Session Number: Second Poster Session
Time: Thursday, August 26, PM

Paper Prepared for the 31st General Conference of
The International Association for Research in Income and Wealth

St. Gallen, Switzerland, August 22-28, 2010

Income Inequality in Turkey, 1994-2006: Decomposition of the Change

Asena Caner

For additional information please contact:

Name: Asena Caner
Affiliation: TOBB University of Economics and Technology

Email Address: acaner@etu.edu.tr

Income Inequality in Turkey, 1994-2006: Decomposition of the Change

Asena Caner
July 2010

I. Introduction:

The Turkish economy has been experiencing a major transformation since the 1990s. There are various factors behind this transformation. One is the demographic transition that the country has been going through. The Turkish population has aged as we observe a downward trend in the share of the young and an upward trend in the share of the middle-aged groups.

Another factor is the education reform that made the 8-year primary education compulsory for all children. This reform helped increase the overall level of education in the country. The percentage of those who did not complete primary school was cut drastically. On the upper part of the education distribution the share of those with college or graduate degrees has increased. Along with these changes, a movement out of agriculture has been continuing. Between 1994 and 2006, the share of households that reported agriculture as the main source of income has declined from about 30% to about 20%.

An average Turkish household enjoyed an increase in the amount of economic resources that it has access to during the period that I analyze in this paper. Per capita income of a median Turkish household has gone up by about %5.5 annually between 1994 and 2006. The per capita income of an average household has increased at an annual rate of 4% during this period. Interestingly, as income inequality has been increasing in many developing countries, it has declined in Turkey over the period 1994-2006. Inequality as measured by the Gini index registers an overall decrease in inequality as the values taken by this index has decreased from 0.469 in 1994 to 0.426 in 2006.

In this study, I aim to answer the following question: To what extent can the change in the distribution of per capita income between 1994 and 2006 be explained by changes in household characteristics, namely the age, education and occupation of the household head and the change in the mean industry income? And, what would the income distribution look like had these factors been changed to their 2006 levels?

To answer this question I employ the semi-parametric approach of DiNardo et al. (1996) and Cameron (2000). DiNardo et al.(1996) present decompositions in terms of probability density functions, thereby allowing a more transparent analysis of distributional changes than the traditional method of decomposing opaque summary statistics such as the Gini coefficient. Cameron (2000) presents decompositions in the form of cumulative distribution functions, Lorenz curves and generalized Lorenz curves.

The paper is organized as follows. The next section surveys the literature on decompositions of inequality. Section 3 describes the data. Section 4 compares the 1984 and 1990 distributions. Section 5 introduces the decomposition notation and Section 6 details the decomposition methodology and results. The final section concludes the paper.

II. Literature

In the modern inequality decomposition literature, Shorrocks (1982) and (1984) are probably the most widely cited studies. Shorrocks' (1982) seminal paper shows that, given a set of desired decomposition properties and under several assumptions, there is a unique factor decomposition rule that is independent of the inequality index used. The main advantage of this non-parametric technique lies in the absence of assumptions about structural relationships, i.e. no formal model or econometric estimation is involved. This advantage is, however, the source of its weakness. In the absence of economic structure very little can be said about the economic mechanisms driving the results.

Shorrocks (1984) examines the decomposition of inequality by population subgroup. He shows that a broad class of inequality measures, monotonic transformations of additively decomposable indices can be decomposed into components reflecting only the size, mean and inequality value of each population subgroup. These decomposition rules have been widely used by researchers (see for example Cowell and Jenkins (1995) and Jenkins (1995)).

The decomposition of inequality indices comes with a serious problem. Shorrocks (1982), following Atkinson's (1970) axiomatic approach, identifies six desirable properties of inequality decompositions and examines the constraints these properties impose on the decompositions. He shows that in principle there are an infinite number of possible decomposition rules. Among these he suggests that we use the "natural decomposition" which allocates interaction effects equally across variables. Unfortunately, even the use of the natural decomposition is capable of yielding very different results when used with different inequality indices. In fact, this highlights the primary advantage of the method used in my paper. The advantage is that the decomposition does not rely on specific inequality indices and therefore avoids the indeterminacy problem.

Fields (1998) and Morduch and Sicular (1998) propose regression based methods, which involve estimating standard income generating equations. By writing these equations in terms of covariances, one can express the contribution of the explanatory variables to the distributional changes as a function of the magnitude of the coefficients in the income equation and the size of the change in the variable.

DiNardo et.al.(1996) suggest a semi-parametric approach in which decompositions are presented in terms of probability density functions. This method allows a more transparent analysis of distributional changes than the traditional method of decomposing opaque summary statistics such as the Gini coefficient. Cameron (2000) uses a variant of their method by presenting the decompositions in the form of cumulative distribution functions, Lorenz curves and generalized Lorenz curves. Via these curves one can explicitly see how growth impacts poverty, inequality and social welfare. One can also see which parts of the income distribution has been affected by the various factors considered.

III. Data

The data used in this study consist of the 1994 Household Income and Consumption Expenditures Survey (HICS) and the 2006 Household Budget Survey (HBS), both conducted by the Turkish Statistical Institute (TurkStat). The surveys collect a wide range of information on the different sources of income, household composition, the characteristics of the head and other household members, and on consumption patterns of households.

The 1994 survey consisted of a total of 26,256 households. The survey was not repeated until 2002. Since 2002, TurkStat has been conducting the survey on an annual basis with a smaller sample size. 2006 is the last year for which data are available at the time of this study. The 2006 sample consisted of 8,640 households.

The income question in these surveys asks the respondent his/her income in both the past month and the past year. The income concept used in this analysis is household disposable income. It consists of earned income from all jobs, profit, rent and interest income, pension income, scholarships, the cash value of government assistance as well as non-cash income received by all persons that live in the household. The value of regular payments to other households is subtracted from the total to estimate disposable income.

Also available in the data are the household characteristics such as the gender, age and education of the household head, the occupation and industry category of the primary job of the household head. We also have information on the household size and the rural-urban status of the household.

