
Munich Personal RePEc Archive

The impact of the 2008 crisis on top labor incomes in Turkey: A nonparametric analysis

Tumen, Semih

Central Bank of the Republic of Turkey

21 May 2013

Online at https://mpra.ub.uni-muenchen.de/47112/
MPRA Paper No. 47112, posted 21 May 2013 18:17 UTC

The Impact of the 2008 Crisis on Top Labor Incomes in Turkey: A Nonparametric Analysis*

Semih Tumen †
Central Bank of the Republic of Turkey

May 21, 2013

Abstract

This paper presents a nonparametric analysis of the impact of the 2008 crisis on earnings distribution in Turkey. Using micro-level data from the Household Labor Force Survey (2004 – 2011), I show that the crisis has operated most visibly above the upper quartile of the earnings distribution. I present three main findings: (1) the share of the top decile—especially the top percentile—has increased significantly right after the crisis, (2) the top quartile (i.e., the right tail) of labor incomes in Turkey resembles, on average, a Pareto distribution with a corresponding Gini coefficient of around 0.23, and (3) following the crisis, the earnings differentials have widened above the top quartile and wildly deviated from the Pareto form. I document that the changes observed right after the crisis have mostly been temporary; that is, a normalization process has operated in the aftermath of the crisis. I argue that the health of the domestic banking system might be an important determinant of the effect of large scale financial crises on top labor incomes. A sound banking system can generate income polarization in a country, when global crises lead to asymmetric income reallocations across sectors.

JEL codes: C14, D33, J31.

Keywords: Earnings differentials; inequality; the 2008 crisis; Turkish Household Labor Force Survey.

*I thank Hakan Ercan, Bradley Heim (the editor), Thomas Lange, Hakki Yazici, Tugba Zeydanli, and seminar participants at Hacettepe University for helpful comments and suggestions. The views expressed here are of my own and do not necessarily reflect those of the Central Bank of the Republic of Turkey. All errors are mine.

†semih.tumen@tcmb.gov.tr. Research and Monetary Policy Department, Central Bank of the Republic of Turkey, Istiklal Cad. No:10, 06100 Ulus, Ankara, Turkey.

1 Introduction

Top incomes capture only a small fraction of the population, but they represent a quite significant share of total income and total taxes paid in an economy [Atkinson, Piketty, and Saez (2011)]. Hence, the extent of inequality is very sensitive to movements in top income shares. The same is true when the focus is only the labor income. The labor markets have recently been generating considerable inequality in developed countries, especially in the United States, with the very top earners capturing a large fraction of total income. Estimates based on US earnings data reveal that labor income earned at the top percentile has surged from 5.1 percent in 1970s to 12.4 percent in 2007 [Saez (2008)]. Similar changes have also been recorded in most European countries and Japan, in slightly smaller magnitudes [Piketty, Saez, and Stantcheva (2013)].

In the literature, there is an unsettled issue that I want to highlight: the effect of economic crises on top labor incomes. Although the underlying mechanisms are still unclear, the main consensus is that the degree of earnings inequality generally goes down right after the shock [Atkinson and Morelli (2011)].[1] To check whether this consensus can turn into a general rule and to understand the underlying mechanisms that can connect crises to inequality, a larger volume of empirical evidence from studies based on individual country experiences are needed.

In this study, I investigate the effect of the 2008 crisis on top labor incomes and, more generally, on earnings differentials in Turkey. The Turkish case is interesting, because Turkey has demonstrated one of the most resilient macroeconomic performances as a response to the crisis. Past unpleasant experiences—mainly the crises of 1990s—had revealed severe deficiencies in the Turkish banking system. The burden was so large that all of the financial regulations and policies designed in the last 10 years targeted a strong, stable, and credible banking system so as not to face similar catastrophic experiences again. Having this strong banking system in the background, the negative effects of the 2008 crisis have only been transitory, sector-specific, and moderate.

[1] For alternative explanations on the potential mechanisms, see, for example, Kumhof and Ranciere (2010), Fitoussi and Saraceno (2010), Rajan (2010), and Philippon and Reshef (2012).

