HETEROGENEITY IN LABOR INCOME PROFILES: EVIDENCE FROM TURKEY*

Emrehan Aktuğ[†]

Tolga Umut Kuzubas[‡]

University of Texas at Austin

Boğaziçi University

Orhan Torul§

Boğaziçi University

Abstract

In this paper, we investigate labor income profiles in Turkey. In doing so, we investigate the role of educational attainment, gender, and the public versus private sector employment on labor income profiles by using the Household Budget Survey data from 2002 to 2014. We first report that while average labor income profile in Turkey exhibits a moderate hump-shape over age, there exists an immense degree of heterogeneity in labor income trajectories over education, gender and sector of employment. Second, while the public sector employment is more advantageous for low-educated Turkish employees, university graduates in Turkey's labor market face a risk versus return trade-off in their choice of sectoral employment: the private sector labor 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. Our findings via distributional clustering analysis, ordinary least squares and pseudo-panel estimations all indicate that in attempts to infer about economy-wide average labor income profiles, abstracting away from any of these listed factors could lead to misleading inferences.

Keywords: pseudo-panel analysis; synthetic cohorts; public vs private sector; education; gender **JEL Classification:** D31, I24, R20

^{*}We would like to thank Murat Güray 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 and 13920. All remaining 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

Both the number and the scope of studies addressing income inequality have risen sharply 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 Turkey has the second highest level of income inequality among OECD countries in 2014.² In this paper, we analyze heterogeneity in labor income profiles of Turkish employees, aiming to discern the determinants of this notable labor income inequality in an era of increasing distributional concerns.

The literature examining income profiles over the life-cycle has expanded in recent years, especially for developed economies.³ In particular, the recent comprehensive analysis by Lagakos et al. (2018) extracts labor income profiles in a selected group of both developed and developing countries by concentrating on male workers in the private sector. Our work complements their paper, and extends their analysis by dissecting the role of gender and the public versus private sector of employment in Turkey, while also factoring in the role of education, thereby constituting the first comprehensive research on labor income profiles in Turkey. Using a rich cross-sectional data set by the Turkish Statistical Institute's Household Budget Survey covering 2002 to 2014, our results document that income profiles in Turkey vary immensely over these previously unveiled characteristics.

We first show that the average life-cycle labor income profile in Turkey is moderately hump-shaped over age with a peak around 45, similar as in the case of the United States and Germany. Our decomposition exercises, however, reveal novelly that this pattern is generated by the private sector employees with high school and university backgrounds, since the public sector employees face almost monotonic and ever-increasing labor income trajectories. We further report that cross-

¹See Piketty (2014), Piketty (2015), Atkinson (2015), Milanovic (2016) and Saez and Zucman (2016) for through discussions on recent debates about several dimensions of income and wealth inequality, and Tamkoc and Torul (2018) for a review and comparison of the economic inequalities in Turkey.

²See OECD Income Inequality Database for further details.

³See Lagakos et al. (2018), Kolasa (2012) and Rupert and Zanella (2015), among others.

sectional dispersion of labor income in Turkey is also increasing with age, in accordance with the findings on developing countries, yet contrary to developed economies.⁴ We document that this pattern is also driven by the private sector compensations, as variance-to-mean income ratio over age is ever-increasing for the private sector but stagnant for the public sector employees.

Second, we document that the private sector labor income profiles of university graduates in Turkey display both a higher average level and a higher degree of cross-sectional dispersion compared to their public sector counterparts, thereby implicating a *risk versus return trade-off* for their sectoral choice of employment.⁵ This risk versus return trade-off, however, is not applicable for primary or high school graduate employees, as the public sector employees with below-university education earn more on average and face lower cross-sectional variation than their private sector counterparts, while having to compete for the scarce public sector job positions.

Third, we find strong evidence for a gender pay gap in Turkey, especially prevalent among primary school graduate employees. This observation is consistent with the historically low female labor force participation rate in Turkey, which is only around 30%. Indeed, our results confirm that female labor force participation rate increases with education, which is as high as 70% among female university graduates. However, the fact that 50% of women in the Turkish labor force are primary school graduates, combined with the significant and life-long persistent wage gap for this group is likely to impede a higher female labor force participation in Turkey.⁶

Throughout our investigation, we first use descriptive graphical analyses to shed light on the heterogeneity in labor income profiles in Turkey, and we complement our graphical analyses with ordinary least squares (OLS) and pseudo-panel estimations. We further verify that our results hold true when using alternative data sets. We believe our use of several analytical approaches

⁴See Lagakos et al. (2018), among others.

⁵This finding is particularly relevant and critical for the Turkish economy, as the share of university graduates in the labor force is ever-increasing, and the public sector employment constitutes no less than a non-negligible 12-15% of total employment in Turkey, which is going to increase by a further 3-5% due to a recent decree law (KHK No:696) enacted in late December 2017. See OECD (2015) for further details.

⁶For a detailed discussion on this issue, see Uraz et al. (2010) and Tansel (1994).

and data sets offer both rigor and robustness in our findings on the several critical dimensions of heterogeneity in labor income profiles in Turkey.

The rest of paper is organized as follows: in section 2 we summarize the previous literature; in section 3 we describe the data and provide the detailed description of labor income over various clusters; in section 4 we explain our estimation methodology and present our results, and in section 5 we discuss our findings and conclude.

2 Literature Review

Earlier literature on labor income profiles documents well that average labor income profiles exhibit a hump-shaped pattern over age in many developed countries.⁷ Findings on developing economies are, however, rather limited, and Turkey is no exception.⁸ Among the very few studies on Turkey, Cilasun and Kirdar (2009) investigate average life-cycle income profiles of household-heads in Turkey between 2002-2006 and report that median income profiles display a hump-shaped pattern over life-cycle when controlling for educational attainment. However, authors do not concentrate on labor income of households and do not investigate the role of gender and the public versus private sector of employment, both of which have first-order implications, as we document in this paper.

Turkish Statistical Institute's Household Budget Survey (HBS), the data set we use for our benchmark analysis, has been used widely to address several other questions related to income inequality, precautionary savings, income and expenditure decompositions, but not labor income profiles. Eksi and Kirdar (2015) investigate wage inequality in Turkey by addressing the role of educational attainment for the 2002-2011 period and Bakis and Polat (2015) study the same topic by addressing industries as well for the 2002-2011 period but with Household Labor Force

⁷See Attanasio and Browning (1995) and Alessie et al. (1997), among others.

⁸See Lagakos et al. (2018) for a recent discussion on the developments in developing countries.

