// NO.26-024 | 06/2026

DISCUSSION PAPER

// EFI ADAMOPOULOU, MANOLIS GALENIANOS, NICHOLAS GIANNAKOPOULOS, PANTELIS KAMMAS, AND IOANNIS LALIOTIS

Earnings Dynamics Through Boom, Crisis, and Recovery: Evidence from Greece

Leibniz Association

ZEW

Earnings Dynamics Through Boom, Crisis, and Recovery: Evidence from Greece*

Efi Adamopoulou† Manolis Galenianos‡ Nicholas Giannakopoulos§
Pantelis Kammas¶ Ioannis Laliotis ‖

May 2026

Abstract

This paper documents the evolution of labor earnings and earnings dynamics in Greece during 2002–2023 using administrative matched employer–employee data from the Greek social security system. The data span the expansion of the 2000s, the deep recession of 2009–2013, and the subsequent recovery. We show that earnings, volatility, and inequality closely tracked macroeconomic conditions. Real earnings rose during the pre-crisis expansion, declined sharply during the recession (particularly among lower earners) and only partially recovered afterward. Earnings volatility and downside risk increased substantially during the crisis, while inequality rose both in the cross-section and within cohorts. Workers entering the labor market during the recession experienced unusually weak initial earnings and persistently lower subsequent earnings. Using regional variation in unemployment rates, we show that weaker local labor-market conditions at entry are associated with lower early-career earnings.

Keywords: Labor earnings, Inequality, Volatility, Mobility, Administrative data, Greece
JEL codes: J30, D31, C55

*We thank the editor and two referees for thoughtful comments, as well as the participants at the 2025 World Congress of the Econometric Society for useful discussions. We are grateful to EFKA Director Alexandros Varveris for his instrumental role in securing access to the data. We thank Manos Matsaganis and Panos Tsakloglou for their assistance with the data, and Thanos Bourlos for helping us decipher the evolution of the Greek social security system. We also thank Jiangting Wang for excellent research assistance.

†ZEW, Deutsche Bundesbank, University of Mannheim, and CEPR
‡Royal Holloway, University of London
§University of Patras
¶Athens University of Economics and Business
‖University of Patras and Hellenic Observatory Centre, London School of Economics

1 Introduction

Income inequality and earnings dynamics have become central topics in macroeconomics and labor economics. While cross-sectional measures of inequality provide useful information about the distribution of income at a point in time, they offer only a partial picture of economic heterogeneity. Longitudinal earnings data reveal additional dimensions of labor-market risk and opportunity, including earnings volatility, persistence, mobility, and the extent to which macroeconomic shocks affect workers differently across the earnings distribution and over the life cycle. In recent years, the increasing availability of administrative data has substantially expanded researchers' ability to study these issues in a comparable and high-resolution manner across countries (Guvenen, Pistaferri, and Violante, 2022).

This paper provides new evidence on the evolution of labor earnings and earnings dynamics in Greece during 2002–2023 using administrative matched employer–employee data from the Greek social security system. The paper is part of the Global Repository of Income Dynamics (GRID) project, which aims to construct harmonized statistics on earnings inequality, earnings risk, and earnings mobility across countries using comparable administrative data sources. Despite the severity of the Greek economic crisis and the substantial changes in labor-market institutions during the 2010s, there exists relatively limited evidence on the evolution of labor earnings dynamics in Greece using large-scale longitudinal administrative data. Our goal is to help fill this gap.

The dataset that we use consists of the universe of monthly social security records submitted by private-sector employers to the main Greek social security funds during 2002–2023. These records contain detailed information on workers, firms, and employment relationships, including earnings, employment spells, occupations, firm identifiers, sector, location, and worker demographics.[1] The data cover the vast majority of private-sector salaried employees in Greece over more than two decades and span three distinct macroeconomic periods:

[1]This dataset is essentially untapped by researchers. One of the few papers to use data from the Greek social security system is Saez, Matsaganis, and Tsakloglou (2012) which investigates the effects of a reform of the social insurance system from the early 1990s.

the expansion of the 2000s, the deep recession and labor-market adjustment of 2009–2013, and the subsequent recovery.

The Greek case is particularly informative for studying earnings dynamics. Between 2008 and 2013, Greece experienced one of the deepest economic contractions observed in an advanced economy in the postwar era. Real GDP declined by more than 25 percent, unemployment rose from below 8 percent to almost 28 percent, and real wages fell sharply. The crisis also coincided with major labor-market reforms, including large reductions in the minimum wage and a substantial decentralization of collective bargaining. These developments generated unusually large movements in employment, wages, and labor-market inequality, making Greece a useful setting for studying how earnings distributions evolve during periods of extreme macroeconomic stress.

Our analysis documents several broad patterns. First, labor earnings evolved closely with aggregate economic conditions. Real earnings increased across most of the distribution during the expansion of the 2000s, declined sharply during the crisis (particularly among below-median and median earners) and partially recovered during the following decade. Second, earnings volatility and downside earnings risk increased substantially during the recession and gradually returned toward pre-crisis levels afterward. Third, inequality rose sharply during the crisis, both in the cross-section and within cohorts, before declining during the recovery period. Fourth, the crisis had especially strong effects on younger workers and on cohorts entering the labor market during the recession, whose initial earnings were unusually low and whose subsequent earnings trajectories remained persistently weaker than those of adjacent cohorts.

In addition to documenting these patterns, we present an application that examines how labor-market conditions at the time of labor-market entry are associated with subsequent earnings outcomes. Using regional variation in unemployment rates at entry, we show that cohorts entering the labor market during periods of high unemployment experienced substantially lower earnings during the early years of their careers, with larger losses among

workers with lower levels of education. These findings are consistent with the broader literature on recession-entry cohorts and labor-market scarring (Oreopoulos, von Wachter, and Heisz, 2012; Schwandt and von Wachter, 2019).

The paper contributes to several strands of literature. First, it contributes to the growing cross-country literature on earnings dynamics based on administrative data, including recent work associated with the GRID project. Second, it contributes to the literature on labor-market adjustment during deep recessions by documenting the evolution of wages, inequality, volatility, and mobility during the Greek crisis. Third, it provides descriptive evidence on the labor-market experiences of recession-entry cohorts in a setting characterized by exceptionally large macroeconomic shocks and substantial institutional change.

The remainder of the paper is organized as follows. Section 2 presents the macroeconomic background and describes the data and sample construction. Section 3 documents the evolution of earnings distributions, volatility, mobility, and inequality. Section 4 examines the earnings trajectories of recession-entry cohorts. Section 5 concludes.

2 Macroeconomic Background and Data

Our dataset covers the period 2002-2023. This section provides a brief description of the key macroeconomic developments during that period and describes the data that we use for our analysis.

2.1 Macroeconomic background

The years 2002-2023 were extremely volatile for the Greek economy. Panel A in Figure 1 shows the evolution of real GDP per capita in Greece in 2002-2023. These years can be separated into three distinct periods. In 2002-2008, the economy expanded strongly and real GDP grew by 3.7% on average annually. In 2009-2013, the economy collapsed and real GDP contracted by 27% due to three overlapping and interrelated crises: a sovereign debt crisis,

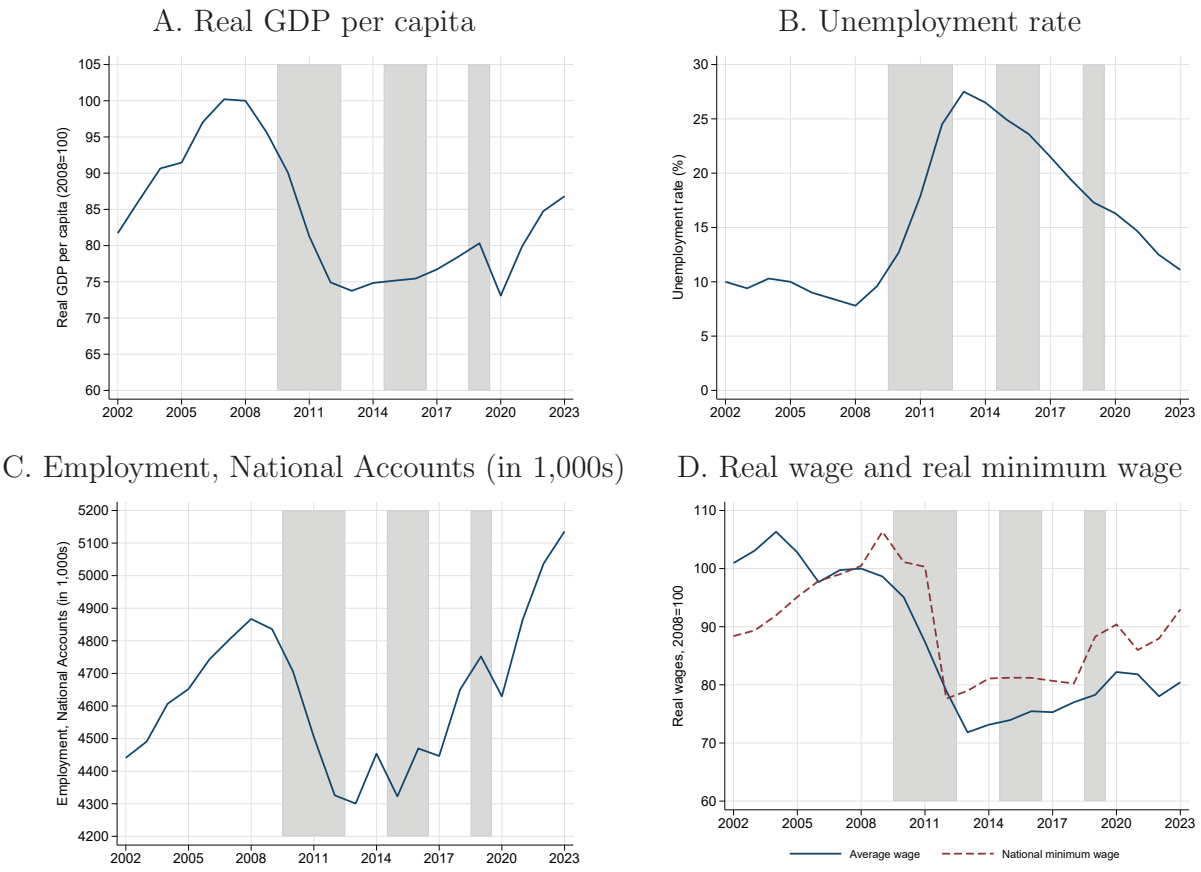


Figure 1: Macroeconomic Developments in Greece in 2002-2023

Source: Eurostat. Real GDP per capita and real wages are normalized to 100 in 2008.

a sudden stop in the balance of payments, and a banking crisis. After 2014, the economy stabilized and grew at a relatively steady, if somewhat low, rate of 1.3% annually, punctured by a brief, steep recession during the COVID pandemic and a sharp recovery thereafter.[2]

[2]Giannoulakis, Sakellaris, and Tsoukalas (2025) provide a broad overview of economic growth and structural changes in the Greek economy over the past 6 decades using a growth accounting and decomposition framework. Galenianos (2015) uses a comparative cross-country perspective to investigate the role of external imbalances in the origins of the Greek crisis. Gourinchas, Philippon, and Vayanos (2017) develop a quantitative macro-finance framework that combines cross-country benchmarking with a stylized DSGE model to analyze the Greek crisis and decompose the roles of fiscal imbalances, sudden stop, and banking crisis. Chodorow-Reich, Karabarbounis, and Kekre (2023) decompose the sources of the Greek boom and bust and assess the role of fiscal and financial policies in amplifying the depression. Arkolakis, Doxiadis, and Galenianos (2017) use a quantitative international trade framework to analyze the adjustment of the Greek economy to the sudden stop and, particularly, focus on the reallocation of production between the non-tradable and tradable sectors.

Developments in the labor market followed macroeconomic events quite closely. The unemployment rate was on a downward path in the first period, though it remained at a relatively high level of 8% in 2008 after 15 years of continuous and significant growth, suggesting that structural rather than cyclical reasons were responsible for its persistently high level.[3] During the crisis, the unemployment rate more than tripled to a peak of 27.7% in 2013, and it has been steadily falling since then, down to 10.4% at the end of 2023. See panel B in Figure 1.

