HETEROGENEITY IN LABOR INCOME PROFILES: EVIDENCE FROM TURKEY*

EMREHAN AKTUĞ†
University of Texas at Austin

TOLGA UMUT KUZUBAŞ‡
Boğaziçi University

ORHAN TORUL§
Boğaziçi University

February 25, 2020

Abstract

In this paper, we investigate labor income profiles in Turkey. In doing so, we study the role of educational attainment, gender, and the public versus private sector employment on labor income profiles by using the Turkish Statistical Institute's Household Labour Force Survey micro data from 2004 to 2018. We first report that while the 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 the 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 the sector of employment: the private sector labor income profiles display a similar level of average income but 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 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 are grateful to İnsan Tunalı, Mehmet Ali Soytaş, Kjetil Storesletten, Murat Güray Kırdar, Malik Çürük, and other faculty at the Department of Economics, Boğaziçi University for their helpful comments and suggestions. We thank the participants at the 6th All-Istanbul Workshop, and the Center for Economics and Econometrics Annual 2018 Conference for their valuable input. We also thank the editor-in-charge Arthur H. O. van Soest and two anonymous reviewers for their insightful 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 in *labor incomes* and Turkey is not an exception as it has the fifth-highest level of income inequality among OECD countries in 2018.[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 *labor 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 Labour Force Survey covering 2004 to 2018, 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 50, similar to 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-sectional dispersion of labor income in Turkey is also increasing over age, in accordance with the findings on developing countries, yet contrary to developed economies. (Lagakos et al., 2018). We document that this pattern is also driven by the private sector compensations, as the variance-to-mean income ratio over age is ever-increasing only for university graduates in the private sector but stagnant for all other groups including 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

[1] See 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 Tamkoç and Torul (2020) for a review and comparison of the economic inequalities in Turkey.

[2] See OECD Income Inequality Database for further details.

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

public sector counterparts, particularly after age 35, thereby implicating a *risk versus return trade-off* for their choice of the sector of employment.[4] 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% recently. Indeed, our results confirm that the female labor force participation rate increases with education, which is as high as 70% among female university graduates. However, the fact that 35% 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.[5]

Throughout our investigation, we first use descriptive graphical analyses to shed light on the heterogeneity in the distribution of labor income profiles in Turkey. We complement our graphical analyses with ordinary least squares (OLS) and pseudo-panel estimations. We further verify that our results hold when using alternative data sets.[6] We believe our use of several analytical approaches 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 heterogeneities we report in this paper are crucial in understanding the Turkish labor market and plausibly the labor markets of similar economies. For instance, concentrating on the labor income profiles of only male private sector employees, as done in Lagakos et al. (2018), would lead to misleading predictions about the income profiles of Turkey's public sector employees, which increase monotonically over age, contrary to the hump-shape seen in the private sector. Similarly, to propose sound policy recommendations, say to improve Turkey's female labor force participation, neglecting gender differences in the distribution of labor income profiles, especially among the low-educated employees, could lead to miscal-

[4] This 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.

[5] For a detailed discussion on this issue, see Uraz et al. (2010) and Tunalı et al. (2017).

[6] In particular, we confirm that our results hold true via Turkish Statistical Institute's Household Budget Survey (HBS). Our results via HBS are available as an online appendix.

culated predictions. We aim that these robust heterogeneities we document in this paper contributes to a better comprehension and policy analysis of the labor market of a large developing economy, and plausibly beyond.

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 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 the life cycle when controlling for educational attainment. However, the authors do not concentrate on the *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 Labour Force Survey (HLFS), 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. Using HLFS, Ekşi and Kırdar (2015) investigate wage inequality in Turkey by addressing the role of educational attainment for the 2002-2011 period and Bakış and Polat (2015) study the same topic by addressing industries as well. Nazlı (2014), Yükseler and Türkan (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) study income inequality by presenting 90/10 and 90/50 wage ratios by gender, age, education, and the sector of employment by using the Surveys on Income and Living Conditions (SILC).

[7] See Attanasio and Browning (1995) and Alessie et al. (1997), among others.
[8] See Lagakos et al. (2018) for a recent discussion on the developments in developing countries.

Tamkoç and Torul (2020) study the evolution of wage, income and consumption inequality in Turkey after 2002 via a cross-country comparable methodology.

Regarding the labor income profile over the 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. In a recent attempt, relying on the synthetic cohort methodology, Tunalı et al. (2017) scrutinize female labor force participation and labor income profiles of women in Turkey and examine urban/rural differences in detail. 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 synthetic 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 under a pseudo-panel setting.

3 Data and Descriptive Results

We use cross-sectional data from the Turkish Statistical Institute's (TurkStat) Household Labour Force Survey (HLFS) covering the period 2004-2018.[9] HLFS is conducted annually by the TurkStat on a representative sample of approximately 140,000 Turkish households. The dependent variable we use throughout our analyses is individual labor income that consists of cash, income received in-kind and bonus.[10] We restrict our sample to individuals between 20 and 59 years of age due to the limited number of observations beyond this range. The resultant total sample size is 1,105,000.[11] 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 2018; and we exclude workers who work less than 35 hours a week to focus only on full-time employees. TurkStat pro-

[9] Information on the public versus private sector employment is only available for 2009-2018, thus it becomes our working sample when investigating labor income differences due to the public versus private sector employment.

[10] 58% 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.

[11] For further details about the sample, please see Table B.1

vides education data in six 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.[12]

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 compactly.[13]

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 50-54 year age group for the median labor income, coupled with ever-increasing cross-sectional dispersion until the age 60.[14] 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).[15]

[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 the late 40s, followed by a stagnant profile beyond. High school graduates experience a similar upward trajectory over their life-cycle, but 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 exhibit-

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

[13] On 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.

[14] The 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.

[15] The 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.

ing a rather stagnant life-cycle trajectory. On average, high school and university graduates earn 38% and 128% more than primary school graduates over the life-cycle, respectively.

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 versus 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 levels of the public and private sector employees: on average, the public sector employees earn 80% more than the private sector employees over the life-cycle.[16] In the private sector, individuals aged 40-44 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 40-44 at the 25th percentile earn 125% and at the 75th percentile earn 190% of the sample mean. Then, the wage profile stays stagnant until the retirement, in a sizeably narrower band. Overall, 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 60% of the public sector employees are university graduates, only 17% of the private sector employees hold a university degree or above.[17]

[place Figure 3 here]

Even though the private sector jobs pay less on average, the dispersion in the private sector income of employees over 35 years old is larger than that of the public sector. Figure 4 displays that the variance

[16] The public sector employment share among full-time employees is 24% and it stays unchanged over the period 2009-2018. For further details, please see online appendix.

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

of (logarithmic) income in the private sector first monotonically increases over age starting from 0.2, surpasses that of the public sector after age 35 and remains rather stable around 0.25. 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 a 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 versus 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 backgrounds earn less than the mean income, whereas university graduates earn above the mean and median income after the age of 30, coupled with facing higher dispersions towards retirement. In the public sector, however, labor income profiles monotonically increase over age, with 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]

[18] As 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.

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 a similar level of average income but a higher degree of cross-sectional variation compared to their public sector counterparts. As Figure 6 verifies, the average income of university graduates exceeds the average income of their public sector counterparts at age 40, and stays higher until age 56. 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.