Survey (HLFS). Nazli (2014), Yukseler and Turkan (2008) and Ceritoglu (2009) focus on savings decisions of households but not labor income profiles over life-cycle. Tansel (2005) and Tansel (1994) demonstrate the public versus private sector wage differentials and a gender pay gap in Turkey via a single year of individual-level data. Tansel et al. (2018) studies income inequality by presenting 90/10 and 90/50 wage ratios by gender, age, education, and sector by employing the Surveys on Income and Living Conditions (SILC). Tamkoc and Torul (2018) study the evolution of wage, income and consumption inequality in Turkey after 2002 via a cross-country comparable methodology. Regarding labor income profile over life-cycle in Turkey, the number of studies is extremely limited. As a rare exception, Cilasun (2009) analyzes labor income profiles in Turkey via a pseudo-panel data approach by constructing cohorts based on birth-years for household-heads, thereby not allowing the estimation of the role of educational attainment, gender, or the public versus private sector employment. By using a more comprehensive data set and expanding analyses on several fronts, we believe this paper sheds light on the role of the lacking dimensions on labor income profiles, thereby constituting the first comprehensive analysis of labor income profiles in Turkey.

Throughout our econometric analysis, we rely on OLS regressions to estimate life-cycle profiles of labor income in Turkey; and we complement our analysis via pseudo-panel estimation methodology. 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 the marginal effects of the listed heterogeneities in a pseudo-panel design.

3 Data and Descriptive Results

We use cross-sectional data from the Turkish Statistical Institute's (TurkStat) Household Budget Survey (HBS) covering the period 2002-2014. HBS is conducted annually by the TurkStat on a representative sample of approximately 10,000 Turkish households. The dependent variable is individual labor income that consists of cash, income received in-kind and bonus. We restrict our sample to 20-59 year-old individuals due to the limited number of observations beyond this range. The resultant total sample size is 86,666. We convert nominal labor income into real units by deflating via the Turkish consumer price index (CPI), for which we use the base year as 2014; and we exclude workers who earn below the half of minimum wage so as to focus only on full-time employees (Krueger et al., 2010). TurkStat provides education data in eleven ordinal categories, which we re-cluster into three categories: i) primary and secondary school graduates, ii) high-school graduates, and iii) university or post-university graduates.

We start by reporting our results via 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.¹⁴

⁹Information on the public versus private sector employment is only available for 2002-2011, thus it becomes our working sample when investigating labor income differences due to the public versus private sector employment.

¹⁰58% of the sample have positive income received in-kind, which constitutes 12% of the labor income, on average. 23% of the sample receive a bonus, and among them, bonus constitutes 13% of the labor income.

¹¹For further details about the sample, please see Table E.1

 $^{^{12}}$ There are 101,648 observations with positive labor income. After truncation, we have 86,666 observations, which constitute 85% of all positive labor income earners.

¹³We do the re-clustering in order to increase the efficiency of our estimates. Results by alternative education re-clustering are available upon request.

¹⁴On the box plot, the upper bar of the box represents the third quartile and the lower bar displays the first quartile. The line inside the box represents the median. 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 moments.

3.1 Average Labor Income Profile

In Figure 1, we plot the *average* labor income profile in Turkey. We document a moderate hump-shaped pattern with a peak at the 45-49 year age group for the median labor income, coupled with ever-increasing cross-sectional dispersion until the age 60.¹⁵ This hump-shaped pattern is most similar to labor income profiles in Germany and the United States among the countries that are analyzed in Lagakos et al. (2018).¹⁶

[place Figure 1 here]

3.2 Education

We next investigate the role of educational attainment, which is one of the key determinants of labor income differences across age categories. Indeed, in Figure 2 where we condition labor income profiles on education, we report a clear positive correlation between labor income and education. Among the highest earners, i.e. university graduates, we observe a sharp increase in labor income over age until late 30s, followed by a stagnant profile beyond, with the exception of an additional increase for the 55-59 year age group. High school graduates experience a similar upward trajectory over their life-cycle, with a moderate downturn around age 55-59. The upward trajectory of primary school graduates is yet in a much narrower band compared to the two other education categories, thereby exhibiting a rather stagnant life-cycle trajectory. On average, high school and university graduates earn 35% and 120% more than primary school graduates, respectively.

¹⁵The peak of the *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.

¹⁶The sharp decrease after the age of 60 stems mainly from retirement: since 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. In our estimations, we exclude the teen labor and the retired, therefore we omit both the 15-19 and above 60 year age groups.

A cross-country comparison of Turkey reveals that when conditioned on education, labor income profiles in Turkey exhibit similarities again with Germany, where university graduates have non-decreasing labor income profiles, and high school and primary school graduates have slightly hump-shaped patterns, with a peak at around age 50 (Lagakos et al., 2018). 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]

3.3 Public vs Private Sector Employment

In Figure 3 we plot labor income profiles by the public versus private sector employment. Figure 3 reveals stark differences in income profiles of the public and private sector employees: the public sector employees face monotonic upward income trajectories over the life-cycle, whereas the private sector employees are considerably more likely to encounter stagnant labor income trajectories in a sizeably narrower band. In the private sector, individuals aged 35-39 at the 25th percentile earn half of the mean labor income and at the 75th percentile earn the mean. Afterward, we can observe the decline in labor income until retirement for both quartiles. In the public sector, individuals aged 35-39 at the 25th percentile earn 90% and at the 75th percentile earn 170% of the sample mean. Labor income is increasing and before retirement individual at the 25th percentile earns the mean and at the 75th percentile earns the 190% of the sample mean. On average, the public sector employees earn 64% more than the private sector employees. In addition, the public sector employees of any age group earn more than their private sector counterparts. These contradictory patterns can be attributable mainly to the educational backgrounds of employees in the two sectors: while over 45% of the public sector employees are university graduates, only 10% of the private sector employees hold a university degree or above.¹⁷

¹⁷We discuss this issue in more detail in the next section. For further distributional statistics, see Appendix E.

[place Figure 3 here]

Even though the private sector jobs pay less on average, the dispersion in the private sector income of employees over 30 years old is larger than that of the public sector. Figure 4 displays that the variance of (logarithmic) income in the private sector first monotonically increases over age starting from 0.2, surpasses that of the public sector after age 30 and remains rather stable around 0.4. On the contrary, the variance of (logarithmic) income in the public sector remains almost stagnant at 0.2 over the life-cycle. ¹⁸ This dispersion profile for the private sector in Turkey differs from many developed countries, such as Germany, France, the United Kingdom and Canada, where variance moderately decreases or remains almost constant over the life-cycle. A notable exception among developed countries is the United States with its similar variance profile to Turkey (Lagakos et al., 2018). The similar concave pattern in Turkey's variance trajectory is also observed in several developing countries, such as Mexico, Uruguay, and Chile.