Employment followed the opposite trajectory to unemployment: it increased during the growth period of the 2000s, peaked in 2008, dropped significantly during the crisis, and entering a period of recovery since then, eventually surpassing the pre-crisis peak (panel C in Figure 1). Panel D in Figure 1 presents the evolution of the average real wage and the real level of the National Minimum wage, both indexed at 100 in 2008. The average real wage peaked in the mid-2000s, dropped by almost 30% during the crisis, and staged a modest recovery in subsequent years. The real National Minimum Wage rose gradually until the crisis, declined sharply in 2012 as part of a reform described below, and then gradually increased again.

An important institutional development in the Greek labor market was the implementation of a significant reform package in February 2012. The reform package reduced the National Minimum Wage by an unprecedented 22% in nominal terms (see panel D), greatly weakened collective bargaining agreements reducing their coverage from 85% in 2011 to less than 20% four years later (OECD), suspended statutory seniority bonuses (the triennia), and reduced the cost of redundancies.[4] These changes coincided with a sharp decline in

[3]Key features of the labor market included a relatively high minimum wage, consistently around 60% of the mean private-sector wage (Kanellopoulos, 2015) and a "domino effect" from national to sectoral and occupational collective bargaining, leading to a rigid wage structure (Koukiadaki and Kokkinou, 2016).

[4]One component of the reform was to expand the scope for firm-level bargaining to firms with fewer than 50 employees. Firm-level agreements were allowed to set wages below the floors established by sectoral and occupational collective agreements, provided they remained above the National Minimum Wage. Giannakopoulos and Laliotis (2020) document that the reform led to a substantial increase in firm-level agreements and a marked decline in negotiated wage floors in subsequent years.

wages, particularly for the low-paid, as well as the stabilization of employment.[5]

2.2 Data source

The source of the data that we use in our analysis is the Greek Social Security system and, specifically the Social Insurance Institute (IKA) and the National Social Security Fund (EFKA). IKA was the primary social security fund for private-sector employees until the end of 2016. EFKA was established in 2017 through the merger of IKA with other social security funds that insured the self-employed, public-sector employees, and private-sector employees in specific sectors and firms that were not previously covered by IKA. EFKA currently insures all workers in Greece.

Our data is drawn from records of Social Security Contributions (SSCs). Since January 2002, employers who hire workers under private-law contracts and whose employees are insured by the Social Insurance Institute (IKA) or the National Social Security Fund (EFKA) are required to submit a monthly *detailed periodic statement* for each employee. The purpose of these statements is to record social security contributions and determine eligibility for various benefits, and they include detailed information about the worker (e.g. date of birth, gender, citizenship), the firm (e.g. sector, location, establishments), and the employment relationship (e.g. earnings, social security contributions, pay components, occupation). Our data consists of the universe of statements that were submitted to IKA in 2002-2016 and to EFKA in 2017-2023.

All private-sector employees work under private-law contracts and our data cover the vast majority of such employees. In the 2000s, approximately 10% of employees in the economy's business sector were insured in funds other than IKA and in the 2010s that share declines to negligible levels as these funds gradually merge with IKA/EFKA.[6] Public-sector

[5] Gatopoulos, Louka, Polycarpou, and Vettas (2026) find that the reduction in minimum wages and decentralization of collective bargaining contributed to a significant decline in labor costs, helping restore cost competitiveness and leading to rising employment, though with a delay. However, productivity growth remained weak and the social costs, including wage compression and inequality increase, were substantial.

[6] Appendix B describes how we identify workers who enter our data through funds mergers.

employment is approximately 20% of total employment and most public-sector employees work under public-law contracts, so they are not included in our data.[7] Self-employed workers account for 35% of employment on average during the period and that share is on a slow downward trajectory, according to the Labor Force Survey. The self-employed are not included in our data. Overall, our data covers roughly 40-50% of employment and the vast majority of private-sector salaried employees.

2.3 Constructing the analysis panel

We use the IKA/EFKA data to construct an annual panel of private-sector workers from 2002 to 2023. Appendix A describes the steps we take to clean and harmonize the monthly data. The outcome of the cleaning and harmonization process is a dataset where each worker-establishment pair has one entry per month, compensation type (base pay, bonus, overtime, etc.), and work type (full time or part time). Each entry consists of the worker's characteristics (worker ID, age, gender, citizenship), the establishment's characteristics (establishment ID, firm ID, location, sector), and the characteristics of the employment relationship (occupation, compensation, employer contributions, employee contributions, days of insurance, insurance package). To turn the monthly dataset into an annual panel of worker earnings, we merge all entries within a calendar year that have the same worker ID and create an annual earnings variable by summing all the values for compensation across compensation types, across establishments, and across months. The entries for occupation, sector, and insurance package are taken from the monthly entry with the maximum base pay. We deflate earnings using the Consumer Price Index. This is the baseline panel.

We now describe the selection criteria. We identify workers who enter our sample through the merger of a fund with IKA or EFKA, as detailed in Appendix B, and drop these workers from our sample. In each year, we keep prime-age workers between 25 and 55 years old. In each year, we keep workers who are attached to the labor force, defined as having gross

[7]The main source of information about public-sector employment is the Human Resources Management System of the Public Sector (formerly, the Census of Public Employees).

compensation for the full year that exceeds 1.5 times the monthly minimum wage for that year. We keep workers whose employer is in the extended business economy, defined as the NACE 2.0 sectors B-N, P, Q, R, and S (except S94), and we drop workers whose employer is in sectors A (Agriculture), O (Public Administration), S94 (Membership Organizations), T (Households as Employers) and U (Extraterritorial Organizations). Since we do not have data on public-sector employees, workers in sectors P (Education) and Q (Health) are mostly employed at private providers of education and health. Finally, for computational purposes, we keep workers who are born on days 2, 7, 12, 17, 22, or 27 of the month, who account for approximately 20% of the IKA/EFKA population. This is our analysis panel. The analysis panel is a balanced-coverage, representative sample of roughly 20% of the workers in the IKA/EFKA dataset, selected by birth-date.

Appendix C provides a detailed comparisons of employment, firm counts, and inequality measures from this dataset and from external sources (National Accounts, ERGANI administrative data, World Inequality Database). The comparison shows close alignment, validating the dataset's accuracy and representativeness.

3 The evolution of earnings, volatility, and inequality

This section presents statistics on the evolution of earnings, on the distribution of earnings growth, and on earnings inequality.[8] We present a number of different variables and we focus on five stylized facts that relate how these magnitudes evolve during and after the Greek crisis:

Stylized fact 1: The Greek crisis produced exceptionally large and persistent declines in real earnings across the entire distribution.

[8]In the construction of the sample, the computation of the statistics, and their graphic illustration, we adhere to the standardized guidelines provided by GRID.

Stylized fact 2: The earnings collapse was much larger among below-median and median earners and inequality increased significantly.

Stylized fact 3: Downside earnings risk rose dramatically during the crisis.

Stylized fact 4: Within-cohort inequality increased sharply during the crisis.

Stylized fact 5: Cohorts entering the labor market during the crisis experienced unusually weak earnings trajectories that persisted over time.

3.1 Evolution of earnings

We begin by examining the evolution of the cross-sectional distribution of real earnings in Figure 2. Panel A displays the evolution over time of selected percentiles of the logarithmic real earnings distribution, each measured relative to its 2002 value. Panels C and E show the evolution of the same percentiles separately for women and men but, since these are very similar, we focus our discussion on panel A.

In 2002-2008, earnings increased strongly for all percentiles of the distribution, with lower incomes experiencing slightly higher growth rates. Overall growth for the period was roughly 10-15 log points. This pattern stands in contrast to many developed economies which saw income stagnation for low-earners over the same period. For example, below-median earners in Italy experienced real earnings declines of approximately 40 log points (similar in magnitude to Greek low-earners losses during the crisis period), but this happened over a much longer, two-decade period, and the causes are likely very different (Hoffmann, Malacrino, and Pistaferri, 2022).

In 2009–2013, earnings declined sharply across the board and the decline was largest at the bottom of the distribution: it reached nearly 50 log points at the 10th and 25th percentiles, around 40 log points at the median, and about 30 log points at the 75th and 90th percentiles. The breadth of the decline, affecting all percentiles of the earnings dis-

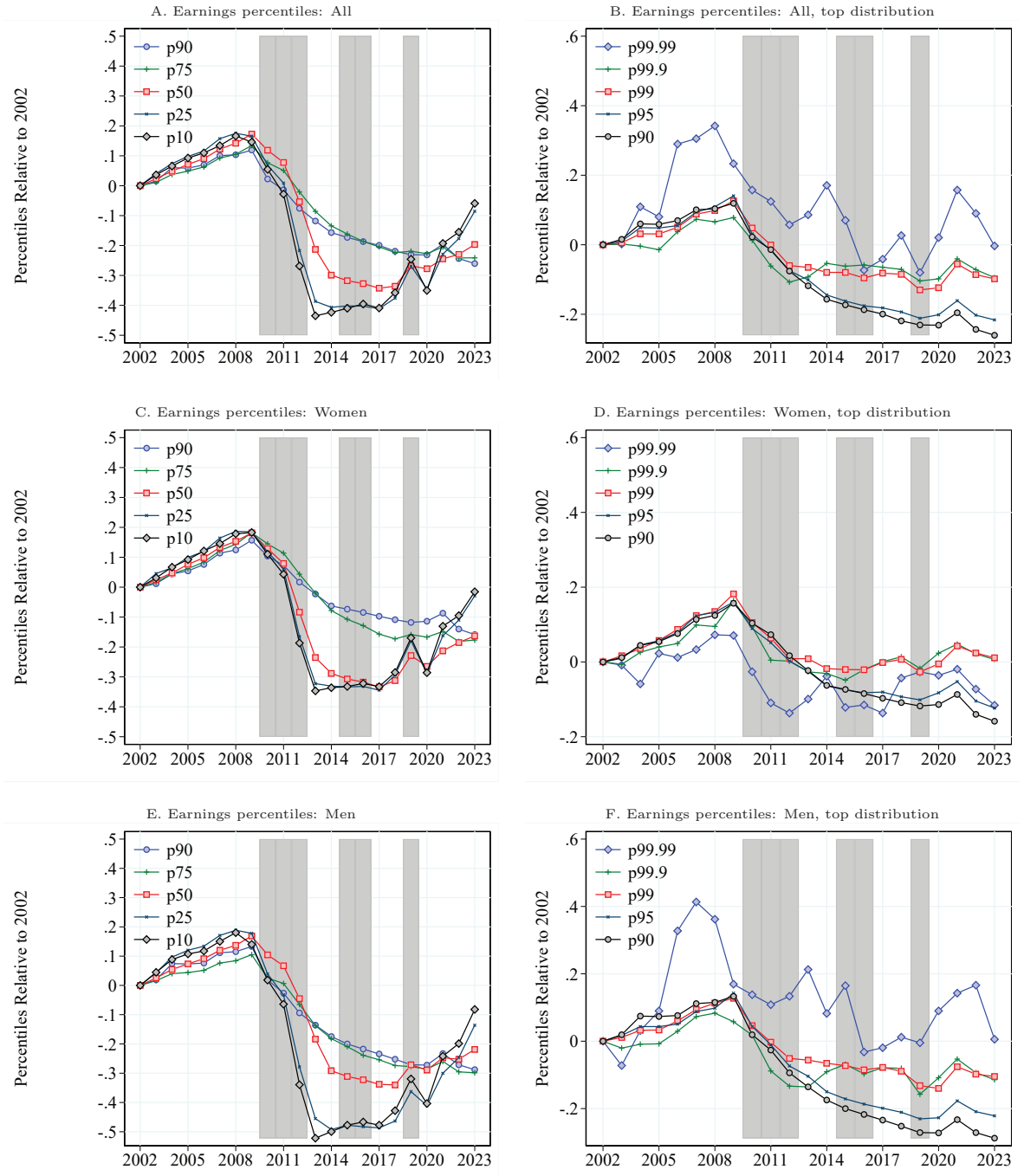


Figure 2: Percentiles of the distribution of log annual earnings

Notes: Percentiles of log annual earnings for all respondents, women, and men. All percentiles are normalized to 0 in 2002. The shaded areas indicate recession years.

tributions, is our first stylized fact and its depth for below-median earners is our second stylized fact. Both stylized facts stand out when compared with other countries' experience.

