An Investigation of Labor Income Profiles in Turkey*

Emrehan Aktuğ†
University of Texas at Austin

Tolga Umut Kuzubaş‡
Boğaziçi University

Orhan Torul§
Boğaziçi University

First Version: August 3, 2017
This Version : August 3, 2017

Abstract

In this paper, we empirically investigate the role of age, gender, educational attainment and public versus private sector employment on labor income profiles in Turkey. We first report that contrary to the observed hump-shaped age profile in most economies, average labor income profile in Turkey is almost monotonic and ever-increasing over age, which stems mainly from aggregation as there is a considerable degree of heterogeneity in income trajectories over the highlighted characteristics. Second, while the public sector employment is more advantageous for low-educated Turkish workers, university-graduate employees in Turkey face a risk versus return trade-off in their choice of sectoral employment: the private sector income profiles display both a higher level of average income and a higher degree of cross-sectional variation compared to their public sector counterparts. Third, we report a significant gender pay gap especially among low-educated workers, which aligns well with historically low female participation rates in Turkey.

‡For the latest version, please visit: http://www.econ.boun.edu.tr/torul/ilipt.pdf

Keywords: labor income, income inequality, pseudo-panel analysis, synthetic cohort

JEL Classification: D31, I24, R20

*We would like to thank Murat Kırdar, Malik Çürük, and other faculty at the Department of Economics, Boğaziçi University for their helpful comments and suggestions. We also thank the participants at the 6th All-Istanbul Workshop in Istanbul for their feedback. Torul acknowledges financial support by Boğaziçi University Research Fund, grant number BAP 9081. All errors are ours.

†Address: University of Texas at Austin, Department of Economics, 2225 Speedway, Austin, TX 78712, USA. E-Mail: emrehanaktug@utexas.edu

‡Address: Boğaziçi University, Department of Economics, 34342 Bebek, Istanbul, Turkey. E-Mail: umut.kuzubas@boun.edu.tr

§Corresponding Author. Address: Boğaziçi University, Department of Economics, 34342 Bebek, Istanbul, Turkey. E-Mail: orhan.torul@boun.edu.tr

1 Introduction

Income and wealth inequality has become more visible and gained notable popularity over the recent decades.[1] Much of the rise in inequality aligns well with a widening dispersion of labor income, and Turkey is not an exception as she has the fourth highest level of income inequality among OECD countries.[2] In this paper, we analyze labor income profiles of Turkish workers, aiming to discern the determinants of this immense labor income inequality in an era of increasing distributional concerns.

The literature which examines income and consumption profiles over life-cycle has expanded in recent years, particularly for developed countries.[3] The most recent and comprehensive analysis by Lagakos et al. (2017) extracts labor income profiles in both developed and developing countries by concentrating on male workers in the private sector. Our work complements their paper, and extends their analysis by incorporating gender and sector of employment in Turkey, while also factoring in education, thereby constituting the first comprehensive research on labor income profiles in Turkey. Using a rich cross-sectional data set, our results verify that income profiles in Turkey vary considerably over the listed characteristics.

We first show that the average life-cycle labor income profile in Turkey is almost monotonic and ever-increasing over age, contrary to the hump-shaped labor income profile with a peak around the age 50 in many countries, with the notable exceptions of the United States and Germany. On the other hand, we observe that variance of labor income in Turkey is increasing with age, in accordance with the findings on developing countries, yet contrary to developed economies.[4]

Second, we document that the private sector income profiles of university-graduates in Turkey display both a higher level of average and a higher degree of cross-sectional variation compared to their public sector counterparts, thereby implicating a risk versus return trade-off for the sectoral choice of employment. This finding is especially relevant for the Turkish economy, as the public sector employment constitutes 32% of total employment, a figure significantly higher than that of the world average.[5] This risk versus return trade-off, however, is not valid for primary, secondary or high school graduates in Turkey, as the public sector employees with below-university education earn more on average and face lower cross-sectional variation than

[1] See Piketty (2014) and Saez and Zucman (2016) for a through discussion on recent debates about income and wealth inequality.

[2] Among 35 OECD member countries as of 2012. See OECD (2012) for further details.

[3] See Lagakos et al. (2017), Kolasa (2012) and Rupert and Zanella (2015), among others.

[4] See Lagakos et al. (2017), among others.

[5] According to our calculations based on Turkish Statistical Institute's Household Budget Survey and Household Labour Force Survey.

their private sector counterparts.[6]

Third, we find evidence for gender pay gap in Turkey, especially prevalent among primary school graduates. This observation is consistent with the historically low female labor force participation in Turkey, which is around 30% on average. Indeed, the participation rate of women increases with education, which is around 70% among university graduates. However, the fact that 50% of women in the labor force are primary school graduates, combined with the significant wage gap for this group, is likely to impede a higher female labor force participation in Turkey.

Our econometric techniques, as well as our use of multiple data sets offer both rigor and robustness to our descriptive results. Using both ordinary least squares (OLS), and a pseudo-panel approach via the construction of artificial cohorts, our findings pass several robustness checks and sensitivity tests, and offer reliability on the results we report on several dimensions of heterogeneity.

The rest of paper is organized as follows: section 2 summarizes the previous literature on labor income profiles; section 3 describes the data and provides the detailed description of labor income over various clusters; section 4 explains our estimation methodology and presents our results; and section 5 concludes.

2 Literature Review

Average labor income profile exhibits a hump-shaped pattern over age in many developed countries, which is well-established in the literature.[7] Findings on developing economies are, however, rather limited, and Turkey is no exception.[8] As a notable exception, Cilasun and Kirdar (2009) investigate *aggregate* life-cycle income profiles of household *heads* by educational attainment in Turkey between 2002-2006, and report that median income profiles display hump-shaped pattern over life-cycle when conditioned on educational attainment. However, they do not concentrate on *labor*-income of households, or gender or sector of employment.

Turkish Statistical Institute's Household Budget Survey data, which is the data set we use for our analyses, has widely been used to address several other questions related to income inequality, precautionary savings, income and expenditure decompositions in Turkey. Nazli (2014), Yukseler and Turkan (2008) and and Ceritoglu (2009) focus on savings decisions of households but not labor income profiles. Eksi and Kirdar (2015) investigate wage inequality in

[6]Therefore, the main challenge low-educated employees face in Turkey is the scarcity of jobs in the public sector.
[7]See Attanasio and Browning (1995) and Alessie et al. (1997), among others.
[8]See Lagakos et al. (2017) for a recent discussion on the developments on developing countries.

Turkey for the 2002-2011 period by addressing the role of educational attainment. Cilasun (2009) studies labor income profiles in Turkey via a pseudo-panel data approach, where he constructs cohorts based on birth-years for household heads, thereby not allowing the estimation of the effects of educational attainment, gender, and the public versus private sector employment. By using a more comprehensive data set, we are able to explore the role of the lacking dimensions on labor income profiles.

We use OLS to estimate life-cycle profile of labor income and complement our analysis with a pseudo-panel methodology.[9] To this end, we construct cohorts based on birth-year, educational attainment, gender and the public versus private sector employment, which allows us to identify their effects on labor income profiles in a pseudo-panel approach.

3 Data and Descriptive Results

We use cross-sectional data from the Turkish Household Budget Survey (HBS) covering the period 2002-2014.[10] This survey is conducted annually by the Turkish Statistical Institute (TurkStat) on a representative sample of approximately 10,000 Turkish households.[11] We convert nominal labor income into real units by deflating via the Turkish consumer price index (CPI), and exclude workers who earn below the minimum wage so as to focus on full-time employees. TurkStat provides education data in eleven ordinal categories, which we re-cluster into three categories: i) primary or secondary school graduates, ii) high-school graduates, and iii) university or post-university graduates.

We start with providing descriptive graphical analysis of household income profiles over the life-cycle with various clusters, using box-plots to provide visualization of level and dispersion of income profiles in a compact manner.[12]

[9] Instead of constructing birth-year cohorts over a 1-year span, we use 5-year spans so as to enlarge cohort sizes, reduce erraticity and minimize measurement errors.

[10] Information on the public versus private sector employment is only available for 2002-2011, thus it becomes our working sample when investigating sectoral employment differences.

[11] Income variable in the survey refers to total labor income consisting of cash, income received in-kind and premiums.

[12] On the box plot, the end of the whiskers represents the lowest observation within 1.5 times the interquartile range of the lower quartile and the highest observation within 1.5 times the interquartile range of the upper quartile (Tukey, 1977). We follow this approach, because we believe visual distributional illustration is more informative than reporting merely on first moments.

3.1 Labor Income

Figure 1 plots the distribution of labor income over age groups. We observe a peak at around age 45-49 for the median labor income, coupled with ever-increasing cross-sectional dispersion. Average labor income over life-cycle is not hump-shaped, but quite flat, especially between ages 35 and 55.[13] This flat pattern is most similar to labor income profiles in Germany and the United States among the countries that are analyzed in Lagakos et al. (2017).[14] The sharp decrease after the age of 60 stems from retirement: since after 1999, the retirement age in Turkey is 60, which was even lower prior to 1999, therefore the oldest cluster corresponds to the individuals who work after retirement and possibly settle for relatively lower wages.[15]

[place Figure 1 here]

3.2 Education

Educational attainment is one of the key determinants of labor income differences across age categories. Figure 2 displays a clear positive correlation between labor income and education. Among the highest earners, i.e. university graduates, we observe a sharp increase until the age of 35-39, followed by a stagnant profile, with the exception of an additional increase for the age 55-59 group. High school graduates experience a similar upward trajectory over their life-cycle, with a slight downturn around age 55-59. The upward trajectory of primary school graduates is relatively limited compared to the two other education categories, thereby displaying rather stagnant life-cycle trajectories.

A comparison of Turkey with other countries reveals that when conditioned on education, labor income profiles in Turkey exhibit a similar pattern to the ones in Germany, where university graduates have non-decreasing labor income profile, and high school and primary school graduates have slightly hump-shaped patterns, with a peak at around age 50 (Lagakos et al., 2017). Furthermore, these patterns are at odds with the evidence for developing countries such as Brazil, Chile, and Mexico, where education premium is relatively lower.

[place Figure 2 here]

[13] The peak of mean labor income is reached at around 55-59, which is further in the life-cycle compared to the result of previous studies (Cilasun and Kirdar, 2009). The main reason behind this difference is that we concentrate only on full-time employees, whereas Cilasun and Kirdar, 2009 consider all positive labor income earners.

[14] Note that most of the countries Lagakos et al. (2017) investigate differ from Turkey, as they have clear hump-shaped labor income profiles over age.

[15] In our estimations, we exclude the child-labor and the retired, therefore we omit both the 15-19 and 60-64 age groups.