[place Figure 4 here]

3.4 Education and Public vs Private Sector Employment

In Figure 5 we depict labor income profiles by education and the public versus private sector 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 educational background earn less than the mean income, whereas university graduates earn above the mean and median income after the age 30, coupled with facing higher dispersions towards retirement. In the public sector, however, labor income profiles

¹⁸As one of the anonymous referees rightfully points out, the lower dispersion of labor income profiles in the public sector is at least to a certain extent attributable to the public sector payment structure that is pre-determined by law: promotion criteria in the public sector are usually pre-defined and mostly tenure-based. These salary gains as a result of promotions are often confined to particular rates, thereby putting a bound on the cross-sectional dispersion of the public sector labor incomes.

monotonically increase over age, with an almost constant variance and a limited education premium. Further, contrary to the ever-increasing income profile in the public sector, we observe a hump-shaped pattern in the private sector with different trajectories over education.

[place Figure 5 here][place Figure 6 here][place Figure 7 here][place Figure 8 here]

Figure 5 reveals that university-graduate employees in Turkey face a risk versus return tradeoff 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. 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 the limited number of the public sector jobs.

Figure 8 displays the histograms of labor incomes in the public and private sectors. The two histograms affirm the stark differences in incomes over the public versus private sector employment: while the distribution of income 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 half of minimum wage. The main reason behind the pattern exhibited in the private sector is, to a large extent, due to the employment of low-skilled workers earning in the close proximity of the minimum wage.

¹⁹The step-wise increases on the left-end of the private sector distribution stem from the adjustments to the *real* minimum wage over time, i.e. as we convert nominal income into real income by deflating via the consumer price index, the survey year's inflation induces step-wise departures from the lower bound: the half of minimum wage.

In order to display the first-order role of education on income dispersion, in Figure 9 we abstract from the sector of employment and present the variance of labor income over education. Figure 9 verifies that educational background indeed plays a critical role in income dispersion over the life-cycle, and the increasing income dispersion over age is generated predominantly by high school and university graduates.²⁰ In brief, our results summarize that labor income profiles by education and sector of employment differ significantly from one another, and heterogeneity in labor income profiles due to these factors are immense.

[place Figure 9 here]

3.5 Gender

In Figure 10 we display gender differences in average labor income profiles. We report that both the mean and the median labor incomes of male employees are slightly lower than those of their female counterparts until age 35, but higher afterward. Further, while income profiles of male employees exhibit an upward trajectory until age 40-44 and remain stagnant afterward, labor income profiles of female employees increase over age until early 30s, beyond which income remains stagnant. In addition, cross-sectional income dispersion of both male and female employees exhibit monotonically upward trends over the life-cycle.

[place Figure 10 here]

In order to elaborate more on gender differences in income profiles, we first focus on the role of education and display labor income profiles by gender and education in Figure 11. We report that while labor income profiles of high school and university graduates possess similar lifecycle trajectories for both genders, labor income trajectories of primary school graduates exhibit

²⁰In addition to educational background, the public versus private sector of employment is also decisive in delivering the dispersion patterns in Figure 4 and Figure 9. For further details, see Appendix B, where we condition employees with respect to their educational backgrounds and the sector of employment, and plot their variance-to-mean income ratios over the life-cycle.

striking gender differences: for males, we report a hump-shaped pattern with considerable income dispersion, whereas for females we document an age-independent income profile with low mean, median and variance levels. Male primary school graduates earn 35% more than their female counterparts. The same ratio is 29% for high school graduates and 24% for university graduates. We argue that these stark gender differences for low educated employees align well with the historically low female labor force participation in Turkey (around 30%), as Turkish women's labor market prospects are significantly worse than those of men's. 21,22

[place Figure 11 here]

While Panel (b) in Figure 11 reveals limited gender differences in labor income profiles over the public versus private sector of employment. Male workers in the private sector earn 14% more than their female counterparts and earn 10% more in the public sector, on average. Figure 12 reveals that conditioning further on education unveils a gender pay gap for the below-university graduates in the private sector: primary school graduate female employees 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 employees 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 labor income prospects in the private sector. The gender pay gap among university graduates in the private sector is yet rather limited.

For the public sector, on average we observe a moderate gender pay gap among the primary school graduates, which diminishes by further educational attainment.

²¹In particular, this observation is valid for female employees with primary education (and below) as they earn labor incomes significantly lower than that of their male counterparts. This gap narrows down over educational attainment. See Uraz et al. (2010) for further discussion.

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

[place Figure 12 here]

[place Figure 13 here]

4 Estimation Methodology and Results

We next complement our descriptive graphical analysis with econometric regressions. We first rely on OLS estimation of pooled cross-sections of labor income profiles for different subcategories. Our main estimation equation is as follows:²³

$$\log(y_{it}) = \alpha + \sum_{j=1}^{7} \beta_j \operatorname{age}_{ij} + \sum_{k=1}^{5} \gamma_k \operatorname{edu}_{ik} + \delta \operatorname{public sector}_i + \sum_{l=1}^{5} \xi_l \operatorname{public sector}_i \times \operatorname{edu}_{ik} + \theta \operatorname{gender}_i + \mu \operatorname{area}_i + \phi \operatorname{tenure}_i + \rho_t + \varepsilon_{it}$$
(1)

where $\log(y_{it})$ refers to natural logarithm of labor income of person i in year t; age refers to age categories of 5 year intervals captured by dummy variables: ages 20 to 24, 25 to 29, ..., 55 to 59; edu refers to years of schooling categories: 6 to 8 years (primary school graduates or dropouts), 9 to 12 years, 13 to 14 years, 15 to 16 years and 17 years and above; $public\ sector$ 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; 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. $ext{24}$

For robustness purposes, we next conduct a pseudo-panel estimation. Since the Turkish HBS

²³Our data is a repeated cross-section, and not a longitudinal one and we are pooling cross-sections i.e. we observe each individual i only for one point in time and y_{it} denotes the observation on individual i at time t. Age and education categories are denoted as j and k, respectively.

 $^{^{24}}$ We take ages 20 to 24 as the baseline age group. HBS measures *tenure* as is the number of years worked in the main job

data is composed of independent cross-sections, it is not possible to track the same individuals over time. Thus, we construct cohorts based on their common characteristics, such as educational attainment, the sector of employment, gender and year of birth to create a synthetic cohort panel, following the approach by Deaton (1985).²⁵

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:

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

where c denotes cohorts, \bar{y}_{ct} denotes cohort income averages and \bar{x}_{ct} denote the vector of variables generated by cohort averages, and $\bar{\theta}_c$ stands for the cohort-fixed effects.^{26,27}

We present our OLS estimation results under alternative specifications in Table 1. We report that the regression coefficients for age categories verify a hump-shaped pattern in labor income over the life-cycle, with a peak around age 40 to 44.²⁸ Further, we provide evidence on education premium over years of schooling, and higher average income levels in the public sector. In order to shed light on role of the public versus private sector employment on education premium, in column (2) of Table 1, we incorporate the interaction of education and the public versus private 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 further robust evidence of a significant gender pay

²⁵We discuss the advantages of pseudo-panel estimation over OLS in more detail in Appendix D.