I group the empirical results under three main categories. **First**, I find that the share of the top decile, especially the top percentile, has increased significantly right after the crisis. In particular, I show that the share of the top percentile is stabilized around 5 percent for the period 2004–2008, but it jumped to 8 percent in 2009. A normalization process has operated thereafter, pushing the share of the top percentile back to around 5 percent. This is in contrast with what the empirical literature on the link between crises and inequality reports. For example, Roine, Vlachos, and Waldenstrom (2009) find using cross-country data covering the period 1900–2000 that a financial crisis—domestic or international—would permanently reduce the share of the top 1 percent by about 0.2 percentage points for each year of the crisis. Using data from the US, Morelli (2011) shows that the top percentile has been negatively affected for three years following the financial crisis, before starting to recover.

Although this result sharply deviates from what the literature documents, it nevertheless communicates important insights regarding the mechanisms that may be potentially affecting the changes in the income share of the top percentile during financial crises. First of all, based on real macroeconomic indicators, the effect of the 2008 crisis on the Turkish economy looked severe. The GDP growth was -5 percent in 2009 and the rate of unemployment in the non-agriculture sector jumped from 12 percent to 19 percent. But the banking sector was still healthy and the repercussions were never felt like the waves of a banking shock. Profits in the banking sector were quite high, which kept the salaries high and generated large bonuses for top employees.[2] Based on this observation, I argue that the health of the domestic banking system might be an important determinant of the effect of global financial crises on top one percent earners [see Bell and van Reenen (2010) for a similar insight]. This is an interesting case, because it suggests that a sound banking system can generate income polarization in the society, when global crises lead to asymmetric income reallocations across sectors.

The main mechanism at work might have been operating through the structure and the composition of the financial assets. For example, in the U.S., the "bad" assets have been the major problem; they pulled the returns and profits down in the entire economy. The negative

[2]This is in sharp contrast with the case in the US and most of the other developed countries. Lee and Rose (2010) document that bank profitability was damped and the whole banking system was under significant liquidity pressure in 2009.

consequences of the crisis have been persistent and severe. The Turkish banking system, on the other hand, did not have a significant amount of bad assets. The crisis hit the economy through temporary capital outflows—i.e., outflow of the so-called hot money. The resulting temporary increase in the returns of interest bearing assets opened up huge profit opportunities for asset holders. This in turn inflated earnings, including performance pay and bonuses.

The 1997–98 Asian crisis provides an example similar to the Turkish experience. In Singapore, top income shares increased significantly following the crisis [Atkinson, Piketty, and Saez (2011)], which was basically a regional shock impinging on individual countries. The rise in top income shares in Singapore after the crisis was remarkable but only temporary. Just as in the Turkish case, Singapore had a strong banking system relative to a majority of the countries in the region and their government performed an active crisis management strategy.

Second, the top quartile (i.e., the right tail) of the earnings distribution in Turkey resembles, on average, a Pareto distribution with a corresponding Gini coefficient of roughly 0.23. This is an important finding for two reasons. (1) The Pareto distribution has the property that the distance between observations is constant for all points over the distribution. If these distances define the extent of "deprivation" in the society, than the dispersion parameter of the Pareto distribution is a good measure of social distress due to inequality.[3] (2) A Gini coefficient of 0.23 (against an overall Gini coefficient of 0.40) implies that deprivation below the top quartile is much higher than that above the top quartile. In other words, income differentials above the top quartile have been much smaller than those below the top quartile.

Third, 2009 has been the only year within the 2004–2011 period in which the earnings distribution above the top quartile has widely deviated from the Pareto form. In this year, the earnings differentials followed a U-shaped pattern meaning that the differentials have been much higher at the top and the bottom than they are at the middle of the distribution. This suggests that deprivation in the society has gone up significantly in 2009, given the definition above. This is interesting because, as I discuss above, improved stability of the banking system has led the Turkish economy to develop a resilient macroeconomic performance following the

[3] See Cowell (2011) for extensive information on the properties of the distributional forms often used in the study of inequality.

global financial crisis, but it has also led to increased levels of distress in the society. This is also related to the pay comparisons literature. For example, Clark and Oswald (1996) find that a deterioration in the relative income perceptions in the society can lead to decreased job satisfaction and satisfaction from life in the overall. I conclude that a strong banking system can itself generate income polarization and deprivation in a country, when global crises lead to asymmetric income reallocations across sectors.