3.3 Sector of Employment

Figure 3 plots labor income profiles over sector of employment. We report a monotonic upward trajectory over age in labor incomes for the public sector employees and a rather stagnant trajectory for their private sector counterparts. Further, the public sector employees of any age group earn more than their private sector counterparts. These contradictory patterns arise mainly due to the educational background compositions of employees in the two sectors: while approximately half of the public sector workers are university graduates, only 13% of the private sector employees hold a university degree or above.[16]

[place Figure 3 here]

Even though the private sector jobs pay less on average, the dispersion of the private sector incomes is larger than the public ones. Figure 4 displays that the variance of incomes in the private sector surpasses that of the public sector after age 30, and keeps increasing until age 40, accompanied by a stable dispersion level afterwards. On the other hand, the variance of incomes in the public sector remains almost constant over the life-cycle. This dispersion profile for the private sector in Turkey differs from many developed countries such as Germany, France, the U.K. and Canada, where variance moderately decreases or remains almost constant over the life-cycle. The similar concave pattern in the variance trajectory in Turkey is also observed in several developing countries, such as Mexico, Uruguay, and Chile.[17]

[place Figure 4 here]

3.4 Education and Sector of Employment

Figure 5 depicts labor income profiles over age and sector of employment. We observe substantial variation in labor incomes over the three education categories in the private sector, yet limited differences in the public one. In the private sector, almost all workers with below-university education earn less than the mean income, whereas university graduates earn above the mean and median income after age 30, coupled with higher dispersion towards retirement. In the public sector, however, income profiles monotonically increase over age, with an almost constant variance and a limited education premium. Further, contrary to the ever-increasing income

[16] We discuss further on this issue in the next section on educational differences within sectors, illuminate on the pay gap between sectors.

[17] A notable exception among developed countries is the Unites States with its similar variance profile to Turkey. See Lagakos et al. (2017) for details.

profile in the public sector, we observe a hump-shaped pattern in the private sector with different trajectories.

[place Figure 5 here]

[place Figure 6 here]

Figure 5 reveals that university-graduate employees in Turkey face a risk versus return trade-off in their sectoral choice of employment: the private sector income profiles display both a higher level of average income and a higher degree of cross-sectional variation compared to their public sector counterparts.[18] However, for primary and high-school graduates, this trade-off disappears, making the public sector jobs more appealing, where this additional demand is rationed due to limited number of the public sector jobs.

Figure 6 displays the histograms of labor incomes in the public and private sectors. While the distribution of incomes in the public sector resembles a normal distribution except for its long right tail, the distribution in the private sector is close to a Pareto distribution, i.e. left-skewed with a mass around the minimum wage.[19] The main reason behind this pattern in the private sector, to a large extent, is due to employment of low-skilled workers earning the minimum wage.[20]

[place Figure 7 here]

3.5 Gender

Figure 8 displays gender differences in average labor income profiles. Both mean and median labor incomes are slightly higher for males at each age category. For male workers, income profiles exhibit an upward trajectory until age 40-44, and remains stagnant afterwards. On the other hand, for female workers, we observe an increase in labor income until age 30, after which median income remains stable. Further, we document an increasing cross-sectional dispersion for both males and females over age.

[place Figure 8 here]

[18] As we highlighted, this observation is vital for the Turkish economy, as the public sector employment constitutes 32% of total employment, a figure significantly higher than that of the world average.

[19] The step-wise increases on the left-end of the private sector distribution stems from the adjustments to the real minimum wage.

[20] In order to display the isolated role of education on the income dispersion, in Figure 7 we abstract from the sector of employment and present the variance of labor income over education.

In order to elaborate further on gender differences in income profiles, we first focus on the role of education, and display labor income profiles over age and education in Figure 9. Even though labor income profiles of high school and university graduates possess similar trajectories for both genders, labor income trajectories of primary school graduates exhibit striking gender differences: for males, we report a hump-shaped pattern with high income dispersion, whereas for females we document an age-independent income profiles with low mean, median and variance levels.[21]

[place Figure 9 here]

While Panel (b) in Figure 9 presents limited gender differences over the sector of employment, Figure 10 shows that conditioning further on education reveals gender pay gap for below-university graduates in the private sector: primary school graduate female workers in the private sector are the lowest income-earning group of all with no prospects of earning nearly the mean income throughout their life-cycles. Further, only a select group of high-school graduate female workers in the private sector earn above the mean income, and they do so only when their income profiles peak. Their male counterparts of similar educational backgrounds have noticeably better income prospects in the private sector. Gender pay gap in the private sector incomes for university graduates are rather limited.

For the public sector, we observe a slight gender pay gap among the primary school graduates, but not for better-educated employees. Overall, we report a significant gender pay gap in favor of male workers especially among low-educated employees, which aligns well with the historically low female participation rates in Turkey.[22]

[place Figure 10 here]

4 Estimation Methodology and Results

We rely on OLS estimation of pooled cross-sections of labor income profiles for different sub-categories.[23] Our main estimation equation is given as:

[21] Labor incomes of female workers with primary education and below are significantly lower, as they contribute to household income by taking low-paying jobs as caretakers, cleaners or factory laborers. They are predominantly employed in the informal sector with shorter working hours, which rationalizes the observed wage gap for the less educated groups.

[22] Female labor participation rate actually increases over years of schooling, up to 70% among university graduates, and only 25% among primary school graduates. Further, more than 50% of women in the labor force are primary school graduates and the significant pay might be a factor affecting the female labor force participation rates.

[23] For robustness purposes, we also use TurkStat's Household Labour Force Survey, which offer qualitatively similar results, as we display in Appendix A).

$$
\begin{aligned}
\log(y_{it}) =& \alpha + \sum_{j}^{7} \beta_j \text{ age}_{ij} + \sum_{k}^{5} \gamma_k \text{ edu}_{ik} + \delta \text{ sector}_i + \sum_{l}^{5} \xi_l \text{ sector}_i \times \text{edu}_{il} \\
&+ \theta \text{ gender}_i + \lambda \text{ union}_i + \mu \text{ area}_i + \phi \text{ tenure}_i + \rho t + \varepsilon_{it} \qquad (1)
\end{aligned}
$$

where $\log(y_{it})$ refers to natural logarithm of labor income of person i in year t; *age* refers to ordinal age categories of 5 year intervals: ages 20 to 24, 25 to 29, ..., 55 to 59; *edu* refers to ordinal years of schooling categories: 6 to 8 years, 9 to 12 years, 13 to 14 years, 15 to 16 years and 17 years and above; *sector* is a dummy variable which equals 1 if individual i works in the public sector; *gender* is a dummy variable which equals 1 if individual i is female; *union* is a dummy variable which equals 1 if individual i is member of a union; *area* is a dummy variable which equals 1 if individual i resides in an urban location; *tenure* stands for years of job experience of individual i, and ρ captures the year-fixed effect.[24]

For robustness purposes, and in order to have comparable results with the previous literature, we also conduct a pseudo-panel estimation. Since the Turkish HBS data is composed of independent cross-sections, we are not able to track the same individuals over time. Thus, we construct cohorts based on their common characteristics, such as educational attainment, sector of employment, gender and year of birth to create a synthetic cohort panel, following the approach by Deaton (1985).

We construct groups by birth year with a 5-year span starting from 1950-1954 to 1985-1989, for a total of 8 groups. We use a static linear model with cohort fixed-effects as follows:

$$
\bar{y}_{ct} = \bar{x}_{ct}\beta + \bar{\theta}_c + \bar{\epsilon}_{ct} \qquad (2)
$$

where c denotes the cohorts, $\bar{y}_{ct}$ are the cohort income averages and $\bar{x}_{ct}$ are the vector of variables generated by cohort averages.[25]

We present our OLS estimation results under alternative specifications in Table 1. We observe that the coefficients of age categories indicate a hump-shaped pattern of labor income over the life-cycle, with a peak at around age 40 to 44.[26] Further, we find evidence of education

[24] We take ages 20 to 24 as the basis age group.

[25] Since each cohort consists of different members in each year, the cohort effect is time varying as $\bar{\theta}_{ct}$. According to Verbeek and Nijman (1992) with a sufficiently large cohort size the time varying $\bar{\theta}_{ct}$ can be treated as constant over time, which is $\bar{\theta}_c$ in regression. The reason is that clustering similar individuals into cohorts tends to homogenize individual effects among individuals grouped in the same cohort, so that average individual effect is approximately time-invariant (Ziegelhofer, 2015). Thus, it is possible to use conventional estimation tools such as fixed-effects estimator (see Appendix E for further details.

[26] See Appendix D for conditional marginal effects of age evaluated at the means of other covariates.

premium over years of schooling, and higher average income in the public sector. In order to shed light on the sectoral differences of education premium, in column (2) of Table 1, we incorporate the interaction of education and sector of employment to our baseline regression, and report that education premium in the private sector surpasses that of the public sector for each education category. In addition, as discussed in the descriptive analysis, Table 1 provides evidence for a significant gender pay gap.

[place Table 1 here]

In order to further elaborate on the role of gender and education, we estimate a modified version of our baseline regression by subsamples based on gender. Table 2 reveals that all education groups but the female primary school graduates follow a hump-shaped trajectory over the life-cycle, yet with different peak ages. For the female primary school graduates, our estimation results confirm age-independent labor income profiles, as we discussed in the previous section.

[place Table 2 here]

In Table 3 we repeat this exercise for the subsamples based on sector of employment. We observe a similar robust hump-shaped pattern of labor income profiles over age, except for the primary school graduates in the public sector. Gender difference in favor of males is still present for both sectors.

[place Table 3 here]

Table 4 presents our estimation results using the pseudo-panel methodology, which provides results consistent with the results by OLS and our descriptive analysis.[27] Table 4 demonstrates that labor income profile tends to be hump-shaped over age groups as coefficients of age and age-squared are positive and negative, respectively. Moreover, the presence of the education premium is statistically confirmed for both high school and university graduates.

[place Table 4 here]

[27]We conduct a fixed-effects regression without any sector specification at the cohort level. That is why it is infeasible to estimate the effect of the public versus private sector employment on labor incomes under this specification. Since the explanatory variables should be time-varying in fixed-effects estimator, we interact the age variable with time-invariant characteristics in the regressions. We take the natural logarithm of labor income for regressions.

5 Conclusion

In this paper, we explore on the labor income profiles in Turkey. We report that labor income profile trajectories of Turkish workers vary considerably over education, gender and sector of employment. We first document that contrary to the observed hump-shaped pattern in most economies, average labor income profile over age in Turkey monotonically increases, which yet arises due to aggregation, as we further document that the isolated effect of age on labor income profile induces a hump-shaped pattern. We complement our descriptive analysis with OLS and pseudo-panel estimations, which confirm our observations on the nature of labor income profiles over the life-cycle. Second, we empirically illustrate that the public sector employment provides higher average incomes and lower dispersions for low-educated employees. However, university-graduates face a risk versus return trade-off in their sectoral choice of employment: the private sector income profiles display both a higher level of average at the cost of a higher degree of cross-sectional variation compared to their public sector counterparts. We document that especially when low-educated, female workers in Turkey have worse income prospects than their male counterparts, which aligns well with the historically low female labor force participation rates of predominantly low-educated Turkish women.

While our findings shed light on several dimensions of heterogeneities in Turkish labor income profiles with a comparable perspective, we believe a full-fledged panel-data analysis would be illuminating. Given data limitations, we leave this to future research.

References

Alessie, R., Lusardi, A., and Aldershof, T. (1997). Income and Wealth over the Life Cycle Evidence from Panel Data. *Review of Income and Wealth*, 43:1–32.

Attanasio, O. and Browning, M. (1995). Consumption over the life cycle and over the business cycle. *American Economic Review*, 85(4):1118–1137.

Ceritoglu, E. (2009). *The Empirical Importance of Precautionary Saving in Turkey*. PhD thesis.

Cilasun, M. (2009). *Income and, Consumption and Saving Behavior of Turkish Household*. PhD thesis.