²⁶Since each cohort consists of different members in each year, the cohort effect is time varying: $\bar{\theta}_{ct}$. 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 takes the form $\bar{\theta}_c$ in our regression equation. The reason behind the constancy 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 D for further details).

²⁷Instead of constructing birth-year cohorts over a single-year span, we use 5-year spans so as to enlarge cohort sizes, reduce erraticity and minimize measurement errors. We discuss more on the efficiency versus bias trade-off in Appendix D.

²⁸See Appendix C for marginal effects of age evaluated at the means of other covariates.

gap. These findings provide further support on the risk-versus-return trade-off faced by university graduates in their sectoral choice of employment.

[place Table 1 here]

In order to elaborate further on the role of gender and education, we estimate a modified version of our baseline regression by conditioning on gender and education, and we display our findings in Table 2. Table 2 reveals that marginal effects of age indicate a hump-shaped labor income trajectory over the life-cycle for all gender and education groups (albeit with different peak ages), but the female primary school graduates. 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 the above exercise by this time conditioning on the public versus private sector of employment. We document a similar robust hump-shaped pattern in labor income profiles, except for the primary school graduates in the public sector. We also verify gender differences in favor of male employees both in the public and the private sector.

[place Table 3 here]

We next turn to the pseudo-panel methodology and summarize our estimation results in Table 4. Table 4 reveals that the pseudo-panel estimation offers results consistent with those by OLS and our descriptive analysis, thereby contributing further robustness to our findings.²⁹ Table 4 confirms that the labor income profile is hump-shaped over age groups, as coefficients for age and

²⁹We conduct a fixed-effects regression without any sector specification at the cohort level, therefore 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 use labor income in its natural logarithm form in regressions.

age-squared are positive and negative, respectively.³⁰. Further, Table 4 verifies the presence of a significant education premium valid to both high school and university graduates.

[place Table 4 here]

5 Discussion and Concluding Remarks

In this paper, we explore labor income profiles over the life-cycle in Turkey. In doing so, we study the role of education, gender and the public versus private sector employment, all of which we document matter starkly and heterogeneously. In brief, our findings first reveal that the average life-cycle labor income profile in Turkey is moderately hump-shaped over age, similar as in the case for the United States and Germany. Our decomposition exercises, however, elucidate novelly that this pattern in averages is driven by the private sector employees, as the Turkish public sector employees encounter ever-increasing labor income profiles over their life-cycle. Second, we report that the public versus private sector income profiles of university graduates in Turkey display sizable differences: labor income profiles of the private sector employees exhibit both a higher average level and a higher degree of cross-sectional dispersion compared to their public sector counterparts, thereby implicating a risk versus return trade-off for their sectoral choice of employment in an economy with ever-increasing share of university graduates in the labor force. Third, we find strong evidence for a gender pay gap in Turkey, especially prevalent among primary school graduate employees, which we argue is consistent with the historically low female labor force participation rate in Turkey.

Throughout our investigation, we first use descriptive graphical analyses to shed light on the heterogeneity in labor income profiles in Turkey, and we complement our graphical analyses with OLS and pseudo-panel estimations. We further verify that our results hold true when relying on

³⁰This result is consistent with Tansel (2005).

alternative data sets (i.e. HLFS). We believe our use of several analytical approaches, as well as our use of multiple data sets offer both rigor and robustness to our findings on the several dimensions of heterogeneity in labor income profiles in Turkey.

While our analysis in this paper is confined to the study of the Turkish economy, we believe our findings offer lessons beyond. The frontier state-of-the-art research investigating labor income profiles across countries by Lagakos et al. (2018) concentrates solely on male employees in the private sector. Our findings indicate the limitations of inferring such figures as representative of countries of interest. Specifically, our results connote that in countries with sizeable public sector employment, concentrating only on the private sector income profiles to infer about economy-wide averages would be misleading, and the same misinference concern would be valid when focusing only on male employees for economies where gender differences in the labor market are sizeable, as in the case of most developing economies. We believe the missing gender and the public versus private sector employment dimensions would plausibly play a seminal role in household decisions in the labor market, as well as in their portfolio choice and risk-sharing decisions, thereby preserving decisive implications on the effectiveness of policy decisions.

While our findings shed light on several dimensions of heterogeneities in Turkish labor income profiles with a comparable methodology, 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.
- Atkinson, A. B. (2015). *Inequality*. Harvard University Press.
- Attanasio, O. and Browning, M. (1995). Consumption over the life cycle and over the business cycle. *American Economic Review*, 85(4):1118–1137.
- Bakis, O. and Polat, S. (2015). Wage inequality in turkey, 2002-2010. *Economics of Transition*, 23(1):169–212.
- Ceritoglu, E. (2009). The Empirical Importance of Precautionary Saving in Turkey. PhD thesis, University of Nottingham.
- Cilasun, M. (2009). Income and, Consumption and Saving Behavior of Turkish Household. PhD thesis, Middle East Technical University.
- 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.
- Krueger, D., Perri, F., Pistaferri, L., and Vioalente, G. (2010). Cross sectional facts for macroe-conomists. *Review of Economic Dynamics*, 13(1):1–14.
- Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life-Cycle Wage Growth Across Countries. *Journal of Political Economy*, 126(2):797–849.
- Milanovic, B. (2016). Global inequality: A new approach for the age of globalization. Harvard University Press.
- 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, Bilgi University.
- OECD (2015). Government at a Glance 2015.
- Piketty, T. (2014). Capital in the Twenty-First Century. Harvard University Press.
- Piketty, T. (2015). About capital in the twenty-first century. American Economic Review, 105(5):48–53.
- 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.

- Tamkoc, M. N. and Torul, O. (2018). Cross sectional facts for macroeconomists: Wage, income and consumption inequality in turkey. Boğaziçi University, Department of Economics Working Papers. Available at http://www.econ.boun.edu.tr/torul/csfm.pdf.
- Tansel, A. (1994). Wage employment, earnings and returns to schooling for men and women in Turkey. *Economics of Education Review*, 13(4):305–320.
- Tansel, A. (2005). Public-Private Employment Choice, Wage Differentials, and Gender in Turkey. Economic Development and Cultural Change, 53(2):453–477.
- Tansel, A., Dalgic, B., and Güven, A. (2018). Wage Inequality and Wage Mobility in Turkey. Social Indicators Research.
- Tukey, W. (1977). Exploratory Data Analysis. Reading, PA: Addison-Wesley.
- Uraz, A., Aran, M., Hüsamoğlu, M., Şanalmış, D. O., and Capar, S. (2010). Recent trends in female labor force participation in turkey.
- 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.