During the same time period, Spain also experienced a severe recession, but earnings losses were concentrated on below-median male earners, while the earnings declines of female and median-or-above earners were relatively modest (Arellano, Bonhomme, De Vera, Hospido, and Wei, 2022). The depth of the earnings decline is consistent with the combination of an exceptionally severe recession and substantial changes in wage-setting institutions during the crisis years (see section 2.1 for a description).

In 2014-2023, the earnings of median and below-median earners were initially stable and started growing after 2018, recovering some (though far from all) of the losses incurred during the crisis. By contrast, above-median earners experienced further, though small, declines in their labor earnings. The long duration of the earnings losses is unprecedented. A full decade after the stabilization of GDP, real median earnings remain more than 35 log points below their 2008 peak and above-median earnings are even lower. It is only below-median earnings that have recovered strongly, which coincided with successive rises in the minimum wage after 2019.

Panels B, D, and F plot the evolution of earnings for all, women, and men top earners. Apart for the highly volatile 0.01 percent, the patterns for the other top earning groups (5th, 1st, and 0.1st percentile) are broadly similar to those of above-median earners.

3.2 Distribution of earnings growth and mobility

This section examines how various moments of earnings *growth* vary over time and by earnings level. Following Guvenen, Pistaferri, and Violante (2022), the key variable of interest is growth in *residualized* annual earnings, defined as the one-year change in log annual earnings after residualizing for worker age and calendar-year effects. We denote one-year growth in residualized earnings for individual i in year t by g_{it}^1.

Figure 3 presents the evolution of volatility measures of earnings growth over time. Panel A plots the P90-P10 percentile gap of the real earnings growth distribution for men and women. The volatility of earnings in Greece is higher for women than for men, a common

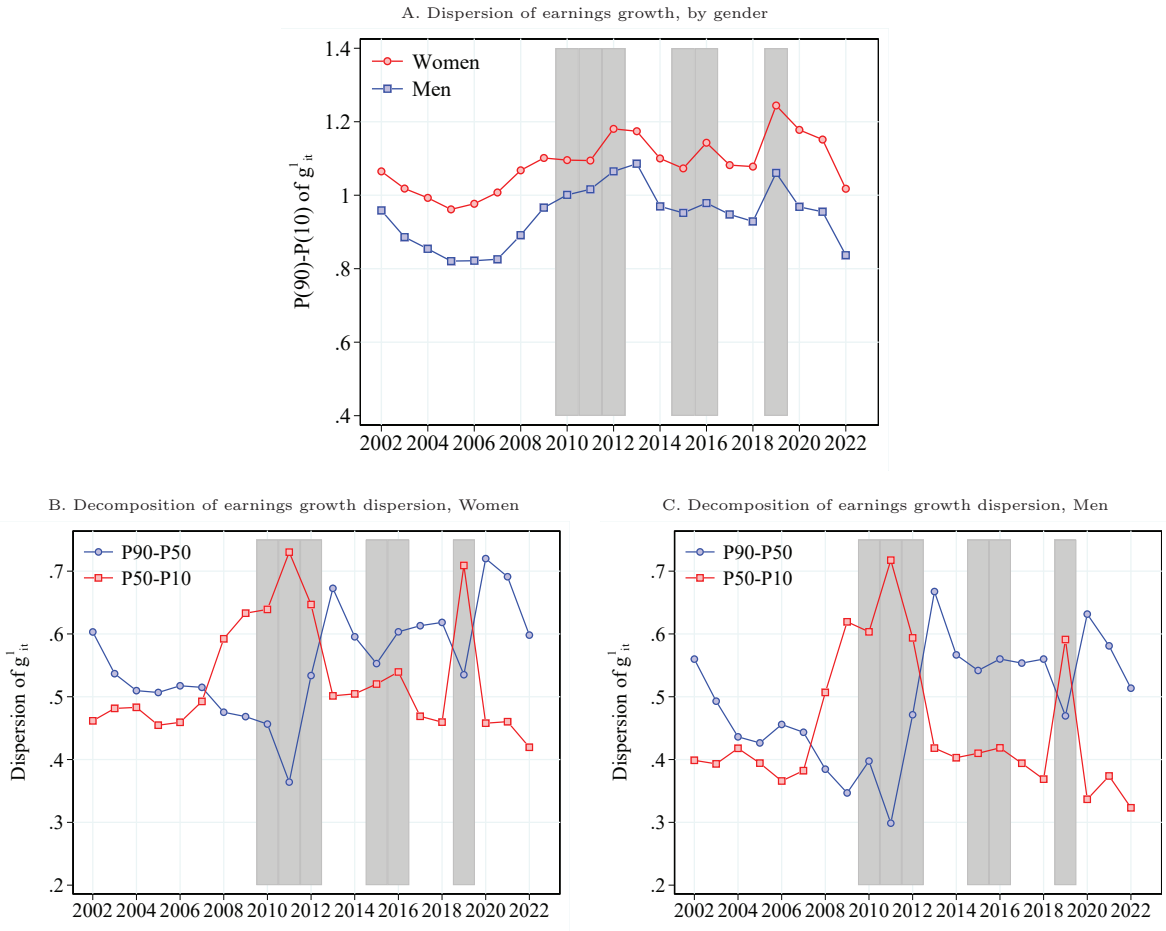


Figure 3: Dispersion of 1-Year Log Earnings Changes

Notes: This figure shows the P90-P50 and P50-P10 differentials of the distribution of 1-year changes in residualized log real annual earnings g_{it}^1 (from t to $t+1$). Shaded areas indicate recessions.

finding in the literature (Drechsel-Grau, Peichl, Schmid, Schmieder, Walz, and Wolter, 2022). Compared to other countries, volatility is somewhat higher in Greece (Hoffmann, Malacrino, and Pistaferri, 2022) and it is likely that the importance of seasonal work in the Greek economy is related to this feature. Noticeably, volatility shows no long-run trend, apart for the sizable but transitory increases during the crises of the early 2010s and COVID.

Panels B and C decompose the volatility of real earnings growth into "upside risk" (P90-P50 gap) and "downside risk" (P50-P10 gap) for women and men, respectively. The downside risk increased dramatically for both women and men during the crisis, almost doubling for

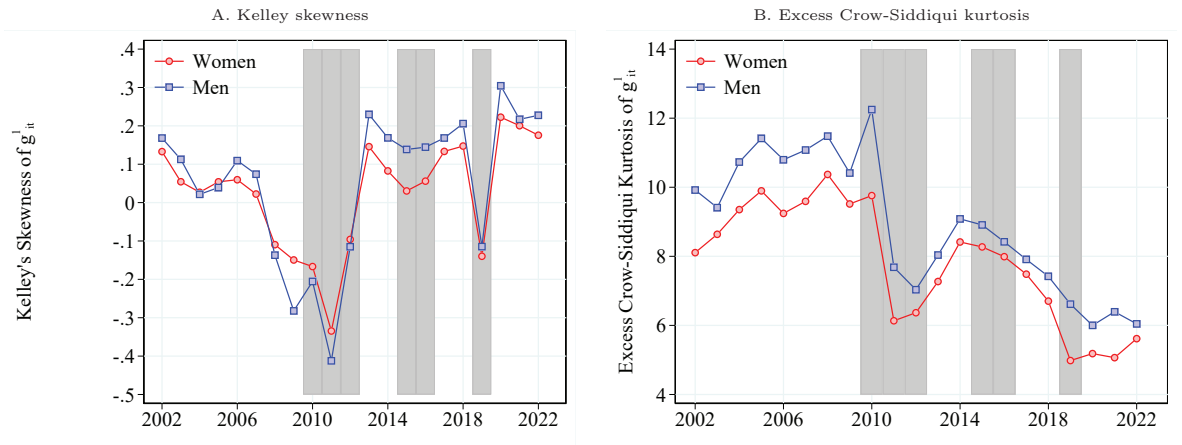


Figure 4: Earnings distribution moments by gender

Notes: Panel A presents estimated Kelley skewness of the income distribution, while panel B presents the excess Crow-Siddiqi kurtosis. Both measures are computed separately by gender. Earnings are expressed in logarithms. See text for details on sample construction and estimation.

men. This is (part of) our third stylized fact. After the crisis subsided, downside risk stabilized, apart for a spike during COVID, at a level similar to that of the mid-2000s. Upside risk declined during the crisis and then increased considerably, so that in 2023 it is higher than in the mid-2000s. These patterns are similar for women and men.

Figure 4, panel A plots the evolution over time of Kelley skewness, which quantifies the asymmetry of a distribution around its mean.[9] The skewness of earnings growth is modestly positive most of the time, meaning that the distribution of earnings growth has a longer right tail. During the crisis, however, it becomes very negative, suggesting that during those years earnings growth had a very long left tail, forming part of our third stylized fact. Comparing with other countries' experience, both the size of the left tail and the duration of the negative-skewness episode are notably large. They are, however, not so dissimilar from the experience of Spain and Italy where, after the financial crisis, skewness was negative for 6 successive years and reached a minimum similar to Greece's (-0.35 for Spain and -0.27 for Italy both in 2008).

Figure 4, panel B plots the excess Crow-Siddiqui kurtosis, which quantifies the magnitude

[9]Kelley skewness is defined as $K = \frac{P_{90}-P_{10}+2P_{50}}{P_{90}-P_{10}}$ and it is equal to zero in a Gaussian distribution.

of extreme values.[10] Kurtosis is high in the 2000s, meaning that there are many extreme values, it drops significantly during the crisis and partially recovers in the mid-2010s before slowly falling in the following years. To understand the unusually sharp decline of kurtosis during the crisis, it is useful to recall that earnings growth is measured for workers who appear in our data in successive years. In 2009-2012, employment declined dramatically and the number of workers in our data dropped by 15% suggesting that there were many job-loss events without new employment in the following year.[11] Hence, it is likely that many of the large negative earning growth events are not recorded due to exit from the dataset while there were few positive earning growth events due to the crisis, leading to a large decline in kurtosis. The persistence of low values for kurtosis during the recovery period is likely related to stable employment growth (less job loss) and low overall earnings growth.

Next, we examine how the higher moments of earnings growth vary across workers with different levels of permanent income. Permanent income is defined as the three-year moving average of residualized log earnings, with percentile ranks computed each year and then pooled across the sample. Each measure is calculated separately for men and women and for three age groups (ages 25–34, 35–44, and 45–55). Remarkably, given the magnitude of the Greek crisis, the patterns are broadly consistent with other countries, with a few exceptions that we point out below.

Figure 5, Panels A and B present the level of earnings volatility (P90-P10 earnings growth ratio) by permanent income for men and women, respectively. Volatility declines with age and in permanent income, with the exception of the very top incomes. At the ages of 25-34, volatility is higher for women and, with age, gradually converges to that of men, a feature observed in other GRID countries that is likely due to intermittent labor force participation. For both genders, volatility is greatest among those in the lowest permanent-

[10] Excess Crow-Siddiqui kurtosis is defined as $CS = \frac{P_{97.5}-P_{2.5}}{P_{75}-P_{25}} - 2.91$ and it is equal to zero in a Gaussian distribution.

[11] The long-term decline in employment is visible in the aggregate data as well: the long-term unemployed's share of the labor force increased from 4% before the crisis to 14% in 2012 according to the Labor Force Survey.