On average, the primary school graduates in the public sector earn 40% more than their private sector counterparts over the life-cycle. This difference is 50% among the high school graduates. Among university graduates, the sector choice creates no difference in terms of the average labor income over the life-cycle. However, as it can be visually seen in Figure 7, the income variation of the university graduates in private sector is higher compared to their public sector counterparts, particularly after age 30-34.

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 minimum wage.[19] 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.

In order to display the first-order role of education on income dispersion, in Figure 9 we present the variance of labor income over education in the private sector. 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 university graduates. However, in the public sector, educational background does not play a decisive role. The variance of labor income profile is stagnant over the life-cycle and the level is nearly the same across education groups (Figure 10).

[place Figure 9 here]

[place Figure 10 here]

[19] The 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 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 is immense.

3.5 Gender

In Figure 11 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 50 and remain stagnant afterward, labor income profiles of female employees increase over age until the 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 11 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 12. We report that while labor income profiles of high school and university graduates possess similar life-cycle trajectories for both genders, labor income trajectories of primary school graduates exhibit 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 27% more than their female counterparts. The same ratio is 12% for high school graduates and 11% 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% recently), as Turkish women's labor market prospects are significantly worse than those of men's.[20,21]

[place Figure 12 here]

While Panel (b) in Figure 12 reveals limited gender differences in labor income profiles over the public versus private sector of employment. Male workers in the private sector earn 9% more than their female

[20] In 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.

[21] In 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, 33% 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.

counterparts and earn 2% more in the public sector, on average and over the period 2009-2018. These values are 13% and 1% in Household Budget Survey (HBS), which has the sector variable only for the 2002-2011 period.

Figure 13 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. On average, male primary school graduates earn 23% more than female primary school graduates. Among high school graduates, this value is 10%. The gender pay gap among university graduates in the private sector is significant: male employees earn 22% more. In the public sector, the gender pay gap persists but diminishes by further educational attainment. The values are 26%, 13% and 11%, respectively.

[place Figure 13 here]

[place Figure 14 here]

4 Estimation Methodology and Results

4.1 Ordinary Least Square (OLS) Estimation

We 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:[22]

$$\begin{aligned}\log(y_{ijkp}) = &\alpha + \sum_{j}^{7} \beta_j \text{ age}_{ij} + \sum_{k}^{5} \gamma_k \text{ edu}_{ik} + \delta \text{ public sector}_i + \sum_{l}^{5} \xi_l \text{ public sector}_i \times \text{edu}_{ik} \\ &+ \theta \text{ gender}_i + \phi \text{ tenure}_i + \rho_p + \varepsilon_{ijkp} \end{aligned} \tag{1}$$

[22] Our 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_{ijkp} denotes the observation on individual i at period p. Age and education categories are denoted as j and k, respectively.

where $\log(y_{ijkp})$ refers to the natural logarithm of labor income of person i in period (year) p; `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 educational attainment categories: primary school graduates or dropouts, middle school graduates[23], high school graduates, and university graduates; `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; `tenure` stands for years of job experience of individual i, and ρ captures the year-fixed effect.[24]

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. 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 public sector surpasses that of the private sector for each education category. When we incorporate middle-aged people into this interaction, as in column (3) of Table 1, we observe that the education premium in the private sector surpasses that of the private sector for university graduates. In addition, as discussed in the descriptive analysis, Table 1 provides further robust evidence of a significant gender pay 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]

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 the 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 men, the peak is at age 40-44 for each education group. For female primary school graduates, the income profile is decreasing over the life-cycle. For high school and university graduates, the income profile is hump-shaped, but the peaks are at age 35-39 and 40-44, respectively.

[place Table 2 here]

[23] This category is only used in regressions. For the graphs, we only keep 3 educational categories.

[24] We take ages 20 to 24 as the baseline age group. HLFS measures `tenure` as the number of years worked in the main job.

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 and high school graduates in the public sector. In the private sector, the peaks are at 35-39, 40-44 and 45-49, for primary, high school, and university graduates, respectively. We also verify gender differences in favor of male employees both in the public and the private sector.

[place Table 3 here]

4.2 Pseudo-Panel Estimation

HLFS data is composed of different cross-sections of observations thus different generations are observed at each date and these generations might have different characteristics or face different labor market conditions which affect their observed income profiles, i.e. part of the profile can be explained by the fact that different generations at the same age observed in the cross-sections. To account for these generation/cohort effects, we follow the pseudo-panel approach and construct cohorts based on the heterogeneities observed in the distributional/OLS analysis.[25] In our estimations, using cohort-fixed effects, we aim to isolate the generation effects from the estimated wage profiles.[26]

The income profiles obtained from OLS estimation using pooled cross-sections of observations does not allow to disentangle observed differences are due to age effects or stem from the differences across separate cohorts. In order to reconcile our findings by taking into account possible cohort effects, we con-

[25] The synthetic cohort methods are well established in the empirical labor economics literature and are commonly in use in recent studies for countries lacking panel data. For instance, in a recent attempt, Tunalı et al. (2017) scrutinize female labor force participation in Turkey using a synthetic-cohort analysis. While synthetic cohort methods have some more desirable properties over genuine panel data analysis, such as lessening measurement error bias due to aggregation of individuals into cohorts, they also come at the cost of a *bias* versus *efficiency* trade-off: increasing cohort size decreases measurement error and bias, but it also decreases the number of cohorts and efficiency. We discuss the advantages and caveats of the pseudo-panel estimation methodology in more detail in Appendix A.

[26] In our analysis, we first construct the wage profiles for different cross-sections i.e. for different periods in the data and observe that the resulting wage profiles exhibit similar trajectories for the entire sample and for the sub-samples based on education, gender and sector of employment (Figures B1, B2, and B3 in the online appendix). In our pseudo-panel estimations, for different constructions of cohorts i.e. based on birth-year and with further clusters for gender, education, and sector of employment, we implicitly assume that the shape of the age-wage profiles obtained for each cohort construction – the coefficients for age and age^2- is identical and focus on the different trajectories for each subgroup. More precisely, we compare the average profiles for each category for the sample period controlling for the presence of different generations in our cross-sections, to make our analysis comparable to the OLS estimations. We plot the estimated cohort effects for different definitions of cohorts (Figures C1-C6 in the online appendix) and observe increasing cohort effects for younger generations, however, we note that the estimated cohort effects are a combination of age, period and generation effects and the identification strategy used by Deaton (1985) decomposition and alternative methods such as intrinsic estimator or maximum entropy method might lead to conflicting results. Therefore, in this paper, we compare the average profiles of different groups i.e. extensive margin rather than aiming to disentangle the evolution of the wage profiles for each group over time and leaving a full-fledged APC analysis for further research.

struct synthetic cohort panel following Deaton (1985).

First, we construct cohorts by birth year 1 and 5-year span starting from 1950-1954 to 1985-1989 and then follow these cohorts in our sample period, 2004-2018. We further refine our construction of cohorts based on education, gender and the sector of employment to capture the heterogeneities revealed in the OLS analysis. The econometric model is specified as a static linear model with cohort-fixed effects and given as:

$$\overline{y}_{ct} = \alpha + \beta_1 \overline{Age}_{ct} + \beta_2 \overline{Age}^2_{ct} + \overline{\theta}_c + \overline{\epsilon}_{ct} \tag{2}$$

where c denotes cohorts, $\overline{y}_{ct}$ denotes cohort income averages and $\overline{Age}_{ct}$ denotes the cohort age averages, and $\overline{\theta}_c$ stands for the cohort-fixed effects.[27,28]

We first estimate equation (2) based on cohorts constructed solely on birth-year intervals and present estimation results in columns (2) and (3) of Table 4 while providing the pooled OLS estimation results in column (1) for comparative purposes. The average profile controlling for the cohort effects is steeper than the one based on the OLS estimation with similar curvatures qualitatively confirming the hump-shaped profile obtained in our distributional analysis/OLS estimation.

