mean and the variance, respectively (Singh, 1998, p. 68). The maximum likelihood estimators for the two parameter lognormal distribution is as follows:

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln(X_i)}{n} \text{ and} \quad (3.7)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \left(\ln(X_i) - \frac{\sum_{i=1}^n \ln(X_i)}{n} \right)^2}{n} \quad (3.8)$$

In the next section, least squares fit of power law proposition to the histogram of firm sizes and the maximum likelihood estimations of power law and lognormal distributions are to be tabulated, provided with a comparison structure between the two fits.

3.5. Fitting to the Firm Size Distribution

As the literature on the size distribution of firms evolved around, lognormal and power law fitting of the Turkish firm sizes should be insightful. To achieve this goal, two different approaches are taken into account. First, a more straightforward method of consolidating the data of millions of firms into logarithmically expanding width of bins and applying least squares estimations to get the Pareto exponent of the log transformed probability mass function. On the other hand, second method dynamically determines the tail length of the fit and use maximum likelihood to estimate the Pareto exponent, as proposed by Clauset et al. (2009) and computerized in R environment by Gillespie (2015). This implementation is flexible enough to also fit lognormal and compare against the power law, so the third part of the section shall provide conclusive evidence about the choice among the fits.

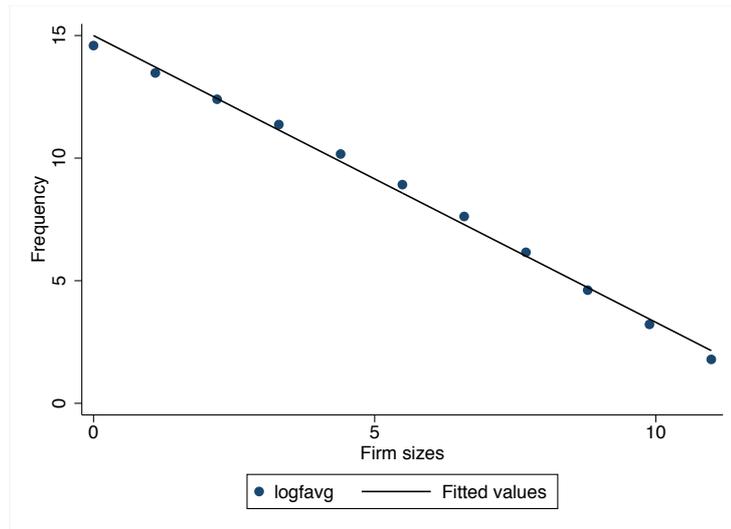


Figure 3.1: Histogram of Averaged log-firm Sizes Between 2005-2016 and Fit

3.5.1. OLS Estimation of the Pareto Exponent

As noted earlier, power law indicates that a simply ordered rank-size plot of a variable depicts a linear relationship, while also coined as Zipf’s law if the Pareto coefficient is 1. Measuring the power law exponent is a rather simple task, if one assumes that such distribution comes from a Pareto type. First, yearly firm size data are binned into size classes of which expands with powers of 3. Then, their histograms are plotted and (3.3) is fitted over. Lacking a formal explanation in published papers utilizing this procedure, the notion behind is keeping the width of size classes equal during logarithmic transformation leading to (3.3).

Figure 3.1 plots the histogram of logarithmically binned data of Turkish firm sizes averaging over the years of 2005-2016. The good fit of a straight line supports the power law regarding the distribution. Further analysis with the absolute value of fitting line shall also conclude about the Zipf’s law. Least squares estimations of Pareto exponents represented with (3.3) and average of firms in each bin are presented in Table 3.2. Size bins and average number of firms in each bin are also reported for extra insight about the data.

Table 3.2: OLS Estimations of the Power Law Exponent

Year	Pareto Coef.	Std. Err.	Adj. R-sq.	Size Bins	Avg. #firms
2005	1.19*	0.04	0.989	0-0	68,906
2006	1.22*	0.049	0.984	1-1	2,169,813
2007	1.20*	0.035	0.992	2-4	710,325
2008	1.08*	0.035	0.989	5-13	242,909
2009	1.21*	0.039	0.989	14-40	86,726
2010	1.20*	0.036	0.991	41-121	26,029
2011	1.16*	0.026	0.995	122-364	7,466
2012	1.26*	0.06	0.978	365-1093	2,040
2013	1.18*	0.042	0.988	1094-3280	473
2014	1.20*	0.041	0.987	3281-9841	101
2015	1.12*	0.036	0.991	9842-29524	25
2016	1.12*	0.037	0.99	29525-88573	6
avg05-10	1.25*	0.053	0.99	88574-265720	0.2
avg11-16	1.28*	0.064	0.972		
avg05-16	1.16*	0.026	0.995		

* $p < 0.01$

According to the results presented, there is a sizeable amount of explanatory power of the Pareto exponent with obtained significance levels for the coefficients and R-squared values. However, estimated exponents are significantly different from 1 which rejects the Zipf's law for Turkish firm size distribution. When the data divided into two parts and averaged during the time span, denoted with variables avg05-10 and avg11-16 and avg05-16 in Table 3.2, it can be seen that there is no significant change between the two parts and Zipf's law remains rejected in medium to long run.

3.5.2. Power Law and Lognormal Distribution Fitting Using MLE

Some serious objections on power law fitting procedures employing binned data and least-squares estimations were from Clauset et al. (2009) and Stumpf and Porter (2012) were noted earlier. In the former, a maximum likelihood estimation procedure for power law exponent is rigorously pictured, which was provided in (3.4). Further comments on these objections could be possible with the following estimations of Turkish firm size distribution. The following lower bounds of firm sizes are calculated before fitting the proposed distributions by minimizing the KS statistic mentioned in (3.5).

Maximum likelihood estimation of power law and lognormal distribution parameters and optimal choice of x_{\min} values are reported in Table 3.3. n_{tail} represents the number of firms of which distribution fits from the data of all firms. Solely depending on these fits, one can say that lognormal distribution does not seem as a better way of describing the firm size distribution in Turkey than power law. Considering the tail length of the distributions, power law fit is better than the lognormal in 2005, 2006, 2007, 2008, 2011, 2013, 2015. The lognormal distribution has extremely poor fits in 2005, 2008, and 2011 which only cover for a handful of firms while the power law does not have such poor fits at all. Pareto, or power law, exponent does not favor the Zipf's law as being significantly different from 1 in all years. Graphical inspection of CDFs and comparing the fits could be seen in Figure 3.2 and Figure 3.3.

Table 3.3: Power Law and Lognormal Fits to the Data

Year	Power Law			Lognormal			
	xmin	alpha	ntail	xmin	par1	par2	ntail
2005	1,240	1.58	451	23,498	-1134.24	30.85	4
2006	109	1.19	7,729	174	0.33	2.25	4,517
2007	113	1.21	8,262	188	-0.38	2.37	4,566
2008	113	1.22	8,676	62,355	-837.05	22.79	6
2009	512	1.44	1,317	143	-7.11	3.24	6,637
2010	147	1.31	6,533	100	-4.96	2.95	10,352
2011	175	1.32	6,036	3,778	7.56	1.44	93
2012	418	1.45	2,149	164	-8.38	3.32	7,388
2013	152	1.29	9,826	161	-16.43	4.2	9,163
2014	316	1.33	4,612	80	-3.66	2.86	23,021
2015	23	1.14	95,436	79	-4.35	2.96	23,292
2016	196	1.3	8,256	100	-4.01	2.88	18,062

As noted before, mainly the tail behavior of distributions is investigated with power law experiments. Considering the CDF of tails of the firm size distribution in Turkey between 2005-2010, there is an evident disturbance on the upper tail of the distribution while headed and leaving 2008. A strange –possibly reverting action can be reported. After both two strangely poor fits of lognormal distributions in 2005 and 2008, following two years are nicely approximated by lognormal. Moreover, being specific to the 2008 and 2009 the slope of power law fit underestimates the slope of the upper tail meaning there is more similarity among larger firms in terms of their employee counts.

Considering the CDFs of years 2005-2016, the best fit of the power law is observed in 2015 which is not confirmed visually. This brings the importance of such analysis is better be done numerically, despite the efforts in literature depending on visual assessments of fits. Figure 3.3 further provides the reverting action after the disturbance observed in lognormal fit to 2011 data. Such extreme and periodical movements are not reported in power law fits. After the fail of fit in 2011, lognormal distribution better approximates the tail of the firm size distribution than power law for the rest of the years, except an exceptional fit of the latter in 2015.

