Supplementary Material

Supplement to "Earnings dynamics and labor market reforms: The Italian case"

(*Quantitative Economics*, Vol. 13, No. 4, November 2022, 1637–1667)

Eran B. Hoffmann
Department of Economics, Hebrew University of Jerusalem

Davide Malacrino
International Monetary Fund

Luigi Pistaferri
Department of Economics, Stanford University

Appendix

A.1 *Other forms of employment: Public sector and self-employment*

One of the limitations of the INPS data is that it does not cover public sector workers or the self-employed. In this section, we use the Survey of Household Income and Wealth (SHIW) to explore these other forms of employment and to gauge their potential impact were our statistics derived on the full population of workers. The SHIW is conducted by the Bank of Italy and is a representative cross-section of households and workers. Since 1989, SHIW follows a rotating sample of panel households. These households are sampled once every 2 years.

To complement our analysis, we use three main features of the survey data. First, the survey identifies self-employment and self-employment income. The data further indicate whether the worker is a sole proprietor, a free-lance worker, or a "member of the professions." Second, the data include a variable for the industry of the worker. These are only broad industry but they indicate whether the worker is a private or a public employee. Third, for all three categories of workers there is a variable measuring the self-reported status of the worker. Fourth, there is a variable on contract type and whether work is supplied on a full-time or part-time basis.

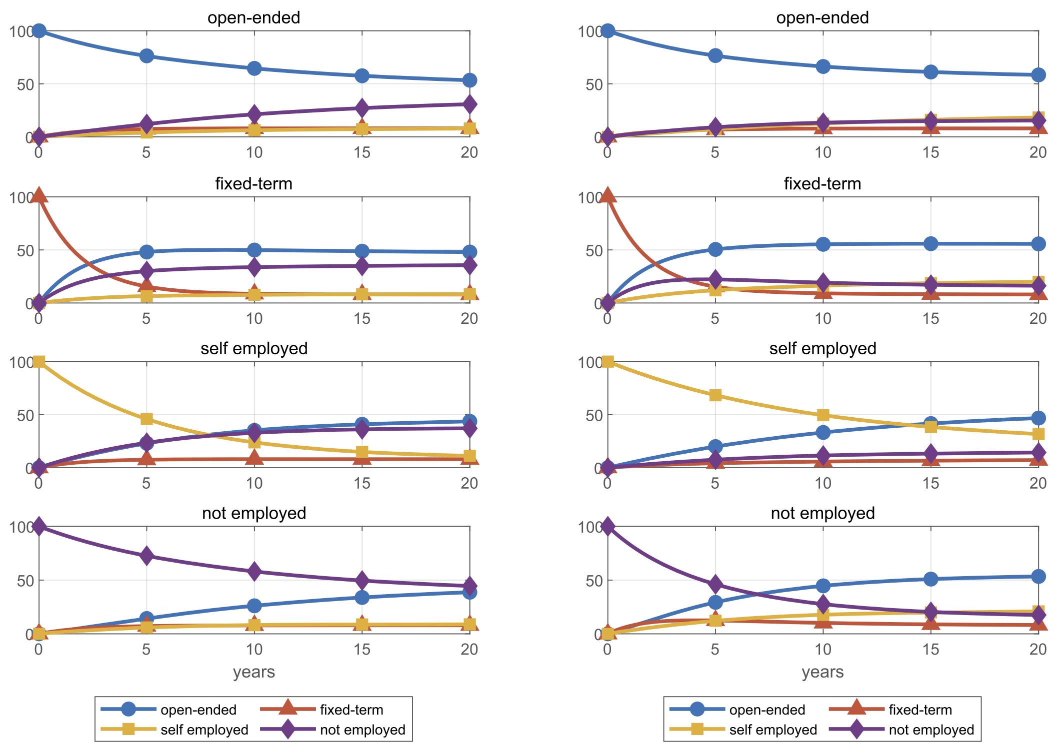
The first question is how often do workers transition between four labor market states: employment with a permanent contract, employment with a fixed-term contract, self-employment, and nonemployment. We measure the 2-year probability of observing a worker in each one of these states switching to any other states. We estimate the probabilities on workers 25–55 years old, and separately for men and women but pooling over all the years since 2000. Using these estimates, we find the unique "generator

Eran B. Hoffmann: eran.hoffmann@mail.huji.ac.il
Davide Malacrino: dmalacrino@imf.org
Luigi Pistaferri: pista@stanford.edu

Available at http://qeconomics.org. https://doi.org/10.3982/QE1865

matrix" that matches the observed transition probabilities. A generator matrix is a matrix in which each row includes the Poisson arrival rate of jumps from the row state to the column state, and has negative values on the diagonal so that the rows add up to one. This matrix defines a continuous time Markov chain and allows us to make predictions on the conditional state distribution of workers at any horizon given initial conditions.