FIGURES

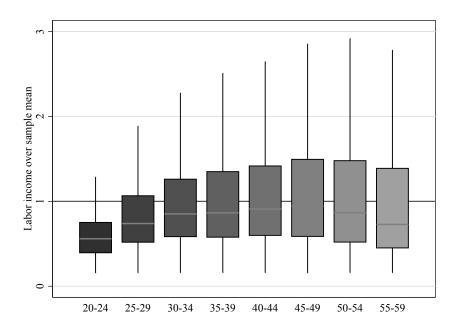


Figure 1: Average labor income profile over age groups

Notes: The sample includes individuals aged 20-59 years in the Household Budget Survey from 2002 to 2014. The sample size is 86,666. The horizontal line at 1 indicates the sample mean. Labor income of each individual is divided by the sample mean.

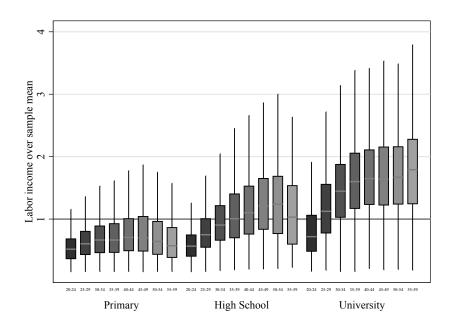


Figure 2: Labor income profile by education

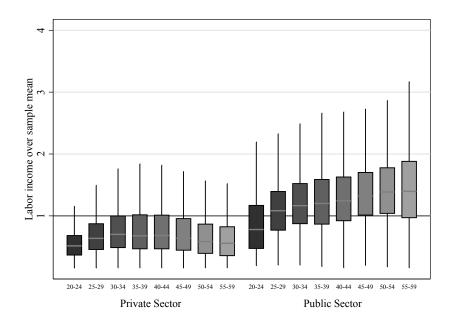


Figure 3: Labor income profile by the public vs private sector employment *Notes:* Sector variable exists from year 2002 to 2011. Therefore, the sample size is reduced to 66,100 for the figures that include sector variable.

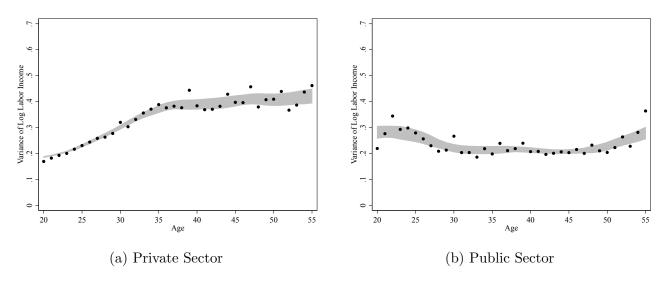


Figure 4: Variance of log labor income by sector

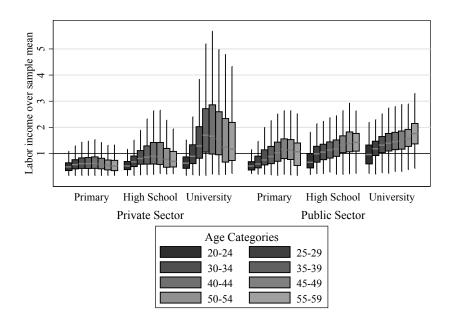


Figure 5: Labor income profile by education and the public vs private sector employment

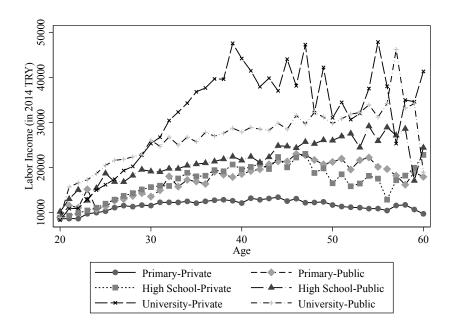


Figure 6: Labor income profile by education and the public vs private sector employment

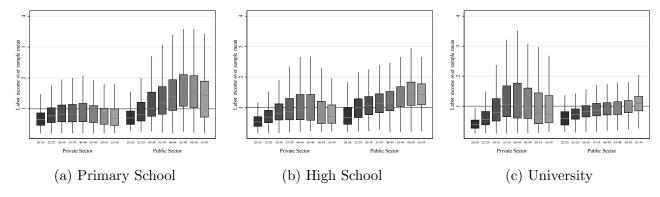


Figure 7: Labor income profile by education

Notes: In this figure, labor income of each individual is divided by the average labor income of the corresponding education group.

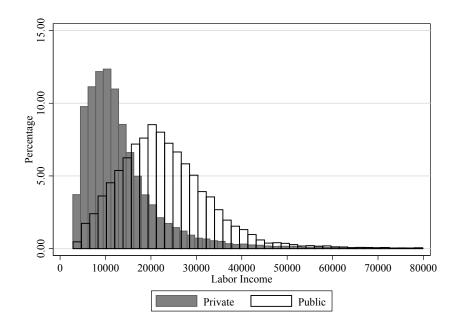


Figure 8: Labor income histograms by the public vs private sector employment

Notes: In this histogram, the individuals who earn more than 80,000 TRY (in 2014 prices) in a year are excluded. They constitute only 0.6% of the sample.

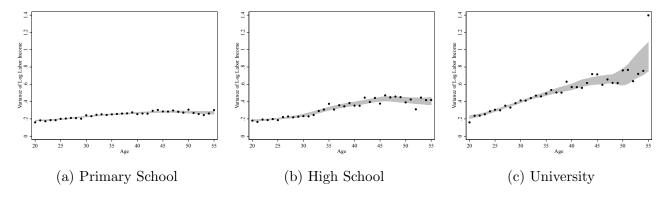


Figure 9: Variance of log labor income in the private sector by education

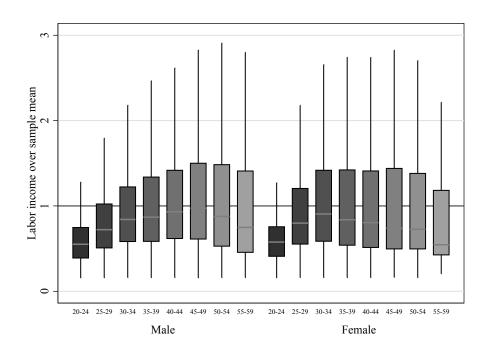


Figure 10: Labor income profile by gender

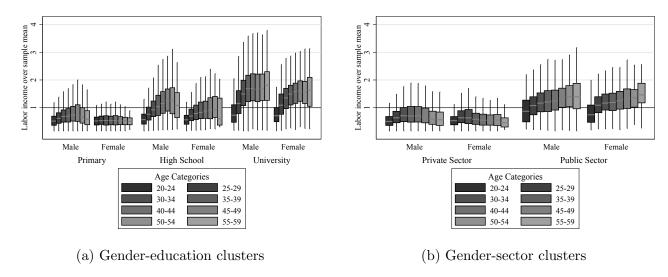


Figure 11: Labor income profile by education, gender and the public vs private sector employment