[place Table 4 here]

Next, we move to the construction of refined cohorts based on gender, education, sector of employment. In column (4), we define cohorts based on birth-year, and education thereby controlling for differential impacts of education across cohorts. The resulting wage profile is less steep than the average profile obtained in columns (2) and (3). In column (5), we repeat the same analysis based on gender. Controlling for the gender and cohort effects, we obtain an almost identical profile to the average one suggesting that cohort effects exhibit similar patterns for males and females. Column (6) presents the estimation controlling for the sector of employment and reveal less steep and less concave profile.[29] The columns (7), (8) and

[27] Since 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 A for further details).

[28] We use 5-year spans to enlarge cohort sizes, reduce erraticity and minimize measurement errors. However, since we a very large number of observations with HLFS, we also construct birth-year cohorts over a single-year span and the results are quite robust. We discuss more on the efficiency versus bias trade-off in Appendix A.

[29] Estimated cohort effects for cohorts based on birth year, birth year and education, birth year and gender; birth year and sector of employment, are provided in Figures C1-C6 of the online appendix.

(9) repeat the same exercise with 5-year span for the birth year and yield similar results.

In Table 5, we compare the average effect of the sector of employment. To this end, we construct cohorts based on birth-year, gender, education, and the sector of employment and estimate equation (2) for birth-years clusters with 1 and 5 year spans. Column (1) and (4) provides the average profiles for 1 and 5 year intervals for the birth year. Next, we divide our sample over the private and public sector employment and include cohorts employed in the public and private sectors separately in our estimations. Controlling for generation, gender, and education differences, we observe that wage trajectories are steeper with a higher curvature for the employees in the public sector, which qualitatively accords well with the results by OLS estimation.

[place Table 5 here]

Table 6 presents an analysis in the same spirit, but focuses particularly on the effect of education on the average wage profile. We control for the generation and gender effects and start again with the benchmark estimation isolating the effect of education and provide the results in column (1) and (5) for different definitions of generation. We then proceed with the estimations for the subsamples and show that as education increases the average wage profile becomes steeper with a higher concavity reconciling our findings for the education premium. Furthermore, our distributional analysis and OLS estimations uncover stark differences in the wage trajectories of male and female employees.

[place Table 6 here]

Controlling for the generation effects, in Table 7, we present pseudo-panel estimations for cohorts constructed by birth-year, gender and education. Our estimations for the subsamples reveal average wage profiles confirming the wage gap between male and female employees for each education level, especially for females with primary education.

[place Table 7 here]

Figure A1 and A2 provide a visualization of the average predicted profiles based on the heterogeneities regarding gender, education and sector of employment derived from the OLS and pseudo-panel estimations. We observe that wage profiles are qualitatively similar for both methods implicating that controlling

for the presence of different generations in our cross-sections, despite small quantitative differences, yield to similar wage trajectories.

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 a nearly same average level but 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 when relying on alternative data sets. 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 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 sizable, as in many 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.

Bakış, 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.

Ekşi, O. and Kırdar, 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. (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.

Nazlı, 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.

Tamkoç, M. N. and Torul, O. (2020). Cross-sectional facts for macroeconomists: Wage, income and consumption inequality in turkey. *The Journal of Economic Inequality*, Forthcoming.

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., Dalgıç, 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.

Tunalı, İ., Kırdar, M. G., and Dayıoğlu, M. (2017). Female labor force participation in turkey: A synthetic cohort analysis, 1988–2013. *TÜBITAK-Project No. 112K517*.

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.

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

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

FIGURES

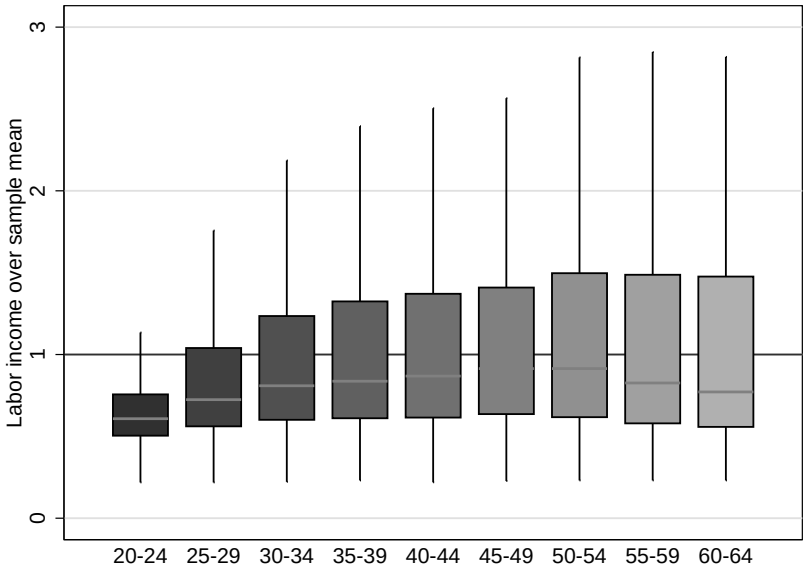


Figure 1: Average labor income profile over age groups

Notes: The sample includes individuals aged 20-59 years in the Household Labour Force Survey from 2004 to 2018. The sample size is 1,105,600. The horizontal line at 1 indicates the sample mean. Labor income of each individual is divided by the sample mean.