Figure A.1 shows the conditional probability of being in each labor market state at every horizon up to 20 years, for women and men separately (panel (a) for women, panel (b) for men). The figures reveal several facts. One is that a fixed-term contract status is more likely to lead to nonemployment. For both women and men, the probability of a fixed-term contract holder to have an open-ended contract within 5 years is 50%, and to have a fixed-term contract 15%. But women with fixed-term contract have a 30% probability of being nonemployed within 5 years, compared to 22% for men. The differ-



(a) women (b) men

FIGURE A.1. Transition probabilities by horizon. *Notes*: Each graph shows the probability of a worker in a certain state (open-ended, fixed-term, self-employed, not employed), which is specified in the title, to be in each of the states after a given number of years. Probabilities calculated using a continuous-time Markov chain with a fixed transition matrix. The transition matrix is estimated using 2-year changes in work status for 25–55-year-olds in the SHIW pooled sample, 2000–2016.

ence is captured by men's higher probability of being self- employed. Self-employment is considerably less persistent status for women. The probability of being self-employed after 5 years for a self-employed woman is only 50% compared to 75% for men. Similarly, nonemployment status is more likely to persist for women: 75% of nonemployed women are nonemployed 5 years later compared to only 50% for men. We conclude the following: workers in open-ended jobs, both men and women, are not likely to transition to a fixed-term contract, while workers in fixed-term employment often end up in an open-ended contract. Furthermore, the "temporary" contract seems to serve its purpose better for men than for women. Lastly, transitions into self-employment are rare. Self-employed men are likely to transition directly into a permanent contract, while self-employed women are equally likely to be employed in a fixed-term contract which suggests that self-employment serves a different purpose for men and women.

The second question is: how often do workers transition between private and public sector jobs? We apply the same method of measuring transition probabilities and extrapolating under the continuous-time Markov chain assumption, but this time to employment in the private sector, the public sector, self-employment, and nonemployment. Figure A.2 shows the estimated probabilities. Both women and men employed in the privates sector have a low probability of transitioning to the public sector. The probability of a worker employed in the private sector to be employed in the public sector after 20 years is 21% for women and 16% for men. Transitions in the opposite direction are much more common: 22% of women and 24% of men in public sector jobs transition to private sector jobs after 5 years, and this number grows to 30% and 44% after 20 years. This is likely due to the gradual reduction in the size of the Italian public sector over the sample period.

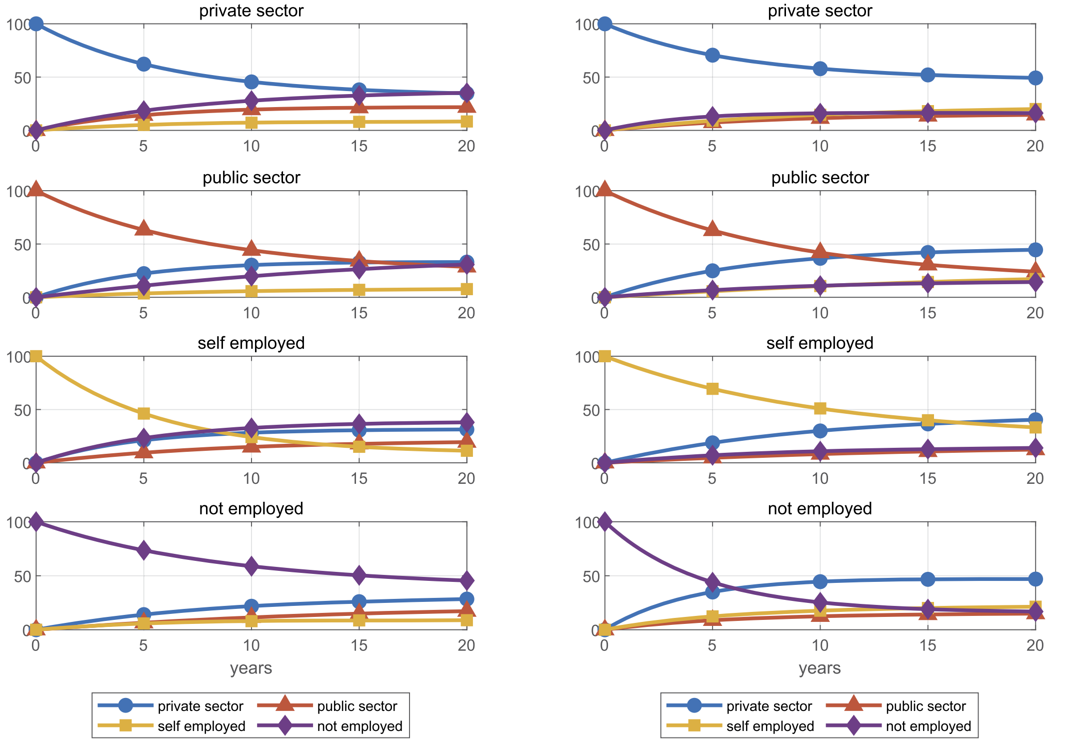
A.2 *Top-income share estimation in right-censored data*

This section describes the procedure of estimating top-income shares in top-coded data and implements it to the INPS data. The basic idea is the following. Let $F(y)$ be the true cumulative distribution function of earnings. A distribution $F(y)$ has a Pareto tail if there is a value $\underline{y}$ such that for all earnings y, y' such that $y \geq y' \geq \underline{y}$, the earnings are Pareto distributed,

$$1 - F(y) = \left(1 - F(y')\right)\left(\frac{y}{y'}\right)^{-\alpha}.$$

The parameter $\alpha > 0$ is called the Pareto-tail index. We assume that there exist a value $y^* \geq \underline{y}$, which is below the top-coding threshold, and thus the probability $P^* = F(y^*)$ can be estimated as the share of observation smaller than y^*, and the conditional mean income of workers $E[y|y < y^*]$, can be reliably estimated as a simple mean. Inverting the expression for the tail distribution, we can define the quantile function for all probabilities $P > F(y^*)$ as

$$Q(P) = y^*\left(\frac{1 - P}{1 - F(y^*)}\right)^{-\frac{1}{\alpha}}.$$



(a) women (b) men

FIGURE A.2. Transition probabilities between sectors by horizon. *Notes*: Each graph shows the probability of a worker in a certain state (private sector, public sector, self-employed, not employed), which is specified in the title, to be in each of the states after a given number of years. Probabilities calculated using a continuous-time Markov chain with a fixed transition matrix. The transition matrix is estimated using 2-year changes in work status for 25–55-year-olds in the SHIW pooled sample, 2000–2016.