In terms of the statistical techniques used, this paper is closest to Saez (2001). Saez develops a method to measure income differentials over the earnings distribution to study optimal income taxation. I use his method to analyze the impact of the 2008 crisis on earnings differentials in Turkey. This method can be classified as a nonparametric method, because it enables the use of the raw data for the purpose of mapping the empirical earnings distribution into a certain set of formulas, using which one can estimate the effects of the crisis on several distributional outcomes. By using a non-parametric technique, I avoid imposing *a priori* functional form assumptions on the structure of earnings differentials.

The plan of the paper is as follows. Section 2 describes the details of the statistical model. Section 3 summarizes the data, presents the results, and elaborates the findings. Section 4 concludes.

2 The Statistical Model

The statistical model builds on the method developed by Saez (2001). The main object that I focus attention throughout the paper is a truncated mathematical expectations operator. Let $X \in \mathcal{X}$ be a random variable with support set $\mathcal{X}$ and let x denote its realizations. I also let $F_X(x)$ and $f_X(x)$ denote the cumulative distribution function (cdf) and the probability distribution function (pdf) of this random variable, respectively. For algebraic simplicity, I assume that these functions are absolutely continuous in the support set $\mathcal{X}$. Then,

$$\mathbb{E}[X|X > x] = \int_x^\infty \frac{z}{\int_x^\infty dF_X(z)} dF_X(z) \tag{2.1}$$

is the (upper) truncated mean of the random variable X, where x is the truncation reference point.

Let $d(x)$ be a twice continuously differentiable—ensured by the continuity of F_X—function of x as follows:

$$d(x) = \frac{\mathbb{E}[X|X > x]}{x}. \tag{2.2}$$

The intuition is the following. $d(x)$ measures the distance between the observations on the right tail. A higher $d(x)$ corresponds to a case in which these distances are larger. To put this statistical definition into specific context, suppose now that x refers to real earnings observations in a nationally representative micro-level dataset. If the depression stemming from being poor is proportional to the income gap, then $d(x)$ can be interpreted as a measure of deprivation in the society. In other words, it is a comparison of the specific real earnings observation x to the mean real earnings above x. The functional form of $d(x)$ depends on the shape and curvature of the cdf F_X.

A particular functional form that is used in the literature to study top incomes is the Pareto distribution. Let $(\sigma - 1)X^{-\sigma}$ be the pdf of a Pareto random variable X over the support $[1, \infty)$. I assume $\sigma > 2$ to ensure a finite variance. Lower σ means that the dispersion (and, therefore, inequality) is higher. The following proposition holds under the Pareto distribution assumption.

Proposition 1. *If X is a Pareto random variable with a dispersion parameter $\sigma > 2$, then $d(x)$ is a constant. Specifically,*

$$d(x) = \frac{\sigma - 1}{\sigma - 2}. \tag{2.3}$$

Proof: The proof simply follows applying the integration by parts formula on the integral

$$\frac{1}{x} \int_x^\infty z dF_X(z). \blacksquare$$

In words, when F_X is of the Pareto form, then $d(x)$ is a constant described perfectly by the dispersion parameter σ.[4] In an application related to optimal high-income taxation in the US economy, Saez (2001) shows that the upper tail of the US earnings distribution resembles a Pareto distribution because $d(x)$ is a constant.

In this framework, taking the truncated mean $\mathbb{E}[X|X > x]$ to data without invoking a specific functional form in order to identify the properties of $d(x)$ corresponds to a nonparametric identification strategy. Remember that $d(x)$ is a direct measure of income differentials at the top of the earnings distribution. Parallel to this definition, I employ such a strategy to study the impact of the 2008 crisis on the differentials at the top of the earnings distribution. Specifically, I use the raw data to calculate $d(x)$ for each potential realization x of the random variable X. The next section describes the data I use and presents the main findings in detail.