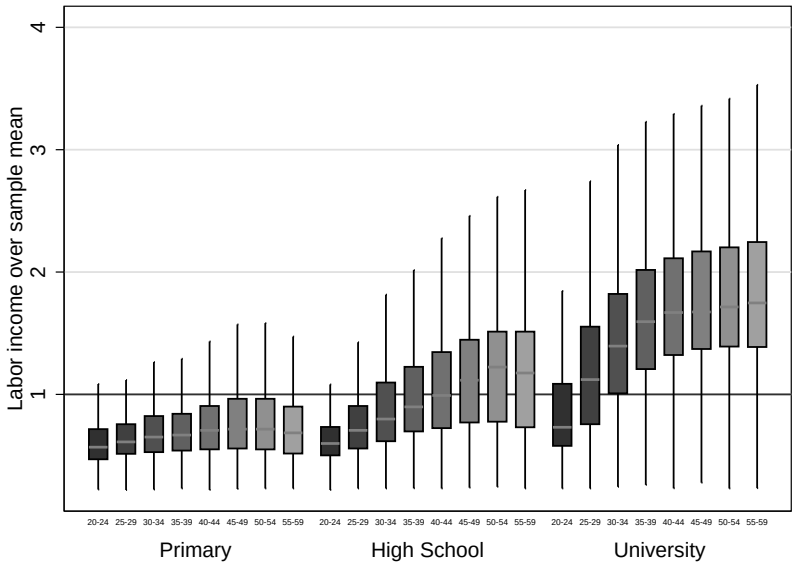


Figure 2: Labor income profile by education

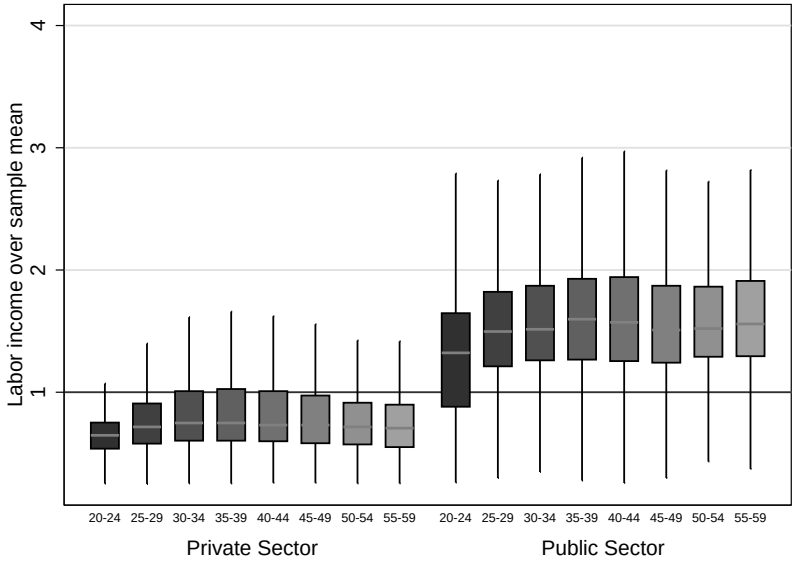
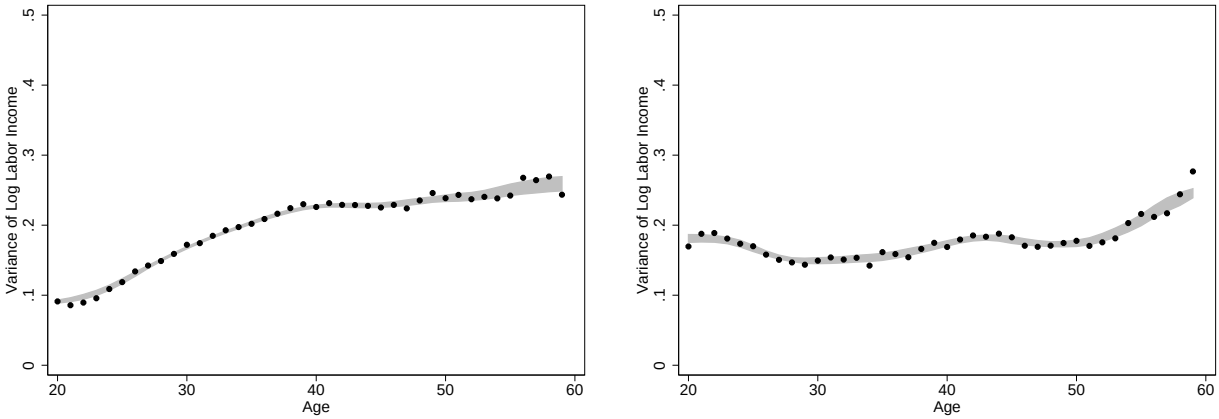


Figure 3: Labor income profile by the public vs private sector employment

Notes: Sector variable exists from year 2009 to 2018. Therefore, the sample size is reduced to 788,000 for the figures that include sector variable.



(a) Private Sector **(b)** Public Sector

Figure 4: Variance of log labor income by sector

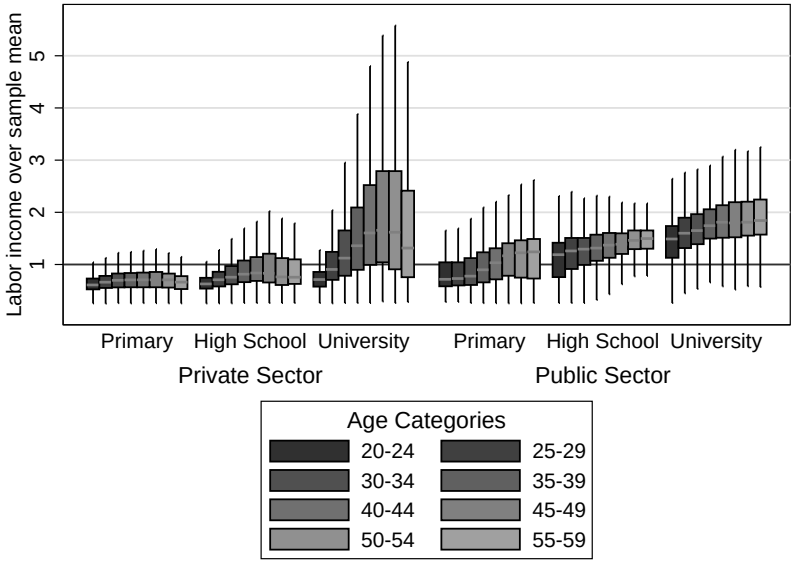


Figure 5: Labor income profile by education and the public vs private sector employment

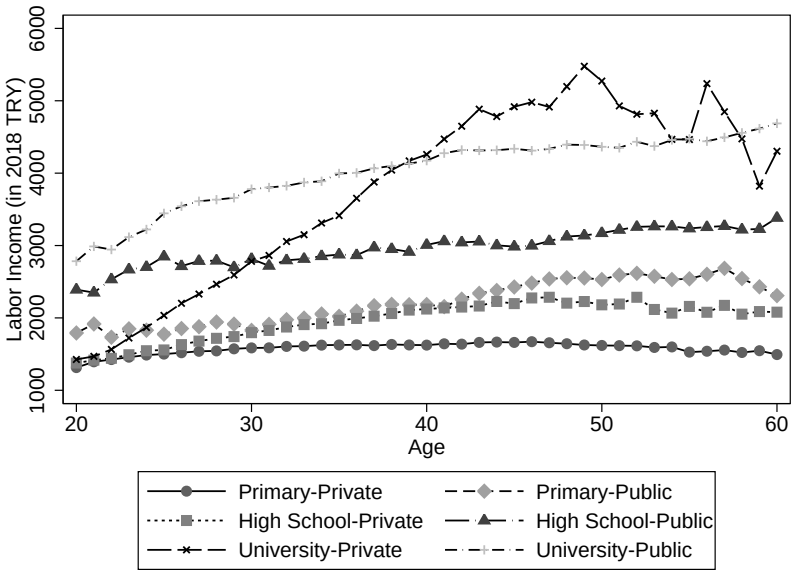
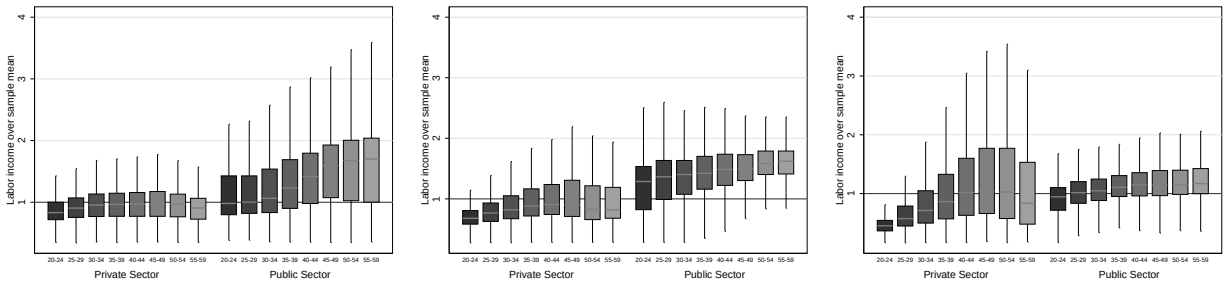