Furthermore, if $\alpha > 1$, the mean income above the $P > P^*$ quantile is

$$E\big[y|y > Q(P)\big] = \frac{\alpha}{\alpha - 1} Q(P).$$

The share of earnings that goes to the top $(1 - P)$ quantiles can then be expressed as

$$\begin{aligned} S_{(1-P)} &= \frac{(1-P)E\big[y|y > Q(P)\big]}{P^* E\big[y|y \le y^*\big] + (1-P^*)E\big[y|y > y^*\big]} \\ &= \frac{\alpha(1-P)y^*}{(\alpha-1)P^* E\big[y|y \le y^*\big] + \alpha(1-P^*)y^*} \left(\frac{1-P}{1-P^*}\right)^{-\frac{1}{\alpha}}. \end{aligned} \tag{A.1}$$

Equation (A.1) provides a closed-form expression for the top-income share for any choice of quantiles above y^* given an estimate of the Pareto parameter α, and the es-

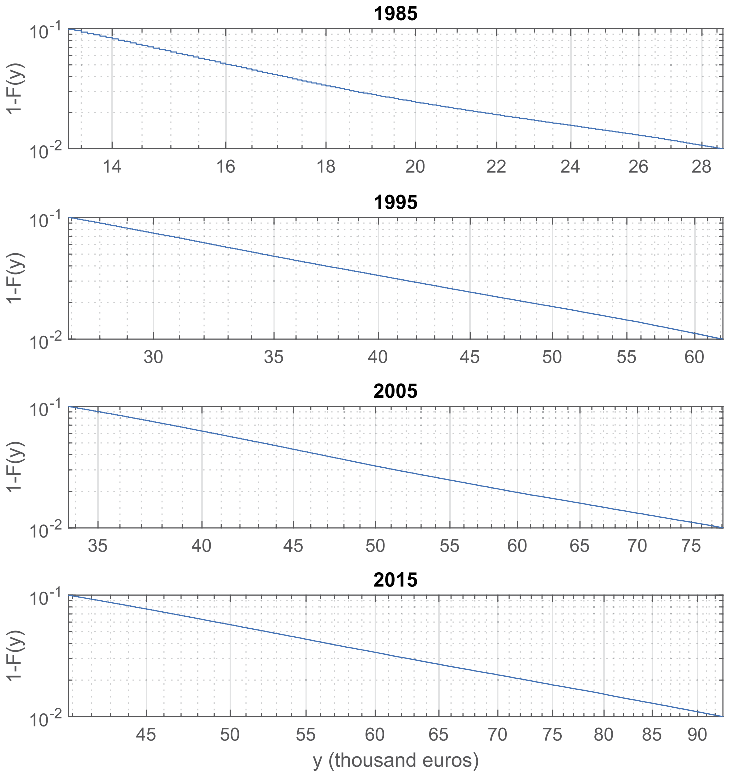


FIGURE A.3. The tail of the income distribution by year. *Notes*: The graphs show the logarithm of the survival function for the years 1985, 1995, 2005, and 2015. The graphs show the distribution from the 0.9 sample quantile to the 0.99 sample quantile. The horizontal axis is the logarithm of earnings in 2010 euros.

timated probability $F(y^*)$ and conditional mean $E[y|y \leq y^*]$, which are obtained non-parametrically.

However, the validity of equation (A.1) relies on the Pareto-tail property of the distribution. Is this assumption reasonable in the Italian context? A typical approach for visually inspecting tail distributions is a Zipf (or log-log) plot: a plot of the logarithm of the survival function $(1 - F(y))$ against $\log(y)$. If the data generating distribution has a Pareto tail, the relationship should be approximately linear at high incomes. Moreover, the absolute value of the slope of the linear line is an estimate of α, the tail index of the Pareto distribution. Figure A.3 shows the distribution of earnings for selected years (1985, 1995, 2005, and 2015) between the 90th and 99th earnings percentiles (which are both below the top-coding threshold and, therefore, do not suffer from bias). The Zipf plots indeed appear to be close to linear, and thus support the Pareto-tail assumption.

A.2.1 *Estimating the Pareto-tail index* We adapt four methods of estimating Pareto-tail index from right-censored data:

1. Maximum Likelihood (ML)
2. Quantile Slope (QL)
3. Kernel Density Slope (KDS)
4. Probability Integral Transform Statistic (PITS)

For all methods, we assume that earnings $\tilde{y}$ are Pareto distributed for values above a lower threshold, $\tilde{y} \geq \underline{y}$, and censored above an upper threshold, $\overline{y}$, so that the observed data is

$$y = \begin{cases} \tilde{y} & \text{if } \tilde{y} < \overline{y}, \\ \overline{y} & \text{if } \tilde{y} \geq \overline{y}. \end{cases}$$

This implies that conditional on earnings being above the lower threshold $y \geq \underline{y}$, the cumulative distribution function is

$$F(y) = \begin{cases} 1 - \left(\frac{y}{\underline{y}}\right)^{-\alpha} & \text{if } y < \overline{y}, \\ 1 & \text{if } y \geq \overline{y}. \end{cases}$$

We consider estimators based on $\underline{N}$ independent observations y_i, $i = 1, \ldots, \underline{N}$, which are greater or equal to the lower threshold. Without loss of generality, we also assume that the earnings observations are in increasing order, and the last $\overline{N}$ are censored.