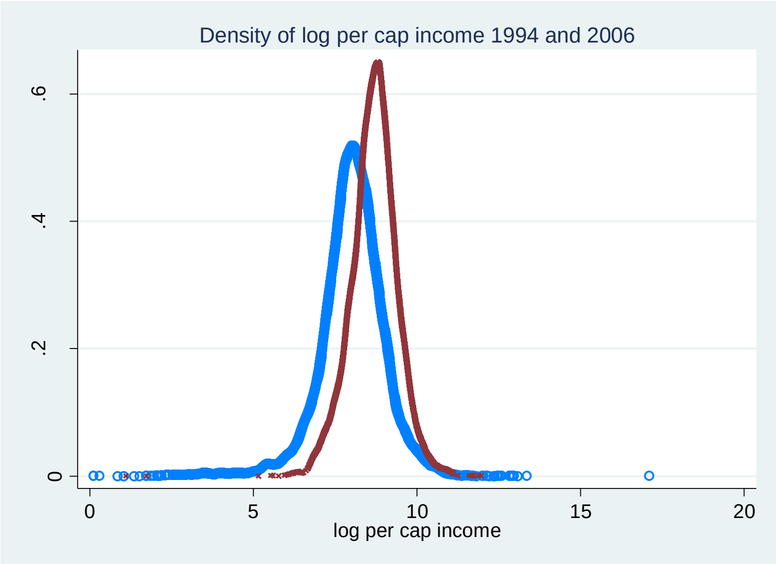
The unit of measurement in this study is the household. Household disposable income is divided by the household size to compute per capita income. The OECD equivalence scale is used to estimate the household size rather than the household size defined as the number of persons in the household. This scale assigns a value of one to the first adult in the household, a value of 0.5 to all other persons 14 or older and a value of 0.3 to those younger than 14. Although some studies prefer to use the household size directly based on the argument that there is no consensus on an ideal equivalence scale, I have chosen to use an equivalence scale. My motivation has been the existence of a non-negligible number of large households that benefit from significant economies of scale. Dividing the household income by the number of individuals leads to a meaninglessly small per capita income, which does not reflect the economic circumstances of these households.

To allow for comparability across time, I inflate 1994 figures to 2006 Turkish Liras by using the Consumer Price Index, estimated by the TurkStat. Moreover, I use the household level sample weights in all calculations.

IV. Comparison of 1994 and 2006

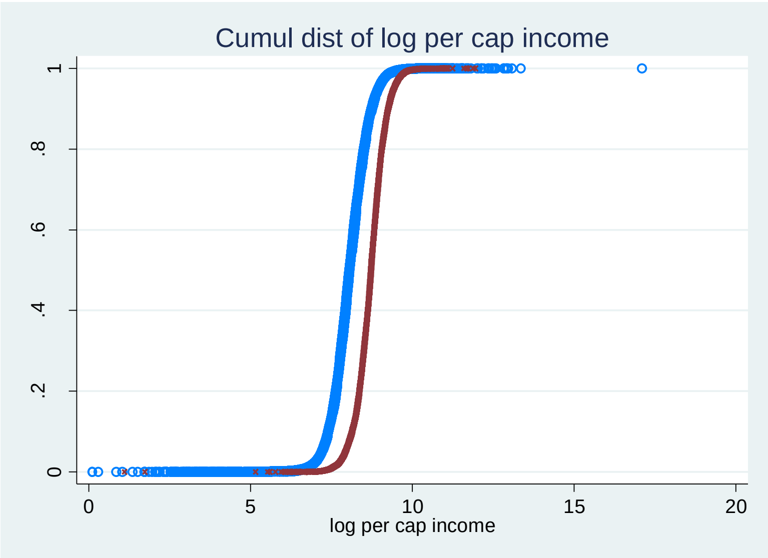
Before I explain and conduct the decomposition analysis, some preliminary investigation of the data in years 1994 and in 2006 is in order. In Figures 1a and 1b, I show the density and cumulative distribution functions of log per capita income in 1994 and 2006. Evident from these figures are the increase in per capita income and the increase in the concentration of the data between the two years.

Figure 1a: The probability density estimates of log per capita income in 1994 and 2006 (kernel density estimates)



Source: Author's calculations.
Note: The density shown in the right belongs to 2006 data.

Figure 1b: The cumulative distribution function estimates of log per capita income in 1994 and 2006 (kernel density estimates)

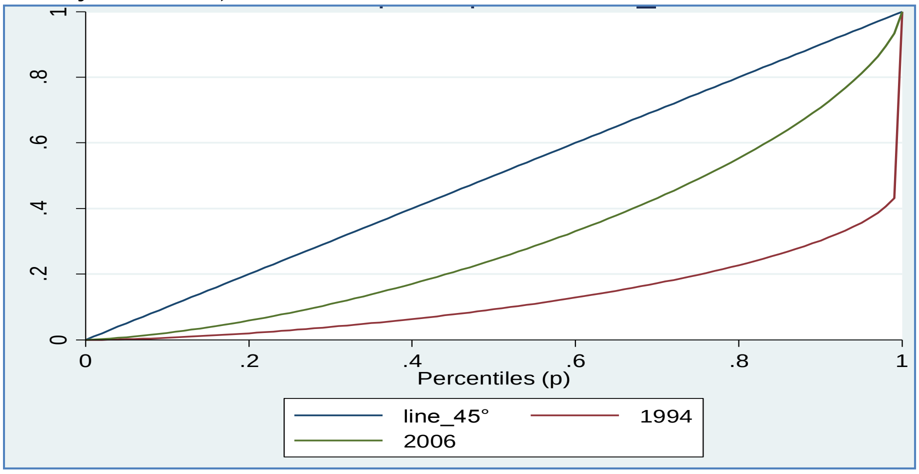


Source: Author's calculations.
Note: The curve shown in the right belongs to 2006 data.

We know that if one Lorenz curve lies everywhere to the left of (or higher than) another it is possible to rank two distributions unequivocally. The higher Lorenz curve is said to Lorenz dominate the lower curve and all summary measures that respect the principle of transfers will show inequality to be lower in the higher curve. Evidently, this is the case in Turkey.

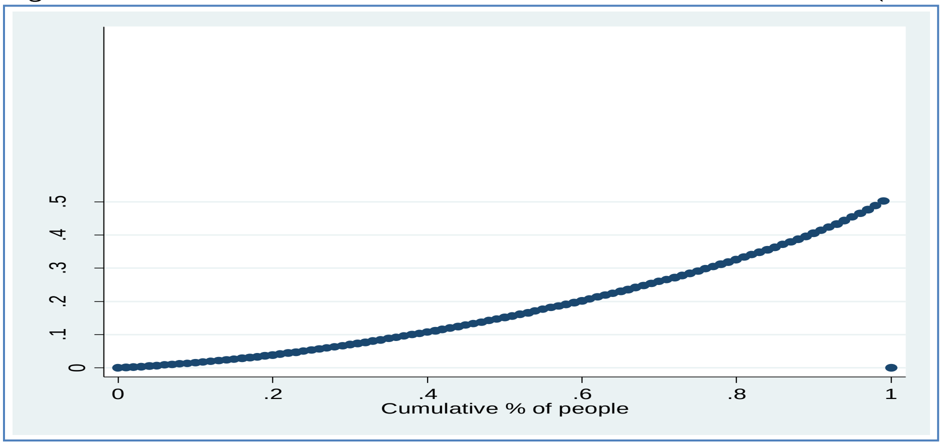
Figure 2a shows the Lorenz curves for years 1994 and 2006. It is clear from this figure that the 2006 distribution Lorenz dominates the 1994 distribution. Movements of Lorenz curves are typically not very large so it is often useful to plot the difference between the curves, as shown in Figure 2b.