Cilasun, S. and Kirdar, M. (2009). A Cross-Sectional Analysis of Income, Consumption and Saving Behavior of Turkish Households. *Turkish Economic Association*, 24(280):9–46.

Deaton, A. (1985). Panel Data from Time Series of Cross Sections. *Journal of Econometrics*, 30:109–126.

Eksi, O. and Kirdar, M. (2015). Emek Gelirlerinin ve Eşitsizliğinin Türkiye için bir Analizi: 2002-2011. *Turkish Economic Association*, 30(353):9–44.

Gardes, F., Duncan, G., Gauber, F., Gurgand, M., and Starzec, C. (2005). Panel and Pseudo-Panel Estimation of Cross-Sectional and Time Series Elasticities of Food Consumption: The Case of American and Polish Data. *Journal of Business; Economic Statistics*, 23:242–253.

Inoue, A. (2008). Efficient Estimation and Inference in Linear Pseudo-Panel Data Models. *Journal of Econometrics*, 142(1):449–466.

Kolasa, A. (2012). Life Cycle Income and Consumption Patterns in Transition. *National Bank of Poland Working Paper No. 133*.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2017). Life-Cycle Wage Growth Across Countries. *Journal of Political Economy (Forthcoming)*.

Moffitt, R. (1993). Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections. *Journal of Econometrics*, 59:99–124.

Nazli, S. (2014). *Labor Income and Consumption Age Profiles: The Economic Lifecycle for Turkey*. PhD thesis.

OECD (2012). Inequality in Labour Income – What are its Drivers and How Can it be Reduced? *OECD Economics Department Policy Notes No. 8.*

Piketty, T. (2014). *Capital in the Twenty-First Century.* Harvard University Press.

Rupert, P. and Zanella, G. (2015). Revisiting Wage, Earnings, and Hours Profiles. *Journal of Monetary Economics*, 72:114–130.

Saez, E. and Zucman, G. (2016). Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578.

Tukey, W. (1977). *Exploratory Data Analysis.* Reading, PA: Addison-Wesley.

Verbeek, M. (2008). Pseudo-Panels and Repeated Cross-Sections. *The Econometrics of Panel Data*, Part II:369–383.

Verbeek, M. and Nijman, T. (1992). Can Cohort Data be Treated as Genuine Panel Data. *Empirical Economics*, 17:9–23.

Yang, Y. and Kenneth, L. (2008). Age–Period–Cohort Analysis of Repeated Cross-Section Surveys. *Sociological Methods & Research*, 36(3):297–326.

Yang, Y. and Kenneth, L. (2013). Age–Period–Cohort Analysis: New Models, Methods, and Empirical Applications. *Chapman & Hall/CRC Interdisciplinary Statistics.*

Yukseler, Z. and Turkan, E. (2008). Türkiye'de Hanehalkı: İşgücü, Gelir, Harcama ve Yoksulluk Açısından Analizi. *TUSIAD - T/2008-03/455.*

Ziegelhofer, Z. (2015). Food Prices and Household Welfare: A Pseudo Panel Approach. *Ferdi Policy Brief B138.*

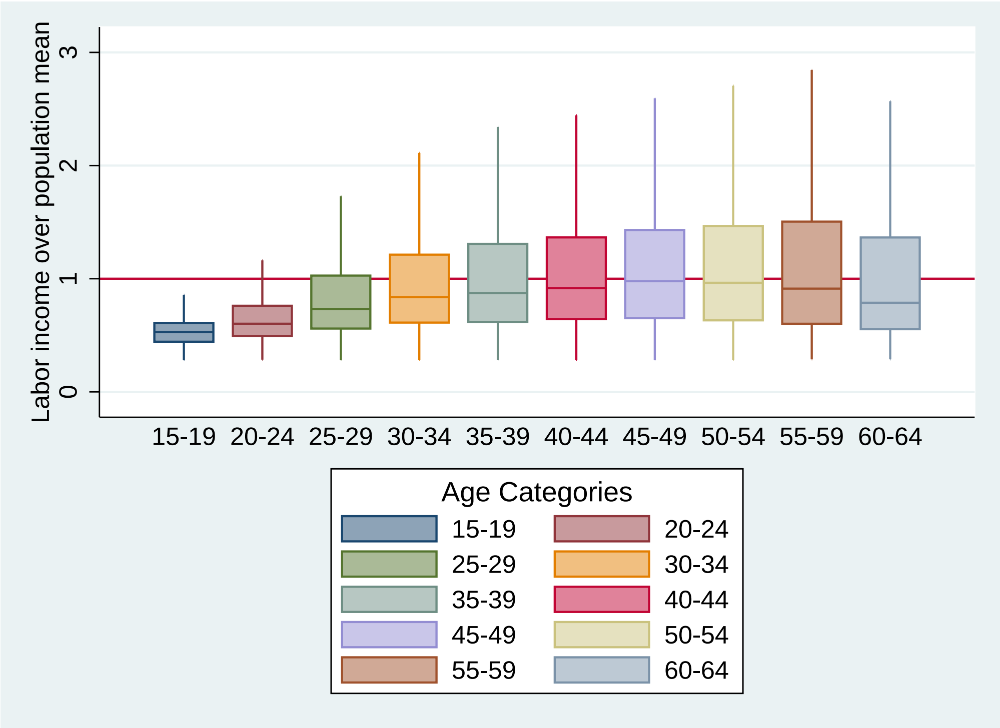


Figure 1: Labor income distribution over age

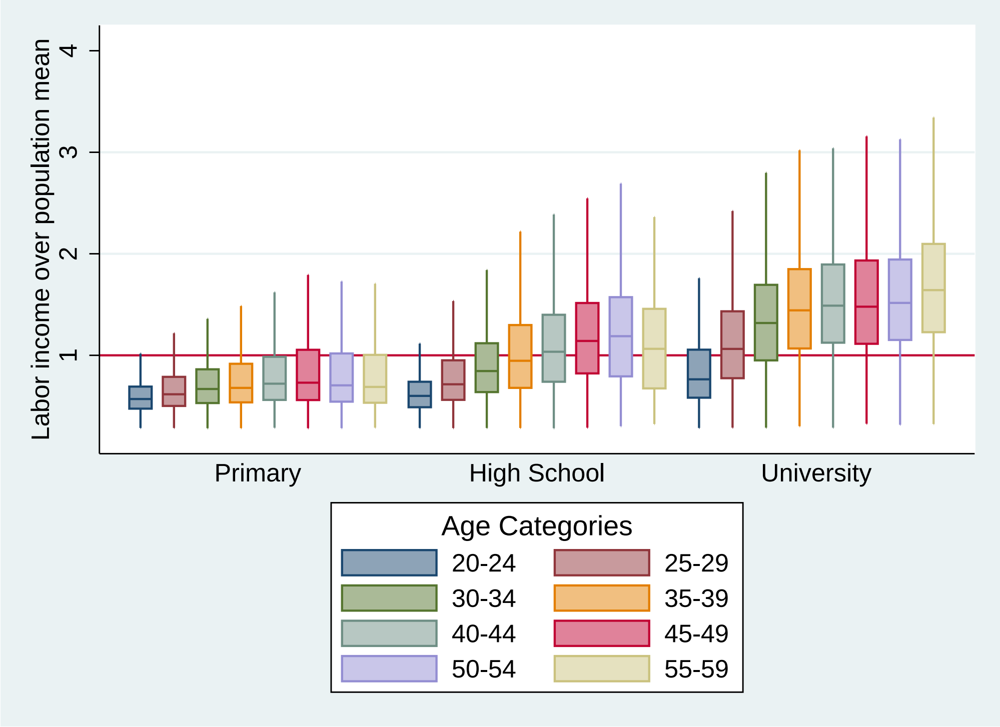


Figure 2: Labor income distribution based over age and education

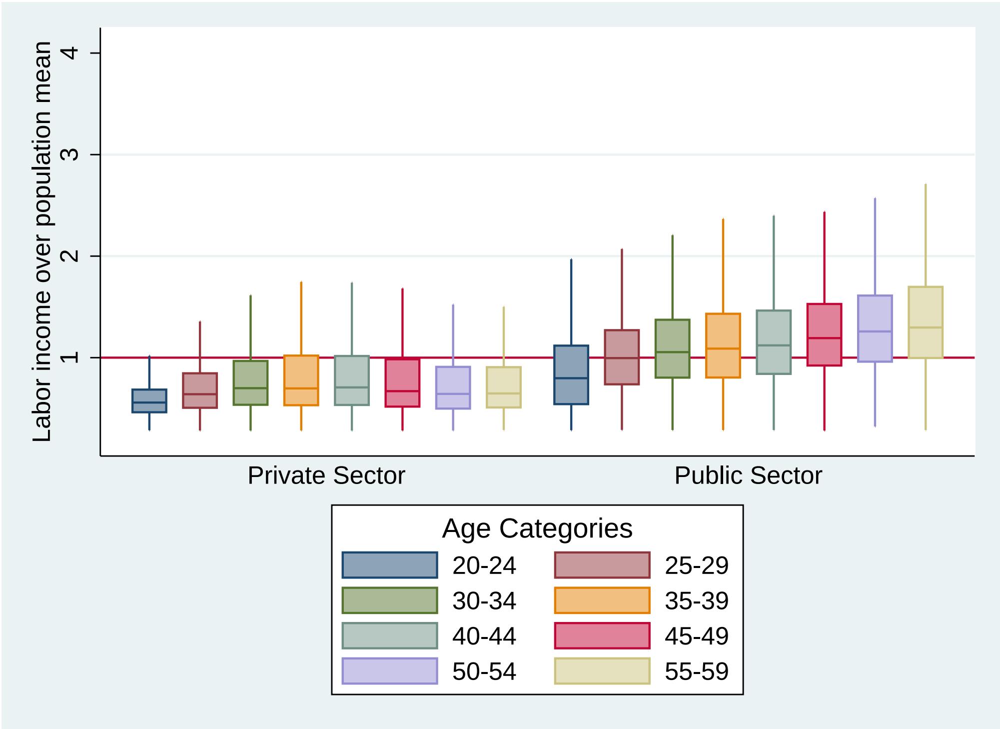
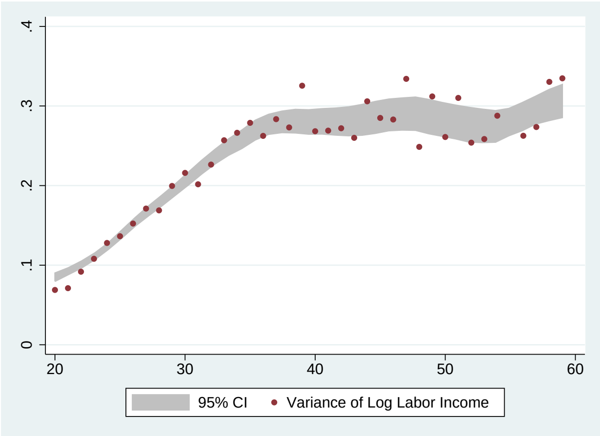
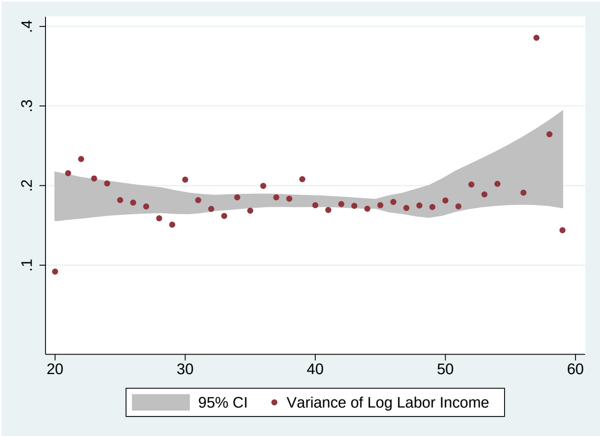