3.5.3. Comparing the Fits

Making assessments about the which distribution fits the data better is a complicated task. Until now, power law and lognormal distributions are separately fit to the data via different procedures. Visual assessment and making comparisons regarding the two distributions solely depending on Figures 1-2-3 could be misleading, yet widely resorted to until recent years.

Table 3.4 provides two p-value scheme. The first depends on goodness of fit tests as suggested by Clauset et al. (2009), computed using bootstrapping technique in order to determine whether the proposed distribution is a good fit to the data. Next, a comparison procedure employing Kullback-Leibler Divergence (Kullback and Leibler, 1951) with likelihood ratios proposed by Vuong (1989) is followed. This method measures the loss of information from an observed distribution when approximated by another. Apparently, the log-ratio of likelihoods are positive if the power law distribution is better, negative if reverse holds and zero if their likelihoods are similar. This log-ratio of likelihoods are subject to variation and if statistically significant, Vuong's method provides a small p-value.

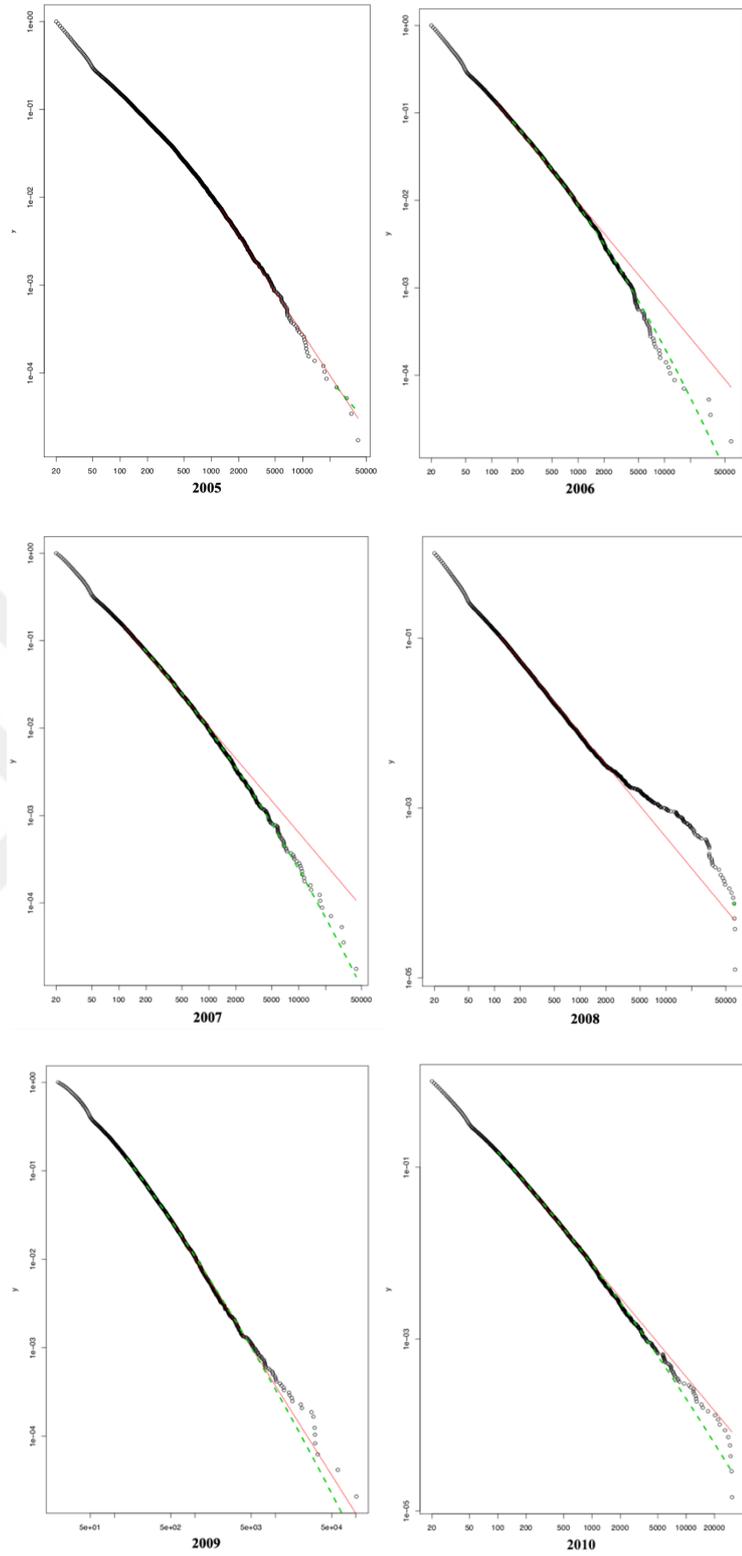


Figure 3.2: Power Law (Solid) and Lognormal (Dashed) Fits Between 2005-2010

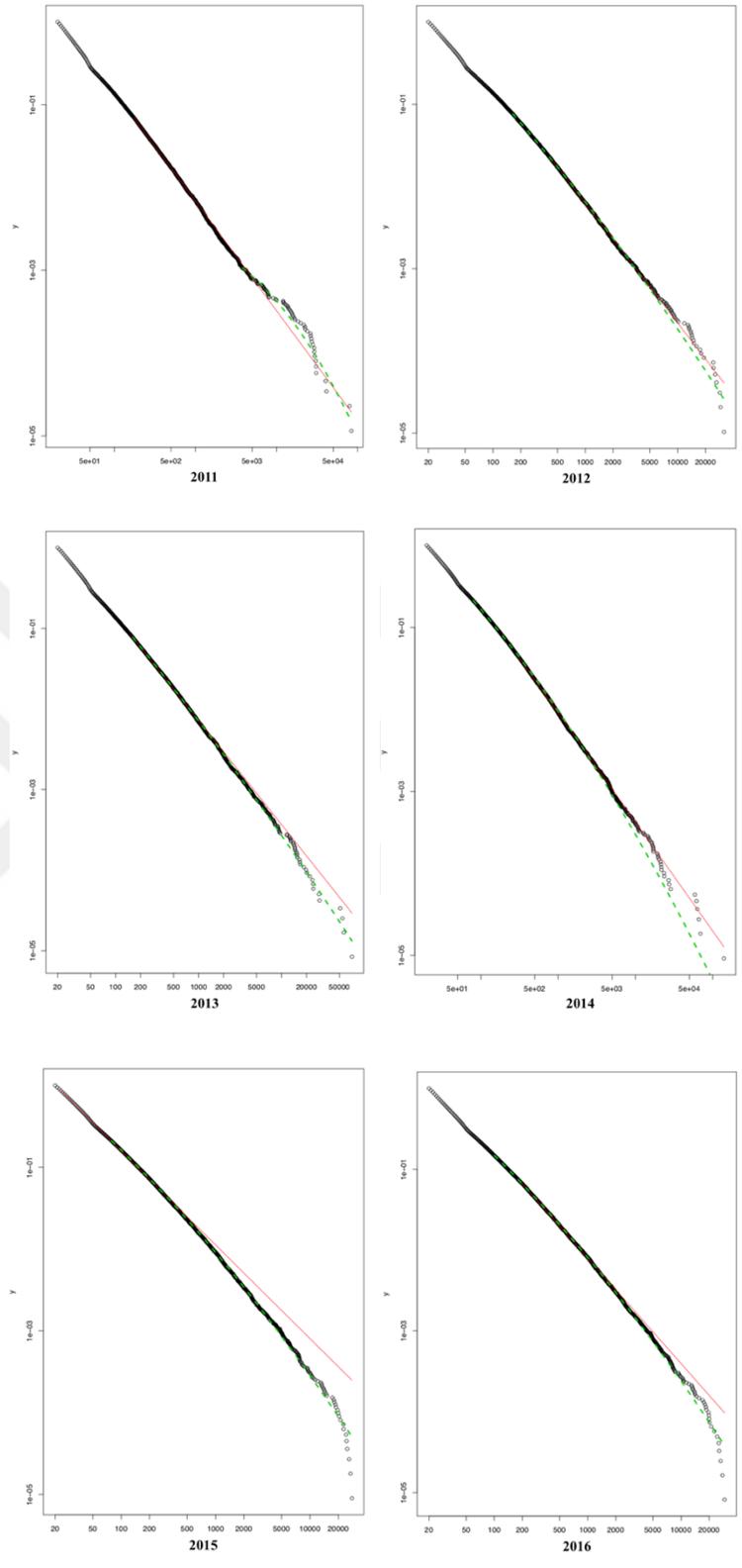


Figure 3.3: Power Law (Solid) and Lognormal (Dashed) Fits Between 2011-2016

The null hypothesis for panel (a) is that whether the data is generated from the subjected distribution, namely power law or lognormal. It can be seen that power law fit is found useful for years 2005, 2009, 2012. The goodness of fit test further favors lognormal for years 2006, 2007, 2010, 2015, 2016. Pointing out the inconclusiveness of an analysis conducted with a predetermined aim on a single distribution, both distributions are statistically confirmed for years 2011, 2014, and rejected for 2008 and 2013. Bootstrapping procedure is implemented via continuously generating datasets with same parameters as reported in Table 3.4, then re-measuring KS statistic, i.e. distance from proposed distribution, therefore obtaining a statistic for plausibility of the fit. During bootstrapping, some values of the scaling parameter make the first moment of the variable becomes infinite so the procedure failed for some years.