Figure 2a: The Lorenz curve estimates of log per capita income in 1994 and 2006 (kernel density estimates)



Source: Author's calculations.
Note: The Lorenz curve for 2006 lies to the left of the curve for 1994.

Figure 2b: The difference between Lorenz curves in 1994 and 2006 (kernel density estimates)

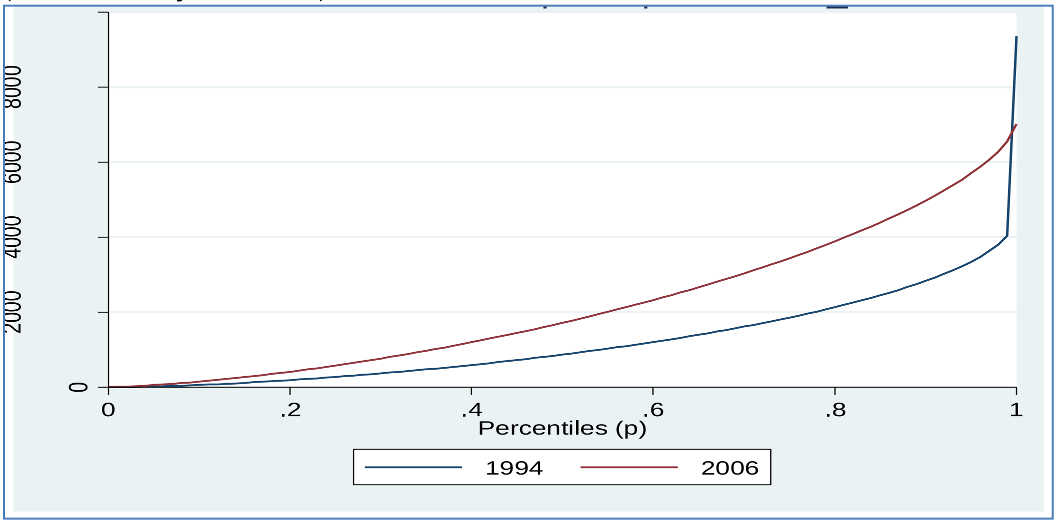


Source: Author's calculations.

To evaluate social welfare, we investigate Generalized Lorenz curves. These curves scale up the vertical axis of standard Lorenz curves by multiplying the cumulative share of per capita income by mean per capita income. We know that if one generalized Lorenz curve lies everywhere above (or to the left of) another, this means that every percentile of the distribution on the left has access to more resources than the other distribution. In such as case, any equity respecting social welfare function will prefer the distribution with the higher curve.

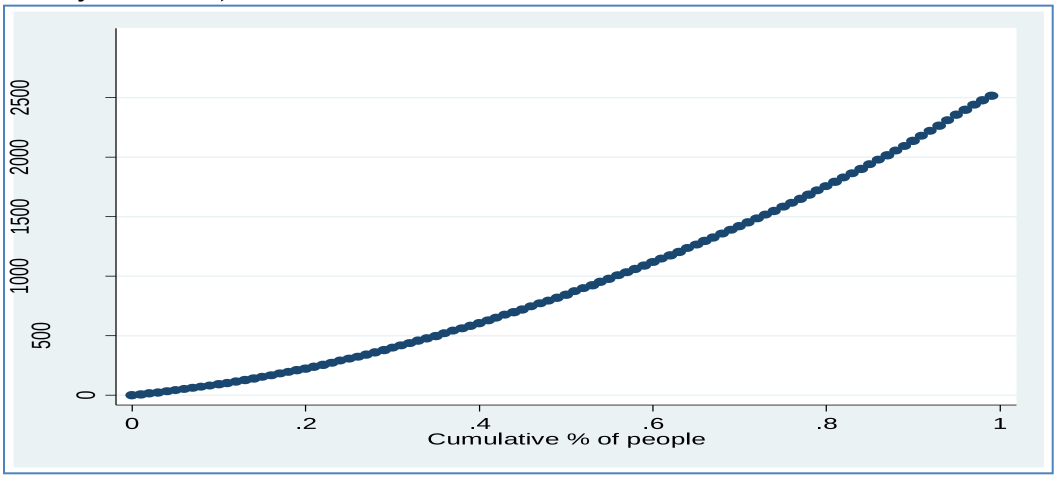
Figure 3a shows the Generalized Lorenz curve kernel estimates of log per capita income in 1994 and 2006. Evidently, the curve for 2006 lies to the left of the curve for 1994, except at the very top of the distribution where the two curves intersect. Hence we cannot rank the two distributions unambiguously according to welfare. Figure 3b shows the difference between the 2006 and 1994 Generalized Lorenz curves. The difference is positive at all percentiles except for the top, where it turns negative due to the small number of very high incomes in year 1994 (only the positive section of the figure is shown).

Figure 3a: The Generalized Lorenz curve estimates of log per capita income in 1994 and 2006 (kernel density estimates)



Source: Author's calculations.
Note: The curve on the left belongs to year 2006.

Figure 3b: The difference between the Generalized Lorenz curves in 1994 and 2006 (kernel density estimates)



Source: Author's calculations.

I also report some indicators of inequality in per capita income in years 1994 and 2006. There is a decrease in the p90/p50 ratio, indicating that the top of the distribution got closer to the median over time. Yet, both the top of the distribution anf the median moved farther from the bottom of the distribution as indicated by the changes in the p90/p10 and p10/p50 ratios respectively.

As indicators of inequality, I also report Gini and Atkinson indices. The Gini index registers an overall decrease in inequality as the values that the index take decreases from 0.469 to 0.426.

The Atkinson index has the ability to gauge movements in different segments of the income distribution depending on the value that the "e" parameter takes. As the level of inequality aversion falls (that is, as e approaches 0) the Atkinson index becomes more sensitive to changes in the upper end of the income distribution. As e grows larger, the index becomes more sensitive to changes in the lower end of the income distribution. When e=0.5 and when e=0.5, we observe a decline in inequality as shown in Table 1. When e takes the value of 2, it registers an increase in inequality, which hints that the lower end of the distribution might not have benefitted from the overall increase in mean income and an overall decrease in inequality.