Figure 3: Labor income distribution based over age and sector



(a) Private Sector



(b) Public Sector

Figure 4: Variance of log labor income over age and sector

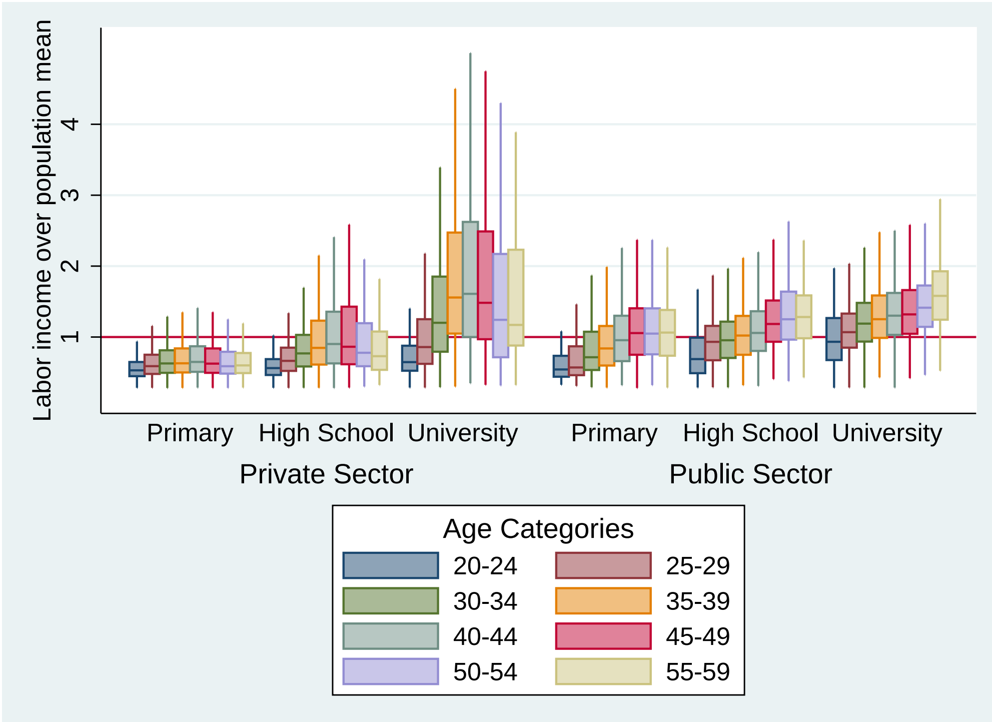


Figure 5: Labor income distribution based over age, education and sector

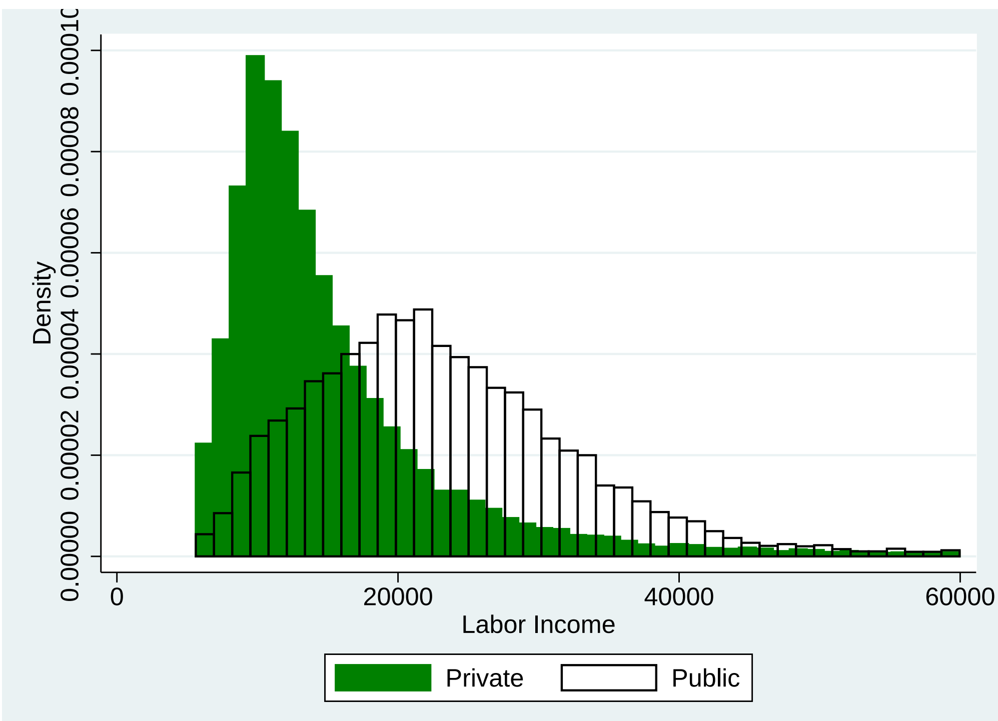
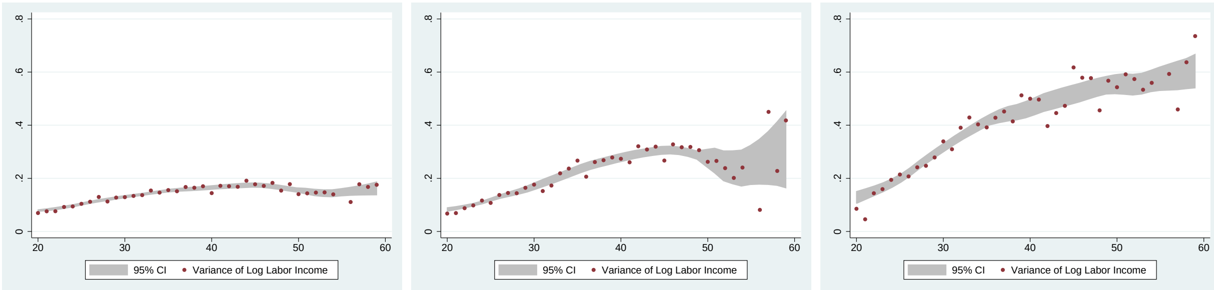


Figure 6: Labor income density over sector



(a) Primary School (b) High School (c) University

Figure 7: Variance of log labor income in private sector over age and education

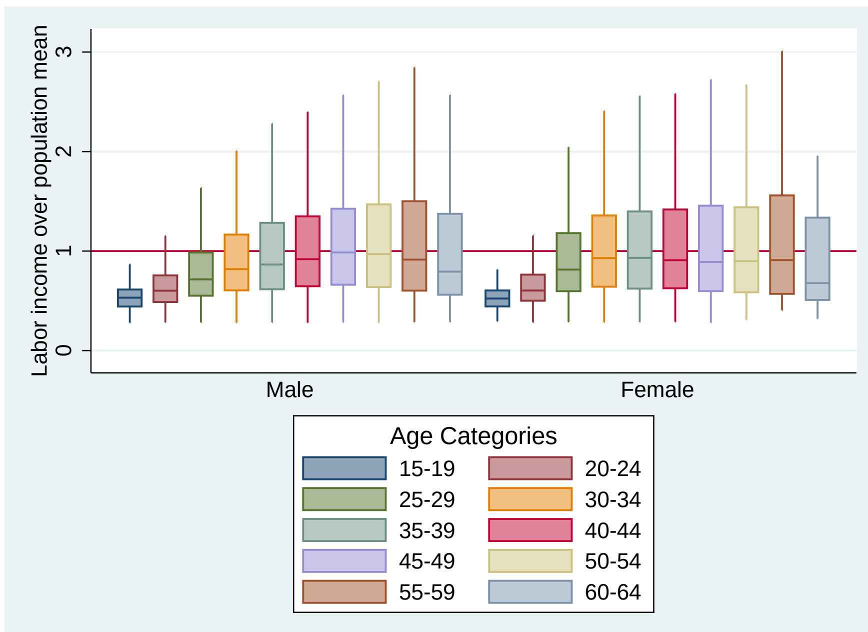
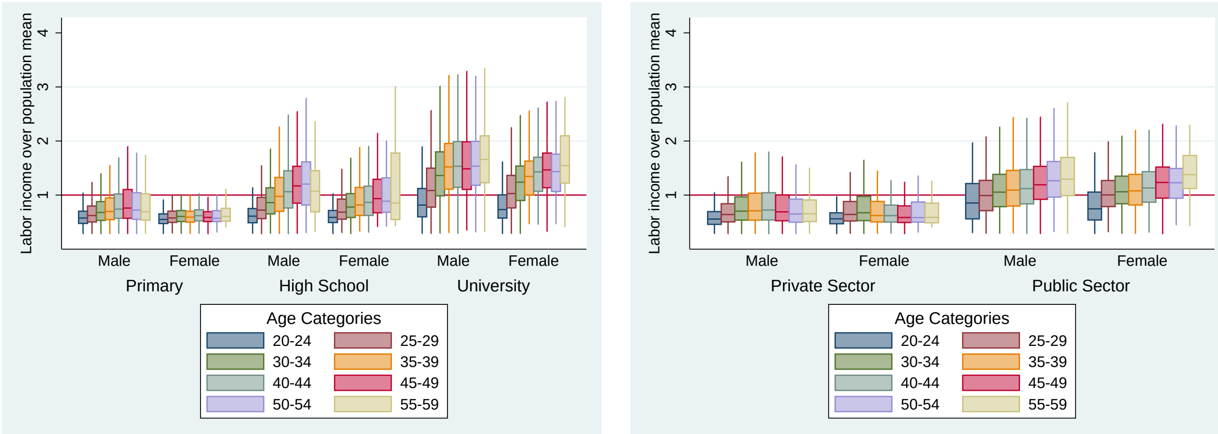


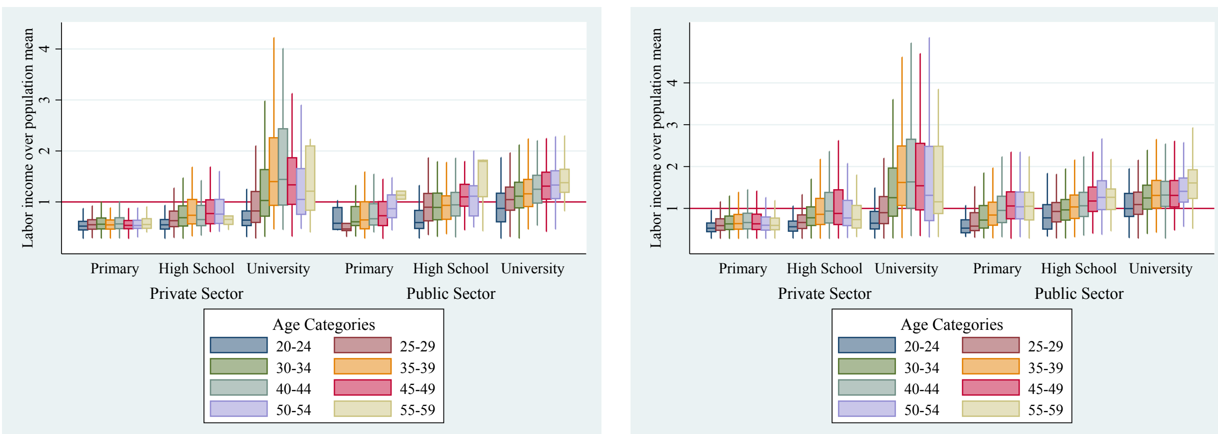
Figure 8: Labor income distribution over age and gender