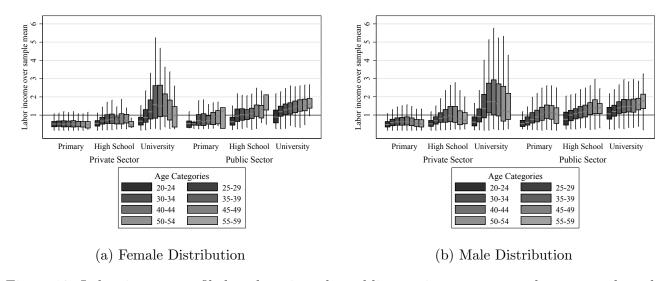


Figure 12: Labor income profile by education, the public vs private sector employment and gender *Notes:* In this figure, labor income of each female is divided by the average labor income of female employees and labor income of each male is divided by the average labor income of male employees.

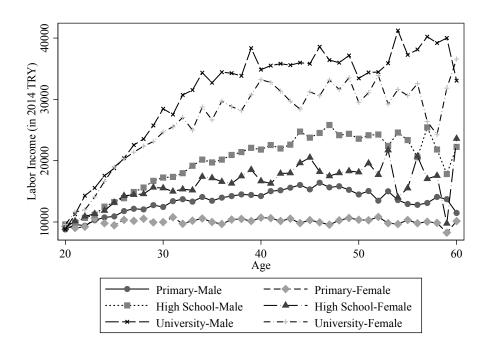


Figure 13: Labor income profile by gender

Notes: In this figure, for every age group the average labor income is plotted based on education and gender clusters.

TABLES

Table 1: OLS Estimates for Labor Income

	(1) $log(Labor Income)$	(2) $log(Labor Income)$
Age		
25 to 29	0.187***	0.183***
	(0.007)	(0.007)
30 to 34	0.291***	0.288***
00 00 01	(0.007)	(0.007)
35 to 39	0.321***	0.315***
	(0.007)	(0.007)
40 to 44	0.320***	0.308***
10 00 11	(0.008)	(0.008)
45 to 49	0.292***	0.279***
40 10 43	(0.009)	(0.009)
50 to 54	0.213***	0.195***
30 to 34	(0.011)	,
55 to 59	0.121***	$(0.011) \\ 0.107***$
55 to 59	(0.016)	(0.016)
Vooms of Education	(0.010)	(0.010)
Years of Education	0.156***	0.164***
6-8		
0.12	$(0.006) \\ 0.302***$	(0.006) $0.305***$
9-12		
10.14	(0.005)	(0.006)
13-14	0.441***	0.435***
1 × 10	(0.008)	(0.013)
15-16	0.648***	0.863***
	(0.008)	(0.013)
17+	1.044***	1.332***
	(0.023)	(0.049)
Education-Sector Interaction		
$6-8 \times \text{Public Sector}$		-0.120***
		(0.015)
$9-12 \times \text{Public Sector}$		-0.127***
		(0.012)
$13-14 \times \text{Public Sector}$		-0.131***
		(0.017)
$15-16 \times \text{Public Sector}$		-0.474***
		(0.017)
$17+ \times \text{Public Sector}$		-0.562***
		(0.054)
Sector(Public=1)	0.271***	0.436***
	(0.006)	(0.010)
Gender(Female=1)	-0.188***	-0.182***
	(0.005)	(0.005)
Area(Urban=1)	0.168***	0.170***
,	(0.005)	(0.005)
Tenure	0.012***	0.012***
	(0.000)	(0.000)
Year Dummies	Yes	Yes
N	66100	66100
R-squared	0.415	0.426
F-statistic	2387.326	2087.542

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

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

	Male			Female			
	Primary	High School	University	Primary	High School	University	
Age							
25 to 29	0.157***	0.245***	0.354***	0.010	0.212***	0.326***	
	(0.010)	(0.013)	(0.031)	(0.021)	(0.020)	(0.026)	
30 to 34	0.223***	0.371***	0.625***	-0.027	0.209***	0.482***	
	(0.010)	(0.014)	(0.032)	(0.022)	(0.023)	(0.029)	
35 to 39	0.234***	0.400***	0.717^{***}	-0.043*	0.189***	0.530***	
	(0.010)	(0.016)	(0.033)	(0.020)	(0.030)	(0.033)	
40 to 44	0.246***	0.386***	0.681***	-0.056**	0.131***	0.511***	
	(0.011)	(0.018)	(0.035)	(0.022)	(0.033)	(0.041)	
45 to 49	0.204***	0.352***	0.665***	-0.096***	0.105^{*}	0.533***	
	(0.012)	(0.020)	(0.038)	(0.024)	(0.046)	(0.047)	
50 to 54	0.113***	0.262***	0.606***	-0.109***	0.070	0.457***	
	(0.014)	(0.026)	(0.042)	(0.033)	(0.065)	(0.058)	
55 to 59	0.015	0.203***	0.635***	-0.190***	-0.291**	0.384**	
	(0.019)	(0.044)	(0.055)	(0.051)	(0.095)	(0.126)	
Sector(Public=1)	0.472***	0.226***	0.012	0.320***	0.170***	0.096***	
,	(0.009)	(0.011)	(0.016)	(0.032)	(0.022)	(0.020)	
Area(Urban=1)	0.211***	0.134***	0.115***	0.144***	0.099***	0.056*	
,	(0.006)	(0.011)	(0.016)	(0.019)	(0.024)	(0.024)	
Tenure	0.007***	0.022***	0.010***	0.014***	0.030***	0.010***	
	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	
N	31029	14480	8621	4118	3415	4437	
R-squared	0.235	0.328	0.208	0.185	0.319	0.255	
F-statistic	554.249	474.720	121.473	44.903	104.048	89.764	