Figure 6: Labor income profile by education and the public vs private sector employment



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

Figure 7: Labor income profile by education and the public vs private sector employment

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

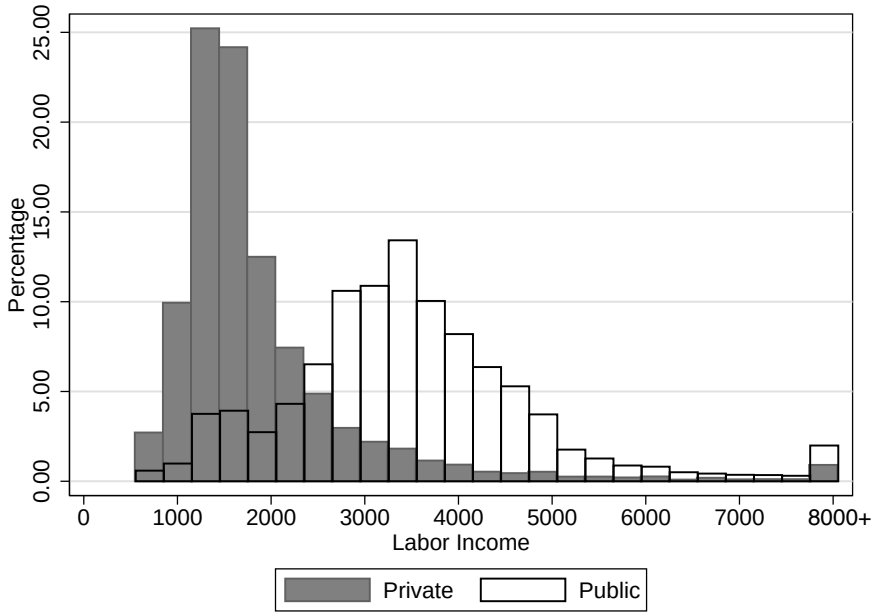
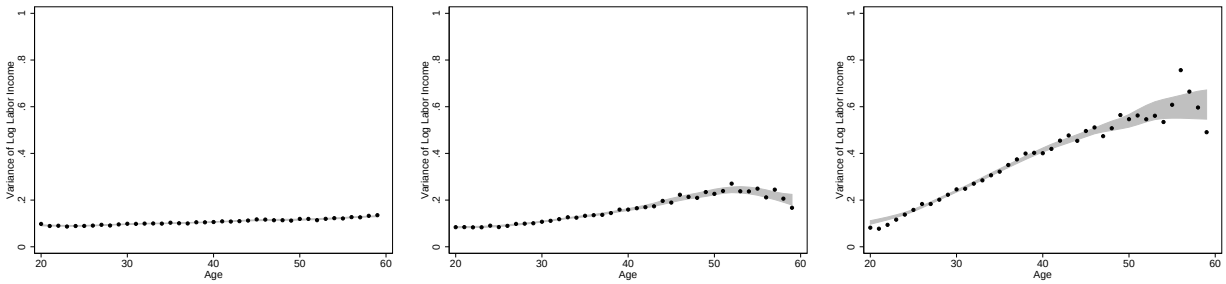


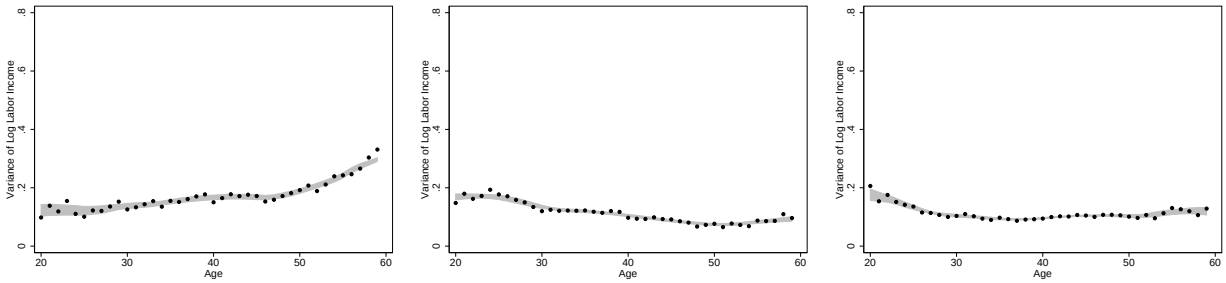
Figure 8: Labor income histograms by the public vs private sector employment

Notes: In this histogram, the individuals who earn more than 8,000 TRY/month (in 2018 prices) are represented as winsorized at the right end. They constitute only 1% of the sample.



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

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



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

Figure 10: Variance of log labor income in the public sector by education

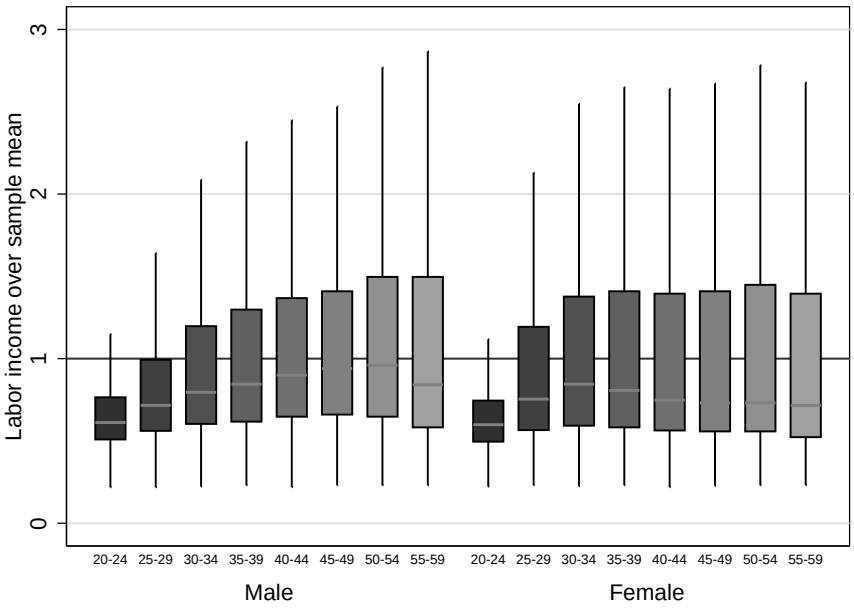
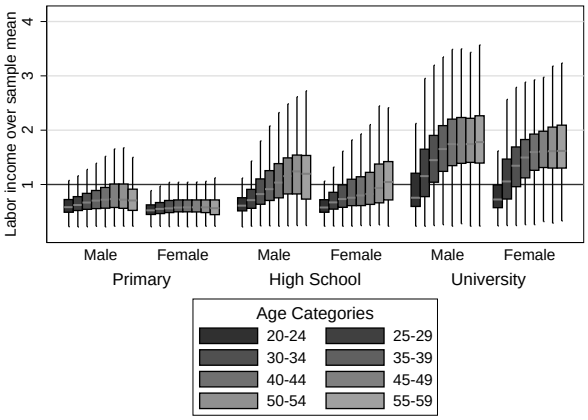
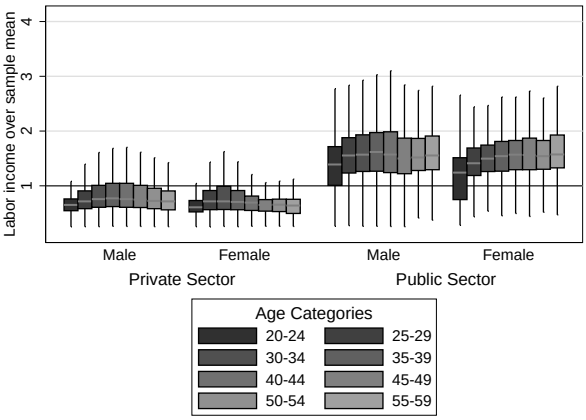