Table 1: Some measures of inequality of per capita income

			1994	2006
	p90/p10		7.893	7.904
Percentile ratios	p90/p50		2.789	2.509
	p10/p50		0.353	0.317
	Gini		0.469	0.426
Inequality Measures	Atkinson	e=0.5	0.188	0.150
		e=1	0.321	0.279
		e=2	0.526	0.545

Source: Author's calculations.

V. Notation:

The notation used in this paper follows DiNardo et.al (1996) and Cameron (2000). Each observation in the analysis can be represented as a vector (y, z, t) of per capital income y, household attributes z, and time t. The joint probability distribution of these variables is $F(y, z, t)$. The density of per capita income at a time t, denoted by $f_t(y)$, is the integral of the density of per capita income conditional on household attributes and a time t_y, shown as $f(y|z, t_y)$, over the distribution of household attributes at time t_z, shown as $F(z|t_z)$.

$$f_t(y) = \int_{z\in\Omega_z} f\left(y\middle|z, t_y = t; d_t\right) dF(z|t_{z=t})$$

$$= f(y; t_y = t, t_z = t, d_t)$$

In the above, Ω_z is the set of household attributes and d_t denotes other distributional characteristics whose details are given below.

We define z to consist of three household attributes, namely the age of the household head (a), education of the household head (e), and the occupation of the household head in the primary job (o).

As the decomposition of the overall change between years 1994 and 2006 requires us to express the timing of distributional attributes, the notation described above is very useful. The density of income in 1994 is denoted by $f(y; t_y = 94, t_z = 94, d_{94})$. The density that we would observe in 1994 if the household attributes changed to their 2006 distribution (but all else remained the same as in 1994) is denoted by $f(y; t_y = 94, t_z = 06, d_{94})$.

I can express the difference in the densities of income per capita in 1994 and in 2006 as composed of five components, as follows:

$$
\begin{aligned}
&f_{94}(y) - f_{06}(y) = \\
&\{f(y; t_y = 94, t_a = 94, t_e = 94, t_o = 94, d_{94}) \\
&- f(y; t_y = 94, t_a = 06, t_e = 94, t_o = 94, d_{94})\} && \text{(i)} \\
&+ \{f(y; t_y = 94, t_a = 06, t_e = 94, t_o = 94, d_{94}) \\
&- f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 94, d_{94})\} && \text{(ii)} \\
&+ \{f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 94, d_{94}) \\
&- f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 06, d_{94})\} && \text{(iii)} \\
&+ \{f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 06, d_{94}) \\
&- f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 06, d_{06})\} && \text{(iv)} \\
&+ \{f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 06, d_{94}) \\
&- f(y; t_y = 06, t_a = 06, t_e = 06, t_o = 06, d_{06})\} && \text{(v)} \qquad \text{(1)}
\end{aligned}
$$

Here, term (i) represents the change in the density that results from the changing age composition in the population, term (ii) represents the change in the density that results from the changing education composition in the population, term (iii) represents the change in the density that results from the changing occupation composition in the population, term (iv) represents the change in the density that results from the change in the mean income earned in various industries, and term (v) represents the residual change that is related to other factors.

VI. Decomposition

In order to implement the decomposition in equation (1), we need to know how to express the counterfactual densities involved. The term (i) in the decomposition shows the effect of changing only the age composition of the distribution. The first part of the term is the original density in 1994. The second part is the counterfactual density that we would observe if the age distribution changed to its 2006 level, but all else remained the same as in 1994.

We know that the original density in 1994 can be written as follows:

$$
\begin{aligned}
&f\big(y; t_y = 94, t_a = 94, t_e = 94, t_o = 94, d_{94}\big) \\
&= \iint f\big(y \big| a, e, o, t_y = 94; d_{94}\big) dF(e, o | a, t_e = 94, t_o = 94) dF(a | t_a = 94)
\end{aligned}
$$

The counterfactual density that we would observe if the age distribution changed to its 2006 level can be written as follows:

$$f\big(y; t_y = 94, t_a = 06, t_e = 94, t_o = 94, d_{94}\big)$$

$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(e,o|a,t_e = 94, t_o = 94) dF(a|t_a = 06)$$
$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(e,o|a,t_e = 94, t_o = 94) dF(a|t_a = 06) \frac{dF(a|t_a = 94)}{dF(a|t_a = 94)}$$
$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(e,o|a,t_e = 94, t_o = 94)\ \omega_a\ dF(a|t_a = 94),$$

where $\omega_a = \frac{dF(a|t_a=06)}{dF(a|t_a=94)}$.

In other words, the counterfactual density is a reweighted version of the original 1994 density. An estimate of the weighting factor, ω_a, is the ratio of the share of the age group a in 2006 sample to its share in 1994 sample.

Term (ii) in the decomposition involves the counterfactual density that we would observe if both the age and the education distributions changed to their 2006 levels. This density can be expressed similarly as follows:

$$f(y; t_y = 94, t_a = 06, t_e = 06, t_o = 94, d_{94})$$
$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(o|a,e,t_o = 94) dF(a,e|t_a = 06, t_e = 06)$$
$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(o|a,e,t_o = 94)\ dF(a,e|t_a = 06, t_e = 06) \cdot \frac{dF(a,e|t_a = 94, t_e = 94)}{dF(a,e|t_a = 94, t_e = 94)}$$
$$= \iint f(y|a,e,o,t_y = 94; d_{94}) dF(o|a,e,t_o = 94)\ \ \omega_{ae}\ dF(a,e|t_a = 94, t_e = 94),$$

where $\omega_{ae} = \frac{dF(a,e|t_a=06,t_e=06)}{dF(a,e|t_a=94,t_e=94)}$.

An estimate of ω_{ae} is the ratio of the share of age-education cells in the 2006 distribution to the share in 1994 distribution. The term (iii) involves estimating the counterfactual density that would prevail if age, education and occupation distributions were changed to their 2006 levels, which can be done in a similar fashion. The weighting factor in that case, ω_{aeo}, involves the shares of age-education-occupation cells.

Next, I estimate terms (i) through (v) one by one. In the analysis described below, the characteristics mentioned belong to the household head.

1. Changes in the age distribution

The change in the age composition of the population is shown in Table 2. In Turkey, between 1994 and 2006, the share of households headed by a person between ages 40 and 59 increased whereas the share of other age categories decreased. Turkey has a young population, however the country has been going through a demographic transition in which the share of the young has been decreasing whereas the share of the middle-aged has been increasing. The numbers in Table 2 are compatible with this transition.