3 Data and Findings

I use micro-level data from the Turkish Household Labor Force Survey (THLFS) conducted by the Turkish Statistical Institute (TURKSTAT) for the period 2004–2011. It is a large, survey-based, and nationally-representative dataset based on which the official unemployment and earnings statistics have been calculated and published. The sample I choose consists of employed individuals of age 16 and above. The earnings series are deflated taking 2004 as the base year; that is, I work with real earnings. The survey asks the after-tax monthly earnings (i.e., wage income plus other income including performance pay and bonuses). I annualize the monthly data in the analysis. Appropriate frequency weights are used in all calculations.

Figure (1) plots the main findings. This is basically a plot of $d(x)$ (i.e., the coefficient) versus x, where x denotes the real earnings (plotted in the log scale). The plot is based on a year-by-year estimation of the coefficient, $d(x)$. The figure says that the distances are large for

[4] Consider a Weibull example. Then $\alpha\lambda X^{\lambda-1}e^{-\alpha X^{\lambda}}$ is the pdf of the Weibull random variable over the support $[0,\infty)$. $\lambda > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter. Then

$$d(x) = \alpha^{-1/\lambda}\frac{e^{\alpha x^{\lambda}}}{x}\left[\Gamma\left(1+\frac{1}{\lambda}\right)+\gamma\left(1+\frac{1}{\lambda},\alpha x^{\lambda}\right)\right], \tag{2.4}$$

where $\Gamma(\cdot)$ is a Gamma function and $\gamma(\cdot)$ is an incomplete Gamma integral. In this more general form, $d(x)$ is a function of x, rather than a constant.

the low earners, while they become smaller as real earnings increase. After some point, which roughly corresponds to the cutoff for the top quartile, the differentials are stabilized for all years in the data horizon, except 2009. In 2009, however, earnings differentials above the top quartile of the distribution have widened nonnegligibly contributing to increased polarization, inequality, and deprivation in the society.

Leaving out the year 2009, I observe a pretty stable and consistent pattern: in the upper quartile of the earnings distribution, $d(x)$ roughly corresponds to a constant, which is approximately 1.57. From the Proposition 1, when $d(x)$ is a constant, the earnings distribution resembles the Pareto form. Formally, $d(x) = (\sigma - 1)/(\sigma - 2)$, where $\sigma > 2$ is the dispersion parameter of the earnings distribution. $d(x) = 1.57$ means that $\sigma \approx 3.75$. From Aaberge (2007), the Pareto distribution has a direct relationship with the Gini coefficient, $\mathcal{G}$, *via* the following formula:

$$\mathcal{G} = \frac{1}{2(\sigma - 1) - 1}. \tag{3.1}$$

Plugging $\sigma = 3.75$ into this formula yields $\mathcal{G} \approx 0.23$. Considering that the estimated Gini coefficient for the entire earnings distribution is around 0.40 (see TURKSTAT), I conclude that the degree of inequality is much lower within the top quartile in Turkey, excluding 2009. Moreover, the level of deprivation in the upper quartile is a constant.

In 2009, however, I observe a wild deviation from the Pareto form, i.e., $d(x)$ is no more constant. In fact, an eyeball test reveals that $d'(x) > 0$ in the top quartile. This means that distances among high earners (i.e., earnings above the top quartile) have increased systematically following the crisis. Although, the distances have gone back to pre-crisis levels within a year or so, the reaction of top incomes in 2009 has been quite remarkable. If earnings differentials describe the extent of deprivation (or social distress) in the society, then the estimates suggest that deprivation and polarization has gone up in 2009.

Another interesting observation is related to the movements in the share of the top percentile. Figure (2) plots the time-series behavior of the income share of the top one percent earners.

The earnings share of the top percentile sharply increased to 8 percent in 2009, while it remained close to 5 percent before and after 2009. Same is true when one looks at the mean earnings in the top percentile rather than shares [see Figure (3)]. Clearly, there is a jump in 2009. The mean annual earnings in the top quartile have become around 66,000 Turkish Liras (in 2004 terms), whereas this amount stays around 40,000 Turkish Liras before and after 2009. The reported gap in mean earnings in the top percentile is more than 50% when the pre- and post-crisis periods are compared. This suggests that widening earnings differentials in the top quartile comes mainly from the developments in the top percentile. However, as Table (1) suggests, there is a widening in earnings differentials in the top decile. This means that top earnings have moved up altogether, but the movement is larger in the top percentile.