Table 3.4: Assessments and Comparisons of the Fits
(a) (b)

Year	(a) Separate fits			(b) Comparing fits		
	p_pwr	p_logn	Inference	p_value	L-ratio	Inference
2005	0.82	na	powerlaw	0.7	0.38	neither
2006	0	0.68	lognormal	0	-3.68	one is better
2007	0	0.98	lognormal	0	-3.24	one is better
2008	0.07	na	neither	error	error	na
2009	0.67	na	powerlaw	0.58	-0.54	neither
2010	0	0.48	lognormal	0.05	1.92	one is better
2011	0.45	0.4	both	0.15	-1.42	neither
2012	0.98	na	powerlaw	0.8	0.25	neither
2013	0.02	na	neither	0.12	-1.54	neither
2014	0.16	0.12	both	0.49	0.67	neither
2015	0	1	lognormal	0	-5.6	one is better
2016	0.01	0.59	lognormal	0.02	2.28	one is better

For panel (b) the null hypothesis is being equally away from the true distribution, so one needs to reject the null hypothesis to favor one of the proposed distributions with a small p-value.

It is seen that the likelihood ratio tests favor the lognormal distribution for years 2006, 2007, 2010, 2015 and 2016. For the rest of the years, firm size distribution in Turkey could not be successfully represented by neither power law not lognormal distribution.

Table 3.5: Algorithm to Compute p-value (Clauset et al., 2009)

Steps	Description
1	Estimations for x_{\min} and α values are calculated.
2	Kolmogorov-Smirnov statistic (KS_{data}) with the estimated values is obtained.
3	Let $P = 0$.
4	For i in $1 : \#$ of iterations
4-a	Generate a sample with $n - n_{\text{tail}}$ elements from $U(1; x_{\min})$,
4-b	Generate a sample with n_{tail} elements from powerlaw with $\hat{\alpha}$,
4-c	Compute the simulated KS statistic (KS_{sim}),
4-d	If $KS_{\text{data}} > KS_{\text{sim}}$ then $P = P + 1$,
4-e	end iterations.
5	$p = P/\#$ of iterations.

To capture the nature of the probabilities obtained, bootstrapping procedure from Clauset et al. (2009) is explained in Table 3.5 and Figure 3.4 exemplifies the output for the year 2009, obtained from R procedures from Gillespie (2015)¹. Top row of Figure 3.4 presents the movements of estimated values through iterations, while the last panel depicts the p-value. Bottom row depicts standard deviations of the estimates. In panels, the 95% confidence intervals for the parameters are also exhibited above and below of the lines. It is seen from Figure 3.4 that p-value of having a power law at the tail in 2009 is not stable even after 5000 iterations.

¹ Figures for the remaining years are supplied in the Appendix.

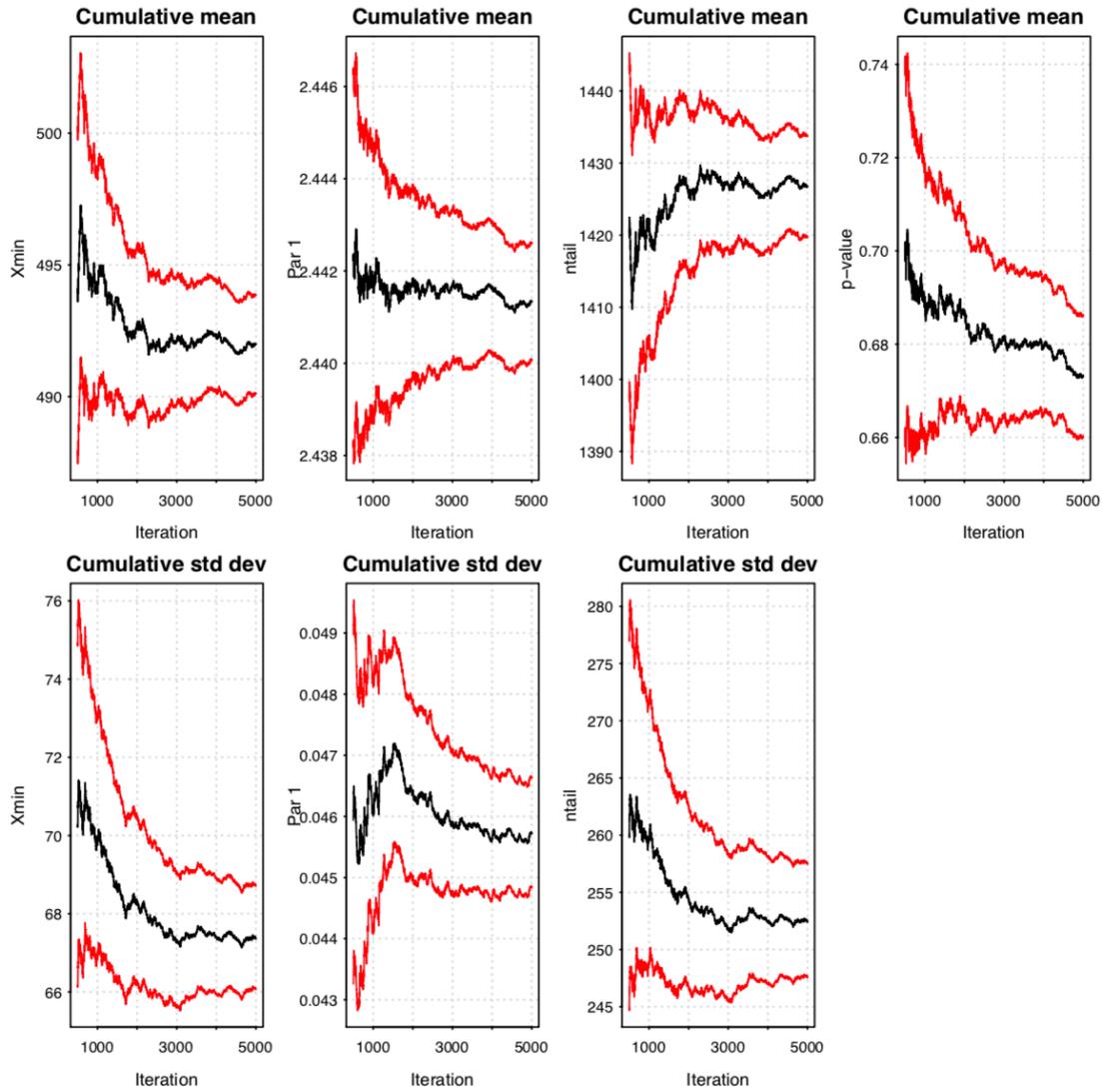


Figure 3.4: Bootstrap Output for the Year 2009

3.6. Conclusions

In this paper, 3.5.1 and 3.5.2 provide evidence that it is possible to fit a power law if intended, for all the subjected years. This fit could be also investigated and supported by visual assessments as depicted by Figures 3.1-3.2-3.3. As 3.5.3 shows, during the period under question, Turkish firm sizes are approximated well enough only by lognormal distribution for the years 2006, 2007, 2010, 2015 and 2016. None of the years are proven to be fit better by a power law. It turns out that computational procedures make difference and are essential to confirm distributional fits to the data.

Moreover, there are three years that observed distribution largely diverges from lognormal, namely 2005, 2008 and 2011. After all these divergent times, mildly except 2005, following years depict reversion to the lognormal distribution while also outperforming the power law fit. If we consider some sort of shocks as unusual boom and bust representing before and after those years, periodical lognormal fit to distribution of firm sizes shows great prospect of being related to possible macroeconomic relations.