A.2.2 *Maximum likelihood* This is the theoretically most efficient method. The likelihood of observation y_i is

$$l(\alpha, y_i) = \begin{cases} \frac{\alpha}{y}\left(\frac{y_i}{\underline{y}}\right)^{-\alpha} & \text{if } y_i < \overline{y}, \\ \left(\frac{\overline{y}}{\underline{y}}\right)^{-\alpha} & \text{if } y_i = \overline{y}. \end{cases}$$

The log-likelihood function is

$$L(\alpha, y_i) = (\underline{N} - \overline{N}) \log \alpha - \sum_{i=1}^{\underline{N}-\overline{N}} \log y_i - \alpha \sum_{i=1}^{\underline{N}} (\log y_i - \log \underline{y}).$$

The ML estimator is the value of α that satisfies the first-order condition,

$$\hat{\alpha}_{\text{ML}} = \frac{\underline{N} - \overline{N}}{\sum_{i=1}^{\underline{N}} (\log y_i - \log \underline{y})} = \frac{\underline{N} - \overline{N}}{\sum_{i=1}^{\underline{N}-\overline{N}} \log y_i + \overline{N} \log \overline{y} - \underline{N} \log \underline{y}}.$$

A.2.3 *Quantile slope* The Quantile Slope (QS) estimator uses the sample quantiles and the slope of the Zipf plots to estimate the Pareto-tail index. The quantile function for a given probability P is defined as

$$Q(P) = \inf\{y : F(y) \geq P\}.$$

In the case of a censored random variable with a Pareto tail, we can express the quantile function as

$$Q(P) = \begin{cases} \underline{y}(1-P)^{-\frac{1}{\alpha}} & \text{if } P < F(\overline{y}), \\ \overline{y} & \text{if } P \geq F(\overline{y}). \end{cases}$$

Let $P_1 < P_2 < \cdots < P_J$ be a set of numbers $P_j \in (0, F(\overline{y}))$. The sample quantile is Q_j is defined as

$$Q_j = y_{\lfloor \underline{N} P_j \rfloor + 1},$$

and converges in probability to $Q(P_j)$ as $\underline{N}$ increases.

We assume that P_J is small enough so that $Q_J < \overline{y}$. The quantile slope estimator is the absolute value of the slope coefficient in the regression of $\log(1-P_j)$ on $\log Q_j$. A closed-form expression for the quantile slope estimator of α is

$$\hat{\alpha}_{\text{QS}} = -\frac{\sum_{j=1}^{J}\left(\log Q_j - \frac{1}{J}\sum_{k=1}^{J}\log Q_k\right)\left(\log(1-P_j) - \frac{1}{J}\sum_{k=1}^{J}\log(1-P_k)\right)}{\sum_{j=1}^{J}\left(\log Q_j - \frac{1}{J}\sum_{k=1}^{J}\log Q_k\right)^2}.$$

A.2.4 *Kernel density slope* The Kernel Density Slope (KDS) estimator operates on a similar principle as the QS estimator, but uses the slope of the density function instead of the quantile function. Let $x = \log y$ be the logarithm of earnings. The CDF of x is

$$F_X(x) = \begin{cases} 1 - \underline{y}^{\alpha} e^{-\alpha x} & \text{if } x < \log \overline{y}, \\ 1 & \text{if } x \geq \log \overline{y}. \end{cases}$$

The density of x, if $x < \log \overline{y}$, is given by

$$f(x) = \alpha \underline{y}^{\alpha} e^{-\alpha x}.$$

For a given $x < \log \overline{y}$, the density function can be estimated nonparamterically using a kernel function K and bandwidth h, according to

$$\hat{f}(x) = \frac{1}{\underline{N} - \overline{N}} \sum_{i=1}^{\underline{N}-\overline{N}} K\left(\frac{x - x_i}{h}\right).$$

Note that this kernel density estimator is generally inconsistent for censored data because the density does not exist at the thresholds $\underline{y}$ and $\overline{y}$. However, for values of x far enough from the thresholds and a small enough bandwidth is can perform well.

Let $x_1 < x_2 < \cdots < x_J$ be as set of points in which the density can be reliably estimated. Then the KDS estimator of the Pareto-tail index is the slope of a regression of $\log \hat{f}(x)$ on x,

$$\hat{\alpha}_{\text{KDS}} = -\frac{\sum_{j=1}^{J}\left(x_j - \frac{1}{J}\sum_{k=1}^{J} x_k\right)\left(\log \hat{f}(x_j) - \frac{1}{J}\sum_{k=1}^{J}\log \hat{f}(x_k)\right)}{\sum_{j=1}^{J}\left(x_j - \frac{1}{J}\sum_{k=1}^{J} x_k\right)^2}.$$

A.2.5 *Probability integral transform statistic* The Probability Integral Transform Statics (PITS) uses the method of moments to estimate the Pareto index. Finkelstein, Tucker, and Veeh (2006) suggest that a PITS estimator for non-censored data and discuss its properties. They show that under a certain type of deviation this estimator is more robust than the ML estimator and other commonly used estimators. Here, we adapt the estimator for the use in censored data.