I also check whether or not any compositional shifts in the observable worker characteristics can explain the jump in the top percentile in 2009. Table (2) summarizes the demographic composition of the top percentile individuals in the sample from 2004 to 2011. I see no obvious demographic compositional change pattern that may directly lead to a jump in the top percentile earnings shares in 2009. Nevertheless, I see slightly more females than males, younger than older, married than nonmarried, urban than rural, tenured than nontenured, and educated than uneducated workers within the top percentile in 2009.

The literature investigating the determinants of the movements in top income shares focuses on individual country experiences to be able to put together a coherent story that would enhance our understanding of the main mechanisms at work. In this study, I concentrate on the effect of the 2008 crisis on top labor income shares and on earnings differentials in Turkey. The consensus in the existing literature—based on country experiences—is that crises, especially banking crises, tend to reduce top income shares. Roine, Vlachos, and Waldenstrom (2009) show that the share of top incomes is positively correlated with GDP growth. Specifically, they find using cross-country data that a financial crisis permanently reduces the share of top percentile by about 0.2 percentage points for each year of the crisis.

However, as my findings suggest, the share of the top labor incomes has increased right after

the crisis in Turkey. The increase is observed above the top decile, but the most notable movement is in the top percentile. This is in stark contrast with the well-documented positive relationship between the GDP growth rates and top income shares. What is the explanation? I propose the health of the local banks as an important determinant of the link between crises and top income shares. The Turkish banks were quite "healthy" when the crisis hit. Interest rates went up following the initial shocks. This lead to tightening in the credit markets, which led to deterioration in the financial conditions of the small firms in the real economy; but, at the same time, banks made enormous profits through their holdings of interest bearing assets (mostly treasury bills). The temporary increase in the returns of interest bearing assets opened up huge profit opportunities for asset holders, which were large firms and banks. This in turn inflated earnings, including performance pay and bonuses. As a result, in the Turkish example, small firms and entrepreneurs suffered, but large firms, banks, and high earners benefited the crisis.

I conclude that in assessing the effects of crises on inequality, a well-founded link between the financial sector and the individual-level earnings has to be established. The Turkish case study and its comparison with the US/European experiences suggest that how the crises affect the financial sector is the key to understanding how a negative shock can diffuse into the income distribution in an economy. Theoretical macro models studying the effect of crises on income inequality should, therefore, focus on the financial sector and its role in transmitting the shock waves to the rest of the economy.

4 Concluding Remarks

In this paper, I use a nonparametric technique to estimate the effect of the 2008 crisis on earnings differentials in the Turkish economy. The Turkish case is particularly interesting, because the strong banking system in Turkey has led the economy to react much more efficiently to the crisis than most of the developed countries did. The focus is on labor income. I show that earnings differentials above the top quartile have widened within the year following the crisis. Moreover, the share of the top percentile has jumped to 8 percent, which was around

5 percent before the crisis. I also find that a normalization process has operated and the pre-crisis standards have been reached from 2010 on.

These results are contrary to the emerging consensus that income differentials shrink following the financial crises. The main observation is that the existence of a strong domestic banking system can lead to increased income polarization in the economy, when global crises lead to asymmetric income reallocations across sectors. Increased profitability of the banking system, which brings together higher performance pay and bonuses to top employees may have contributed to this outcome. Finally, I argue that the feeling of deprivation in the society may have gone up significantly in 2009, if one defines earnings differentials as a direct measure of deprivation.

These results inform the theoretical studies of the connection between crises and income inequality. Specifically, this study suggests that a theoretician who is trying to understand how crises affect the structure of inequality in the society should focus on the role and state of the financial sector—i.e., the banks and other intermediaries.

References

Aaberge, R. (2007): "Gini's Nuclear Family," *Journal of Economic Inequality*, 5, 305–322.

Atkinson, A. B. and S. Morelli (2011): "Economic Crises and Inequality," Unpublished manuscript, University of Oxford.