Figure 11: Labor income profile by gender

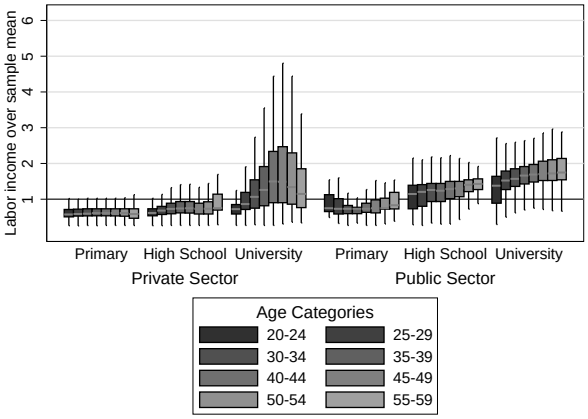


(a) Gender-education clusters

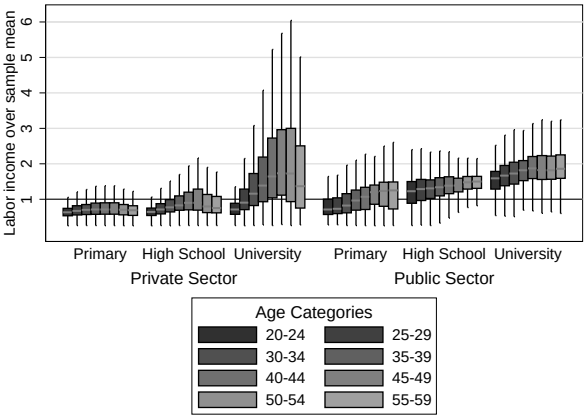


(b) Gender-sector clusters

Figure 12: Labor income profile by education, gender and the public vs private sector employment



(a) Female Distribution



(b) Male Distribution

Figure 13: 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.

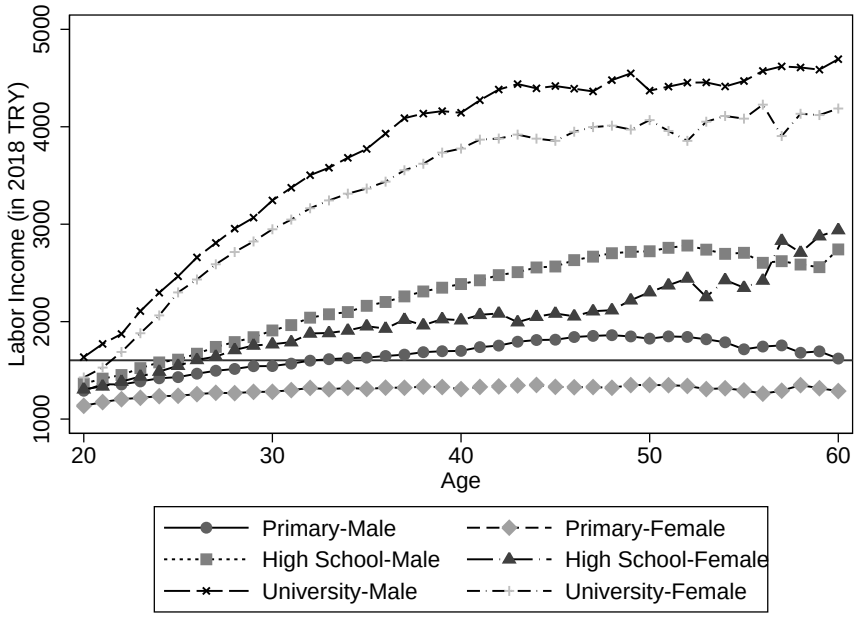


Figure 14: Labor income profile by education and gender

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)	(3) *log*(Labor Income)
Age			
25 to 29	0.118***	0.117***	0.117***
	(0.001)	(0.001)	(0.001)
30 to 34	0.197***	0.196***	0.190***
	(0.001)	(0.001)	(0.001)
35 to 39	0.229***	0.227***	0.216***
	(0.002)	(0.002)	(0.002)
40 to 44	0.232***	0.231***	0.216***
	(0.002)	(0.002)	(0.003)
45 to 49	0.213***	0.212***	0.202***
	(0.002)	(0.002)	(0.003)
50 to 54	0.176***	0.175***	0.166***
	(0.002)	(0.002)	(0.003)
55 to 59	0.131***	0.131***	0.116***
	(0.003)	(0.003)	(0.003)
Education			
Middle School	0.098***	0.095***	0.090***
	(0.001)	(0.001)	(0.001)
High School	0.198***	0.176***	0.158***
	(0.001)	(0.001)	(0.001)
University	0.569***	0.563***	0.503***
	(0.002)	(0.002)	(0.002)
Education-Sector Interaction			
Middle School × Public Sector		0.042***	0.057***
		(0.004)	(0.007)
High School × Public Sector		0.156***	0.246***
		(0.003)	(0.006)
University × Public Sector		0.082***	0.206***
		(0.004)	(0.005)
Mid Age (40-54)-Education-Sector Interaction			
Mid Age × Elementary School × Private Sector			-0.022***
			(0.002)
Mid Age × Elementary School × Public Sector			0.059***
			(0.006)
Mid Age × Middle School × Private Sector			-0.046***
			(0.003)
Mid Age × Middle School × Public Sector			0.045***
			(0.006)
Mid Age × High School × Private Sector			0.028***
			(0.003)
Mid Age × High School × Public Sector			-0.071***
			(0.004)
Mid Age × University × Private Sector			0.323***
			(0.006)
Mid Age × University × Public Sector			-0.075***
			(0.003)
Sector(Public=1)	0.294***	0.209***	0.157***
	(0.001)	(0.003)	(0.005)
Gender(Female=1)	-0.140***	-0.141***	-0.139***
	(0.001)	(0.001)	(0.001)
Tenure	0.012***	0.012***	0.013***
	(0.000)	(0.000)	(0.000)
Year Dummies	Yes	Yes	Yes
N	785,605	785,605	785,605
R-squared	0.52	0.52	0.53
F-statistic	48,683	43,629	33,950