For directing future research, some notes on firm level data and computational methods deserve attention. Working with big datasets is a major problem for firm size distribution analysis, in which case binned data becomes attractive for short term work. When the data is binned surely a great amount of information could get lost in distributional perspective. Despite the increase in CPU power in recent years, methods implemented in this paper, *e.g.* bootstrapping, are still being computationally burdensome and require workarounds with such datasets. Millions of data points with thousands of re-sampling is out of bounds of the commercial computers. Unfortunately, as in need for employing supercomputers, such possibilities are not available to most of the researchers even primarily due to the exclusiveness of the firm level data to institutions. Whether binning procedure gets labeled as “censoring”, “summarizing” or “compressing”, it is at least understandable due to the physical limitations on methodological implementations available yet.

3.7. Bibliography

- Adamic, Lada, Huberman, Bernardo. 2000. Power-law distribution of the world wide web. **Science**. v. 287. i. 5461: 2115-2115.
- Almaas, Eivind, Kovacs, Bela, Vicsek, Tamas, Oltvai, Zoltan, Barabási, Albert. 2004. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. **Nature**. v. 427. i. 6977: 839-843.
- Arshad, Sidra, Hu, Shougeng, Ashraf, Badar. 2018. Zipf's law and city size distribution: A survey of the literature and future research agenda. **Physica A: Statistical Mechanics and its Applications**. v. 492: 75-92.
- Artige, Lionel, Reginster, Alexandre. [13.07.2018]. On the Stylized Fact of the Firm Size Distribution. <http://dx.doi.org/10.2139/ssrn.3219439>
- Axtell, Robert. 2001. Zipf distribution of US firm sizes. **Science**. v. 293. i. 5536: 1818-1820.
- Barabási, Albert, Albert, Reka. 1999. Emergence of scaling in random networks. **Science**. v. 286. i. 5439: 509-512.
- Bee, Marco, Riccaboni, Massimo, Schiavo, Stefano. 2013. The size distribution of US cities: Not Pareto, even in the tail. **Economics Letters**. v. 120. i. 2: 232-237.
- _____, _____, _____. (2017). Where Gibrat meets Zipf: Scale and scope of French firms. **Physica A: Statistical Mechanics and its Applications**. v. 481: 265-275.
- Cabral, Luis, Mata, Jose. 2003. On the evolution of the firm size distribution: Facts and theory. **American Economic Review**. v. 93. i. 4: 1075-1090.
- Chen, Yanguang. 2012. The mathematical relationship between Zipf's law and the hierarchical scaling law. **Physica A: Statistical Mechanics and its Applications**. v. 391. i. 11: 3285-3299.
- Chen, Yanguang. 2012. Zipf's law, 1/f noise, and fractal hierarchy. **Chaos, Solitons and Fractals**. v. 45. i. 1: 63-73.
- Chen, Yanguang. 2016. The evolution of Zipf's law indicative of city development. **Physica A: Statistical Mechanics and its Applications**. v. 443: 555-567.
- Clauset, Aaron, Shalizi, Cosma, Newman, Mark. 2009. Power-law distributions in empirical data. **SIAM Review**. v. 51. i. 4: 661-703.
- Del Castillo, Joan, Puig, Pedro. 1999. The best test of exponentiality against singly truncated normal alternatives. **Journal of the American Statistical Association**. v. 94. i. 446: 529-532.
- Delli Gatti, Dominico, Desiderio, Saul, Gaffeo, Eduardo, Cirillo, Pasquale, Gallegati, Mauro. 2011. **Macroeconomics from the Bottom-up**. Springer Science & Business Media.
- Fazio, Giorgio, Modica, Marco. 2015. Pareto or log-normal? Best fit and truncation in the distribution of all cities. **Journal of Regional Science**. v. 55. i. 5: 736-756.

- Fragkias, Michail, Seto, Karen. 2009. Evolving rank-size distributions of intra-metropolitan urban clusters in South China. **Computers, Environment and Urban Systems**. v. 33. i. 3: 189-199.
- Gabaix, Xavier. 2009. Power laws in economics and finance. **Annual Review of Economics**. v. 1. i. 1: 255-294.
- Gabaix, Xavier, Ibragimov, Rustam. 2011. Rank- $1/2$: a simple way to improve the OLS estimation of tail exponents. **Journal of Business and Economic Statistics**. v. 29. i. 1: 24-39.
- Gangopadhyay, Kausik, Basu, Banasri. 2009. City size distributions for India and China. **Physica A: Statistical Mechanics and its Applications**. v. 388. i. 13: 2682-2688.
- Gibrat, Robert. 1957. **On Economic Inequalities**. trans. E. Henderson. International Economic Papers, The Macmillan Company.
- Gillespie, Colin. 2015. Fitting heavy tailed distributions: the powerLaw package. **Journal of Statistical Software**. v. 64. i. 1: 1-16.
- Gini, Giuseppina. 1922. **Indici di concentrazione e di dipendenza in Biblioteca dell'Economista**. Torino.
- Hart, Peter. 1962. The size and growth of firms. **Economica**. v. 29. i. 113: 29-39.
- Hart, Peter, Oulton, Nicholas. 1997. Zipf and the size distribution of firms. **Applied Economics Letters**. v. 4. i. 4: 205-206.
- Hart, Peter, Prais, Sigbert. 1956. The analysis of business concentration: a statistical approach. **Journal of the Royal Statistical Society**. v. 119. i. 2: 150-191.
- Hopenhayn, Hugo. 1992. Entry, exit, and firm dynamics in long run equilibrium. **Econometrica**. v. 60. i. 5: 1127-1150.
- Ioannides, Yannis, Skouras, Spyros. 2013. US city size distribution: Robustly Pareto, but only in the tail. **Journal of Urban Economics**. v. 73. i. 1: 18-29.
- Johnson, Norman, Kotz, Samuel, Balakrishnan, Narayanaswamy. 1994. **Continuous Univariate Probability Distributions**. Wiley.
- Jovanovic, Bojan. 1982. Selection and the Evolution of Industry. **Econometrica**. v. 50. i. 3: 649-670.
- Kalecki, Michal. 1945. On the Gibrat distribution. **Econometrica**. v. 13. i. 2: 161-170.
- Kesten, Harry. 1973. Random difference equations and renewal theory for products of random matrices. **Acta Mathematica**. v. 131. i. 1: 207-248.
- Kleiber, Christian, Kotz, Samuel. 2003. **Statistical size distributions in economics and actuarial sciences**. Wiley.
- Kullback, Solomon, Leibler, Ronald. 1951. On information and sufficiency. **The annals of mathematical statistics**. v. 22. i. 1: 79-86.
- Luckstead, Jeff, Devadoss, Stephen. 2014a. A comparison of city size distributions for China and India from 1950 to 2010. **Economics Letters**. v. 124. i. 2: 290-295.
- _____, _____. 2014b. Do the world's largest cities follow Zipf's and Gibrat's laws? **Economics Letters**. v. 125. i. 2: 182-186.

- _____, _____. (2017). Pareto tails and lognormal body of US cities size distribution. **Physica A: Statistical Mechanics and its Applications**. v. 465: 573-578.
- _____, _____, Danforth, Diana. 2017. The size distributions of all Indian cities. **Physica A: Statistical Mechanics and its Applications**. v. 474: 237-249.
- Luttmer, Erzo. 2007. Selection, growth, and the size distribution of firms. **The Quarterly Journal of Economics**. v. 122. i. 3: 1103-1144.
- _____. 2011. On the mechanics of firm growth. **The Review of Economic Studies**. v. 78. i. 3: 1042-1068.
- Newman, Mark. 2005. Power laws, Pareto distributions and Zipf's law. **Contemporary Physics**. v. 46. i. 5: 323-351.
- Niwa, Hiro-Sato. 2003. Power-law versus exponential distributions of animal group sizes. **Journal of Theoretical Biology**. v. 224. i. 4: 451-457.
- Okuyama, Kenji, Takayasu, Misako, Takayasu, Hideki. 1999. Zipf's law in income distribution of companies. **Physica A: Statistical Mechanics and its Applications**. v. 269. i. 1: 125-131.
- Pareto, Vilfredo. 1897. **Cours d'Economie Politique**. Lausanne.
- Peng, Guohua. 2010. Zipf's law for Chinese cities: rolling sample regressions. **Physica A: Statistical Mechanics and its Applications**. v. 389. i. 18: 3804-3813.
- Schumpeter, Joseph. 1949. Vilfredo Pareto (1848-1923). **The Quarterly Journal of Economics**, v. 63. i. 2: 147-173.
- Simon, Herbert, Bonini, Charles. 1958. The size distribution of business firms. **The American economic review**. v. 48. i. 4: 607-617.
- Singh, Vijay. 1998. **Two-Parameter Lognormal Distribution. In Entropy-Based Parameter Estimation in Hydrology**. Springer.
- Steindl, Joseph. 1965. **Random processes and the growth of firms: A study of the Pareto law**. Hafner.
- Stumpf, Michael, Porter, Mason. 2012. Critical truths about power laws. **Science**. v. 335. i. 6069: 665-666.
- Toda, Alexis. [22.09.2018]. Wealth distribution with random discount factors. <https://doi.org/10.1016/j.jmoneco.2018.09.006>
- Urzúa, Carlos. 2011. Testing for Zipf's law: A common pitfall. **Economics Letters**. v. 112. i. 3: 254-255.
- Veneri, Paolo. 2016. City size distribution across the OECD: Does the definition of cities matter?. **Computers, Environment and Urban Systems**. v. 59: 86-94.
- Vuong, Quang. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. **Econometrica**. v. 57. i. 2: 307-333.
- Zipf, George Kingsley. 1949. **Human behaviour and the principle of least-effort**. Addison-Wesley.