Atkinson, A. B., T. Piketty, and E. Saez (2011): "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49, 3–71.

Bell, B. and J. van Reenen (2010): "Bankers Pay and Extreme Wage Inequality in the UK," Unpublished manuscript, London School of Economics.

Clark, A. E. and A. J. Oswald (1996): "Satisfaction and Comparison Income," *Journal of Public Economics*, 61, 359–381.

Cowell, F. A. (2011): *Measuring Inequality*, Oxford, UK: Oxford University Press, 3rd ed.

Fitoussi, J. P. and F. Saraceno (2010): "Europe: How Deep Is a Crisis? Policy Responses and Structural Factors Behind Diverging Performances," *Journal of Globalization and Development*, 1, 1–17.

Kumhof, M. and R. Ranciere (2010): "Inequality, Leverage, and Crises," IMF Working Paper 10/268.

Lee, S. J. and J. D. Rose (2010): "Profits and Balance Sheet Developments at U.S. Commercial Banks in 2009," Federal Reserve Bulletin, May 2010, A1–A37.

Morelli, S. (2011): "Banking Crises and Stock Market Crashes in US: The Response of Top Shares in Historical Perspective," Unpublished manuscript, Oxford University.

Philippon, T. and A. Reshef (2012): "Wages and Human Capital in the U.S. Finance Industry: 1909–2006," *Quarterly Journal of Economics*, 127, 1551–1609.

Piketty, T., E. Saez, and S. Stantcheva (2013): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, forthcoming.

Rajan, R. G. (2010): *Fault Lines: How Hidden Fractures Still Threaten the World Economy*, Princeton, NJ: Princeton University Press.

Roine, J., J. Vlachos, and D. Waldenstrom (2009): "The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data?" *Journal of Public Economics*, 93, 974–988.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rate," *Review of Economic Studies*, 68, 205–229.

——— (2008): "Striking It Richer: The Evolution of Top Incomes in the United States," *Pathways Magazine*, Stanford Center for the Study of Poverty and Inequality, Winter 2008, 6–7 (updated March 2012).

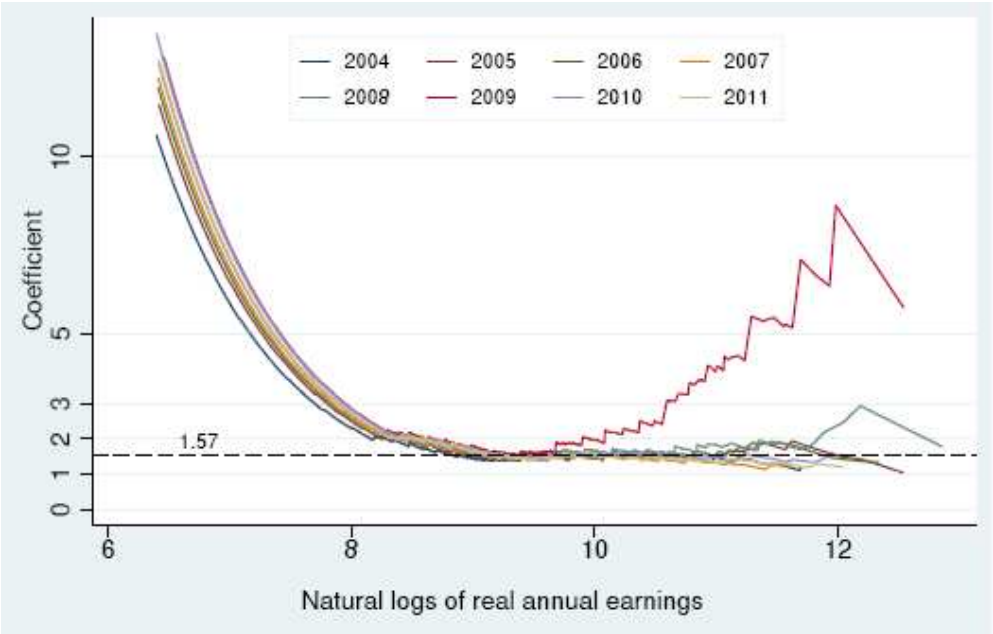

Figure 1: ESTIMATES. The vertical axis plots $d(x)$ versus x in the horizontal axis for each year. Note that x denotes log real earnings. The deviant plot with a red color is for year 2009.