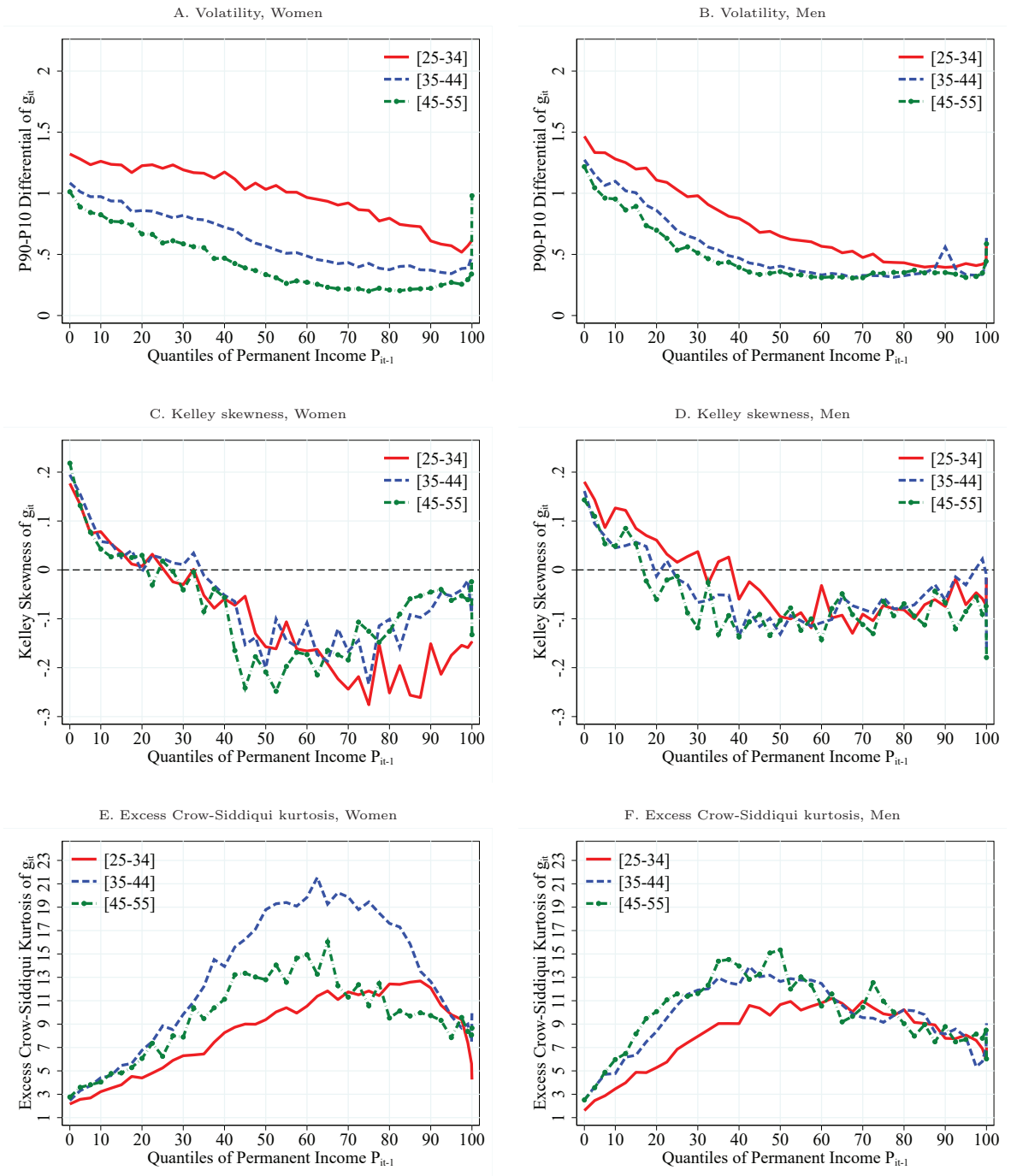


Figure 5: Volatility and asymmetrical and central tendencies of earning growth by income

Notes: Panels A and B present income volatility measured as the P90–P10 percentile difference, panels C and D present Kelley skewness, and panels E and F present excess Crow-Siddiqi kurtosis. Results are shown separately for women and men.

income percentiles.

Panels C and D plot Kelley skewness across permanent-income percentiles. For both men and women, skewness is highest at low permanent incomes, initially declines with permanent income, and then rises slightly. There is little variation among age groups except for younger women who display substantially more negative skewness at higher permanent-income levels. The variation of skewness with permanent income is similar to other countries but the level of skewness is lower in Greece than in, say, Italy or France (though it is similar to Spain's, at least for men). Panels E and F present excess Crow–Siddiqui kurtosis by permanent-income percentile, again for women and men separately. Both genders show an inverted-U pattern, but the age gradients are steeper and the peak central concentration is higher for women.

Finally, we turn to mobility. Unlike volatility, which captures short-term fluctuations in earnings, mobility describes shifts in the persistent component of earnings over longer horizons. To assess earnings mobility in our data, we compute statistics that measure how strongly a worker's current permanent earnings predicts their permanent earnings at a future date.

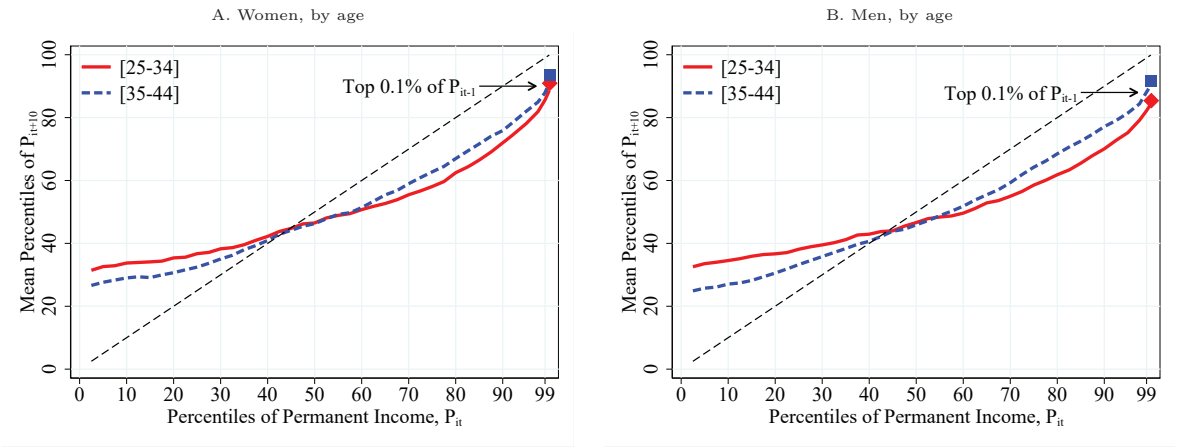


Figure 6: Evolution of 10-year earnings mobility over the life cycle

Notes: This figure shows average rank-rank 10-year earnings mobility. The various curves on the graph correspond to different age groups measured at time t: solid corresponds to 25–34 and dashed corresponds to 35–44. The squares and diamonds correspond to the top 0.1 percentile of the distribution of permanent income at t. Results are shown separately for women and men.

Figure 6 displays, for each initial permanent-income percentile in year t, the average percentile the same workers occupy in year $t+10$ for two age cohorts (25–34 and 35–44), for women (panel A) and men (panel B). Under perfect mobility—where today's earnings rank are uncorrelated with tomorrow's—the curve would be a flat line at the 50th percentile. Conversely, under complete stagnation, where workers remain in the same percentile throughout their working lives, the curve would lie exactly on the 45° line (shown here as the black dashed line).

Panels A and B reveal that in Greece there is a clear reversion toward the mean (as in Spain—see Arellano, Bonhomme, De Vera, Hospido, and Wei (2022)): individuals at the lower end of the permanent income distribution experience upward mobility, while those at the top experience downward mobility, on average. Labor income mobility is slightly higher among the younger cohort, suggesting that mobility is greater early in one's working life. Low-income workers can climb as much as 30 percentage points in the permanent-earnings ranking. These patterns are broadly similar for women and men.

3.3 Inequality in the cross-section, within cohorts, and across cohorts

This section presents measures of inequality over time, in the cross-section, within cohorts, and across cohorts.

Figure 7, panels A, C, and E present the evolution of two measures of inequality for all workers, men, and women, respectively. The two measures are (i) the ratio between the earnings of the 90th and the 10th percentile (P90-P10) and (ii) 2.56 times the standard deviation of earnings.[12] Both measures in all three panels present the same qualitative picture: earnings inequality declined slowly in 2002-2008, increased substantially in 2009-2013, and declined in 2014-2023, ending up at a lower level in 2023 than in the pre-crisis period. These findings complement our second stylized fact.

[12] In a Gaussian distribution, these two measures are approximately equal.

Greece's inequality is relatively high compared to other European countries, as reported in the first GRID wave. The average P90-P10 ratio is considerably higher than in Scandinavian countries and France (where it ranges from 1.4 to 1.9) and slightly higher than Italy and Spain (1.9 and 2.1, respectively) while only Germany (at 2.4) has higher inequality, at par with the US. The counter-cyclicality of inequality is a consistent feature in other countries, though the magnitude of the changes in Greece are larger, likely due to the severity of the recession.

Figure 7, panels B, D, and F present the evolution of the ratio between the earnings of the 90th and the 50th percentiles (top inequality) and the ratio between the earnings of the 50th and the 10th percentiles (bottom inequality) for all workers, men, and women, respectively. The decomposition in panel B shows that bottom inequality increased earlier than top inequality and that it began to decline later in the recovery period. The larger decline in bottom inequality coincides with the significant increase in the minimum wage after 2019. Comparing panels D and F shows that the qualitative patterns for the evolution of top and bottom inequality were similar for women and men, except that inequality, and particularly bottom inequality, increased somewhat more for men than for women.

Panels A and B in figure 8 present the P90–P10 inequality measure for women and men within each of four selected birth cohorts — those entering the labor market in 2003, 2008, 2013 and 2018 — over their life cycle, with each line starting at age 25. Within-cohort inequality follows a similar pattern for all four cohorts, and most variation coincided with changes in aggregate economic conditions.

Focusing on the two cohorts that entered the labor market before the crisis (2003, 2008), within-cohort inequality increased gradually with age, a pattern that is consistent with other countries, and then rose very sharply during the crisis. The cohort that entered during the crisis (2013) experienced much higher inequality from the start, similar or higher to what the 2003 cohort experienced at age 34. The rise in within-cohort inequality even for very young groups is particularly striking because it suggests that even narrowly defined groups