APPENDIX

Appendix to Chapter 2

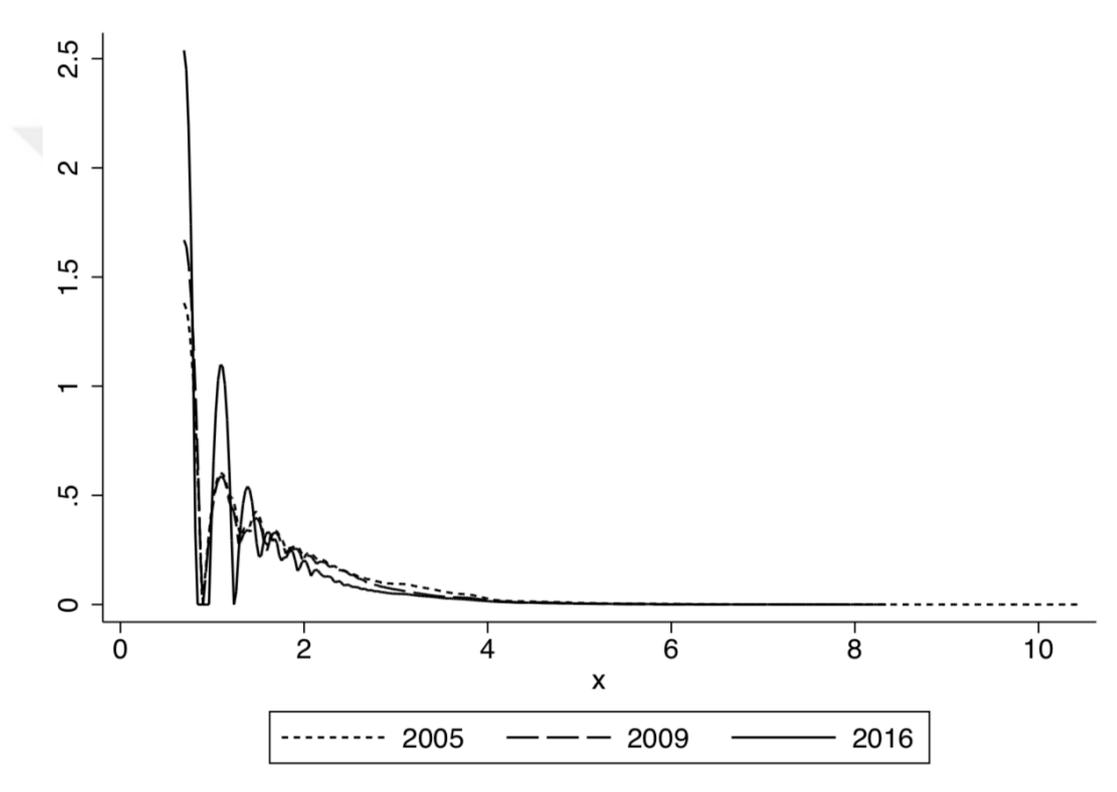


Figure A.1: Size Distribution of All Entrants, 2005, 2009, 2016

Table A.1: Sectoral Composition Between 2005-2008

NACE Rev. 1.1 Sections	2005	2006	2007	2008
Mining and Quarrying	40,349	5,260	5,057	5,845
Manufacturing	423,215	405,882	401,567	418,756
Electricity, gas and water supply	787	946	3,792	1,283
Construction	145,230	148,538	151,621	161,942
Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	1,362,515	1,270,194	1,254,967	1,279,102
Hotels and restaurants	272,923	245,219	250,139	255,004
Transport, storage and communication	606,197	574,004	567,392	566,387
Financial intermediation	62,653	59,566	58,147	48,525
Real estate, renting and business activities	238,496	227,043	201,616	207,206
Public administration and defence; compulsory social security	10,148	-	-	-
Education	15,808	16,268	17,369	12,579
Health and social work	52,071	46,898	46,692	45,698
Other community, social and personal service activities	213,332	206,130	211,684	215,651
Activities of households	8,721	-	-	-
Extra-territorial organizations and bodies	903	-	-	-

Table A.2: Sectoral Composition Between 2009-2012

NACE Rev. 2 Sections	2009	2010	2011	2012
Manufacturing	378,659	345,959	402,298	390,246
Electricity, gas, steam and air conditioning supply	1,443	1,389	1,891	1,551
Water supply; sewerage, waste management and remediation activities	867	685	1,539	2,194
Construction	168,894	156,576	204,764	194,657
Wholesale and retail trade; repair of motor vehicles and motorcycles	1,186,352	1,111,911	1,236,311	1,156,925
Transportation and storage	569,548	536,082	582,017	561,329
Accommodation and food service activities	260,488	242,814	289,388	293,705
Information and communication	21,018	18,674	25,997	31,112
Financial and insurance activities	49,677	46,515	40,768	24,290
Real estate activities	31,985	30,518	41,861	43,808
Professional, scientific and technical activities	153,084	144,677	172,043	169,861
Administrative and support service activities	36,907	30,935	39,474	38,735
Education	12,862	11,608	15,816	17,565
Human health and social work activities	42,109	36,066	39,962	36,424
Arts, entertainment and recreation	39,118	34,335	39,156	32,693
Other service activities	229,410	223,273	224,244	206,741
Activities of extraterritorial organisations and bodies	120	94	271	238

Table A.3: Sectoral Composition Between 2013-2016

NACE Rev. 2 Sections	2013	2014	2015	2016
Manufacturing	457,174	457,650	436,930	471,027
Electricity, gas, steam and air conditioning supply	4,578	5,955	2,256	10,932
Water supply; sewerage, waste management and remediation activities	3,612	4,550	4,447	5,362
Construction	253,747	266,080	216,158	280,484
Wholesale and retail trade; repair of motor vehicles and motorcycles	1,241,728	1,222,454	1,167,324	1,256,580
Transportation and storage	561,779	542,004	524,632	533,099
Accommodation and food service activities	305,526	307,698	305,361	323,030
Information and communication	40,842	42,639	37,517	44,047
Financial and insurance activities	25,964	24,118	20,830	23,779
Real estate activities	51,399	54,324	52,558	65,882
Professional, scientific and technical activities	193,791	202,260	194,551	222,444
Administrative and support service activities	52,148	58,132	53,504	64,446
Education	28,821	30,187	27,081	32,578
Human health and social work activities	42,838	44,050	39,447	47,386
Arts, entertainment and recreation	34,722	35,512	30,564	35,970
Other service activities	228,748	222,636	196,527	233,314
Activities of extraterritorial organisations and bodies	480	454	340	460