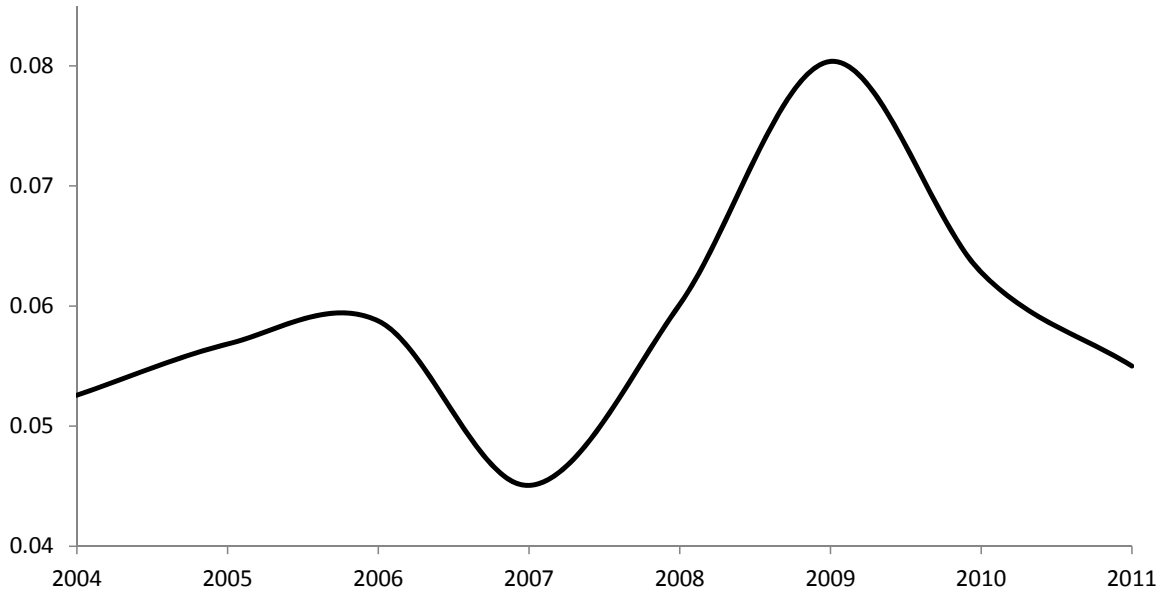

Figure 2: Earnings share of the top percentile. The vertical axis denotes the share of top percentile earnings in total earnings. Appropriate frequency weights are used.