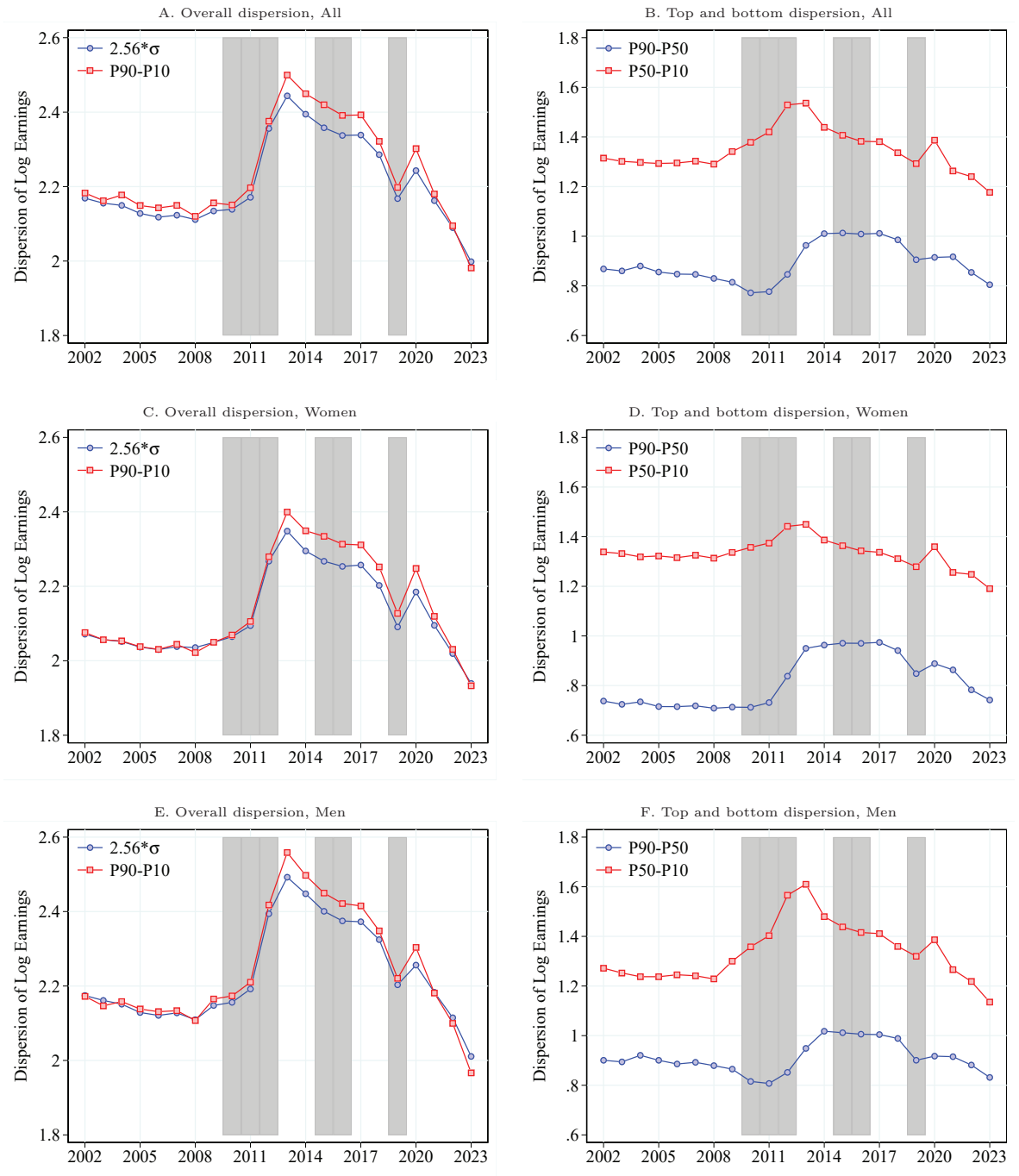


Figure 7: Evolution of inequality

Notes: Left panels present the overall (P90–P10) dispersion and the rescaled standard deviation (using a scaling factor of 2.56, in order to facilitate comparison between the two measures) based on residualized log earnings, separately for all respondents, women, and men. Right panels present top (P90–P50) and bottom (P50-P10) dispersion based on residualized log earnings, separately for all respondents, women, and men. The shaded areas indicate recession years.

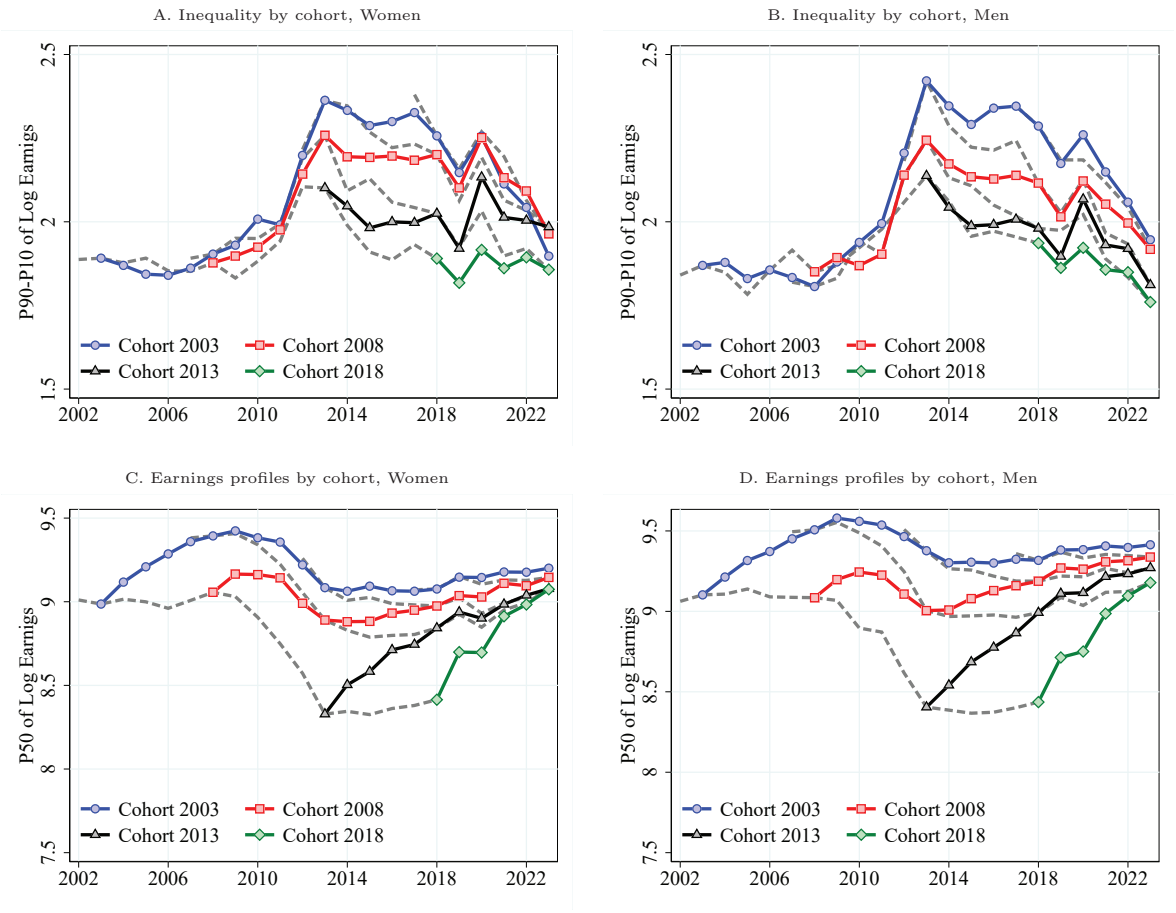


Figure 8: Inequality and earnings profiles by cohort

Notes: This figure shows the evolution of the P90-P10 differential and P50 of the log real annual earnings distribution over time by gender. Grey dashed lines correspond to earnings of 25, 30, and 35 year olds in each year. Each solid line corresponds to an individual cohort, where "cohort t" represents the cohort aged 25 in year t.

had heterogeneous exposure to the crisis. This is our fourth stylized fact.

In the following decade (2014-2023), within-cohort inequality fell, with a somewhat more pronounced decline for men. The 2018 cohort (the graph's final cohort) experienced considerably lower inequality in its early labor market years than the 2013 cohort did. Within-cohort inequality tends to rise as a cohort ages (Guvenen, 2009), so the decline observed since the mid-2010s across all cohorts is, also, noteworthy.

Panels C and D of Figure 8 show the median earnings profiles for different cohorts over time and describes inequality across cohorts. In the pre-crisis years, workers experienced

strong earnings growth in their early years of employment (e.g. cohort 2003) while the median earnings at age 25 (the gray dashed line) were relatively flat over time. During the crisis, earnings declined for all cohorts and the drop was substantially larger for 25-year-old workers when compared to older cohorts (e.g., cohort 2013 vs cohort 2008). After the crisis, the earnings at age 25 recovered, though at a slow pace. The cohorts that entered the labor market during the crisis experienced fast earnings growth after 2013, which partially offset the losses associated with their low starting point. However, the subsequent cohorts exhibited even stronger earnings growth (e.g., cohort 2018 vs cohort 2013). This pattern suggests that the "crisis cohorts" did not experience as complete a recovery as those entering after the recession. This forms our fifth stylized fact.

Motivated by this finding, Section 4 provides a detailed examination of how long-run earnings outcomes for each cohort are shaped by macroeconomic conditions at the time of labor market entry.

4 Unlucky cohorts in the Greek crisis

The evidence in Section 3 showed that workers who reached the age of 25 during the crisis experienced substantially weaker earnings outcomes than cohorts that entered before or after the recession. In this section, we expand our analysis to compare more systematically the paths of workers who entered the labor market under worse macroeconomic conditions with those who entered in (relatively) more benign conditions.

We create cohorts based on workers' entry year in the labor market (rather than age) to document the stark deterioration of the earnings profile of cohorts that began their careers during the crisis. We then proceed to document systematic differences in the earnings trajectories of workers who entered the labor market under different local labor market conditions, proxied by regional unemployment rates, and characterize how those differences evolve over the first nine years of workers' careers.

4.1 Data and Sample Construction

To examine the earnings trajectories of different cohorts, we construct a sample of workers entering the labor market between 2003 and 2014 who have exactly nine years of potential experience for each of Greece's 13 administrative regions. We also impute workers' level of education, as our dataset does not include information on education.

Our starting point is the annual worker panel described in Section 2. For the purposes of this analysis, we extend the age range to include workers aged 20–24, since many labor-market entrants first appear in the data before age 25.[13] We define a worker's entry year l as the first year in which the worker appears in the social security records and define potential experience e in calendar year t as the number of years since first appearance in the data, so that $e = t - l$. We further define each worker's entry region r as the region where the employer of the worker's entry year is located.

We restrict attention to workers who first appear in the data before age 35 and exclude workers already present in the initial year of the dataset (2002), since their true labor-market entry date is unobserved.[14] We exclude the earnings of the first observed year for each worker because initial observations may correspond to partial-year employment spells. We focus on workers observed for up to nine years after entry and, therefore, restrict the sample to cohorts entering between 2003 and 2014 and drop worker-years with 10 or more years of experience.

Table 2 presents descriptive statistics for the entry-cohort sample. The average entry age is approximately 24 years old and the sample is balanced between women and men. The number of entrants declines substantially during the peak crisis years, reflecting the sharp deterioration in labor-market conditions during the recession.

We construct an imputed measure of educational attainment using workers' occupational histories combined with information from the Labor Force Survey (LFS). We compute av-

[13] The dataset that we describe in Section 2 consists of workers between 25 and 54 years old. The data that we use in this section satisfies all the other selection criteria from Section 2, for example regarding attachment to the labor force and the selection of specific birth-dates.

[14] In a robustness exercise, available upon request, we also drop those who first appear in 2003 and 2004. The estimates are qualitatively similar.

Table 1: Descriptive Statistics for Labor Force Entrants

Variable	**Mean**	**Std. Dev.**
Annual Earnings (log)	8.923	0.802
Entry age (Years)	23.843	3.758
Age (Years)	28.984	4.634
Male	0.505	
Entry year		
2003	0.138	
2004	0.113	
2005	0.092	
2006	0.094	
2007	0.091	
2008	0.086	
2009	0.070	
2010	0.059	
2011	0.049	
2012	0.050	
2013	0.071	
2014	0.087	
Observations	1,514,027	

Source: EFKA.
Notes: The table reports descriptive statistics for entrants to the labor force in 2003-2014. Entry year denotes the first year a worker appears in the EFKA dataset.

erage years of schooling by 2-digit occupation in the LFS and assign to each worker the educational level associated with the occupation with the highest estimated schooling requirements observed during the worker's employment history. We then classify workers into three (roughly) equal-sized broad education groups based on this measure. This measure should be interpreted as a coarse proxy for relative educational attainment rather than a precise measure of years of schooling. Nevertheless, it allows us to examine broad differences in the earnings trajectories of workers with different levels of labor-market skills and occupational attainment.

4.2 Entry cohorts during the Greek crisis

We examine the earnings trajectories of cohorts entering the labor market between 2003 and 2014. Figure 9 plots median log real annual earnings by years of potential experience for each entry cohort.

Several patterns emerge clearly from the figure. First, cohorts entering before the crisis, the 2003–2008 cohorts, began their careers under relatively similar labor-market conditions and exhibited broadly similar earnings profiles during their early working years. Initial earnings levels were relatively stable across these cohorts and earnings increased steadily with experience.

Second, cohorts entering during the crisis experienced substantially lower initial earnings. After 2008, entry cohorts display progressively lower earnings at labor-market entry, reflecting the severe deterioration in labor-market conditions during the recession years. The decline in entry earnings is economically very large and mirrors the sharp rise in unemployment and decline in aggregate labor demand documented in Section 2.