Table 2: Summary statistics by age, education, occupation and industry categories

	% of households		
	1994	**2006**	**Change**
Age category			
< 30 years	11.87	7.95	-3.92
30-39 years	28.56	25.05	-3.51
40-49 years	23.51	28.76	5.25
50-59 years	16.41	19.49	3.08
> 59 years	19.65	18.75	-0.9
Education category			
< Primary school	20.87	11.71	-9.16
Primary school	61.55	63.86	2.31
High school	10.89	15.72	4.83
College or above	6.70	8.70	2
Occupation category			
Professionals	10.85	19.78	8.93
Technicians	4.74	8.42	3.68
Sales	23.03	10.4	-12.63
Agriculture	32.06	18.37	-13.69
Crafts, operators	29.32	43.03	13.71
Industry category			
Agriculture	21.70	20.00	-1.7
Mining	1.60	0.80	-0.8
Manufacturing	15.80	18.00	2.2
Electricity, gas, water	0.60	0.60	0
Construction	8.60	9.60	1
Trade	21.40	20.60	-0.8
Transport	7.40	8.30	0.9
Finance	1.00	4.30	3.3
Public administration	21.90	18.00	-3.9

Source: Author's calculations using 1994 and 2006 data.
Notes: In occupation categories, "Professionals" include legislators, senior officials and managers; "Technicians" include associate professionals,"Sales" includes service workers, shop and market workers, "Agriculture" includes skilled agricultural and fishery workers, "Crafts, operators" category includes plant and machine operators and assemblers and elementary occupations.
In industry categories, "Agriculture" includes hunting, forestry and fishing, "Mining" includes quarrying, "Trade" includes wholesale and retail trade, hotels and restaurants, "Transport" includes storage and communication, "Finance" includes financial intermediation, real estate, renting and business activities, "Public administration" includes defense, social security, education, health and social work and other community, social and personal service activities.

Within a given cohort, we would expect inequality to increase with age. At a given point in time, the level of income inequality will be influenced by both age and cohort effects. If cohort effects are small, we would expect income inequality to increase with age. Table 3 shows some descriptive statistics of per capita income by age, education, occupation and industry groups. Income inequality as measured by the coefficient of variation goes up with age but starts to decline after age 50. Within-group Gini coefficients reported in Appendix Table 2 show that there is a decrease in inequality with age, followed by an increase. However, year 2006 numbers do not support this. The reason might be that there are sizable cohort effects.

In Figures 4 and 5, I present the findings of the decomposition analysis. Figure 4 shows the decompositions presented in terms of movements in probability density functions. Figure 4a shows the initial (1994) and final (2006) densities. Figure 4b shows the counterfactual 1994 density adjusted for 2006 age distribution (the higher curve) and the density of the difference between the 1994 distribution and the counterfactual (the lower curve) on the same graph. The next graph (Figure 4c) shows the counterfactual density adjusted for age and education (the higher curve) and the density of the difference between the counterfactual density on the previous graph (4b) and this graph (the lower curve). Graphs 4d-4f are defined similarly.

On the other hand, Figure 5 shows the decompositions in terms of cumulative distribution functions. For example, Figure 5a shows how the cumulative distribution function of per capita income would change if the age composition in 1994 were adjusted to match the age composition in 2006. Clearly, adjusting for the age composition does not have large distributional consequences. The cumulative distribution function shows an almost negligible increase in the percentage of the population in the middle of the distribution.

2. Changes in the education distribution

In Turkey, there has been a remarkable decline in the share of people who did not complete primary school during the 12 years between 1994 and 2006 (see Table 2). Parallel to this development, there has been an increase in the share of high-school graduates and a small increase in the share of those who received a college degree or graduate degrees.

Interestingly, the incomes of those who have more education increased at a slower rate than the incomes of those with less education, as can be seen in Table 3. The decomposition analysis tells us that adjusting for the education composition has a small effect on the distribution of per capita income. The cumulative distribution function shows a small increase in the percentage of the population in the middle of the distribution (see Figure 4b).

Table 3: Per capita income by age, education, occupation and industry categories, expressed in 2006 prices (US$1 ≈ 1.40 TL in 2006 (approximately))

	1994				**2006**			
	Mean	**SD**	**CV**		**Mean**	**SD**	**CV**	***%Δmean***
Age category								
< 30 years	3956	4361	1.10		6635	5538	0.83	*67.7*
30-39 years	4894	6890	1.41		7108	5879	0.83	*45.2*
40-49 years	5156	6899	1.34		7856	7517	0.96	*52.3*
50-59 years	4889	6244	1.28		8270	7707	0.93	*69.2*

> 59 years	4002	4739	1.18		7149	6144	0.86	*78.6*
Education category								
< Primary school	2825	2960	1.05		4653	3909	0.84	*64.7*
Primary school	4221	5129	1.22		6542	5310	0.81	*55.0*
High school	6595	7294	1.11		9292	7804	0.84	*40.9*
College or above	11407	12211	1.07		15349	10448	0.68	*34.6*
Occupation category								
Professionals	9061	8783	0.97		13249	12455	0.94	*46.2*
Technicians	5181	3646	0.70		10235	7712	0.75	*97.6*
Sales	5899	9298	1.58		6397	4004	0.63	*8.4*
Agriculture	3582	3625	1.01		5180	3969	0.77	*44.6*
Crafts, operators	3969	4173	1.05		6128	3774	0.62	*54.4*
Industry category								
Agriculture	3646	3882	1.06		4995	3885	0.78	*37.0*
Mining	5726	4252	0.74		9077	8796	0.97	*58.5*
Manufacturing	5131	13150	2.56		7910	6539	0.83	*54.1*
Electricity, gas, water	6232	3885	0.62		13716	10931	0.80	*120.1*
Construction	3697	5775	1.56		5637	5340	0.95	*52.5*
Trade	6844	18599	2.72		8934	10391	1.16	*30.5*
Transport	4444	5634	1.27		7828	5201	0.66	*76.1*
Finance	6188	4163	0.67		11467	9792	0.85	*85.3*
Public administration	5270	7420	1.41		9244	7077	0.77	*75.4*

Source: Author's calculations using 1994 and 2006 data.
Notes: See Notes to Table 2.