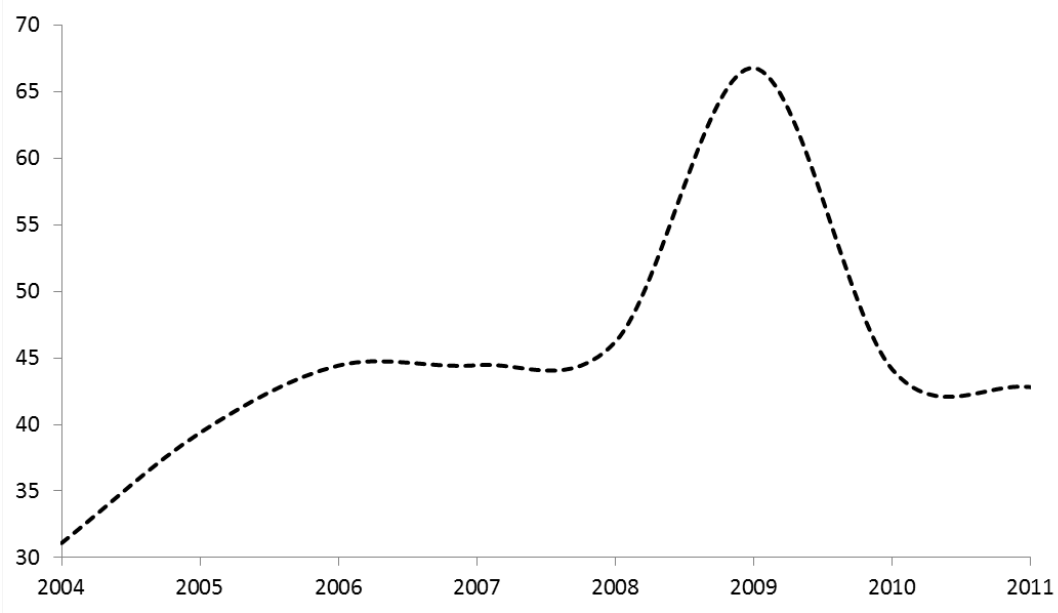

Figure 3: Average earnings in the top percentile. The vertical axis denotes the mean real earnings (in thousand Turkish Liras) in the top percentile for the corresponding years. Appropriate frequency weights are used.

	2004	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
$x \geq 99$	0.053	0.057	0.059	0.045	0.06	**0.08**	0.063	0.055
$x \geq 95$	0.149	0.16	0.165	0.159	0.164	**0.177**	0.164	0.169
$x \geq 90$	0.262	0.254	0.268	0.258	0.269	**0.275**	0.261	0.244
$x \geq 75$	0.463	0.476	0.446	0.49	0.488	0.485	0.483	0.491

Table 1: TOP LABOR INCOME SHARES. The cells of the table describe the top labor income shares for the upper 1%, 5%, 10%, and 25% of the working population, respectively, from 2004 to 2011. For example, the entry 0.463 in 2004 for $x \geq 75$ should be read as follows: in 2004, the upper 25% of the working population earn 46.3% of the total labor incomes in Turkey. The numbers in red color highlight the earnings polarization experienced after the 2008 crisis. Appropriate frequency weights are used.

	2004	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Male	0.803	0.795	0.773	0.816	0.784	0.758	0.748	0.777
Age 15–19	0	0.002	0	0	0.001	0	0	0.001
Age 20–24	0.016	0.011	0.004	0.001	0.002	0.005	0.004	0.006
Age 25–29	0.084	0.103	0.079	0.061	0.083	0.074	0.068	0.043
Age 30–34	0.188	0.199	0.199	0.144	0.194	0.206	0.217	0.152
Age 35–39	0.236	0.244	0.231	0.261	0.233	0.190	0.209	0.212
Age 40–44	0.179	0.181	0.140	0.164	0.172	0.217	0.181	0.272
Age 45–49	0.146	0.121	0.170	0.151	0.132	0.119	0.123	0.150
Age 50–54	0.085	0.062	0.070	0.114	0.095	0.102	0.075	0.091
Age 55–59	0.049	0.052	0.064	0.057	0.056	0.053	0.034	0.038
Age 60–64	0.017	0.020	0.030	0.038	0.020	0.029	0.016	0.025
Age 65+	0.002	0.006	0.013	0.011	0.013	0.007	0.004	0.011
Urban	0.972	0.960	0.966	0.967	0.943	0.961	0.922	0.954
Married	0.828	0.846	0.845	0.857	0.835	0.856	0.805	0.822
Tenure (lj)	6.970	10.075	9.737	9.511	9.651	10.056	11.491	10.331
Large Firm	0.712	0.754	0.745	0.600	0.568	0.713	0.671	0.671
No Degree	0.002	0	0	0	0	0.009	0.001	0
Primary Educ.	0.033	0.029	0.009	0.006	0.011	0	0.029	0.007
Secondary Educ.	0.018	0.010	0.011	0.004	0.011	0.006	0.027	0.011
HS Grad.	0.094	0.077	0.075	0.067	0.050	0.038	0.068	0.024
Voc. Educ.	0.077	0.049	0.047	0.032	0.027	0.022	0.070	0.027
College&+	0.776	0.834	0.858	0.890	0.901	0.924	0.804	0.930
Informal Sector	0.039	0.041	0.061	0.074	0.054	0.033	0.018	0.028

Table 2: Demographic composition within the top percentile. The total number of observations is 194,472. Tenure (lj) refers to "tenure in the last job." Large firm refers to workers employed in firms with 200 employees or more. Appropriate frequency weights are used.