Third, although recession-entry cohorts experienced earnings growth during the recovery period, their earnings trajectories remained persistently below those of earlier cohorts. Some convergence occurred over time as crisis-entry cohorts accumulated labor-market experience and macroeconomic conditions stabilized. Nevertheless, the initial losses associated with

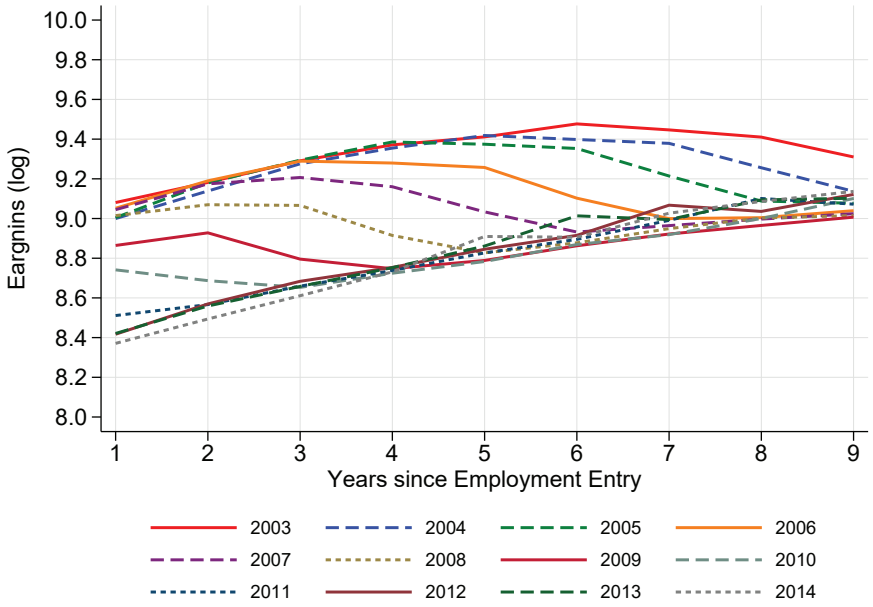


Figure 9: Median log earnings by experience

entering during the recession were only partially offset during the subsequent recovery.

These patterns are consistent with the broader literature documenting weaker early-career earnings outcomes among workers entering the labor market during recessions (Oreopoulos, von Wachter, and Heisz, 2012; Schwandt and von Wachter, 2019). The Greek case is particularly informative because of the unusually large and persistent deterioration in labor-market conditions during the crisis years.

4.3 Earnings trajectories by entry conditions: a regional decomposition

Our final step is to characterize the differences in earning profiles between workers who entered the labor market under different conditions. We use the unemployment rate in Greece's 13 administrative regions as a measure of labor market conditions at the time of entry.[15]. Figure 10 presents the evolution of regional unemployment rates over time together

[15]See Kahn (2010), Oreopoulos, von Wachter, and Heisz (2012), Altonji, Kahn, and Speer (2016), Schwandt and von Wachter (2019), and Rothstein (2023)

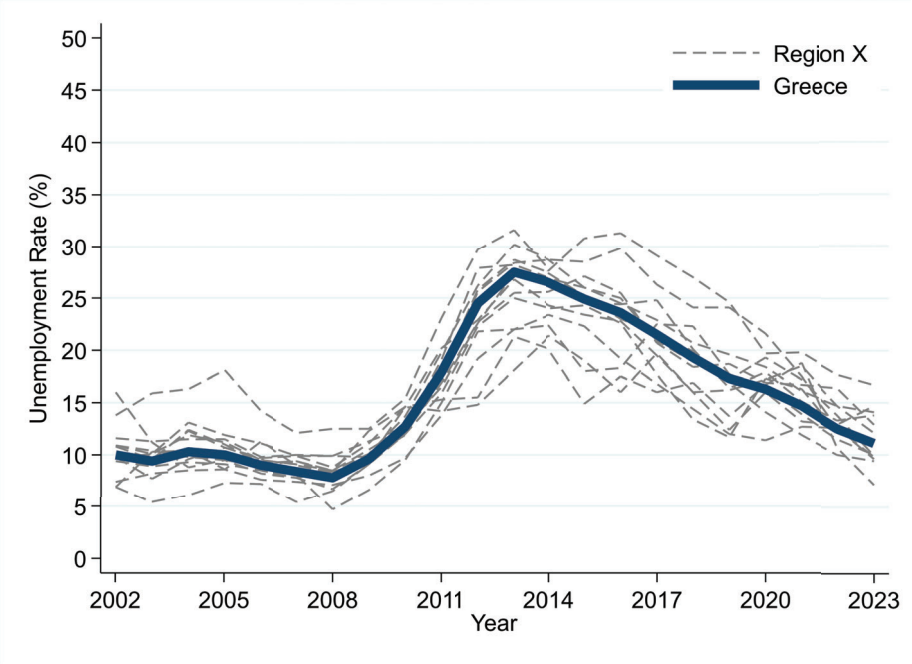


Figure 10: National and regional unemployment rates

with the national unemployment rate (the bold line). Regional labor markets co-move closely over the business cycle but there is substantial cross-regional variation, particularly during the crisis years.

To summarize the differences in earning profiles across cohorts, regions, and experience levels within a unified empirical framework, we estimate a regression that relates average cell-level earnings to the regional unemployment rate at entry, conditioning on region, entry year, calendar year, experience, gender, and education fixed effects.[16] Each cell aggregates individual observations by region of employment entry, year of employment entry, calendar year, gender, and education tercile. The estimated coefficients describe the conditional association between entry unemployment and subsequent earnings within this framework.

Several factors complicate a causal interpretation in this setting. First, labor-market entry timing may respond to macroeconomic conditions if individuals delay entry during recessions. Second, migration between regions may weaken the connection between local

[16]Using cell-level data is appropriate in our case because the analysis does not rely on individual-level control variables and aligns with the regional unemployment rate, which is our main source of variation.

unemployment rates and the opportunities faced by entrants and migration out of the country might introduce further selection concerns. Third, selection into employment is likely particularly important during a period of exceptionally high non-employment such as the Greek crisis. As crisis entrants are likely to be positively selected, the estimates probably underrate the true population-level association. Fourth, our education measure is based on occupational histories and provides only a coarse proxy for educational attainment. Furthermore, educational choices themselves respond to local labor market conditions during recessions (Adamopoulou and Tanzi, 2017; Sievertsen, 2016). For these reasons, the resulting estimates should be interpreted as descriptive evidence about the association between labor-market conditions at entry and subsequent earnings outcomes rather than as causal estimates.

We define $\bar{y}_{r,l,t,d}$ to be the cell's average log real annual earnings and we estimate our regression in the full sample, in two sub-samples that include only men or only women, and in the three sub-samples that include only the low, medium of high educational groups. The baseline specification is:

$$\bar{y}_{r,l,t,d} = \alpha + \beta_e u_{r,l} + \gamma_e + \lambda_r + \delta_l + \theta_t + \pi_d + \phi + \varepsilon_{r,l,t,d} \tag{1}$$

where $u_{r,l}$ is the unemployment rate in region r during the year of employment entry l and λ_r is an indicator variable for entry region r, δ_l for entry year l, θ_t for calendar year t, π_d for educational group d, ϕ for female, and γ_e for potential experience e. The indicator variables absorb persistent spatial, temporal, and compositional heterogeneity.[17] Cell-level observations are weighted by size, and standard errors are clustered by region and year of employment entry to allow for correlation among individuals exposed to similar local labor-market conditions at the time of entry. Given the fixed effects that are included, the coefficient vector β_e captures deviations from the typical experience profiles that correlate

[17]The female fixed effect and the educational group fixed effects are dropped in the regressions that use only one gender and only one educational group, respectively.

with region-specific variation in unemployment rates at labor-market entry.[18]

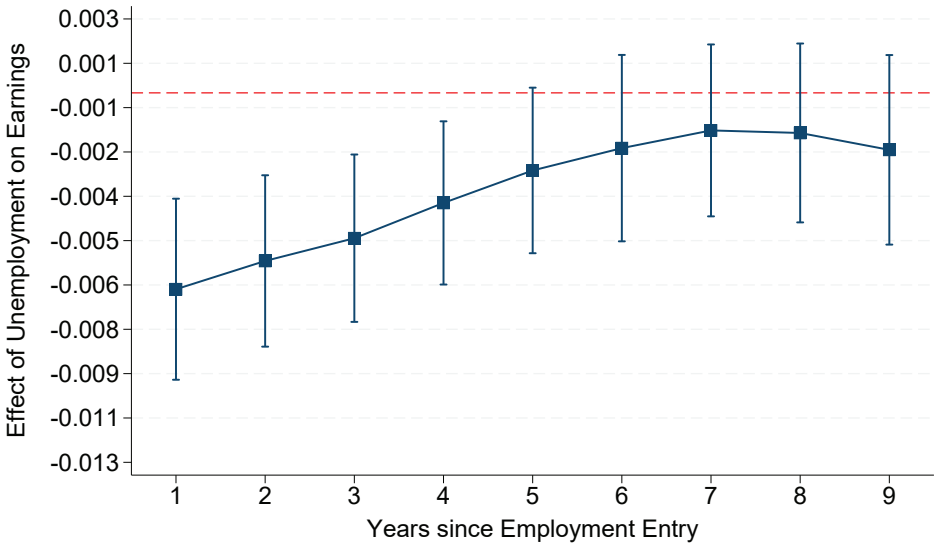


Figure 11: Association of regional unemployment rate at entry on log annual earnings with 95% confidence intervals

Notes: Coefficient estimates and 95% confidence intervals. The dependent variable is the logarithm of average annual earnings at the cell level. The reported coefficients correspond to interactions between the regional unemployment rate and potential experience groups. The regressions include region of entry, entry year, calendar year, potential experience, gender, and educational group fixed effects. Standard errors are clustered at the regional and entry-year levels and are reported in parentheses.

Figure 11 presents estimates of the association between unemployment rates at labor-market entry and subsequent annual earnings by years of potential experience. Table 2 summarizes these estimates for broader experience intervals and for different demographic groups. The estimates show a clear negative relationship between unemployment rates at entry and subsequent earnings during the first years of workers' careers. Workers entering the labor market in regions and years with higher unemployment rates tended to experience substantially lower earnings during their early working years. The magnitude of this relationship gradually fades away with labor-market experience and becomes insignificant 6-7 years after entry.

[18]We do not include the contemporaneous regional unemployment rate in our specification; consequently, β_e captures how earnings profiles vary with labor-market conditions at the time of entry, given the regular subsequent evolution of local labor-market conditions.

Column 1 in Table 2 reports the estimates for the full sample in three three-year groups of potential experience. Taken together, the estimates suggest that workers entering employment under weaker labor-market conditions experienced substantially lower cumulative earnings during the early stages (first six years) of their careers.

Columns 2 and 3 show that the relationship is broadly similar for women and men. In contrast, the differences between education groups in columns 4-6 are more pronounced. The conditional association between entry unemployment and early-career earnings is roughly twice as large for workers in the lowest education tercile as for those in the medium and highest terciles, consistent with greater cyclical exposure among lower-skilled workers, as in Schwandt and von Wachter (2019).

Table 2: The Association of Regional Unemployment Rate at Entry on Annual Earnings by Years of Experience

	(1) Full sample	(2) Men	(3) Women	(4) Low Ed.	(5) Medium Ed.	(6) High Ed.
Years 1–3	-0.0066^{***}	-0.0064^{***}	-0.0068^{***}	-0.0083^{***}	-0.0063^{***}	-0.0058^{**}
	(0.0015)	(0.0019)	(0.0016)	(0.0020)	(0.0019)	(0.0022)
Years 4–6	-0.0030^{**}	-0.0031^{*}	-0.0030^{*}	-0.0043^{**}	-0.0031	-0.0023
	(0.0014)	(0.0018)	(0.0016)	(0.0019)	(0.0019)	(0.0021)
Years 7–9	-0.0002	0.0001	-0.0007	-0.0012	-0.0007	0.0010
	(0.0014)	(0.0018)	(0.0017)	(0.0023)	(0.0020)	(0.0021)
Observations	8,424	4,212	4,212	2,808	2,808	2,808

Source: Own estimations with microdata from EFKA and LFS.
Notes: The dependent variable is the logarithm of average annual earnings at the cell level. The reported coefficients correspond to interactions between the regional unemployment rate and potential experience groups. The regressions include region of entry, entry year, calendar year, potential experience, gender, and educational group fixed effects. Standard errors are clustered at the regional and entry-year levels and are reported in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Overall, the results in this section complement the evidence presented in Section 3. Together, the findings suggest that entrants to the labor market were very negatively affected by the Greek crisis, particularly workers with lower levels of education. More broadly, the results illustrate how large macroeconomic shocks can shape earnings trajectories over extended periods of workers' careers.