Note: Numbers in parantheses are standard errors. * for p < .05, ** for p < .01, and *** for p < .001. 20-24 age category 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	0.127***	0.207***	0.304***	0.088	0.220***	0.245***	
	(0.009)	(0.011)	(0.025)	(0.074)	(0.038)	(0.031)	
30 to 34	0.187***	0.320***	0.590***	0.224**	0.246***	0.371***	
	(0.010)	(0.013)	(0.030)	(0.070)	(0.036)	(0.031)	
35 to 39	0.197***	0.346***	0.775***	0.232***	0.250***	0.417***	
	(0.010)	(0.017)	(0.036)	(0.069)	(0.037)	(0.032)	
40 to 44	0.200***	0.362***	0.700***	0.219**	0.213***	0.437***	
	(0.010)	(0.021)	(0.047)	(0.069)	(0.039)	(0.034)	
45 to 49	0.139***	0.309***	0.672***	0.186**	0.211***	0.460***	
	(0.012)	(0.027)	(0.056)	(0.071)	(0.042)	(0.037)	
50 to 54	0.052***	0.183***	0.437***	0.115	0.170***	0.466***	
	(0.014)	(0.036)	(0.057)	(0.072)	(0.047)	(0.040)	
55 to 59	-0.008	0.123^{*}	0.486***	-0.095	0.071	0.497***	
	(0.019)	(0.055)	(0.091)	(0.079)	(0.074)	(0.051)	
Gender(Female=1)	-0.242***	-0.105***	-0.111* [*] *	-0.282***	-0.190***	-0.133***	
, , , , , , , , , , , , , , , , , , ,	(0.007)	(0.011)	(0.020)	(0.029)	(0.015)	(0.009)	
Area(Urban=1)	0.187***	0.165***	0.223***	0.216***	0.088***	0.065***	
,	(0.007)	(0.014)	(0.039)	(0.015)	(0.013)	(0.011)	
Tenure	0.004***	0.028***	0.034***	0.030***	0.019***	0.004***	
	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	
N	30583	12215	4821	4564	5680	8237	
R-squared	0.167	0.250	0.262	0.311	0.278	0.292	
F-statistic	333.734	208.937	97.859	111.479	114.932	185.274	

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

Table 4: Pseudo-Panel Estimates

	$\log({ m Labor\ Income})$					
	(1)	(2)	(3)	(4)		
Age	0.122***	0.133***	0.116***	0.127***		
	(0.013)	(0.013)	(0.011)	(0.010)		
Age^2	-0.001***	-0.001***	-0.001***	-0.001***		
	(0.000)	(0.000)	(0.000)	(0.000)		
Female \times Age		-0.013*		-0.013***		
		(0.005)		(0.003)		
$High School \times Age$			0.003	0.002		
			(0.005)	(0.004)		
University × Age			0.028***	0.028***		
			(0.004)	(0.004)		
Cohort Dummies	Yes	Yes	Yes	Yes		
N	512	512	512	512		
R-squared	0.779	0.791	0.830	0.843		
F-statistic	148.179	142.236	235.971	209.743		

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

APPENDIX-A: 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 aging 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.³¹ Finally, cohort effects result from the unique experience of each cohort as time goes by. Age-Period-Cohort (APC) analysis allows us to disentangle the 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 both by varying age effects and by period effects, they do not provide an accurate quantitative evaluation of the sources of change. Each graph describes only the variation in the labor income that can be 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 also account for this difference. Thus, there is a need for statistical regression modeling to capture how these three effects work simultaneously.

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 *i*th age group for i=20,...,60 at the *j*th year for j=2002,...,2014. μ stands for the intercept or adjusted mean labor income, α_i is the coefficient for the *i*th age group, β_j is the coefficient for the *j*th year, γ_k is the coefficient for

³¹Examples of social, economic and environmental factors can be wars, natural disasters, and crises.

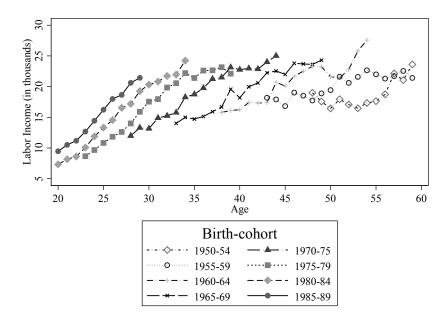


Figure A1: Labor income over life-cycle by birth-cohort

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:

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

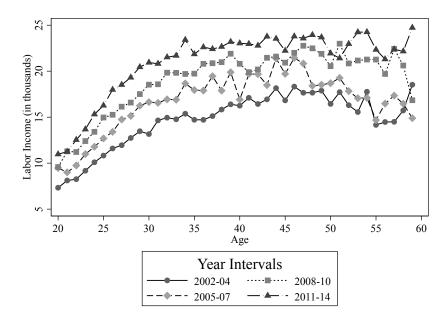


Figure A2: Labor income over life-cycle by time periods

It is important to note that α_{60} , β_{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 solutions³² to the identification problem, but in this paper, we will consider the most recent technique, the intrinsic estimator (Yang and Kenneth, 2013).

The parameter space of the unconstrained model (5) can be decomposed into two orthogonal

 $^{^{32}}$ Reduced two-factor models, constraint generalized linear models, non-linear transformation, and proxy variables are some of the solutions.

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 the 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.³³

For robustness purposes, we use the conventional approach to APC models, i.e. the coefficient 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^TX) 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.³⁴

The age effects are consistent with our findings: the labor income is increasing rapidly until

³³We use the *apc_ie* command in STATA to estimate age, period and cohort effects.

³⁴However, changing the constraint can produce widely different estimates for the effects.

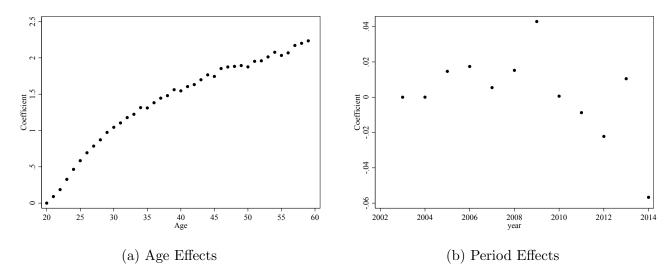


Figure A3: Effect Coefficients

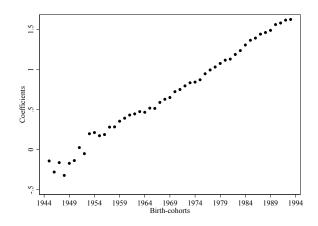


Figure A4: Cohort Effects

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 A3. 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 A4, 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 A1, 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 A2 because the real labor income increases over the years due to positive GDP growth.