Appendix to Chapter 3

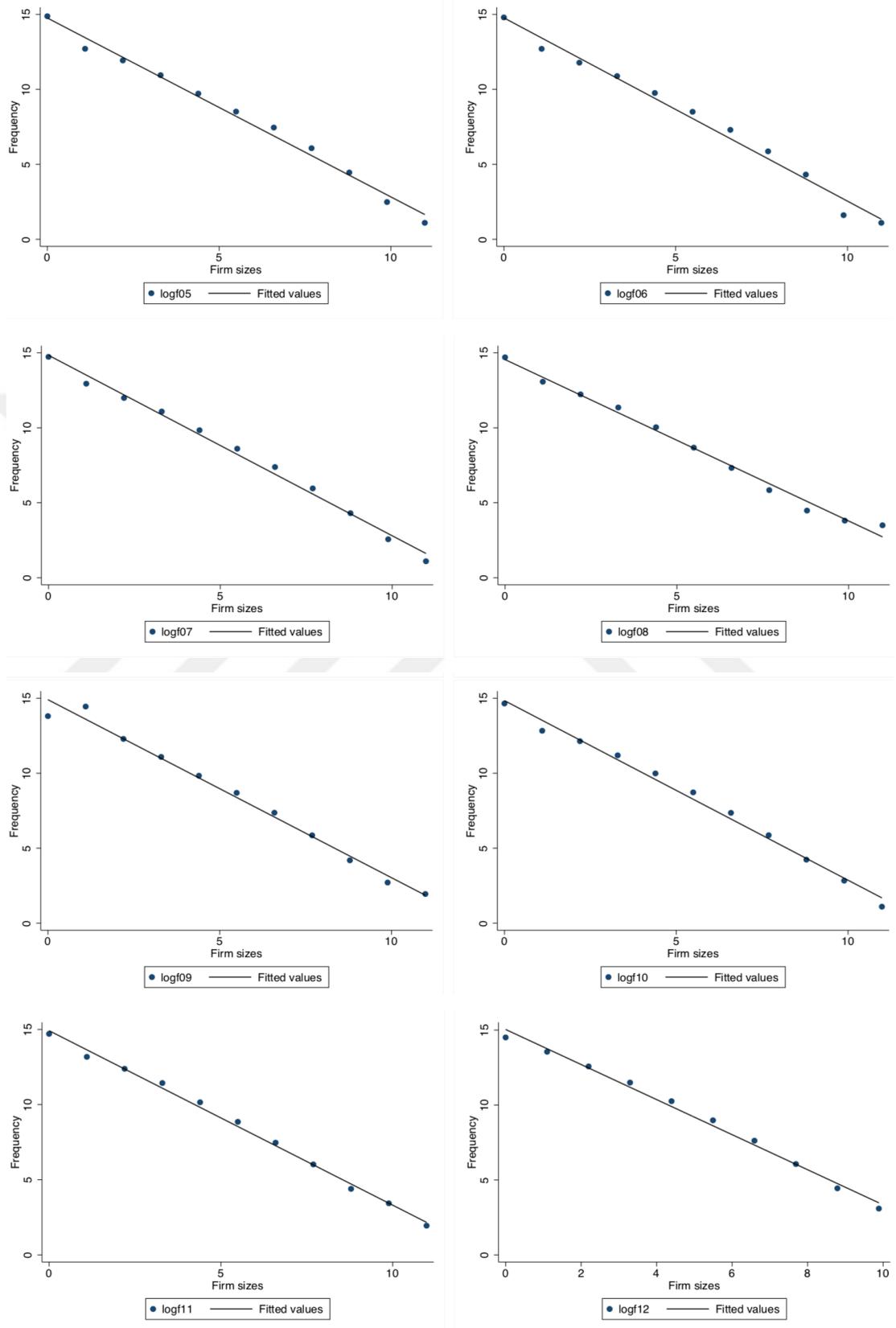


Figure B.1: All Histograms of the log-firm Sizes and Fits-1

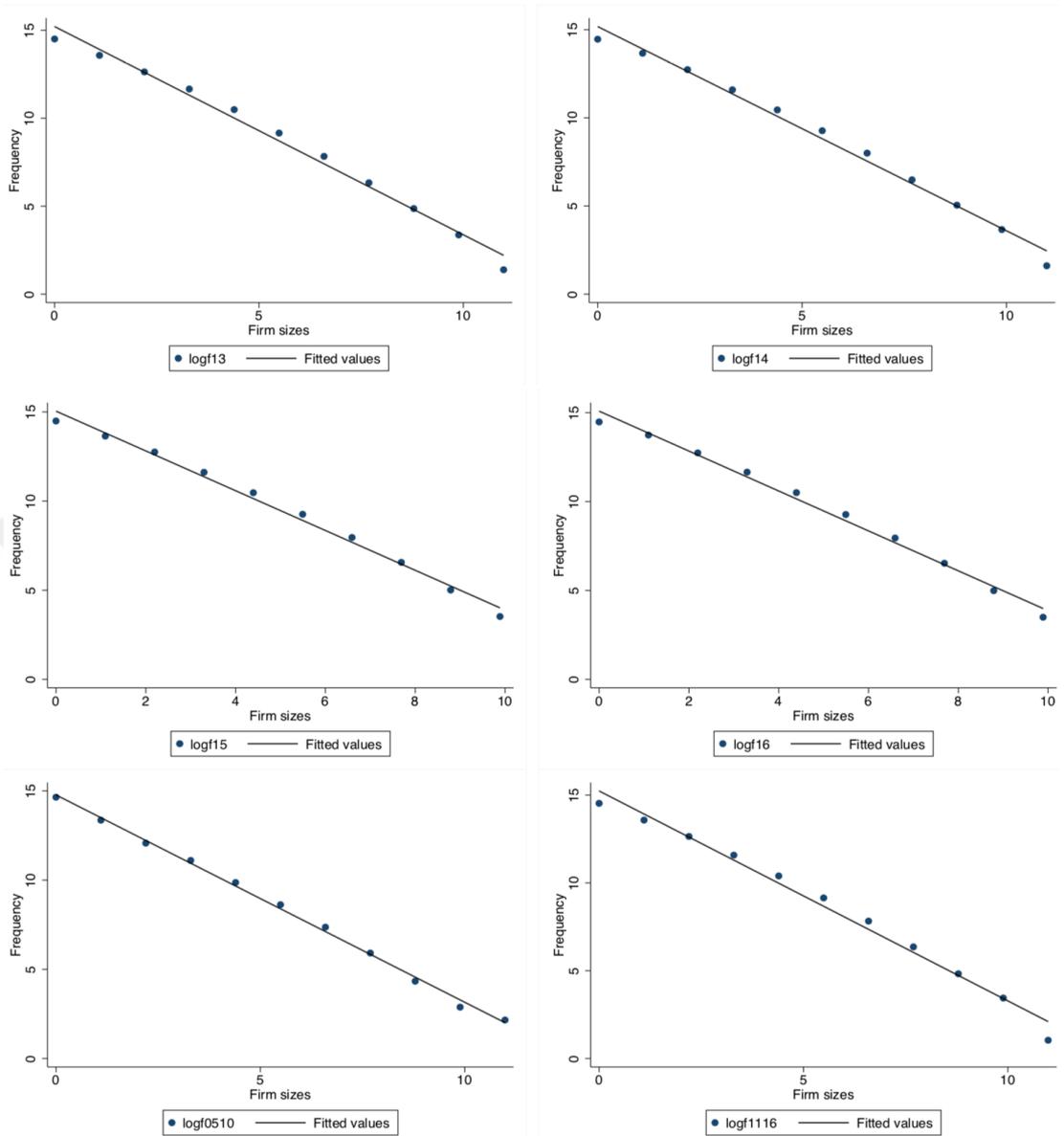


Figure B.2: All Histograms of the log-firm Sizes and Fits-2

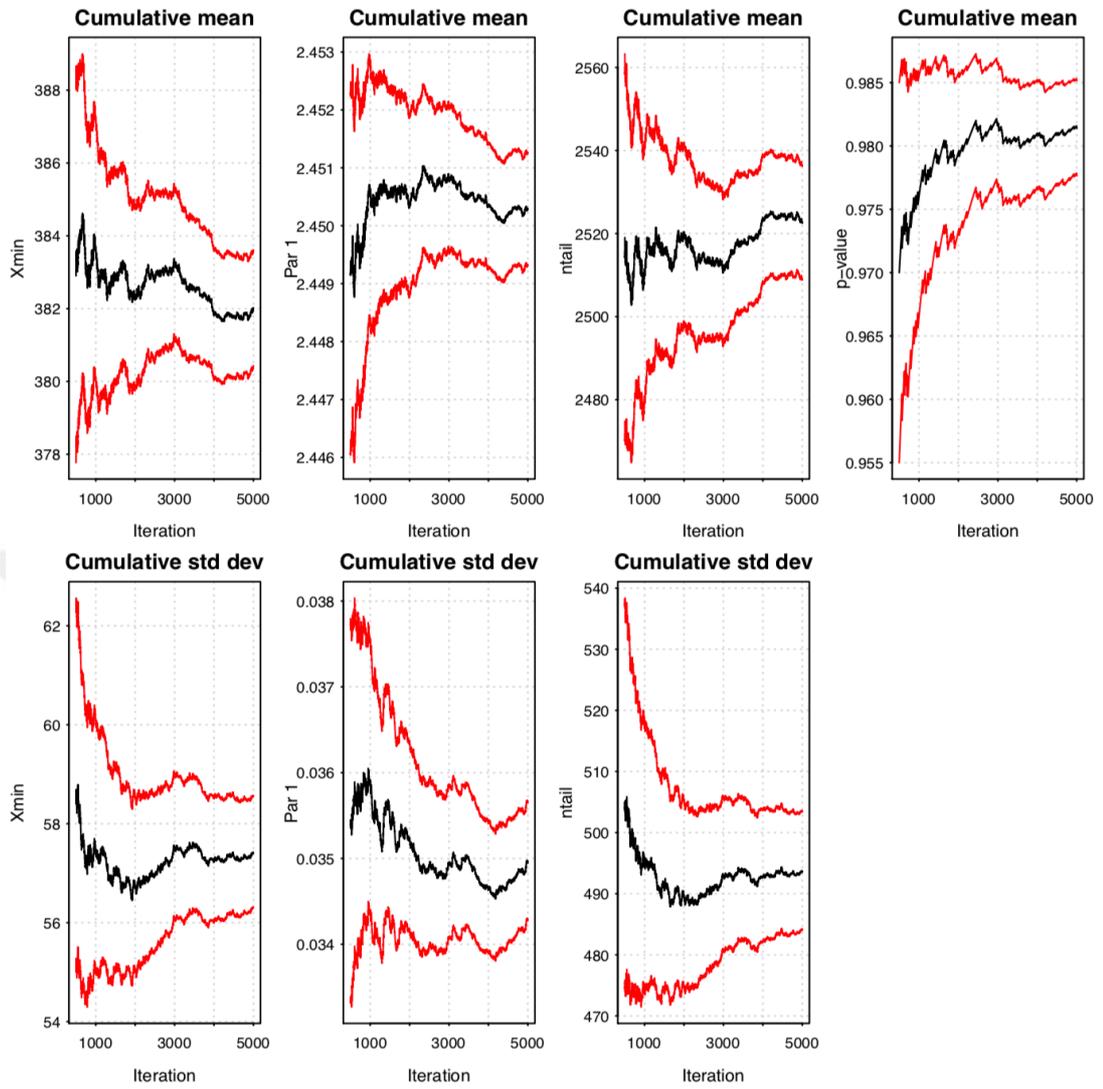


Figure B.3: Bootstrap Results of Powerlaw, 2012

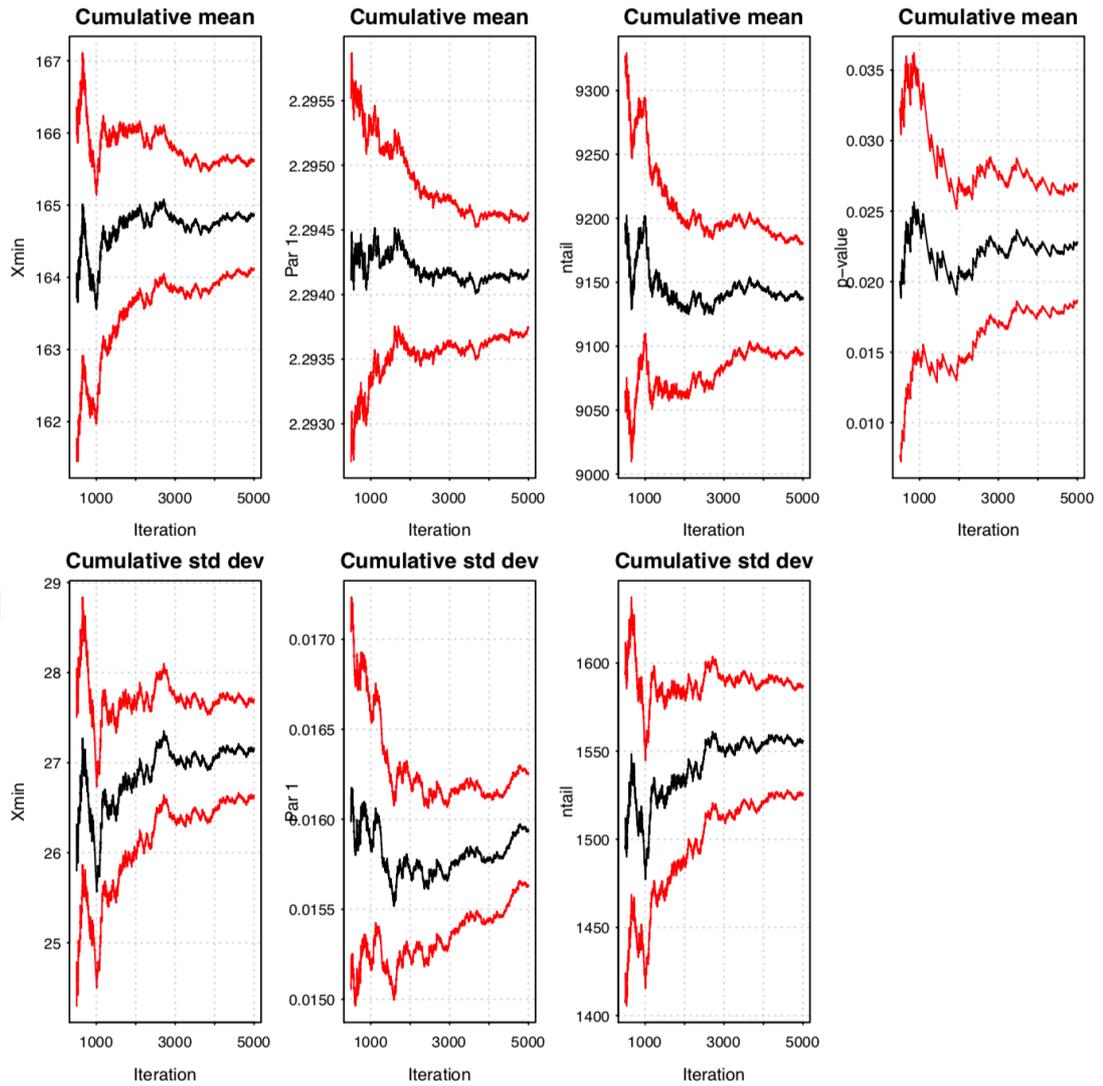


Figure B.4: Bootstrap Results of Powerlaw, 2013

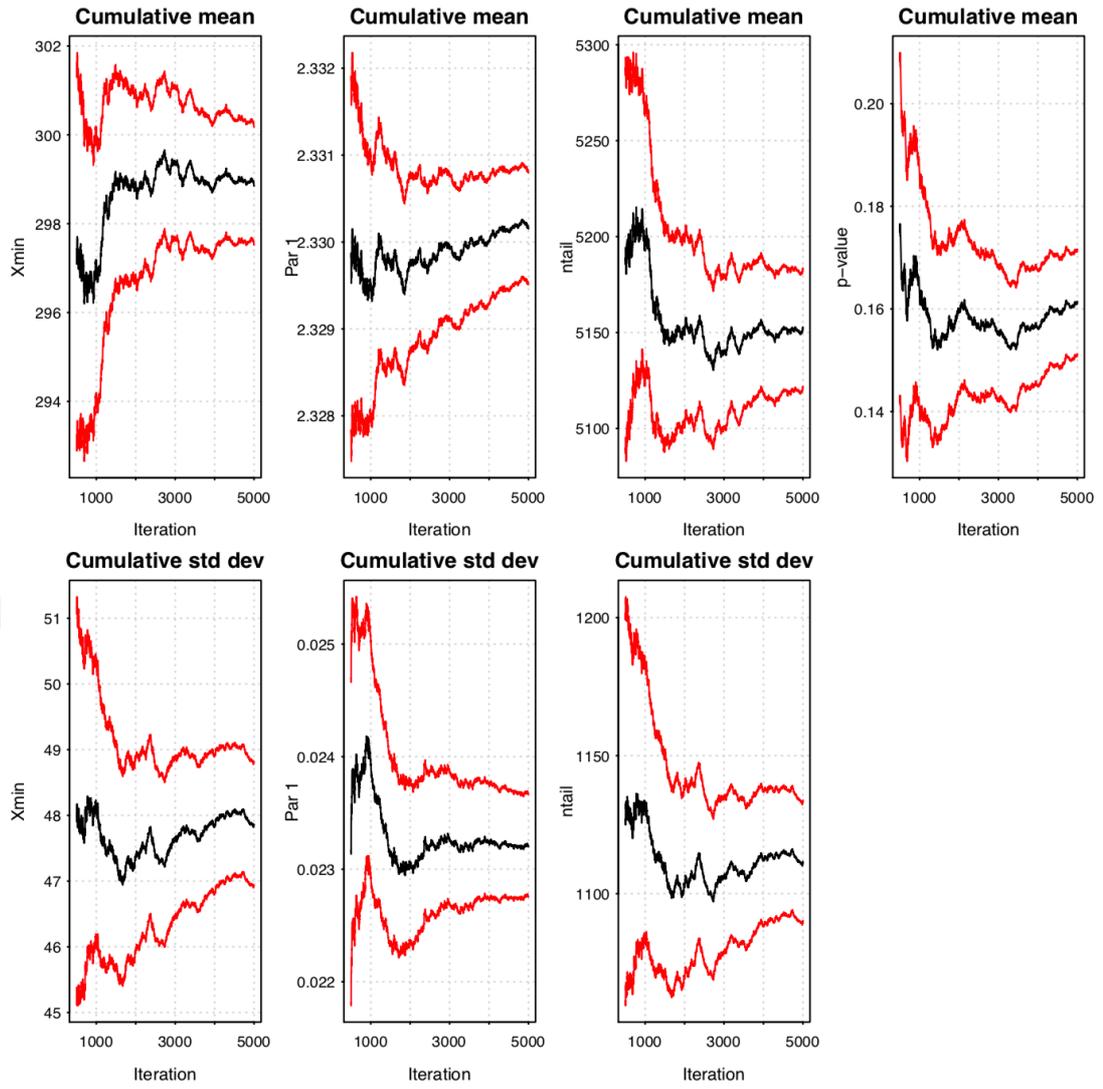


Figure B.5: Bootstrap Results of Powerlaw, 2014

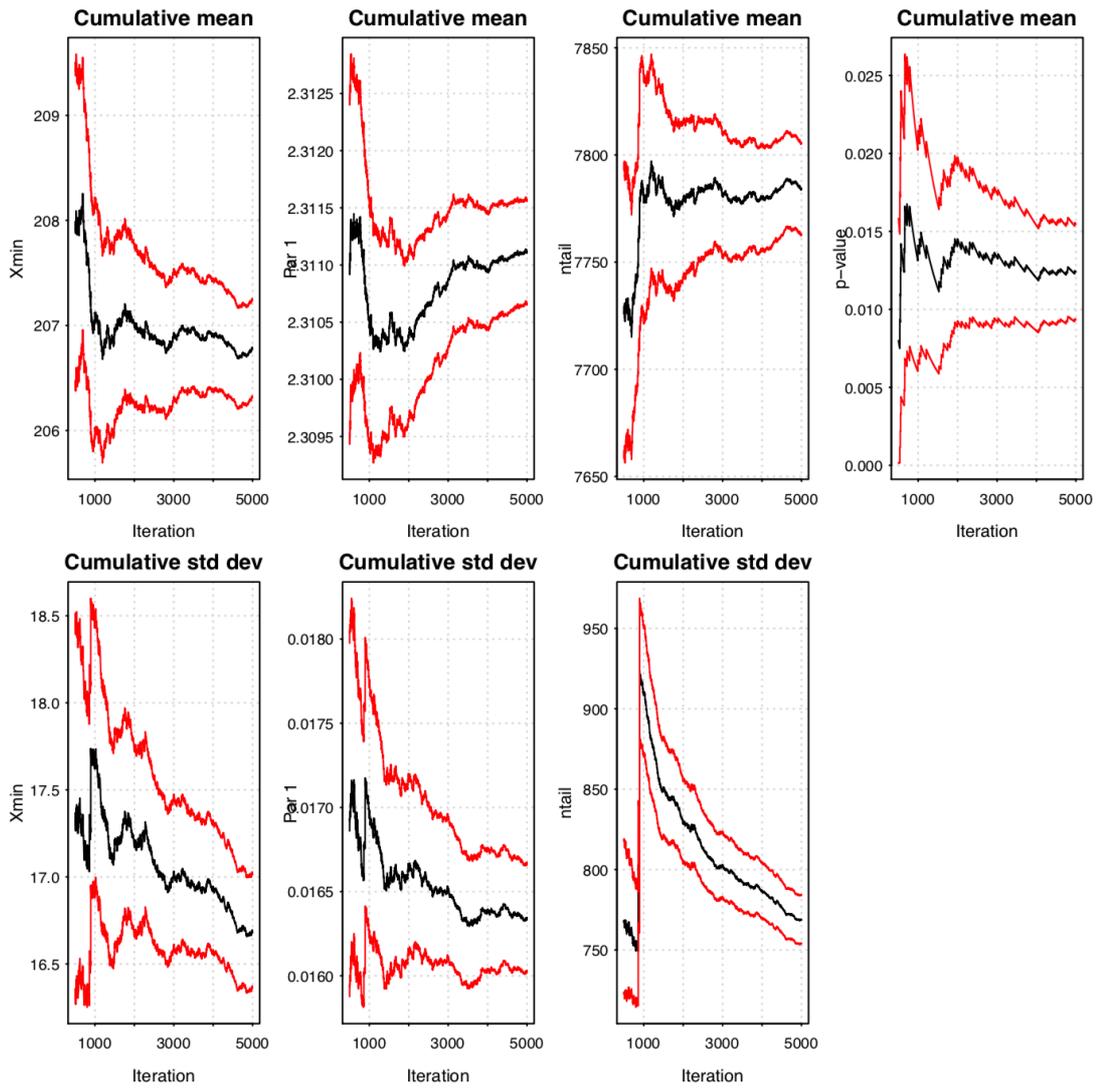


Figure B.6: Bootstrap Results of Powerlaw, 2016

RESUME

Yigit Aydogan

Kirklareli University, Department of Economics

Kirklareli - TURKEY

yigitaydogan@gmail.com

Date and Place of Birth

12/06/1988, Bakirkoy, Turkey.

Education

2013 – 2019:	Yildiz Technical University , Istanbul PhD in Economics (English)
2011 – 2013:	Ege University , İzmir M.A. in Economics