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

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	0.066***	0.097***	0.266***	0.039***	0.072***	0.253***
	(0.002)	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)
30 to 34	0.092***	0.156***	0.429***	0.040***	0.109***	0.381***
	(0.002)	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)
35 to 39	0.098***	0.183***	0.531***	0.034***	0.111***	0.447***
	(0.002)	(0.003)	(0.005)	(0.004)	(0.005)	(0.006)
40 to 44	0.099***	0.197***	0.578***	0.015***	0.082***	0.478***
	(0.002)	(0.004)	(0.006)	(0.004)	(0.005)	(0.008)
45 to 49	0.093***	0.169***	0.571***	-0.013***	0.020**	0.457***
	(0.002)	(0.004)	(0.007)	(0.004)	(0.007)	(0.010)
50 to 54	0.052***	0.111***	0.536***	-0.035***	0.003	0.449***
	(0.003)	(0.005)	(0.008)	(0.005)	(0.011)	(0.012)
55 to 59	-0.001	0.059***	0.507***	-0.072***	0.044*	0.416***
	(0.004)	(0.007)	(0.011)	(0.007)	(0.021)	(0.018)
Sector(Public=1)	0.264***	0.341***	0.277***	0.170***	0.336***	0.303***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)
Tenure	0.011***	0.016***	0.004***	0.014***	0.021***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	296,302	161,101	141,980	60,806	45,448	79,968
R-squared	0.21	0.38	0.28	0.29	0.38	0.33
F-statistic	4,140	7,779	4,035	1,445	1,829	3,294

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	0.066***	0.085***	0.238***	0.004	0.037***	0.140***
	(0.002)	(0.002)	(0.004)	(0.017)	(0.010)	(0.006)
30 to 34	0.094***	0.139***	0.406***	0.010	0.017	0.183***
	(0.002)	(0.002)	(0.005)	(0.015)	(0.010)	(0.006)
35 to 39	0.100***	0.163***	0.532***	0.030*	-0.015	0.216***
	(0.002)	(0.003)	(0.006)	(0.015)	(0.010)	(0.006)
40 to 44	0.099***	0.178***	0.636***	0.011	-0.060***	0.228***
	(0.002)	(0.003)	(0.008)	(0.015)	(0.010)	(0.007)
45 to 49	0.083***	0.172***	0.671***	0.012	-0.129***	0.213***
	(0.002)	(0.004)	(0.011)	(0.015)	(0.010)	(0.007)
50 to 54	0.041***	0.132***	0.636***	-0.049**	-0.175***	0.194***
	(0.003)	(0.007)	(0.014)	(0.015)	(0.011)	(0.008)
55 to 59	-0.014***	0.109***	0.517***	-0.096***	-0.238***	0.182***
	(0.003)	(0.010)	(0.019)	(0.017)	(0.013)	(0.010)
Gender(Female=1)	-0.189***	-0.106***	-0.103***	-0.144***	-0.138***	-0.119***
	(0.001)	(0.002)	(0.003)	(0.005)	(0.004)	(0.002)
Tenure	0.007***	0.019***	0.022***	0.024***	0.019***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	324,263	164,772	104,476	32,845	41,777	117,472
R-squared	0.18	0.22	0.24	0.29	0.23	0.13
F-statistic	3,552	2,416	2,076	947	602	838

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(Labor Income)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	Birth(1-year)	Birth(5-year)	B(1)-E	B(1)-G	B(1)-S	B(5)-E	B(5)-G	B(5)-S
Age	0.0829***	0.1143***	0.1112 ***	0.0974***	0.1135***	0.0825***	0.0974***	0.1163***	0.0769***
	(0.0005)	(0.0062)	(0.0194)	(0.0053)	(0.0064)	(0.0065)	(0.0122)	(0.0154)	(0.0131)
Age^2	-0.0009***	-0.0010***	-0.0095***	-0.0008***	-0.0010***	-0.0007***	-0.0008***	-0.0010***	-0.0006***
	(0.0000)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Constant	5.8013***	4.7725***	4.8795***	5.2504***	4.8394***	5.6304***	5.2318***	4.7636***	5.7081***
	(0.0086)	(0.1325)	(0.4354)	(0.1106)	(0.1334)	(0.1416)	(0.2574)	(0.3257)	(0.2921)
Cohort Dummies	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,049,931	548	76	1,642	1,095	722	345	230	152
R-squared	.052	.911	.913	.812	.773	.773	.843	.817	.828
F-statistics	34,772	412	51	442	293	295	88	52	83

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. B(1) and B(5) refer to birth cohorts created with 1-year span and 5-year span, respectively. E stands for education, G denotes gender and S stands for sector.

Table 5: Pseudo-Panel Estimates

	log(Labor Income)					
	(1)	(2)	(3)	(4)	(5)	(6)
	B(1)-G-E-S	B(1)-G-E-S(1)	B(1)-G-E-S(2)	B(5)-G-E-S	B(5)-G-E-S(1)	B(5)-G-E-S(2)
Age	0.0793***	0.0992***	0.0484***	0.0767***	0.0982***	0.0503***
	(0.0041)	(0.0056)	(0.0037)	(0.0075)	(0.0115)	(0.0059)
Age^2	-0.0006***	-0.0009***	-0.0003***	-0.0006***	-0.0009***	-0.0003***
	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0002)	(0.0001)
Constant	5.7299***	5.1236***	6.6264***	5.7664***	5.1639***	6.4888***
	(0.0830)	(0.1038)	(0.0783)	(0.1421)	(0.1922)	(0.1298)
Cohort Dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3,430	1,909	1,521	850	438	412
R-squared	.514	.528	.556	.549	.556	.602
F-statistics	641	555	473	198	180	167

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. B(1) and B(5) refer to birth cohorts created with 1-year span and 5-year span, respectively. E stands for education, G denotes gender and S stands for sector. S(1) refers to private sector and S(2) refers to public sector.

Table 6: Pseudo-Panel Estimates

	log(Labor Income)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	B(1)-G-E	B(1)-G-E(1)	B(1)-G-E(2)	B(1)-G-E(3)	B(5)-G-E	B(5)-G-E(1)	B(5)-G-E(2)	B(5)-G-E(3)
Age	0.0893***	0.0696***	0.0835***	0.1177***	0.0885***	0.0692***	0.0839***	0.1157***
	(0.0042)	(0.0033)	(0.0050)	(0.0058)	(0.0094)	(0.0075)	(0.0112)	(0.0145)
Age^2	-0.0007***	-0.0005***	-0.0008***	-0.0010***	-0.0007***	-0.0005***	-0.0008***	-0.0009***
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Constant	5.3637***	5.3548***	5.6230***	5.0863***	5.3516***	5.3375***	5.5841***	5.0678***
	(0.0873)	(0.0663)	(0.1047)	(0.1168)	(0.1986)	(0.1588)	(0.2443)	(0.3067)
Cohort Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,104	1,084	987	1,033	681	230	223	228
R-squared	.753	.836	.723	.827	.796	.893	.784	.864
F-statistics	669	1,019	219	642	145	237	45	112

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. B(1) and B(5) refer to birth cohorts created with 1-year span and 5-year span, respectively. E stands for education, G denotes gender and S stands for sector. E(1), E(2) and E(3) refer to primary school, high school and university graduates, respectively.