(a) Gender-education clusters

(b) Gender-sector clusters

Figure 9: Labor income distribution over age, education, sector and gender



(a) Female Distribution

(b) Male Distribution

Figure 10: Labor income distribution over age, education, sector and gender

Tables

Table 1: OLS for Labor Income

	(1) log(Labor Income)	(2) log(Labor Income)
Age		
25 to 29	0.114***	0.108***
	(0.006)	(0.006)
30 to 34	0.199***	0.193***
	(0.006)	(0.006)
35 to 39	0.239***	0.230***
	(0.006)	(0.006)
40 to 44	0.242***	0.228***
	(0.007)	(0.007)
45 to 49	0.232***	0.216***
	(0.008)	(0.008)
50 to 54	0.195***	0.174***
	(0.010)	(0.009)
55 to 59	0.171***	0.156***
	(0.014)	(0.014)
Years of Education		
6-8	0.089***	0.092***
	(0.005)	(0.006)
9-12	0.221***	0.213***
	(0.004)	(0.005)
13-14	0.356***	0.339***
	(0.007)	(0.012)
15-16	0.549***	0.756***
	(0.007)	(0.012)
17+	0.936***	1.200***
	(0.021)	(0.048)
Sector(Public=1)	0.108***	0.228***
	(0.005)	(0.009)
Education-Sector Interaction		
6-8 × Public Sector		-0.071***
		(0.013)
9-12 × Public Sector		-0.057***
		(0.010)
13-14 × Public Sector		-0.070***
		(0.016)
15-16 × Public Sector		-0.406***
		(0.015)
17+ × Public Sector		-0.470***
		(0.052)
Gender(Female=1)	-0.139***	-0.135***
	(0.004)	(0.004)
Union(Member=1)	-0.167***	-0.163***
	(0.005)	(0.005)
Area(Urban=1)	0.091***	0.095***
	(0.004)	(0.004)
Tenure	0.011***	0.011***
	(0.000)	(0.000)
Year Effects	Yes	Yes
N	53878	53878

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Age 20-24 is the basis.

Table 2: OLS for Labor Income Based on Education and Gender

	Male			Female		
	Primary	High School	University	Primary	High School	University
Age						
25 to 29	0.081***	0.145***	0.189***	0.022	0.166***	0.214***
	(0.009)	(0.011)	(0.042)	(0.017)	(0.016)	(0.029)
30 to 34	0.133***	0.248***	0.419***	0.031	0.174***	0.346***
	(0.009)	(0.011)	(0.042)	(0.018)	(0.018)	(0.032)
35 to 39	0.148***	0.297***	0.529***	-0.006	0.197***	0.419***
	(0.009)	(0.013)	(0.043)	(0.016)	(0.022)	(0.036)
40 to 44	0.157***	0.283***	0.519***	0.014	0.156***	0.448***
	(0.009)	(0.014)	(0.046)	(0.017)	(0.025)	(0.045)
45 to 49	0.136***	0.265***	0.523***	-0.038*	0.213***	0.484***
	(0.010)	(0.016)	(0.049)	(0.018)	(0.034)	(0.052)
50 to 54	0.100***	0.217***	0.492***	-0.039	0.163***	0.414***
	(0.012)	(0.020)	(0.052)	(0.028)	(0.043)	(0.069)
55 to 59	0.066***	0.211***	0.536***	-0.051	0.028	0.451***
	(0.017)	(0.034)	(0.062)	(0.046)	(0.108)	(0.111)
Sector(Public=1)	0.191***	0.142***	-0.146***	0.149***	0.179***	-0.049*
	(0.008)	(0.009)	(0.018)	(0.028)	(0.017)	(0.023)
Union(Member=1)	-0.326***	-0.144***	0.040***	-0.237***	-0.130***	-0.002
	(0.010)	(0.008)	(0.012)	(0.037)	(0.016)	(0.015)
Area(Urban=1)	0.113***	0.087***	0.066***	0.027	0.056**	0.039
	(0.006)	(0.009)	(0.019)	(0.017)	(0.018)	(0.027)
Tenure	0.009***	0.016***	0.005***	0.013***	0.015***	0.005*
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	23545	15138	5745	2521	3946	2983

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Age 20-24 is the basis.

Table 3: OLS for Labor Income Based on Education and Sector

	Private Sector			Public Sector		
	Primary	High School	University	Primary	High School	University
Age						
25 to 29	0.068***	0.121***	0.178***	0.028	0.145***	0.172***
	(0.008)	(0.009)	(0.037)	(0.069)	(0.028)	(0.028)
30 to 34	0.118***	0.209***	0.435***	0.128	0.155***	0.288***
	(0.008)	(0.011)	(0.041)	(0.066)	(0.027)	(0.028)
35 to 39	0.132***	0.279***	0.638***	0.138*	0.162***	0.358***
	(0.008)	(0.015)	(0.046)	(0.065)	(0.028)	(0.030)
40 to 44	0.139***	0.293***	0.580***	0.140*	0.142***	0.398***
	(0.009)	(0.018)	(0.057)	(0.065)	(0.029)	(0.032)
45 to 49	0.098***	0.272***	0.590***	0.136*	0.156***	0.439***
	(0.010)	(0.022)	(0.067)	(0.066)	(0.031)	(0.036)
50 to 54	0.058***	0.192***	0.348***	0.117	0.134***	0.475***
	(0.012)	(0.029)	(0.067)	(0.067)	(0.034)	(0.039)
55 to 59	0.051**	0.209***	0.455***	0.022	0.086*	0.525***
	(0.017)	(0.054)	(0.101)	(0.072)	(0.043)	(0.047)
Gender(Female=1)	-0.154***	-0.068***	-0.153***	-0.176***	-0.120***	-0.128***
	(0.006)	(0.009)	(0.024)	(0.025)	(0.010)	(0.009)
Union(Member=1)	-0.305***	-0.242***	-0.152**	-0.303***	-0.091***	0.041***
	(0.015)	(0.014)	(0.055)	(0.012)	(0.008)	(0.009)
Area(Urban=1)	0.082***	0.087***	0.177**	0.173***	0.081***	0.050***
	(0.006)	(0.012)	(0.064)	(0.013)	(0.010)	(0.012)
Tenure	0.006***	0.023***	0.026***	0.018***	0.013***	-0.000
	(0.000)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	21791	11320	2846	4275	7764	5882

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. Age 20-24 is the basis.

Table 4: Pseudo-Panel Estimation

	log(Labor Income)			
	(1)	(2)	(3)	(4)
Age	0.1717***	0.1912***	0.1595***	0.1807***
	(25.61)	(26.57)	(25.29)	(26.69)
Age^2	-0.0017***	-0.0018***	-0.0017***	-0.0019***
	(-18.63)	(-20.34)	(-21.26)	(-23.36)
Female × Age		-0.0205***		-0.0202***
		(-6.02)		(-6.70)
High School × Age			0.0141***	0.0107**
			(3.98)	(3.13)
University × Age			0.0386***	0.0380***
			(10.90)	(11.30)
Cohort Effects	Yes	Yes	Yes	Yes
N	455	455	455	455

Table 5: Variances of Permanent and Transitory Real Labor Income, 2002-2014

	Permanent Variance			Transitory Variance		
Sample Definition	2002-08	2009-2014	Change	2002-08	2009-2014	Change
All	0.311	0.333	7%	0.160	0.094	-41%
Education						
Primary School	0.179	0.134	-25%	0.148	0.067	-54%
High School	0.248	0.227	-8%	0.162	0.098	-39%
University	0.223	0.251	12%	0.172	0.118	-31%
Age						
25-34	0.242	0.256	5%	0.153	0.092	-39%
35-44	0.290	0.340	17%	0.143	0.075	-47%
45-54	0.289	0.358	23%	0.109	0.081	-25%
Earnings %						
Lowest 33 percent	0.086	0.079	-8%	0.127	0.061	-51%
33-66 percent	0.068	0.085	25%	0.072	0.058	-19%
Top 66 percent	0.153	0.175	14%	0.109	0.096	-12%

Variances are for log annual earnings adjusted to 2014 liras

APPENDIX-A

OLS Analysis Using Household Labor Force Survey

Table A1: OLS for Wage

	(1) Wage	(2) Wage
Age		
25 to 29	0.096***	0.094***
	(0.002)	(0.002)
30 to 34	0.169***	0.168***
	(0.002)	(0.002)
35 to 39	0.210***	0.209***
	(0.002)	(0.002)
40 to 44	0.223***	0.221***
	(0.003)	(0.003)
45 to 49	0.219***	0.214***
	(0.003)	(0.003)
50 to 54	0.219***	0.213***
	(0.003)	(0.003)
55 to 59	0.199***	0.195***
	(0.005)	(0.005)
Years of Education		
6-8	0.071***	0.070***
	(0.002)	(0.002)
9-12	0.175***	0.155***
	(0.002)	(0.002)
13+	0.534***	0.564***
	(0.002)	(0.003)
Sector(Public=1)	0.283***	0.280***
	(0.002)	(0.004)
Education-Sector Interaction		
6-8 × Public Sector		0.005
		(0.005)
9-12 × Public Sector		0.082***
		(0.004)
13+ × Public Sector		-0.048***
		(0.005)
Gender(Female=1)	-0.134***	-0.134***
	(0.001)	(0.001)
Tenure	0.009***	0.009***
	(0.000)	(0.000)
Year Effects	Yes	Yes
District Effects	Yes	Yes
N	336484	336484

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$.

Table A2: OLS for Wage Based on Education and Sector

	Private Sector			Public Sector		
	Primary	High School	University	Primary	High School	University
Age						
25 to 29	0.037***	0.064***	0.219***	0.020***	0.052***	0.238***
	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.007)
30 to 34	0.056***	0.123***	0.376***	0.026***	0.110***	0.358***
	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.007)
35 to 39	0.065***	0.176***	0.493***	0.015**	0.123***	0.425***
	(0.003)	(0.004)	(0.008)	(0.005)	(0.008)	(0.009)
40 to 44	0.073***	0.198***	0.536***	0.010*	0.089***	0.469***
	(0.003)	(0.005)	(0.009)	(0.005)	(0.009)	(0.011)
45 to 49	0.084***	0.160***	0.531***	0.002	0.049***	0.475***
	(0.003)	(0.006)	(0.010)	(0.006)	(0.012)	(0.014)
50 to 54	0.085***	0.152***	0.526***	0.008	0.069***	0.452***
	(0.004)	(0.007)	(0.012)	(0.008)	(0.021)	(0.016)
55 to 59	0.084***	0.109***	0.496***	0.030*	0.134**	0.400***
	(0.006)	(0.011)	(0.014)	(0.013)	(0.045)	(0.024)
Sector(Public=1)	0.320***	0.342***	0.250***	0.235***	0.301***	0.226***
	(0.003)	(0.004)	(0.005)	(0.012)	(0.007)	(0.005)
Area(Urban=1)	0.017***	0.038***	0.064***	0.007	0.031***	0.031***
	(0.002)	(0.003)	(0.005)	(0.004)	(0.007)	(0.007)
Tenure	0.007***	0.014***	0.002***	0.009***	0.016***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	125210	73428	66504	16083	18825	3871

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$.