5 Conclusions

This paper provides new evidence on the evolution of labor earnings and earnings dynamics in Greece during 2002–2023 using administrative matched employer–employee data from the Greek social security system. The analysis forms part of the Global Repository of Income Dynamics (GRID) project and contributes harmonized evidence on earnings inequality, earnings volatility, earnings mobility, and cohort dynamics for a country that experienced one of the deepest macroeconomic crises observed in an advanced economy in the postwar period.

Several broad patterns emerge from the analysis. First, labor earnings closely tracked macroeconomic developments over the period. Earnings increased during the expansion of the 2000s, declined sharply during the crisis years 2009–2013, and partially recovered during the subsequent decade. The decline in earnings during the crisis was exceptionally large by international standards and particularly severe for workers at the bottom of the earnings distribution.

Second, the crisis was associated not only with lower earnings levels but also with substantial changes in earnings dynamics. Earnings volatility and downside earnings risk increased sharply during the recession, while inequality rose both in the cross-section and within cohorts. These patterns gradually moderated during the recovery period, though several measures of earnings remained far from their pre-crisis levels even by the end of the sample.

Third, the analysis highlights substantial differences across cohorts. Workers entering the labor market during the crisis experienced unusually weak initial earnings and persistently lower earnings trajectories relative to cohorts entering before or after the recession. The application in Section 4 further shows that cohorts entering employment under weaker local labor-market conditions tended to experience systematically lower earnings during the early stages of their careers, particularly among lower-education workers. These differences narrowed considerably over the first decade of workers' careers, suggesting that the crisis-era disadvantage was persistent but not permanent for most workers.

More broadly, the Greek case illustrates how deep and prolonged macroeconomic disruptions can affect multiple dimensions of labor-market outcomes simultaneously, including earnings levels, inequality, volatility, and cohort dynamics. The richness and longitudinal structure of the administrative data make it possible to document these dimensions jointly and in a harmonized framework comparable to that used in other countries participating in the GRID project.

The paper also points to several directions for future research. The matched employer–employee structure of the data could be used to examine the role of firms, sectors, and worker mobility in shaping earnings dynamics during the Greek crisis. In addition, further work could study employment transitions, labor-force attachment, firm heterogeneity, and the interaction between institutional reforms and earnings adjustment during the recession and recovery periods. More generally, the data provide a useful foundation for future research on labor-market dynamics in Greece and for comparative work on earnings dynamics across countries.

References

Adamopoulou, E., and G. M. Tanzi (2017): "Academic drop-out and the great recession," Journal of Human Capital, 11(1), 35–71.

Altonji, J. G., L. B. Kahn, and J. D. Speer (2016): "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success," Journal of Labor Economics, 34(S1), S361–S401.

Arellano, M., S. Bonhomme, M. De Vera, L. Hospido, and S. Wei (2022): "Income risk inequality: Evidence from Spanish administrative records," Quantitative Economics, 13(4), 1747–1801.

Arkolakis, C., A. Doxiadis, and M. Galenianos (2017): "The Challenge of Trade Adjustment in Greece," in Beyond Austerity: Reforming the Greek Economy, ed. by C. Meghir, C. Pissarides, D. Vayanos, and N. Vettas, pp. 103–135. MIT Press, Cambridge, MA.

Chodorow-Reich, G., L. Karabarbounis, and R. Kekre (2023): "The Macroeconomics of the Greek Depression," American Economic Review, 113(9), 2411–2457.

Drechsel-Grau, M., A. Peichl, K. D. Schmid, J. F. Schmieder, H. Walz, and S. Wolter (2022): "Inequality and income dynamics in Germany," Quantitative Economics, 13(4), 1593–1635.

Galenianos, M. (2015): "The Greek Crisis: Origins and Implications," Crisis Observatory Research Paper 16, Hellenic Foundation for European and Foreign Policy (ELIAMEP), Athens.

Gatopoulos, G., A. Louka, I. Polycarpou, and N. Vettas (2026): "Policy lessons from labor market reforms during crisis: Success and failure of the Greek bailout programs," Journal of Policy Modeling.

Giannakopoulos, N., and I. Laliotis (2020): "Industrial Relations Reform, Firm-Level Bargaining and Nominal Wage Floors," The Manchester School, 88(1), 37–59.

Giannoulakis, S., P. Sakellaris, and J. Tsoukalas (2025): "Economic Growth Dynamics and Productivity Trends in the Greek Economy: A 60-Year Perspective," Working paper, Hellenic Parliament Budget Office.

Gourinchas, P.-O., T. Philippon, and D. Vayanos (2017): "The Analytics of the Greek Crisis," NBER Macroeconomics Annual, 31(1), 1–81.

Guvenen, F. (2009): "An Empirical Investigation of Labor Income Processes," Review of Economic Dynamics, 12(1), 58–79.

Guvenen, F., L. Pistaferri, and G. L. Violante (2022): "Global trends in income inequality and income dynamics: New insights from GRID," Quantitative Economics, 13(4), 1321–1360.

Hoffmann, E. B., D. Malacrino, and L. Pistaferri (2022): "Earnings dynamics and labor market reforms: The Italian case," Quantitative Economics, 13(4), 1637–1667.

Kahn, L. B. (2010): "The long-term labor market consequences of graduating from college in a bad economy," Labour Economics, 17(2), 303–316.

Kanellopoulos, C. N. (2015): "The Effects of Minimum Wages on Wages and Employment," Economic Bulletin, (41), 7–41.

Koukiadaki, A., and C. Kokkinou (2016): "The Greek System of Collective Bargaining in (the) Crisis," in Joint Regulation and Labour Market Policy in Europe during the Crisis, ed. by I. Távora, and M. Martinez Lucio. European Trade Union Institute, Brussels.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012): "The Short- and Long-Term Career Effects of Graduating in a Recession," American Economic Journal: Applied Economics, 4(1), 1–29.

Rothstein, J. (2023): "The Lost Generation? Labor Market Outcomes for Post-Great Recession Entrants," Journal of Human Resources, 58(5), 1452–1479.

Saez, E., M. Matsaganis, and P. Tsakloglou (2012): "Earnings Determination and Taxes: Evidence From a Cohort-Based Payroll Tax Reform in Greece," Quarterly Journal of Economics, 127(1), 493–533.

Schwandt, H., and T. von Wachter (2019): "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets," Journal of Labor Economics, 37(S1), S161–S198.

Sievertsen, H. H. (2016): "Local unemployment and the timing of post-secondary schooling," Economics of Education Review, 50, 17–28.

A Cleaning and harmonizing the data

This section describes how we construct a monthly panel from the raw data. The dataset is monthly and it is organized around social security payments made by employees and employers. Each data entry consists of the month and year, worker characteristics, firm characteristics, and employment relationship characteristics. Worker characteristics are the unique worker ID, date of birth, gender, citizenship, and whether the worker was first insured after 1993 (this characteristic determines some benefits). Firm characteristics are the unique firm ID, unique establishment ID, 8-digit establishment sector code, the establishment's NUTS-3 geographic location,[1] and the identity of the establishment's local IKA/EFKA branch.[2] Employment relationship characteristics are compensation type,[3] insurance package code,[4] number of days of insurance for full-time work, number of days of insurance for part-time work, 6-digit occupation code, worker's gross compensation, social security contributions of employer, and social security contributions of employee. Gross compensation is not top-coded.

Our next step is to clean and harmonize the data. When the same worker ID has multiple dates of birth, we choose the modal date of birth. We harmonize the sectoral codes over the full 2002-2023 period to 4-digit NACE 2.0 codes.

We combine the compensation types into 7 categories: base pay (which is the main category for the compensation of contemporaneous employment), overtime pay, bonus pay, back

[1] There are 74 NUTS 3 geographical units in Greece, largely corresponding to the 52 prefectures, with a further break-down of the most populous prefectures such as Attica and Thessaloniki.

[2] There are approximately 300 separate IKA/EFKA branches in Greece. Knowing an establishment's branch code allows for a more precise identification of its location than its NUTS-3 location.

[3] This refers to the type of payment, such as base pay, bonus, holiday pay, etc. There are approximately 150 compensation type codes and the large number is due to the multitude of ad hoc codes for specific worker groups, e.g. "base pay for crew of tourism boats" and so on.

[4] The insurance package determines the contribution rates for employees and employers and the benefits that are associated with that insurance. There are more than 1,000 distinct insurance packages which differ in the type of insurance (e.g. pension, health), the type of job ("hazardous" or not), the sector of employment, etc. The large number of codes is mostly due to the proliferation of ad hoc packages for highly specific occupations, e.g. "editorial and administrative staff of radio stations, daily newspapers of Athens and Thessaloniki with circulation below 3,000 units" and so on. The vast majority of employees are insured under a small number of main insurance packages.

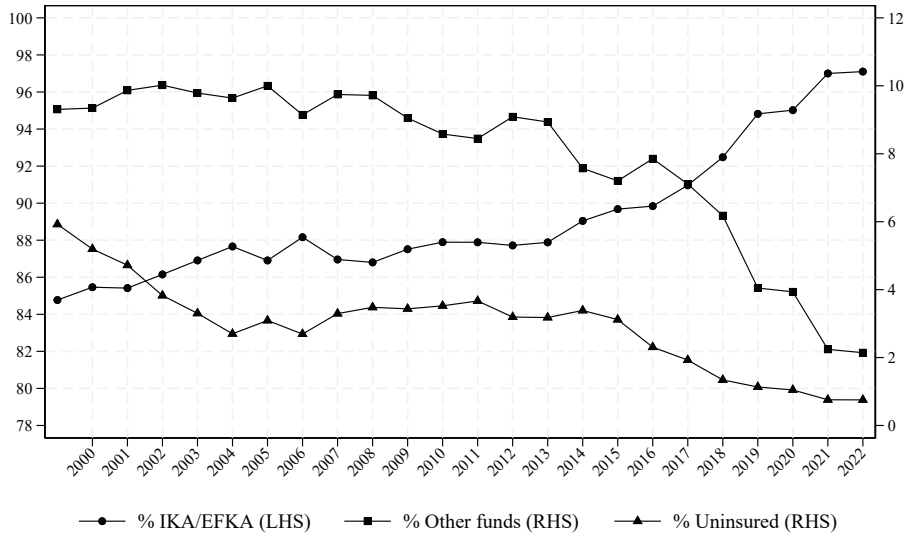
pay, holiday allowance, sick leave allowance, and other payments.[5] We merge entries with the same month-year, worker ID, firm ID, establishment ID, and compensation type that corresponds to "base pay" as follows: we take the maximum gross compensation across entries, we take the modal days of insurance across entries (typically all entries have the same number of days) and we sum the values of the entries for employer and employee contributions. We merge entries with the same month-year, worker ID, firm ID, establishment ID and compensation type that corresponds to the 6 other categories in the same way, except that we sum the values for gross compensation of the entries. For entries of all compensation types, we use the occupation and insurance package of the entry that corresponds to base pay for that month. Finally, if a worker ID has no entries with compensation type "base pay" in a calendar year, then that worker did not engage in paid work and we drop all entries of that worker ID for that calendar year.

Entries with the same month-year, worker ID, firm ID, establishment ID, occupation code, insurance package, payment type and full-time/part-time status refer to separate days of employment for the same job. We merge these entries and sum the days of insurance, gross compensation, employer contributions and employee contributions.

Entries with the same month-year, worker ID, firm ID, establishment ID refer to different types of social security contribution payments associated with the same job and the same gross compensation. We merge the entries with compensation type "base pay" as follows: we take the maximum gross compensation across entries; we take the modal days of insurance across entries (typically, all entries have the same number of days); for employer and employee contributions, we sum the values of the entries. We merge the entries with other compensation types in the same way, except that we sum the values for gross compensation of the entries. For entries of all compensation types, we use the occupation and insurance package of the entry that corresponds to base pay for that month.

[5]Entries where the compensation type is "salary offset" have negative gross compensation and workers with such entries are dropped in the calendar year when these entries appear. We drop entries with compensation type "contributions without pay."

Figure B.1: Insurance fund coverage in the Business Economy



Source: Greek Labor Force Survey, Hellenic Statistical Authority (ELSTAT).