Table 7: Pseudo-Panel Estimates

	log(Labor Income)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	B(1)-G-E	B(1)-G(1)-E	B(1)-G(2)-E	B(1)-G(1)-E(1)	B(1)-G(1)-E(2)	B(1)-G(1)-E(3)	B(1)-G(2)-E(1)	B(1)-G(2)-E(2)	B(1)-G(2)-E(3)
Age	0.0893***	0.1009***	0.0773***	0.0883***	0.0934***	0.1217***	0.0504***	0.0804***	0.1155***
	(0.0042)	(0.0048)	(0.0068)	(0.0022)	(0.0046)	(0.0062)	(0.0034)	(0.0089)	(0.0102)
Age^2	-0.0007***	-0.0009***	-0.0006***	-0.0008***	-0.0008***	-0.0010***	-0.0003***	-0.0008***	-0.0010***
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)
Constant	5.3637***	5.1824***	5.5361***	5.1703***	5.3763***	4.9856***	5.5430***	5.7348***	5.1498***
	(0.0873)	(0.1002)	(0.1377)	(0.0464)	(0.0976)	(0.1286)	(0.0599)	(0.1630)	(0.1944)
Cohort Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3,104	1,641	1,463	548	546	547	536	441	486
R-squared	.753	.828	.679	.918	.858	.856	.791	.608	.800
F-statistics	669	597	234	1,808	380	518	1,260	89	249

Note: Numbers in parantheses are standard errors. * for $p < .05$, ** for $p < .01$, and *** for $p < .001$. B(1) refers to birth cohorts created with 1-year span. E stands for education and G denotes gender. G(1) refers to male and G(2) refers to female. E(1), E(2) and E(3) refer to primary school, high school and university graduates, respectively.

APPENDIX-A: 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 HLFS data is not large enough for Turkey, the minimum cohort size is set at 50[30], 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 681 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

[30] As an alternative, we also use 20 as the minimum cohort size, to be able to get more observations per specification.

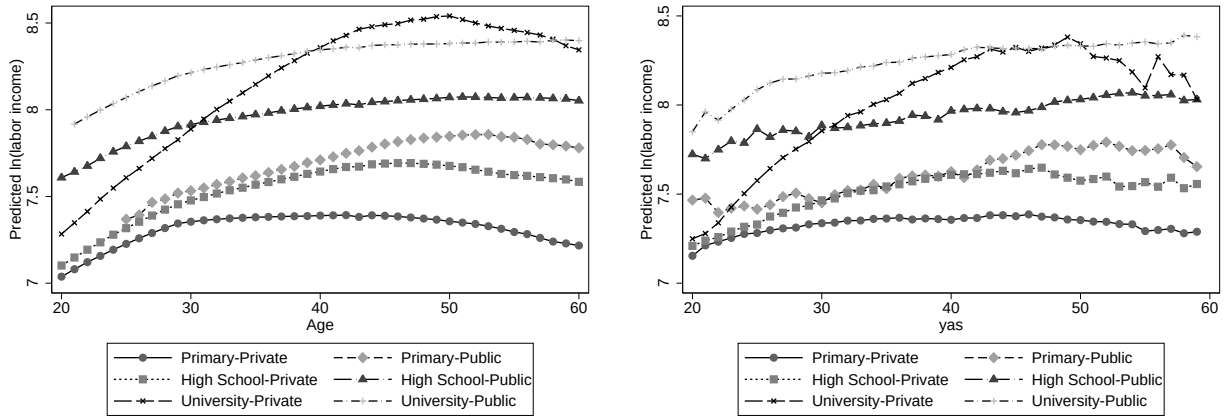
leads to better estimation efficiency but higher measurement error. That is why we have applied some variations in cohort forming such as excluding both the public and the private sectors, but the results do not change much.

For each cohort and each year, we calculate the mean of log income. In our benchmark specification, our synthetic data includes 48 cohorts (8 birth-groups, 2 genders, 3 education categories), 15 time periods and an average cohort size of 1540. There are 681 observations, which is less than 48*15=720, because the pseudo-panel is not balanced owing to an insufficient number of observations for particular groups in some years.

For Figure A1, the synthetic data includes 240 cohorts (40 birth-groups, 2 sector, 3 education categories), 10 time periods (2009-2018) and an average cohort size of 339. However, there are 2101 observations, which is less than 240*10=2400 because of the same reason.

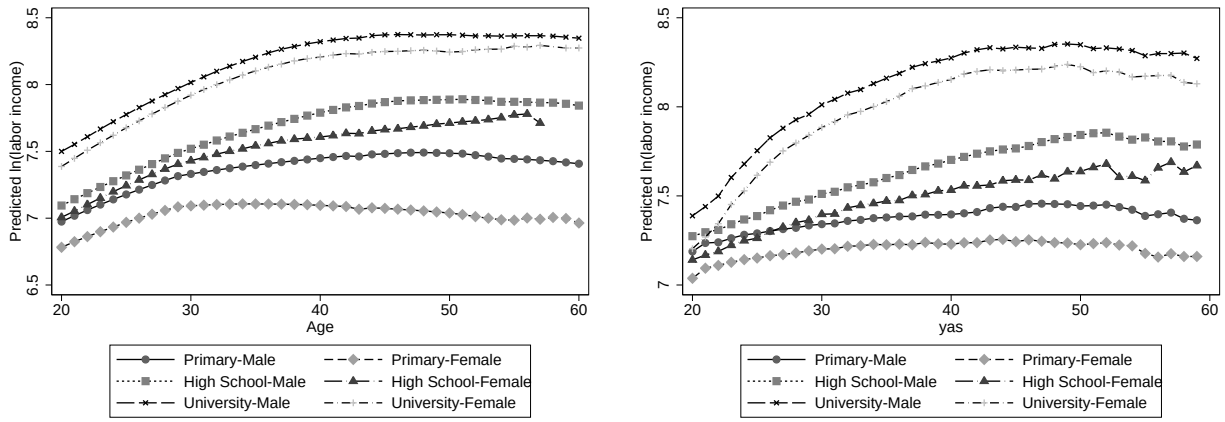
Similarly, for Figure A2, the synthetic data includes 240 cohorts (40 birth-groups, 2 gender, 3 education categories), 15 time periods (2004-2018) and an average cohort size of 337. In this case, there are 3104 observations, which is less than 240*15=3600.

Using the regressions in Table 4, we predict the labor income profiles for the cohorts mentioned above. After predicting the labor income profiles by including the cohort-fixed effects, we take the average of these profiles for each group. To get predicted income profiles with OLS, we use the third regression in Table 1. The predicted labor income profiles, which are obtained with two different methods, are quite similar to each other.



(a) Pseudo-Panel **(b)** OLS

Figure A1: Predicted labor income with pseudo-panel and OLS estimation, education and sector groups



(a) Pseudo-Panel

(b) OLS

Figure A2: Predicted labor income with pseudo-panel and OLS estimation, education and gender groups

APPENDIX-B: 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 1,105,600. The data covers the period 2004-2018. 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 2018; and workers who work less than 35 hours a week are excluded.

There is no sector variable from 2004 to 2008 in the data, so there are many missing values for the sector variable. Therefore, for some analyses the sample size is restricted to 788,239 observations.

Table B1: Descriptive Distributional Statistics

Full Sample (2004-2018)												
	Primary School			High School			University			All Categories		
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	450,589	89,375	539,964	234,463	66,026	300,489	183,110	102,113	285,223	856,830	248,770	1,105,600
Private Sector (2009-2018)												
	Primary School			High School			University			All Categories		
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	267,653	57,870	325,523	127,149	37,812	164,961	64,692	40,003	104,695	459,494	135,685	595,179
Public Sector (2009-2018)												
	Primary School			High School			University			All Categories		
	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total
Observations	30,053	3,094	33,147	34,325	7,668	41,993	77,843	40,077	117,920	142,221	50,839	193,060

Table B2: Descriptive Distributional Statistics

	Gender		Education			
	Male	Female	Primary	High School	University	All
Private	77%	23%	54%	27%	17%	76%
Public	74%	26%	17%	22%	61%	24%
Full-Sample	76%	24%	46%	26%	28%	100%