APPENDIX-B

Age-Period-Cohort (APC) Analysis

In order to unveil the role of time-varying components in the life-cycle income analysis, one needs to explore age, period and cohort effects. Age effects are variations linked to social processes of ageing specific to individuals, but orthogonal to time periods and birth cohorts. Period effects are the sum of all external factors that equally influence all age groups at a certain year.[28] Finally, cohort effects result from the unique experience of each cohort as time goes by. Age-Period-Cohort (APC) analysis allows us to disentangle independent effects of these factors and to estimate the effects of age, period and cohort effects separately.

Figure B1 and Figure B2 are useful for providing insights on temporal patterns. Since the shape of the birth cohort curve is affected by both varying age effects and period effects, they do not provide an accurate quantitative evaluation on the sources of change. Each graph describes only the variation in the labor income that can attributed to factors associated with age or year. From Figure B1 we expect to see a positive cohort effect for younger generations, because in a particular age group, younger cohorts earn more. However, the year effect might be also account for this difference. Thus, there is need for statistical regression modelling to capture how these three effects work simultaneously.

[28] Examples of social, economic and environmental factors can be wars, natural disasters, and crises.

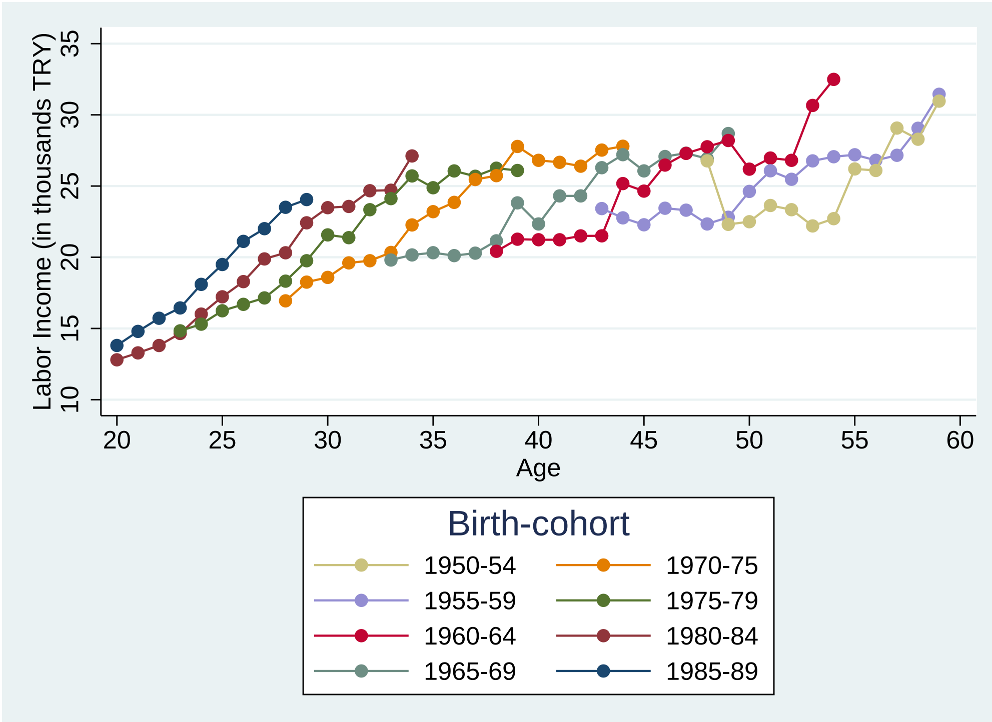


Figure B1: Labor income over life-cycle by birth-cohort

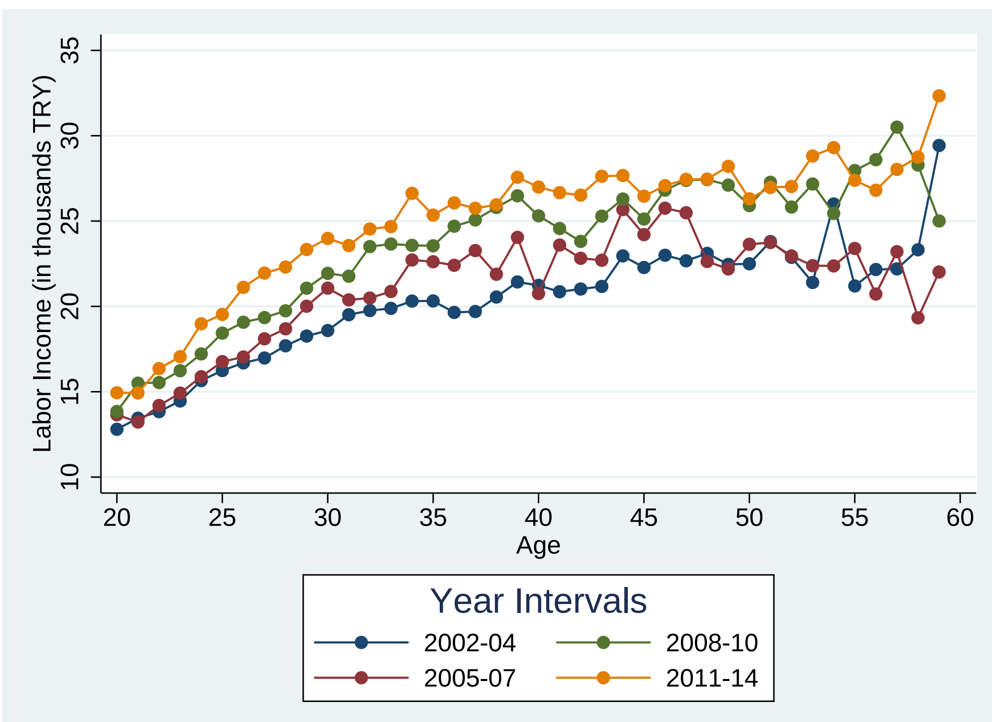


Figure B2: Labor income over life-cycle by time periods

The main impediment to estimate the independent effect of age, period and cohort is the identification problem resulting from these three effects being perfectly collinear (cohort=year-age), i.e. given any two of them, one can precisely determine the third one.

$$W_{ij} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ij} \tag{3}$$

where W_{ij} denotes the observed mean labor income values for the ith age group for i=20,...,60 at the jth year for j=2002,...,2014. μ stands for the intercept or adjusted mean labor income, α_i is the coefficient for the ith age group, β_j is the coefficient for the jth year, γ_k is the coefficient for the kth cohort for k=1942,...,1995 and ϵ_{ij} is of the white noise form.

After re-parametrization as follows,

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0 \tag{4}$$

the model (3) can be written in the following matrix form:

$$Y = Xb + \epsilon \tag{5}$$

where Y is a vector of mean labor income values, X is the design matrix consisting of dummy variable column vectors (Yang and Kenneth, 2008) and ϵ is a vector of random errors with mean zero. Parameter b is defined as follows:

$$b = (\mu, \alpha_{20}, ..., \alpha_{59}, \beta_{2002}, ..., \beta_{2013}, \gamma_{1942}, ..., \gamma_{1992})^T \tag{6}$$

It is important to note that $\alpha_{60}, \beta_{2014}$ and γ_{1993} are excluded from (6) so that constraint (4) can be satisfied. The identification problem is clear when we reformulate (5):

$$\hat{b} = (X^T X)^{-1} X^T Y \tag{7}$$

Due to perfect multicollinearity of age, period and cohort, the design matrix X is one less than full-column rank. Since the inverse of this singular matrix does not exist, it is not possible to estimate age, period or cohort effects without any further restrictions or constraints. That is why the main purpose is to break the linear dependency between these three effects. There are many solutions[29] to the identification problem, but in this paper, we will consider the most

[29] Reduced two-factor models, constraint generalized linear models, non-linear transformation, and proxy variables are some of the solutions.

recent technique, the intrinsic estimator (Yang and Kenneth, 2013).

The parameter space of the unconstrained model (5) can be decomposed into two orthogonal linear subspaces and formalized as follows:

$$b = b_0 + sB_0 \tag{8}$$

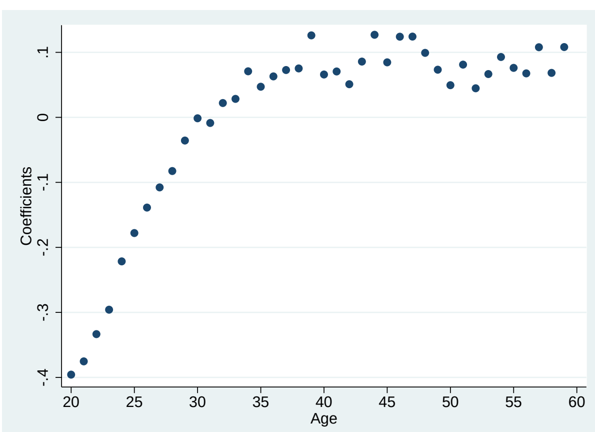
where $b_0 = P_{proj}b$ is the projection of the b to nonnull space of X. B_0 is a unique eigenvector and depends only on matrix X, which is determined by number of age groups and periods. The intrinsic estimator imposes a constraint on the geometric orientation of the parameter b: the eigenvector B_0 in the null space of X has no influence on the parameter b_0. Since B_0 does not depend on observed values, it is a sensible constraint.

$$B = (I - B_0 B_0^T)\hat{b} \tag{9}$$

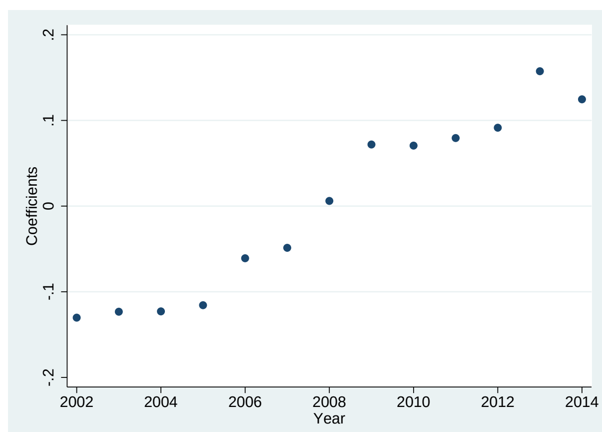
We first estimate $\hat{b}$ of model (4) and project $\hat{b}$ on the intrinsic estimator B by removing the component in the B_0 direction (Yang and Kenneth, 2013).

$$X\hat{b} = X(B + tB_0) = XB + 0 = XB \tag{10}$$

In short, intrinsic estimator allows us to estimate the projection of the unconstrained vector on the nonnull space of the matrix X by removing the influence of null space.[30]



(a) Age Effects



(b) Period Effects

Figure B3: Effect Coefficients

For robustness purposes, we use the conventional approach to APC models, i.e. the coefficient

[30] We use the *apc_ie* command in `STATA` to estimate age, period and cohort effects.

constraints approach. As the identifying constraint on the parameter vector b in equation (6), the equality of the effect coefficients of the first two periods is imposed as the only constraint that makes the matrix $(X^T X)$ in equation (7) non-singular and allows the estimation of the effects separately. The results are consistent with the outcomes of the intrinsic estimator approach.[31]