The outcome of the steps described above is a dataset where each worker-establishment pair has one entry per month, compensation type, and work type (full time or part time).

B Fund mergers

According to the Labor Force Survey, up to 10% of private-sector wage employees self-report to have been insured by funds other than IKA in the 2000s. These workers were mostly employed at banks and utilities and their funds were merged with IKA (before 2016) and EFKA (after 2017). Figure B.1 presents the evolution of the share of employees in the economy's business sector who were insured by IKA/EFKA, insured by a different fund or were uninsured.

To avoid administrative changes in the social insurance system from impacting our panel, we drop drop workers who enter our dataset due to a fund merger. To identify such workers, we create a variable for the first month when a worker ID is observed in the data (this happens in 2002 for 70% of our worker IDs). For each 4-digit sector, we identify months

with an unusually large number of new worker IDs, particularly for workers who are unlikely to be new entrants to the labor force on account of their age. We cross-check these sectors and months with the Government Gazette to confirm that they correspond to publicized fund mergers. When we find a match, we remove from our data all workers who appear in the dataset for the first time on the specific month and at the specific 4-digit sector. The outcome is a balanced-coverage panel.

C Comparison with other data sources

This section compares the data of IKA/EFKA on employment, firms, and inequality with similar data from four other sources: the Labor Force Survey, the National Accounts, ERGANI, and the World Inequality Database.

ERGANI is an electronic platform established by the Ministry of Labor in 2013 to record information about employment relationships under private-law contracts.[1] For this purpose, ERGANI conducts an annual census of firms with private-law employees every October and records the flows (hires and separations) of private-law employees every month. ERGANI covers, essentially, the same population as EFKA and, hence, provides a very useful point of comparison. We use aggregate data from ERGANI's annual census.

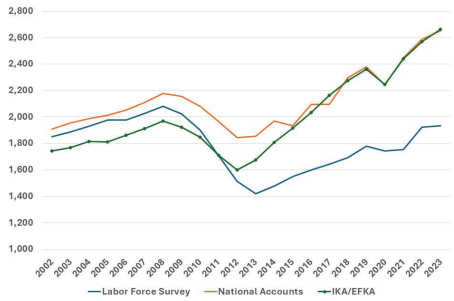
Employment: We compare the series on aggregate employment from IKA/EFKA with those from the Labor Force Survey, the National Accounts and ERGANI.

Panel A in figure C.1 plots the number of employees in IKA/EFKA, the Labor Force Survey, and the National Accounts for all sectors except A (Agriculture), O (Public Administration), P (Education), and Q Health). Most workers in the excluded sectors are not private-law employees and are, hence, not included in the IKA/EFKA data. Specifically, most agricultural workers are self-employed and sectors OPQ have high shares of public-

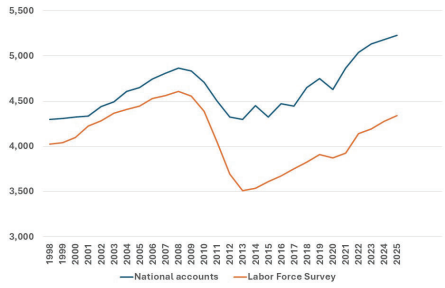
[1] Almost all private-law contracts are for private-sector employees, although some temporary public-sector jobs might also be under private law. The vast majority of employees of the General Government are employed under public-law contracts, e.g. civil servants, public school teachers, police officers, etc.

Figure C.1: Employment in IKA/EFKA, National Accounts, Labor Force Survey (in 1,000s)

A: Employees, except AOPQ



B: Total employment, all sectors



Source: IKA/EFKA, Eurostat. Most workers in Sector A (Agriculture) are self-employed and most workers in sectors O (Public Administration), P (education), and Q (Health) are public-sector employees under public-law contracts

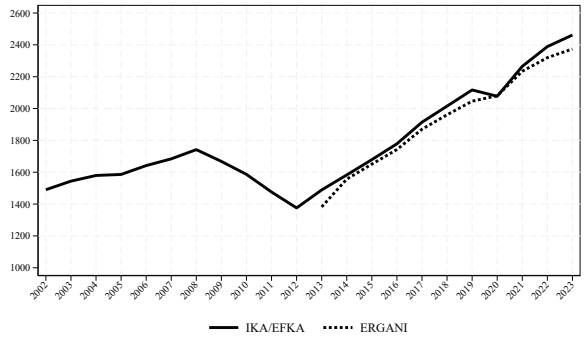
sector workers under public-law contracts.

The IKA/EFKA series lines up very well with the National Accounts series. In 2002-2014, the IKA series has 10% fewer employees, on average, than the National Accounts series, and that difference fluctuates very little from year to year. After 2015 (when independent insurance funds are gradually merged with IKA) the two series become essentially identical. The series for the Labor force Survey moves similarly to the series from IKA and the National Accounts in the 2000s, with consistently 5% fewer employees than the National Accounts. In the 2010s, however, there is significant divergence: the number of employees in the LFS is 22% lower than in the NA on average and, more worryingly, that difference fluctuates a lot from year to year, e.g. from 9% in 2010 to 18% in 2012 to 25% 2014, to 20% in 2015.

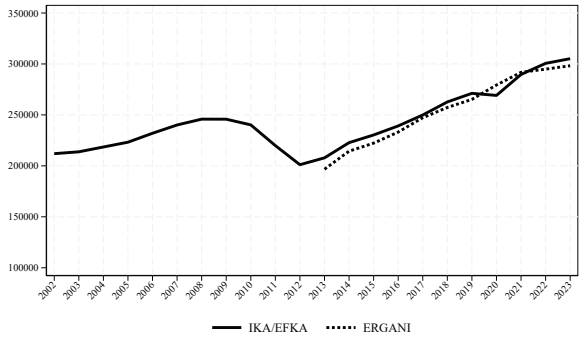
The divergence between the employment figures of the NA and LFS series is not confined on salaried employment or the selected sectors. Panel B in figure C.1 plots total employment (including the employed in sectors AOPQ and self-employed in all sectors) from NA and LFS and a similar picture emerges: the two series co-move quite closely until 2010 and then diverge significantly, where the magnitude of their divergence fluctuates considerably over time. The culprit is likely to be the large reduction in the response rate of the Labor Force

Figure C.2: Employment and Firms in IKA/EFKA and ERGANI

A: Employees in October, all sectors (in 1,000s)



B: Number of firms



Source: IKA/EFKA, ERGANI.

Surveys, a phenomenon that is increasingly visible across many countries and leads to less reliable estimates from such surveys. For this reason, we focus our comparison with sources other than the Labor Force Survey.

Panel A in Figure C.2 plots the number of employees in IKA/EFKA in October and the number of employees in ERGANI's annual census, which takes place in October of each year.[2] ERGANI covers the same population (private-law employees) as IKA/EFKA and, so, we keep all sectors, including those with a minority of employees with private-law contracts (sectors AOPQ). The two series are essentially identical. This is particularly reassuring because they are two independent administrative datasets that measure the same population, employees under private-law contracts.

Number of firms: ERGANI's annual census reports the number of firms that employ workers under private-law contracts. This can be compared with the number of firm IDs that are observed in the IKA/EFKA dataset. Panel B in Figure C.2 plots the two time series and shows that they are, essentially, identical.

[2]Note that Greek employment is highly seasonal due to the large tourism sector and the number of employees in the off-season month of October is considerably lower than for the full year.

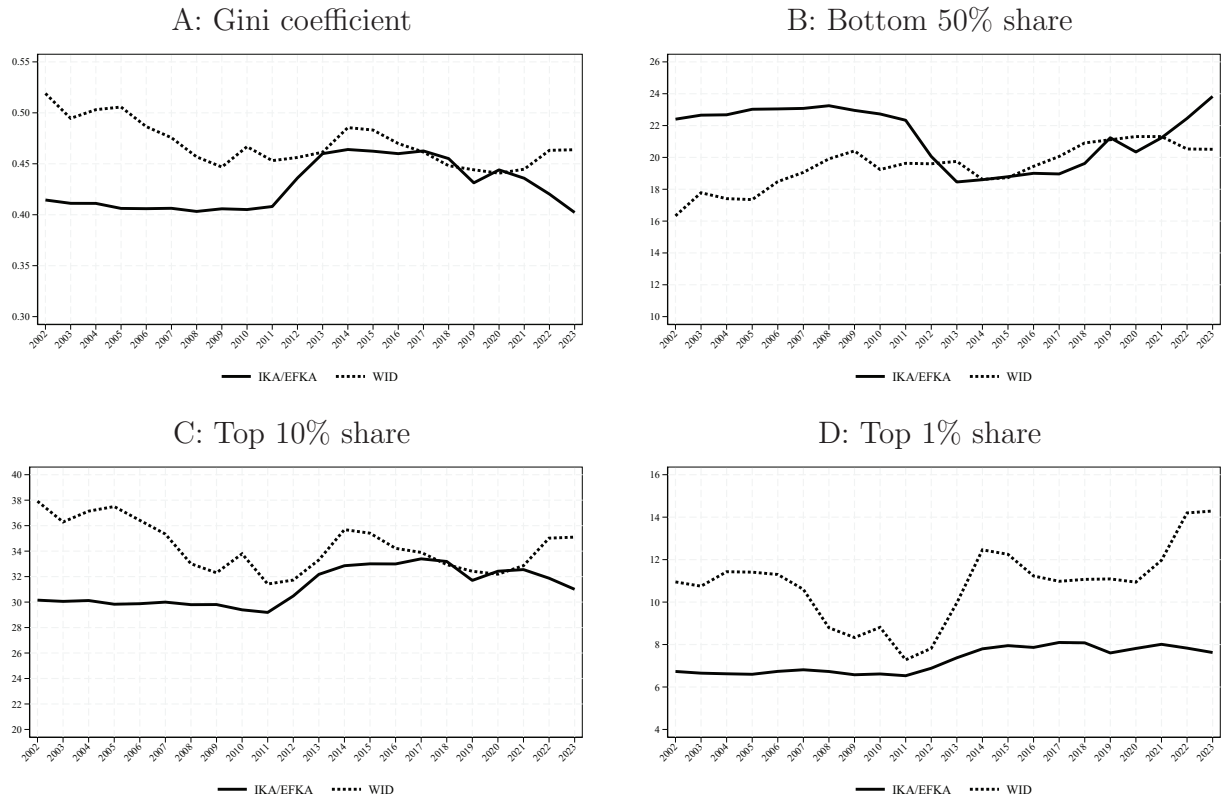
Inequality: We compare various measures of income inequality from IKA/EFKA with data from surveys, accessed from the World Inequality Database (WID). It is important to note that the population is different in the two data sources: IKA/EFKA has data on labor income for (most) private-sector workers, while WID has data on income from any source for the full population. Therefore, there is no a priori reason to expect a close match between the inequality measures computed using these two sources. Nevertheless, it is still useful to observe whether there are large deviations between them.

Panel A in figure C.3 presents the Gini coefficient. Broadly speaking, the data appears reasonably close. The general magnitudes are similar, both series show a gradual decline in inequality in the 2000s, though quantitatively the drop is much more pronounced in the WID data. During the early 2010s, both series show a large increase in inequality, followed by a gradual decline. The data deviates somewhat after 2022, when the WID measures are based on extrapolations.

Panel B in figure C.3 presents the income shares of the bottom 50% of earners. These shares are quite similar: they are somewhat stable in the 2000s, they decline during the crisis (more pronounced in the case of IKA/EFKA) and they rise after the crisis bottoms out in 2013. Panel C presents the top 10% shares which are also somewhat similar. Panel D presents the top 1% which are quite different: the IKA/EFKA share is quite stable, increasing mildly in the early 2010s, while the WID share fluctuates a lot more and rises significantly between the early 2010s and 2023. This deviation is probably to be expected since labor earnings are a lower share of total income for the top 1%. Overall, there is no great discrepancy between the two series.

Summary: Broadly speaking, aggregating the IKA/EFKA micro-data provides time series that are very close to alternative data sources.

Figure C.3: Measure of inequality in IKA/EFKA and World Inequality Database



Source: IKA/EFKA and World Inequality Database

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de