We define a transformation $G(y, a)$ by

$$G(y, a) = \frac{(1+t)}{1 + t\underline{y}^{a(1+t)}\overline{y}^{-a(1+t)}}\left(\frac{y}{\underline{y}}\right)^{-at},$$

where $t > 0$ is a tuning parameter. The expected value of $(\frac{y}{\underline{y}})^{-\alpha t}$ is

$$E\left[\left(\frac{y}{\underline{y}}\right)^{-\alpha t}\right] = \int_{\underline{y}}^{\overline{y}} \alpha \underline{y}^{\alpha(1+t)} y^{-\alpha(1+t)-1}\, dy + \underline{y}^{\alpha(1+t)}\overline{y}^{-\alpha(1+t)} = \frac{t}{1+t}\underline{y}^{\alpha(1+t)}\overline{y}^{-\alpha(1+t)} + \frac{1}{1+t},$$

which means that the expected value of $G(y, \alpha)$ is equal to 1. The PITS estimator relies on this property and is defined implicitly as the solution to the equation

$$\frac{1}{\underline{N}}\sum_{i=1}^{N} G(y_i, \hat{\alpha}_{\text{PITS}}) = 1.$$

A.2.6 *Comparison* All four methods presented here are consistent estimators of the Pareto-tail index. Yet each one has unique characteristics that may make it attractive for a given application. The quantile slope and kernel density slope rely on intuitive methods and map in a simple and transparent way to the data. They effectively ignore all the information that is contained in the censored observations. In addition, the sampling points (P_j's and x_j's) are arbitrary and their choice may affect the estimates. This becomes a bigger concern when the KDS method is applied on sparse data as the estimates may include a considerable bias. The KDS method also require the choice of a kernel function and a bandwidth, which makes it a more complicated procedure.

The maximum likelihood estimator is the most efficient and is also simple to calculate. However, Finkelstein, Tucker, and Veeh (2006), and Beran and Schell (2012) that the ML estimator is not robust when the tail is only approximately Pareto distributed. Finkelstein, Tucker, and Veeh (2006) suggest using the PITS estimator, which is adapted here for the use with censored data. The PITS estimator delivers more robustness than the ML estimator when the Pareto distribution is mixed with another "noise" distribution, at a mild cost in efficiency.

To illustrate these differences, we apply each estimator to 100 bootstrapped samples of the INPS 2015 earnings data. We set the lower threshold so that $F(\underline{y}) = 0.9$ to focus attention on the top 10% of income. INPS top-code daily earnings and report the annual earnings. Since annual earnings records which are smaller than the maximal top-coded threshold (365 times the daily threshold) may be partially top-coded, we set the upper threshold for the estimators at the $F(\overline{y}) = 0.99$. For the QS method, we need to choose the cumulative probability points for estimation of the quantiles. We pick 100 equally distanced points between 0 and 0.85 of the conditional distribution. For the kernel density, we use the Epanechnikov kernel with a bandwidth of 0.05 and 100 equally spaced sample points in log earnings, between the 90th percentile and the 98.5th percentile. For the tuning parameter of the PITS estimator, we pick $t = 0.5$.

Figure A.4 shows bootstrap replication histograms for the four methods. All four methods keep the dispersion of estimates low. Bootstrapped standard deviations range from 0.0089 for the ML estimator and 0.0189 for the KDS estimator. Due to different reasons discussed in this Appendix, the point estimates are statistically different. The economic implications are not large though; the top 1% share is 6.63% based on the PITS estimator (lowest index) is 6.43% based on the QS estimator (highest index).

Our estimate of the Pareto-tail index for all years in the sample is reported in Figure A.5.[32] The shaded area is the 95% confidence interval based on 100 bootstrap replications.

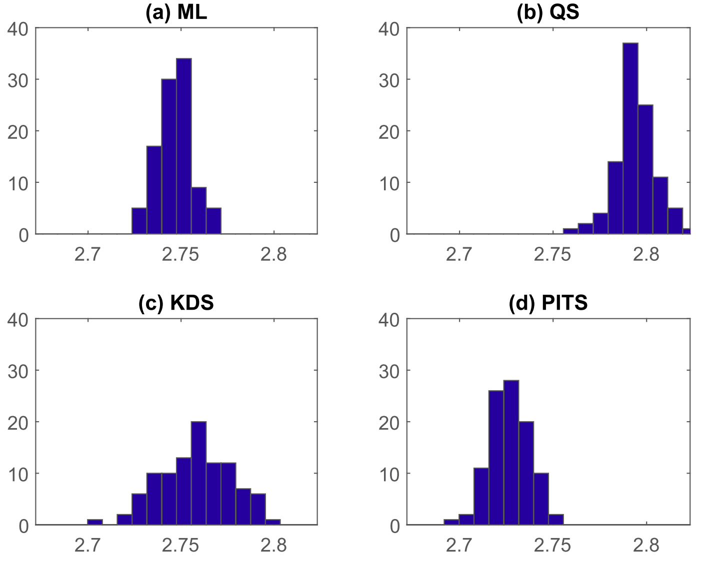


FIGURE A.4. Bootstrap replication histograms for Pareto-tail index estimates—earnings in 2016.

[32]These estimates are higher than in other countries. One explanation is that rules in collective bargaining agreements compress pay at the top relative to lower-ranked employees; alternatively, in small firms (which are the majority in Italy) fairness concerns may be more "visible" and lead to less concentration.