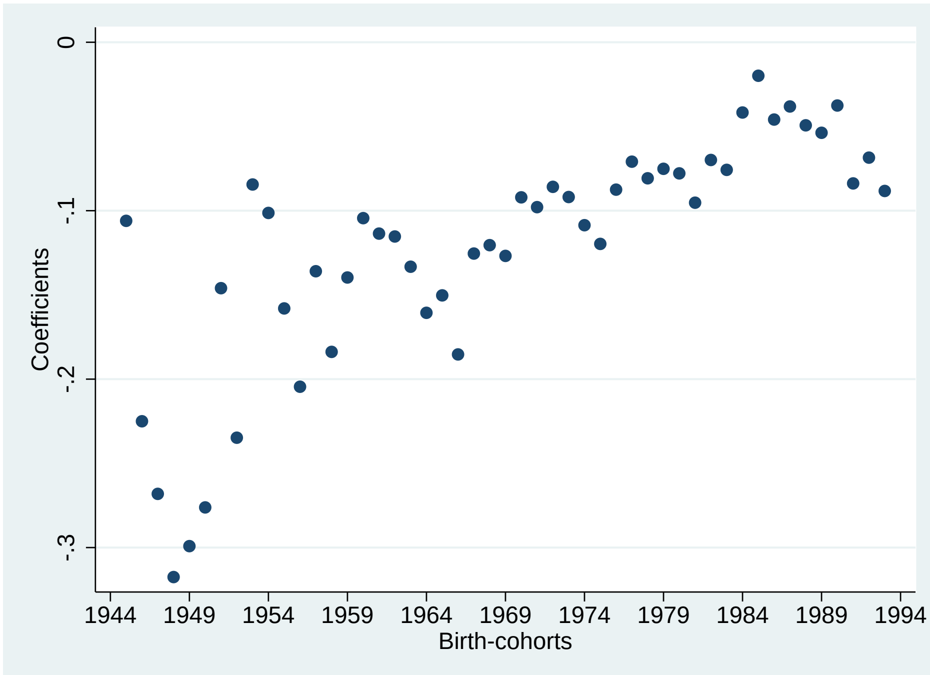


Figure B4: Cohort Effects

The age effects are consistent with our findings: the labor income is increasing rapidly until age 35 and stays nearly the same, with small oscillations until age 60 for the working population. At the same time, we observe a monotonic increase for the period effects in Figure B3. Since the Turkish economy is constantly growing (with the exception of the year 2009), the year effect on labor income increases over time. Finally, as shown in Figure B4, the cohort effect shows an increasing trend, where some cycles are without any particular path. Since the share of university and high school graduates in the population increases with younger cohorts, we expect to see a higher cohort effect for younger cohorts. That is consistent with the data shown Figure B1, because at a particular age, younger cohorts earn more than older ones. The minor fall at the right end of the graph is due to the cohorts that are still in the education process. Similarly, the increasing period effect is consistent with the data in Figure B2 because the real labor income increases over the years due to positive GDP growth.

[31] However, changing the constraint can produce widely different estimates for the effects.

APPENDIX-C

Variance/Mean Profiles over Age

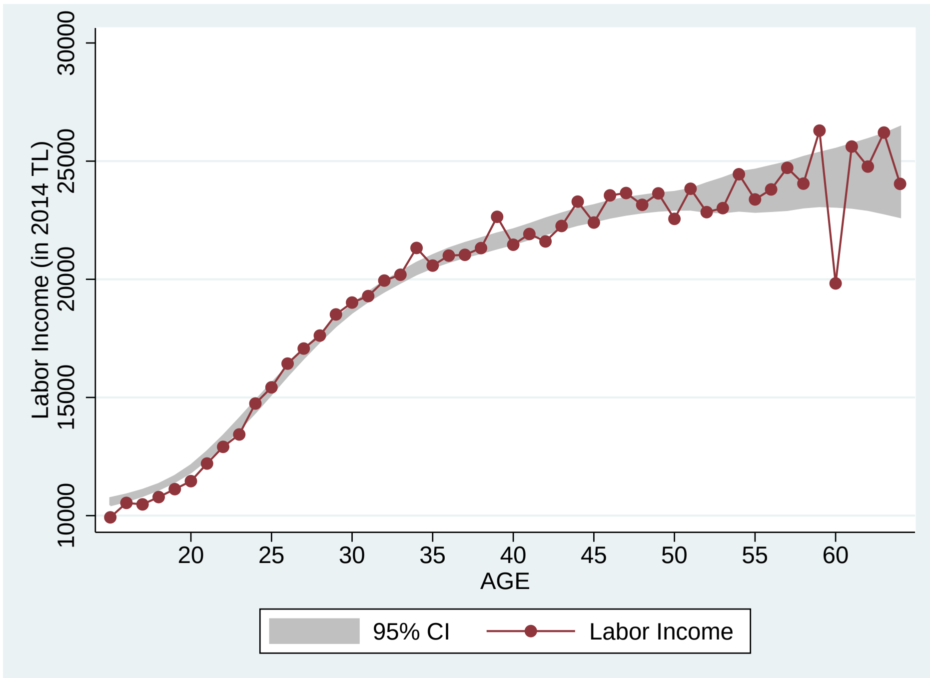


Figure C1: Full sample

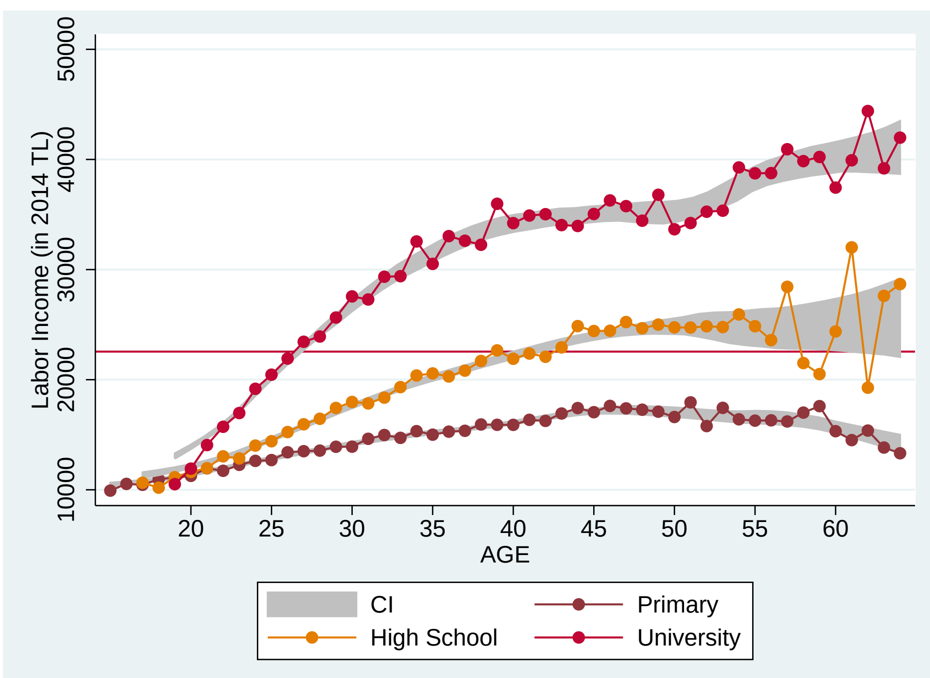
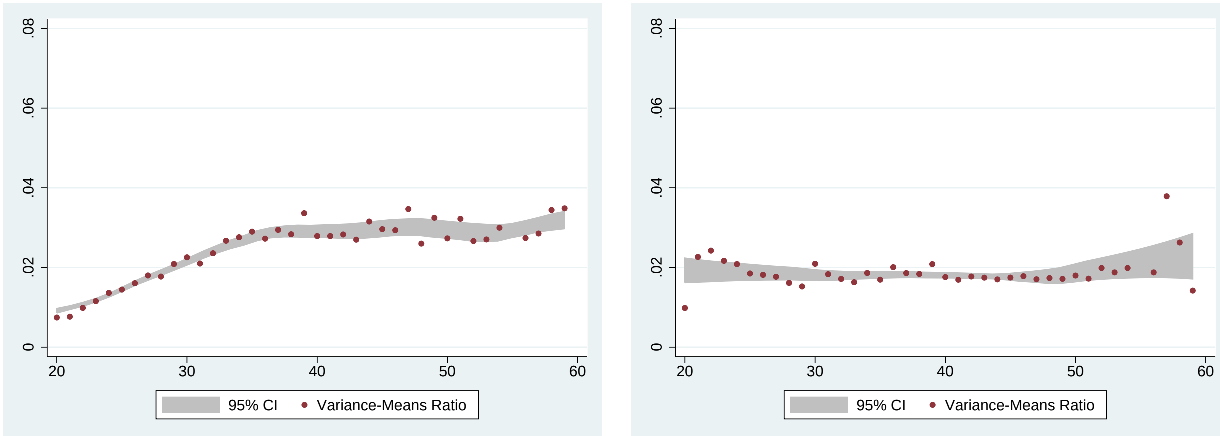
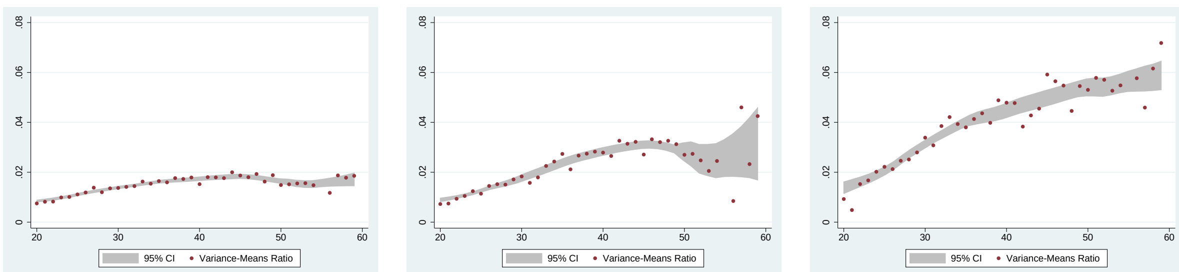


Figure C2: Labor income over education groups



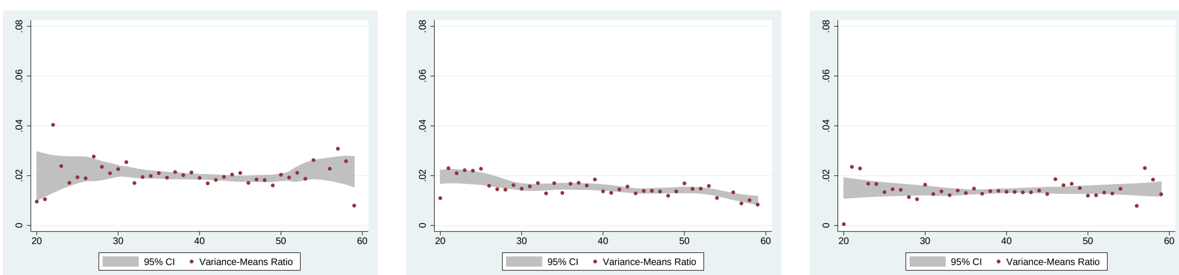
(a) private sector (b) public sector

Figure C3: Variance/Mean



(a) primary school (b) high school (c) university

Figure C4: Variance/Mean (private sector)