2005 – 2010:	Koç University , Istanbul B.A. in Economics (English) (with High Merit Scholarship)
YDS -2017	English, 96.25/100

Publications and Presentations

Articles:

- “Estimations on the Firm Size Distribution in Turkey” with Murat Donduran, *Journal of European Theoretical and Applied Studies*, Vol: 6(2), p. 35-51, 2018.

Presentations:

- “An Analysis of the Firm Size Distribution in Turkey” with Murat Donduran, X. International Balkan and Near Eastern Social Sciences Congress Series – Ohrid/Macedonia, Oct. 27-28, 2018.
- “Assessment of the Current Account Deficit and Economic Growth Relationship within PIIGS Countries”, with Filiz Erataş, Third International Conference in Economics – ECONAnadolu 2013, June 19-21, 2013.

Work Experience

May'13 - present	Kirklareli University , Research Assistant, Economics Department
Oct'12 - May'13	Celal Bayar University , Research Assistant, Economics Department

Jan'12 - Oct'12	Ege University , Research Assistant, Economics Department
Apr'11 - Jul'11	Reckitt Benckiser Turkey , Sales Dept. – PWR Sales Executive, İzmir
Jul'10 - Apr'11	Reckitt Benckiser Turkey , Sales Dept. – PWR Sales Executive, Kayseri Customer&Distributor Relationship Management, full responsibility of brand performances on field.
Mar'10 - Jun'10	Unicredit Group Istanbul – Corporate Finance Full Time Intern/Analyst Mergers&Acquisitions, Investment Financing, IPOs, SPOs
Nov'09 - Dec'09	Peppers&Rogers Group Marketing Consultancy – Research Intern
Sept'09	Koç Holding Headquarters – Corporate Communications and Social Responsibilities Projects Intern
Apr'09 - Aug'09	Avea Headquarters – Marketing Research Full Time Intern
Jun'08 - Jul'08	ING Bank Headquarters Istanbul, Career Academy Internship Program in Cash Management Department