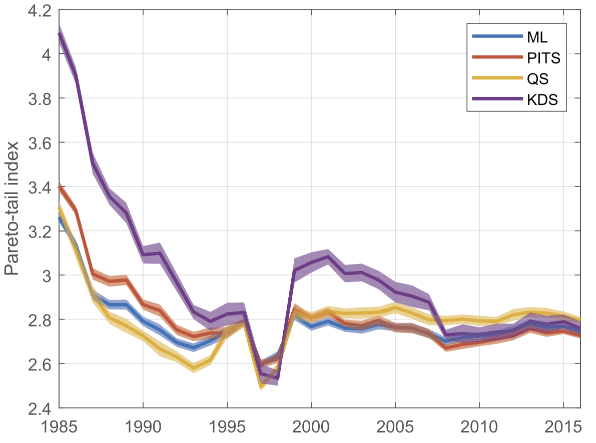


FIGURE A.5. Pareto-tail index estimates.

A.2.7 *Top-income shares* We use the estimated Pareto-tail indexes and equation (A.1) to estimate the top-income share. We pick y^* for each year so that $P^* = 0.99$. Figure A.6 shows the top-income shares for the top 10%, 5%, 1%, and 0.1% of earnings.

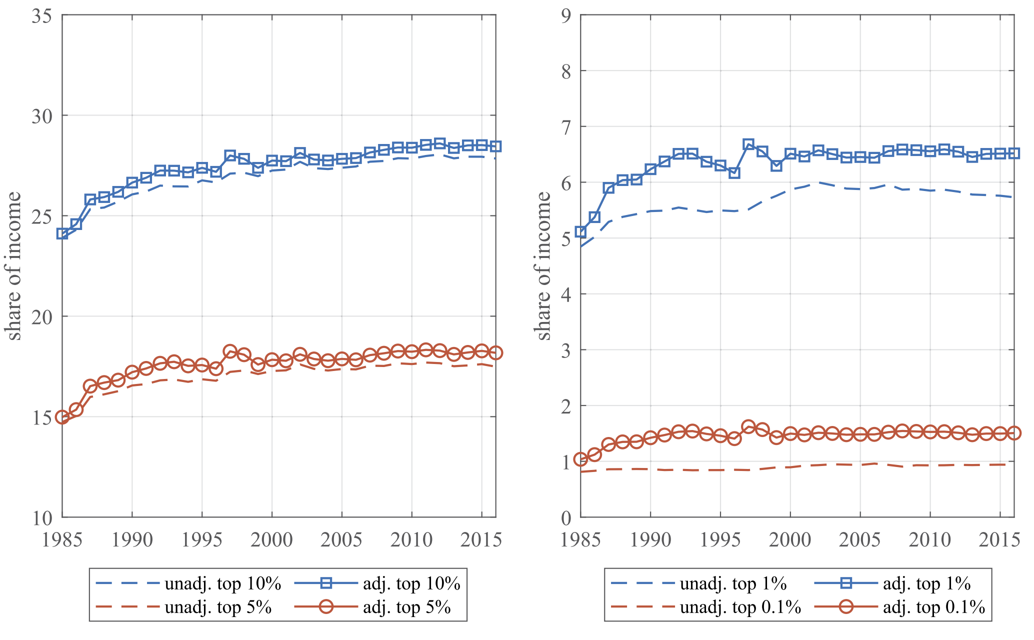


FIGURE A.6. Top-income shares based on adjusted and unadjusted earnings.

A.3 *Additional figures*

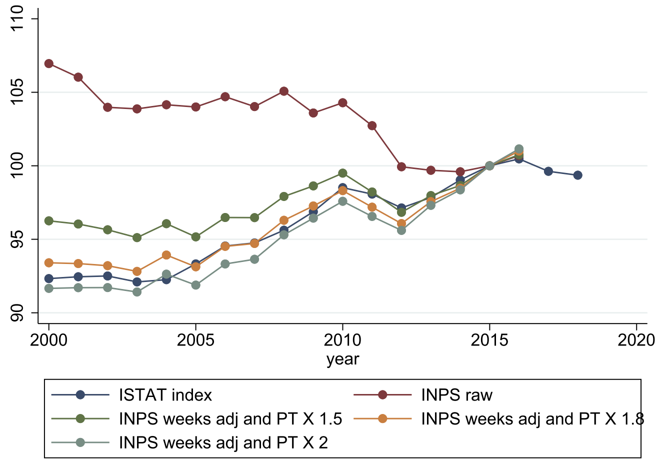
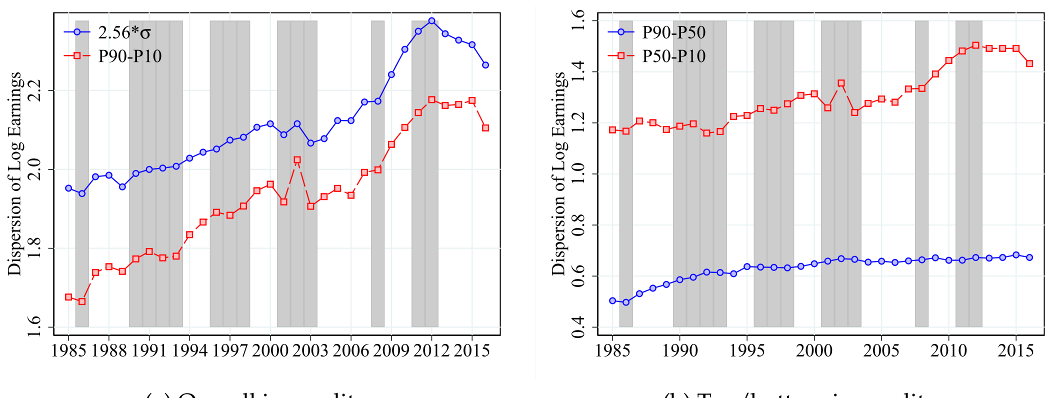


FIGURE A.7. Closing the gap in the evolution of earnings from 2000 to 2016 (index, 2015 = 100).