(a) primary school (b) high school (c) university

Figure C5: Variance/Mean (public sector)

APPENDIX-D

Graphs on Isolated Age Effect

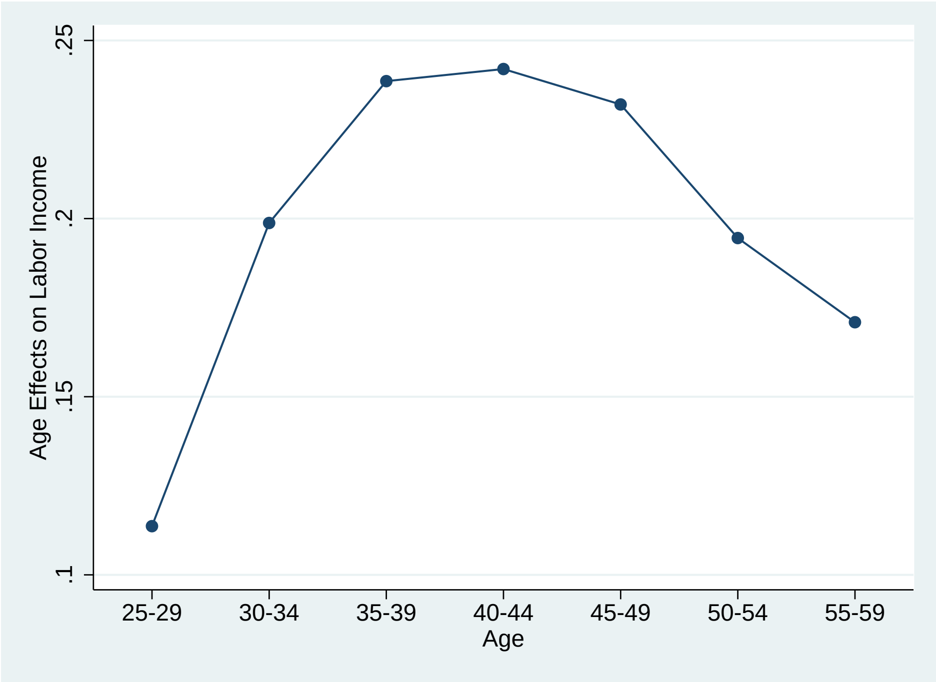


Figure D1: Isolated age effect

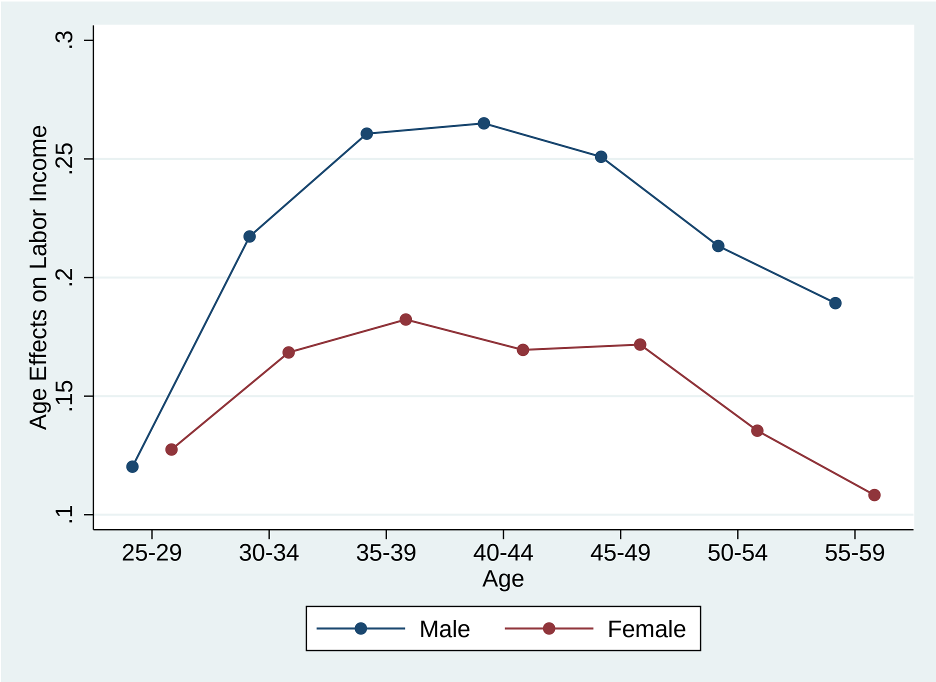


Figure D2: Isolated age effect over gender

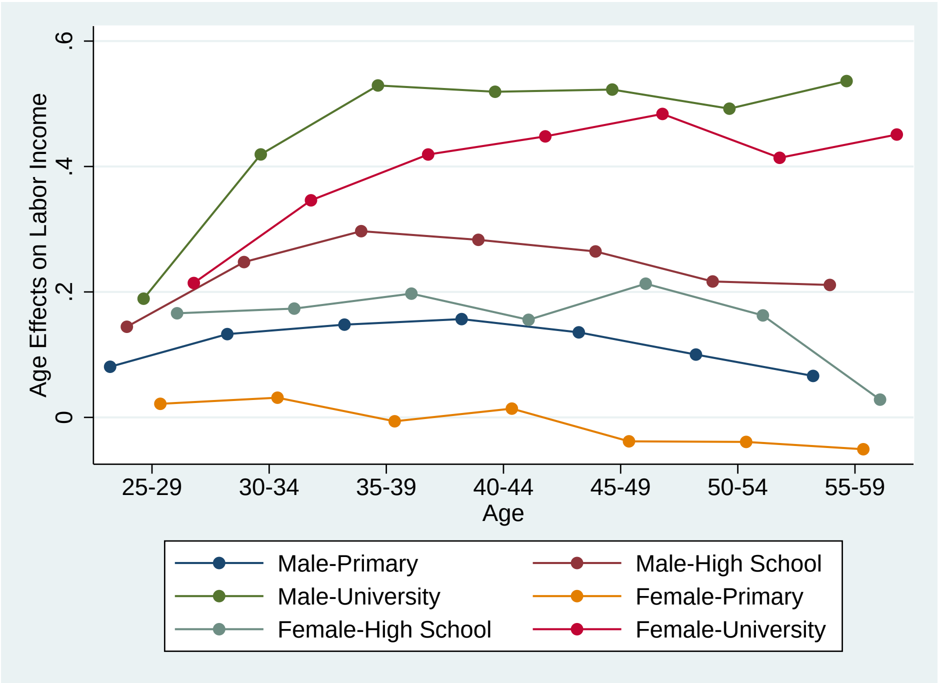


Figure D3: Isolated age effect over education and gender

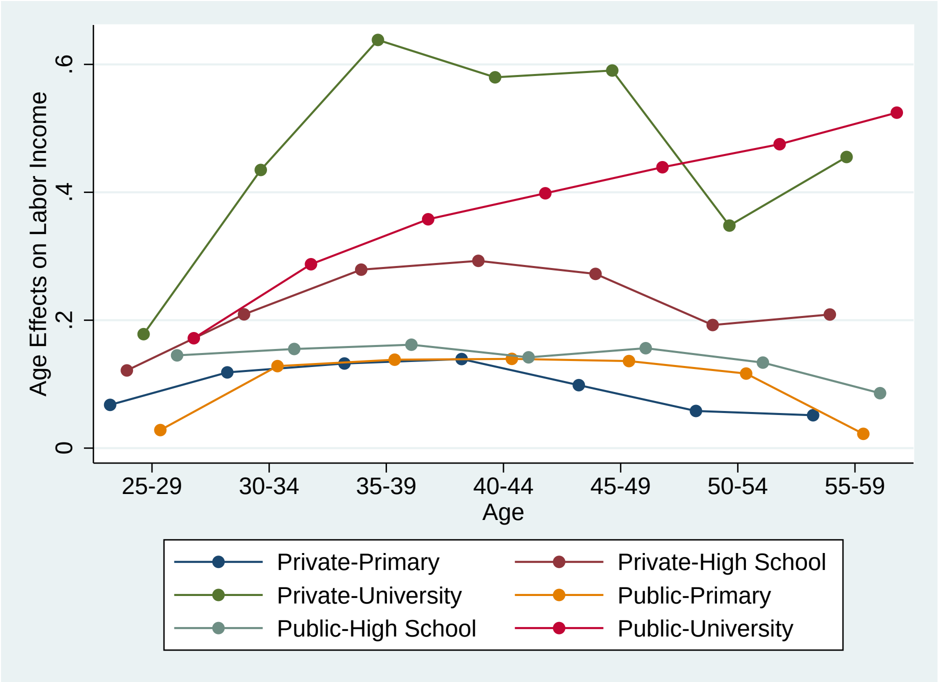


Figure D4: Isolated age effect over education and sector

APPENDIX-E

Pseudo-Panel Method

The pseudo-panel method has many advantages for our purposes. In genuine panel data analysis, the main concern is measurement error. The pseudo-panel approach reduces measurement error bias due to the aggregation of individuals into cohorts. Yet, the bias and efficiency trade-off is critical: increasing cohort size decreases measurement error and bias, but it decreases number of cohorts and efficiency. We optimally define cohorts considering with this trade-off in mind.

The genuine panel data are subject to attrition and non-response bias, and that data spans short time periods such as 3 or 4 years for Turkish data. On the other hand, pseudo-panel data tends to suffer less from attrition and non-response bias, because each individual is observed only once. The data is often larger, both in terms of the number of individuals and in the time period it spans due to simply being repeated cross-sectional data (Verbeek, 2008). Pseudo-panel data may consist of systematic heteroscedasticity via aggregation. To prevent serious estimation errors, we weight each observation by a heteroscedasticity factor that is a function of cell size by following Gardes et al. (2005). Arguably, there are downsides of the pseudo-panel approach as well, such as loss of individual information due to aggregation, but for our purposes and data in hand, this is the best possible empirical approach, which is widely accepted in the literature.

Theoretically, the cohort size needs to go infinity in order to be able to treat pseudo-panel data as though they are genuine panels, so that conventional methods like fixed-effects estimator can be employed (Inoue, 2008). That is why cohort size should be sufficiently large. More than one hundred individuals in each cohort is suggested by Verbeek and Nijman (1992) to reduce the measurement error bias to a negligible degree. Since the measurement error becomes negligible only when cohort sizes are large (Moffitt, 1993) and HBS data is not large enough for Turkey, the minimum cohort size is set at 50, following Ziegelhofer (2015) Monte Carlo Simulation outcomes. Ziegelhofer claims that the increasing bias resulting from decreasing the limit from 100 to 50 is not a significant amount for the estimation. At the same time, the number of total observations has to be large enough so that statistical efficiency can be obtained, which is 455 for our pseudo-panel data. That is to say, there is an obvious trade-off between cohort size and the number of cohorts (Verbeek, 2008). The larger the number of cohorts, the smaller is the cohort size, which leads to better estimation efficiency but higher measurement error. That is why we have applied some variations in cohort forming such as excluding both the public and the private sectors, but the results do not change much.

For each cohort and each year, we calculate the mean of log income. Our synthetic data includes 59 cohorts, 13 time periods and an average cohort size of 158. There are 455 observations, which is less than 59x13, because the pseudo-panel is not balanced owing to an insufficient number of observations for particular groups in some years. If only one year of data exists for a cohort, we exclude this cohort because it lacks the panel property.