3. Changes in the occupation distribution

In the data, occupation classification is made according to ISCO-88 major groups. In Turkey, there has been a non-negligible decline in the share of agriculture as an occupation in the primary job. Table 2 shows that the share decreased from about 32% to about 18%. An increase in a similar magnitude has been observed for those in the "Crafts, operators" occupation category. This category includes plant and machine operators and assemblers and elementary occupations that do not require much skill. Some of those who have lefty agriculture might have moved to this occupation category in the meantime. Another change is the increase in the share of "Professionals" in the population, which is compatible with the overall increase in the level of education.

Although the movement out of agriculture is evident in the data, I do not find a sizable effect of this change on the distribution of per capita income. Figure 4d shows a very small distributional effect.

4. Changes in the mean incomes in industry categories

In order to account for the changes in mean incomes in different industries, I estimate the mean incomes in years 1994 and 2006, compute the change and then make an adjustment for the change in the 1994 distribution. In the data, industry classification is made according to NACE Rev-1 sections. The method that I use follows Cameron (2000) and can be described as follows:

I estimate the regression equation expressed as:
$\log(y) = \alpha + \beta.I + \gamma.A + \varepsilon$,
where I is a vector of dummy variables indicating industries and A is a vector of dummy variables indicating age-education cells as well some other controls including occupation dummies, gender of the household head, rural-urban status and household size. The omitted dummies are agriculture (occupation), <30 years and <Primary school (age-education) and agriculture (industry). Full results are presented in Appendix Table 1.

The aim of estimating this regression equation is to gauge the change in mean industry incomes between the two years, captured by the differences in the coefficients on the industry dummies, $\beta_{06} - \beta_{94}$. This information is presented in Table 4. We observe that all industries experienced a higher income growth relative to agriculture, except for the construction and trade industries. The growth rates of incomes in the finance and public administration industries are noteworthy.

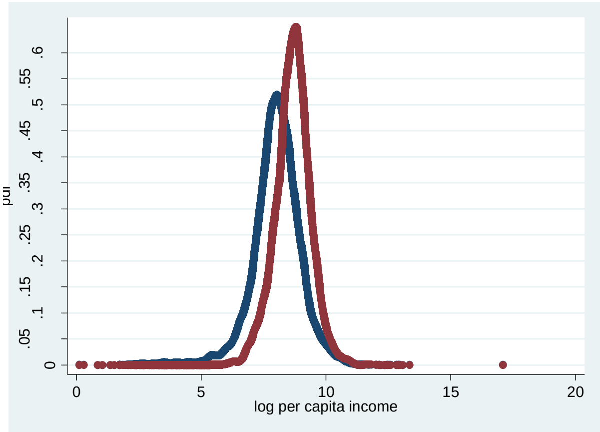
Table 4: Changes in mean incomes within industries

	β_{94}	β_{06}	% change in β between 1994 and 2006
Mining	0.5395	0.6024	11.66
Manufacturing	0.3814	0.5082	33.25
Electricity, gas, water	0.5783	0.8084	39.79
Construction	0.2506	0.2409	-3.87
Trade	0.4539	0.4494	-0.99
Transport	0.4326	0.5214	20.53
Finance	0.3566	0.5363	50.39
Public administration	0.1817	0.3477	91.36

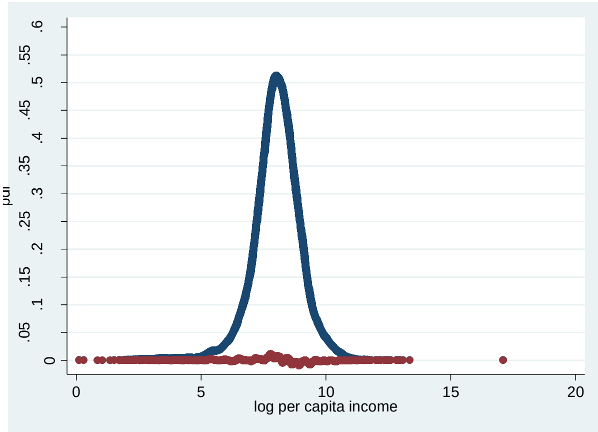
In the decomposition analysis, the counterfactual distribution in which the mean industry incomes were shifted to their 2006 levels is estimated by changing the incomes of households in the appropriate industry category by the amount of $\beta_{06} - \beta_{94}$. As agriculture is the base category, the incomes of those working in the agriculture industry were unchanged. The results, presented in Figure 4e, show that there are little distributional effects of changes in mean industry incomes.

Figure 4: Probability density functions (pdfs) and the differences between probability density functions due to indicated factors.

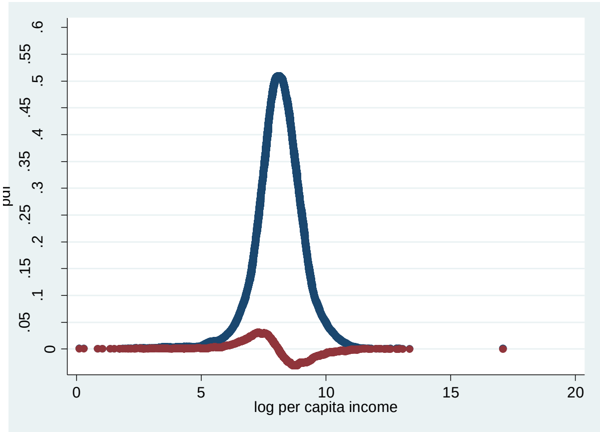
a) 1994 and 2006 pdf's



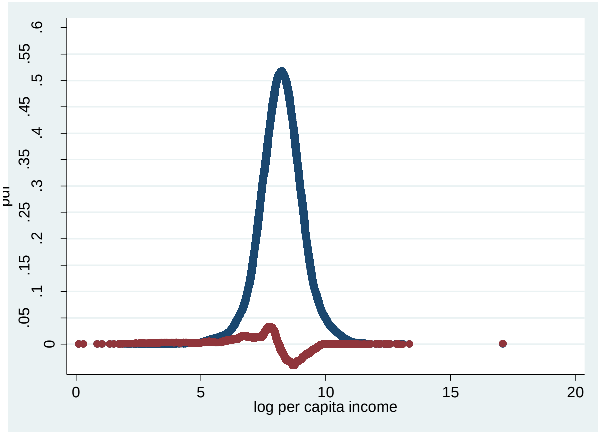
b) 1994 pdf after age adjustment



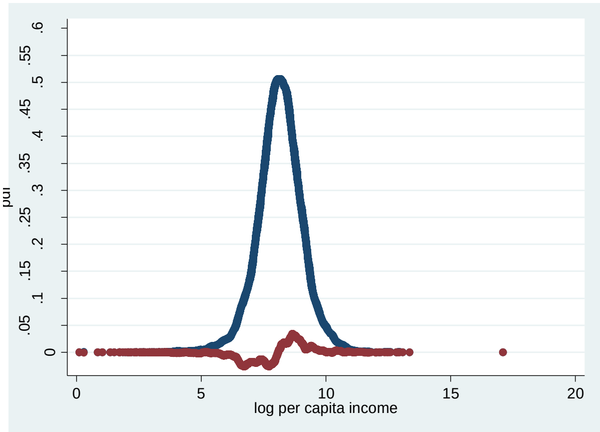
c) 1994 pdf after age-education adj.