APPENDIX-B: Variance-to-Mean Income Ratio over Age

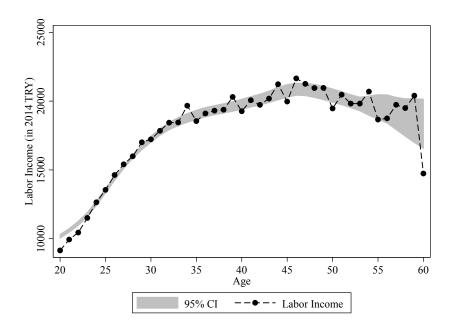


Figure B1: Full sample

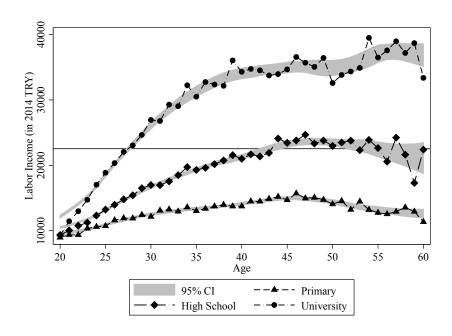


Figure B2: Labor income by education

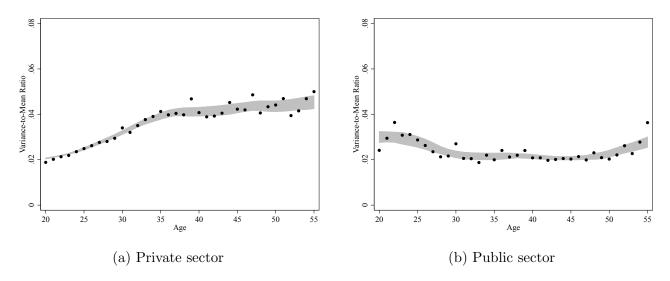


Figure B3: Variance-to-Mean Income Ratio

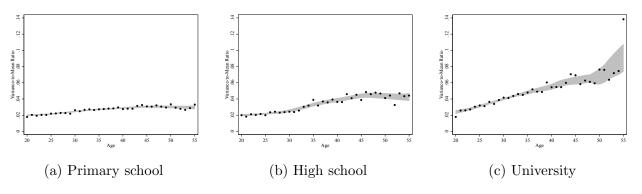


Figure B4: Variance-to-Mean Income Ratio (Private sector)

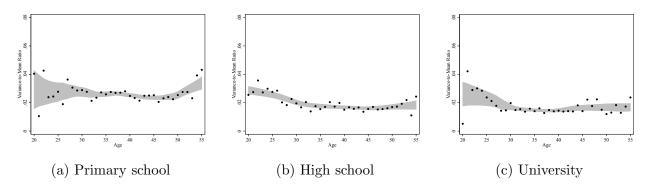


Figure B5: Variance-to-Mean Income Ratio (Public sector)

APPENDIX-C: Graphs on Marginal Age Effect

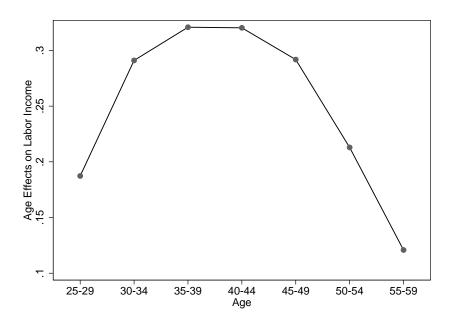


Figure C1: Marginal age effect

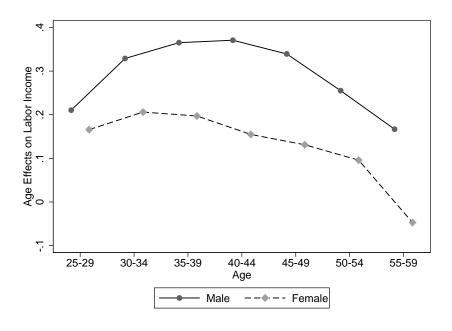


Figure C2: Marginal age effect by gender

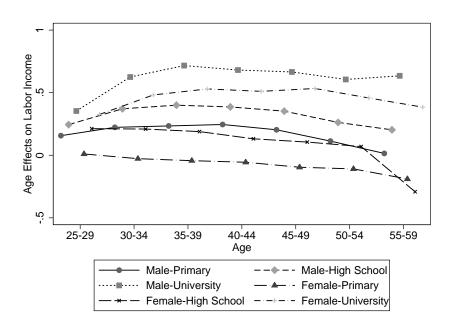


Figure C3: Marginal age effect by education and gender

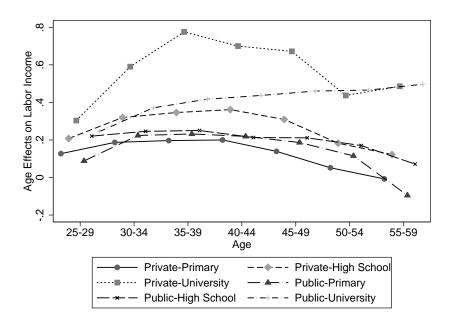


Figure C4: Marginal age effect by education and the public versus private sector employment

APPENDIX-D: PSEUDO-PANEL METHOD

The pseudo-panel method has several advantages over standard genuine panel data estimations. In the standard genuine panel data analysis, the main concern is the 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 also critical: increasing cohort size decreases measurement error and bias, but it also decreases the 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 the 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 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 associated estimation errors, following Gardes et al. (2005), we weight each observation by a heteroscedasticity factor that is a function of cell size. Arguably, there are downsides of the pseudo-panel approach as well, such as the loss of individual information due to aggregation, but for our purposes and data in hand, this technique is one of the best possible empirical approaches, which is also 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), which is why the 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 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 512 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 512 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.

APPENDIX-E: DESCRIPTIVE DISTRIBUTIONAL STATISTICS

The sample includes 20-59 year-old individuals due to the limited number of observations beyond this range. The sample size is 86,666. The data covers the period 2002-2014. We convert nominal labor income into real units by deflating via the Turkish consumer price index (CPI), for which we use the base year as 2014; and workers who earn below the half of minimum wage are excluded.

There is no sector variable from 2012 to 2014 in the data, so there are 20,566 missing values for the sector. Therefore, for some analyses the sample size is restricted to 66,100 observations.

Table E1: Descriptive Distributional Statistics

					Full Sa	mple (2002-20	14)					
	Primary School		High School		University			All Categories				
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	38904	5899	44803	18366	4491	22857	12254	6752	19006	69524	17142	86666
Frequency	44.89%	6.81%	51.70%	21.19%	5.18%	26.37%	14.13%	7.80%	21.93%	80.22%	19.78%	100.00%
					Private	Sector (2002-2	011)					
		Primary School High School			University			All Categories				
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	26690	3893	30583	9749	2466	12215	3243	1578	4821	39682	7937	47619
Frequency	56.05%	8.17%	64.22%	20.47%	5.18%	25.65%	6.81%	3.31%	10.12%	83.33%	16.67%	100.00%
					Public S	Sector (2002-20	011)					
		Primary Schoo	ol	High School		University		All Categories				
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	4339	225	4564	4731	949	5680	5378	2859	8237	14448	4033	18481
Frequency	22.47%	1.21%	24.70%	25.60%	5.13%	30.73%	29.10%	15.47%	44.57%	78.18%	21.82%	100.00%

Table E2: Descriptive Distributional Statistics

	Ger	nder	Education			
	Male	Female	Primary	High School	University	All
Private	83.33%	16.67%	64.22%	25.65%	10.12%	72.04%
Public	78.18%	21.82%	23.84%	30.87%	45.26%	27.96%
Full-Sample	81.89%	18.11%	53.17%	27.07%	19.75%	100.00%