(a) Overall inequality (b) Top/bottom inequality

FIGURE A.8. Inequality, pooled sample.

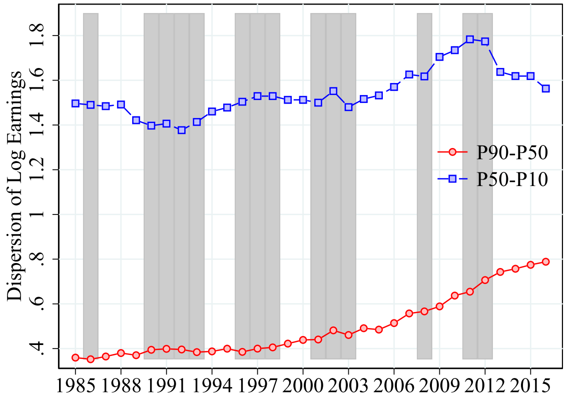


FIGURE A.9. Initial inequality, pooled sample.

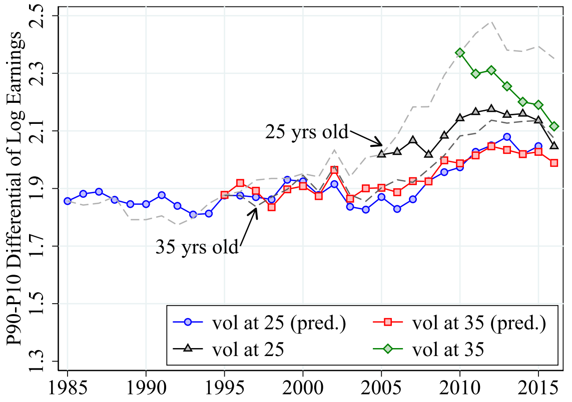


FIGURE A.10. Inequality by cohort, pooled sample.

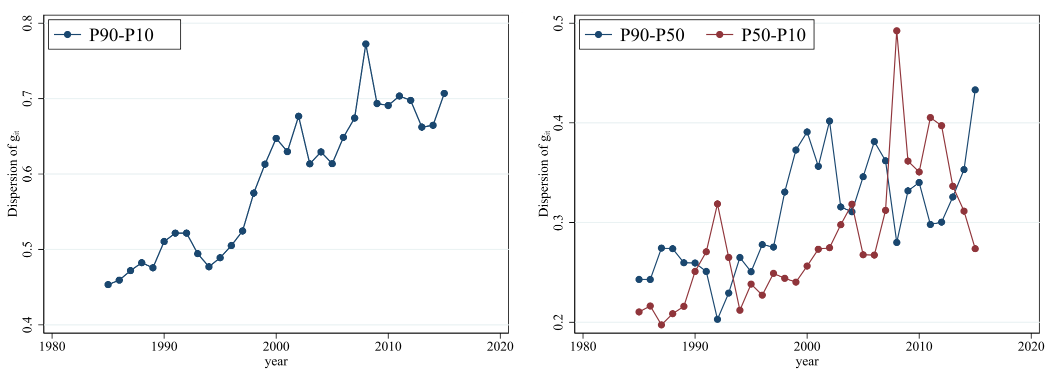


FIGURE A.11. Volatility and its upside/downside decomposition, pooled sample.

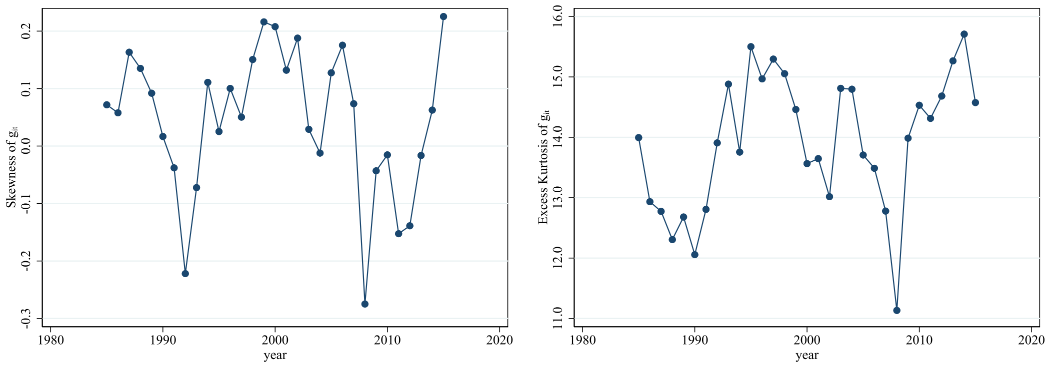


FIGURE A.12. Skewness and kurtosis of earnings growth, pooled sample.