d) 1994 pdf after age-educ-occup. adj.



e) 1994 pdf after age-educ-occup-income adjustment



f) Residual pdf and 2006 pdf

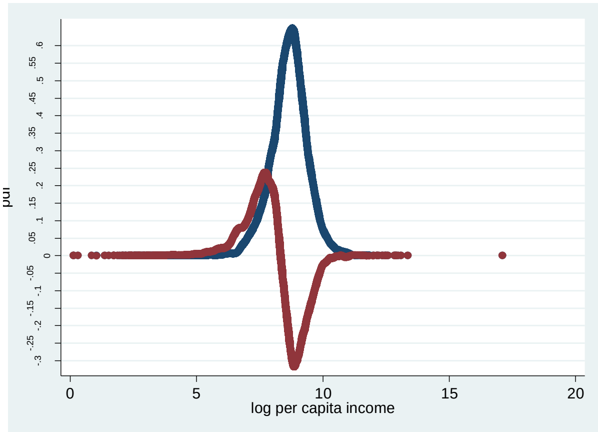
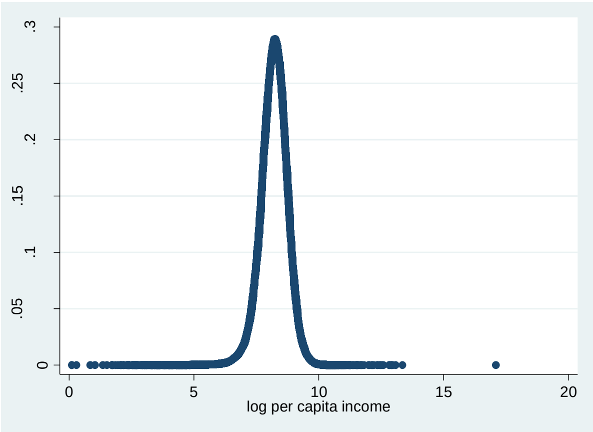
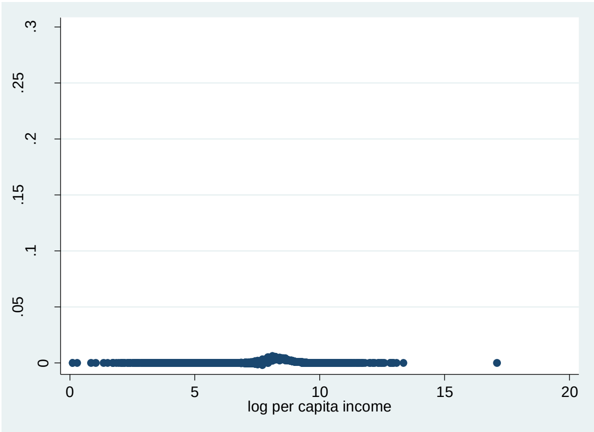


Figure 5: Differences between the cumulative distribution functions of per capita income adjusted for the indicated factors.