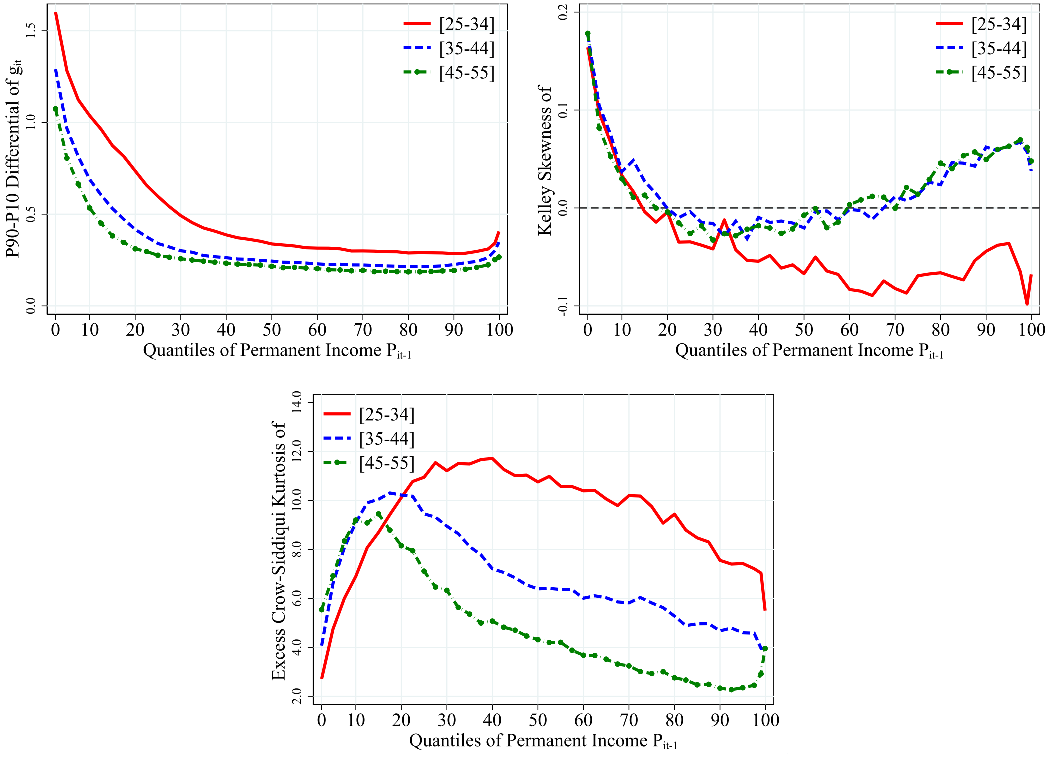


FIGURE A.13. Volatility, skewness, and kurtosis of earnings growth by permanent income, pooled sample.

FIGURE A.14. Inequality for public employees, SHIW.

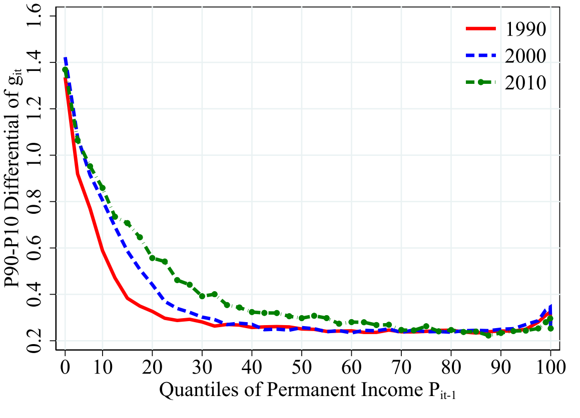


FIGURE A.15. Volatility by permanent income for selected years.

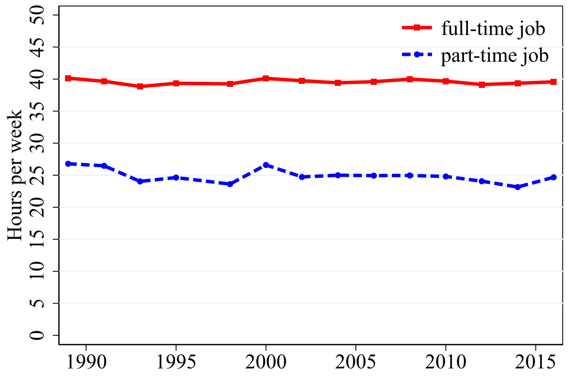


FIGURE A.16. Hours per week by employment status. *Notes*: The lines show average hours of work per week by full- and part-time status based on SHIW. The sample is restricted to 25–55-year-old workers, and it excludes the self-employed.

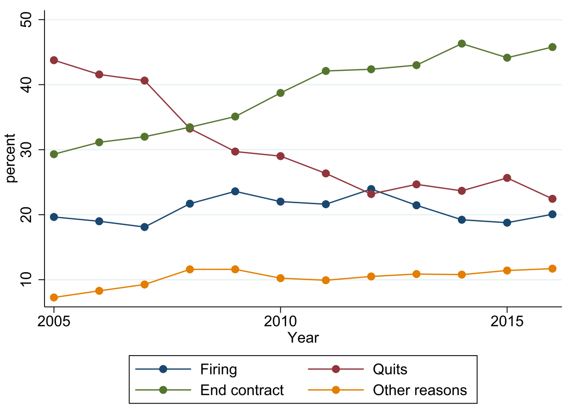


FIGURE A.17. Cause of separations.

REFERENCES

Beran, Jan and Dieter Schell (2012), "On robust tail index estimation." *Computational Statistics & Data Analysis*, 56, 3430–3443. [8]

Finkelstein, Mark, Howard G. Tucker, and Jerry Alan Veeh (2006), "Pareto tail index estimation revisited." *North American Actuarial Journal*, 10, 1–10. [8]

Co-editor Giovanni L. Violante handled this manuscript.

Manuscript received 6 March, 2021; final version accepted 20 April, 2022; available online 8 June, 2022.