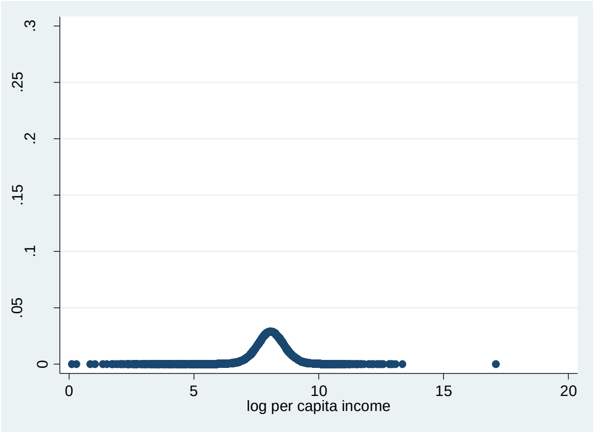
a) Actual difference 1994 to 2006



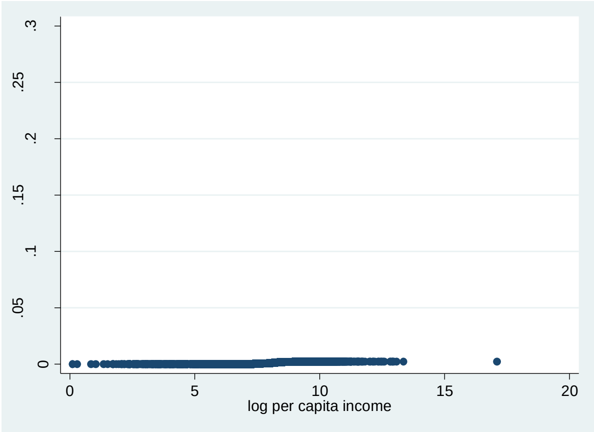
b) Age composition



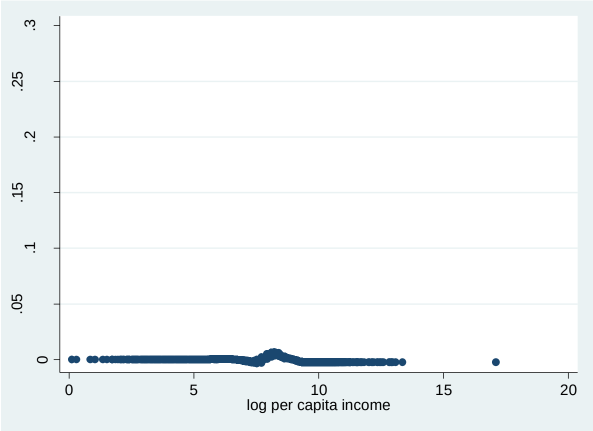
c) Education composition



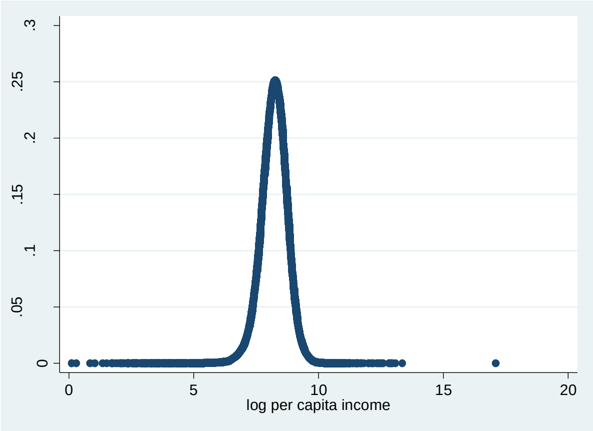
d) Occupation composition



e) Mean industry income



f) Residual



VII. Conclusion:

The Turkish economy has experienced a major transformation since the 1990s. There has been significant growth in average per capita income. Inequality measures indicate an overall decline in income inequality in the past two decades. The demographic transition has changed the age distribution; the education reform has changed the distribution of educational attainment; the movement out of agriculture and the development of some service sectors has changed the occupational and industrial distribution of households. All of these changes which might have affected income inequality in the country. The aim of this paper is to examine exactly this issue. The question that I ask is the following: To what extent can the change in the distribution of per capita income between 1994 and 2006 be explained by changes in household characteristics, namely the age, education and occupation of the household head and the change in the mean industry income? And, what would the income distribution look like had these factors been changed to their 2006 levels?

To answer this question I employ the semi-parametric approach of DiNardo et al. (1996) and Cameron (2000). I present decompositions in terms of probability density functions and cumulative distribution functions, thereby allowing a more transparent analysis of distributional changes than the traditional method of decomposing opaque summary statistics such as the Gini coefficient. My results show that changes in the distribution of age, education and occupation and the changes in the mean industry incomes have had a small effect on the income distribution in Turkey.

References:

Atkinson, A., 1970, "On the measurement of inequality", *Journal of Economic Theory* 2, 244–263.

Cameron, L., 2000, "Poverty and inequality in Java: examining the impact of the changing age, educational and industrial structure", *Journal of Development Economics*, 62, 149-180.

Cowell, F., Jenkins, S., 1995, "How much inequality can we explain? A methodology and an application to the United States", *The Economic Journal* 105, 421–430, March.

DiNardo, J., Fortin, N.M., Lemieux, T., 1996. "Labor market institutions and the distribution of wages, 1973–1992: a semiparametric approach", *Econometrica*, 64 (5), 1001–1044, Sept.

Fields, G.S. (2003) "Accounting for income inequality and its change: A new method with application to the distribution of earnings in the United States", *Research in Labor Economics*, 22, pp. 1-38

Jenkins, S., 1995, "Accounting for inequality trends: decomposition analyses for the UK, 1971–86", *Economica* 62, 29–63.

Morduch, J. and Sicular, T. (2002) `Rethinking inequality decomposition, with evidence from rural China', *The Economic Journal*, vol. 112, pp. 93-106.

Shorrocks, A.F. (1982) "Inequality decomposition by factor components", *Econometrica* 50, pp. 193-211.

Shorrocks, A. F. (1984) "Inequality decomposition by population subgroups", *Econometrica* 52 (6), 1369–1385, November.

Appendix:

Table A1: Dependent variable: Log per capita income (in 2006 TL)

		1994		2006	
		Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Professionals		0.2611	3.45	0.1077	1.98
Technicians		-0.0019	-0.02	-0.1027	-1.79
Sales		0.0230	0.31	-0.3151	-5.58
Crafts, operators		-0.1151	-1.57	-0.2726	-5.42
< 30 years	Primary school	0.3554	4.19	0.5728	4.56
< 30 years	High school	0.8374	9.14	0.7894	6.11
< 30 years	College or above	1.0122	9.5	1.1975	8.5
30-39 years	< Primary school	-0.0756	-0.82	0.3382	2.39
30-39 years	Primary school	0.5454	6.53	0.6596	5.35
30-39 years	High school	0.9412	10.83	0.9564	7.62
30-39 years	College or above	1.3846	15.1	1.3114	10.19
40-49 years	< Primary school	0.3796	4.29	0.5147	3.95
40-49 years	Primary school	0.7258	8.66	0.9087	7.38
40-49 years	High school	1.0622	11.57	1.2043	9.58
40-49 years	College or above	1.4214	15.39	1.3904	10.8
50-59 years	< Primary school	0.5501	6.37	0.8008	6.26
50-59 years	Primary school	0.8601	10.16	1.1161	9.04
50-59 years	High school	1.2582	12.67	1.2439	9.64
50-59 years	College or above	1.8453	17.22	1.6316	12.33
> 59 years	< Primary school	0.3469	4.1	0.7007	5.62
> 59 years	Primary school	0.8104	9.52	1.0699	8.63
> 59 years	High school	1.4495	13.54	1.3947	10.2
> 59 years	College or above	1.9034	16.77	1.7715	12.41
Mining		0.5395	5.65	0.6024	6.27
Manufacturing		0.3814	5.14	0.5082	9.77
Electricity, gas, water		0.5783	4.97	0.8084	7.8
Construction		0.2506	3.29	0.2409	4.43
Trade		0.4539	6.13	0.4494	8.48
Transport		0.4326	5.67	0.5214	9.48
Finance		0.3566	3.55	0.5363	8.62
Public administration		0.1817	2.49	0.3477	6.56
Male head		0.1579	7.04	-0.0369	-1.65
Urban		0.0382	2.95	0.2354	17.64
Hh size		-0.0783	-11.25	-0.2421	-29.2
Constant		7.1748	83.65	8.0318	64.52

Notes: Omitted dummies are agriculture (occupation), <30 years and <Primary school (age-education) and agriculture (industry).

Table A2: Within-group Gini coefficients by age, education, occupation and industry categories.

		1994	2006
		Gini	Gini
Age categories	< 30 years	0.46	0.42
	30-39 years	0.45	0.42
	40-49 years	0.45	0.43
	50-59 years	0.48	0.43
	> 59 years	0.51	0.42
Education categories	< Primary school	0.44	0.42
	Primary school	0.43	0.39
	High school	0.44	0.36
	College or above	0.45	0.36
Occupation categories	Professionals	0.54	0.42
	Technicians	0.36	0.34
	Sales	0.47	0.37
	Agriculture	0.43	0.43
	Crafts, operators	0.38	0.37
Industry categories	Agriculture	0.44	0.43
	Mining	0.32	0.52
	Manufacturing	0.48	0.35
	Electricity, gas, water	0.32	0.32
	Construction	0.48	0.42
	Trade	0.57	0.41
	Transport	0.42	0.42
	Finance	0.33	0.47
	Public administration	0.41	0.42