How Unequally Heavy Are the Tails of the Distributions of Income Growth?

Silvia Sarpietro†
Bologna

Yuya Sasaki‡
Vanderbilt

Yulong Wang§
Syracuse

Abstract

Common moment-based measures of earnings risk, including the variance, skewness, and kurtosis, may be undefined in population under heavy-tailed distributions. In this light, we propose conditional Pareto exponents as novel measures of earnings risk that are robust against non-existence of moments, and develop estimation and inference methods for them. Using these measures with an administrative data set for the UK, the New Earnings Survey Panel Dataset (NESPDP), and the US Panel Study of Income Dynamics (PSID), we quantify the tail heaviness of the conditional distributions of earnings changes given age, gender, and past earnings. Our main findings are that: 1) the population kurtosis, skewness, and even variance may fail to exist for the conditional distribution of earnings growth; 2) earnings risk is increasing over the life cycle; 3) job stayers are more vulnerable to earnings risk, and 4) these patterns appear in both the period 2007–2008 of great recession and the period 2015–2016 of a positive growth among others.

Keywords: earnings & income risk, heavy tail, conditional Pareto exponent.

---

* We gratefully acknowledge the UK Data Service for granting access to NESPDP and their guidance and assistance with regard to this data set. We benefited from useful comments by seminar participants at the University of Bologna. All remaining errors are ours.

† Silvia Sarpietro: silvia.sarpietro@unibo.it. Department of Economics, University of Bologna, 40126 Bologna BO, Italy

‡ Yuya Sasaki: yuya.sasaki@vanderbilt.edu. Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA

§ Yulong Wang: ywang402@syr.edu. Department of Economics, Syracuse University, 110 Eggers Hall, Syracuse, NY 13244-1020, USA
## 1 Introduction

One of the main goals of the literature on income dynamics is to quantify income risk. Recent studies show that the distribution of income changes has heavier tails than the normal distribution. This evidence has relevant implications for how income risk affects economic decisions. Since heavy tails incur large risk premium costs in economies with risk-averse agents, quantifying the tail heaviness, along with features of preferences of economic agents, is necessary to measure the social costs of the uncertainties. However, a rigorous econometric method to quantify the tail heaviness of the income growth distribution conditional on individuals’ attributes is missing.

In this paper, we address this issue by first proposing a new method of estimation and inference about the conditional Pareto exponent, which characterizes the tail heaviness. Second, applying our method to the administrative data set for the UK, the New Earnings Survey Panel Dataset (NESPD), and the US Panel Study of Income Dynamics (PSID), we quantify the tail heaviness in the distribution of earnings changes and discuss how it varies with heterogeneous attributes of individuals. We find that the commonly used measures of income risk, such as sample kurtosis, skewness, and even standard deviation, may be misleading since their population counterparts may not exist (i.e., may be infinite) for the conditional distribution of earnings growth given age, gender, and past earnings. Finally, by interpreting the conditional Pareto exponent as a robust measure of the conditional earnings risk, we show that tail earnings risk increases over the life cycle, is higher for job-stayers, and these patterns appear in both the period 2007–2008 of the great recession and the period 2015–2016 of a positive growth despite some differences.

There is an extensive literature on earnings and income dynamics models, dating back to the 1970s – see the comprehensive survey by Moffitt and Zhang (2018). To explain the heavy tail feature, the existing literature exploits the following strategies: 1) fitting data to mixed normal distributions using mixture approaches (Geweke and Keane 2000; Bonhomme and Robin 2009); 2) conducting nonparametric density estimates and graphically exhibiting heavier tails than normal distributions with deconvolution and spectral decomposition approaches (Horowitz and Markatou 1996).

---

1 See, for instance, Golosov, Troshkin, and Tsyvinski (2016) and references therein.
Bonhomme and Robin (2010); Arellano, Blundell, and Bonhomme (2017); Botosaru and Sasaki (2018); Hu, Moffitt, and Sasaki (2019); 3) reporting the sample moments such as skewness and kurtosis and comparing them with counterparts from normal distribution (Guvenen, Karahan, Ozkan, and Song, 2019; Arellano, Bonhomme, De Vera, Hospido, and Wei, 2021, among others); 4) reporting some quantile-based measures such as Kelly’s measure of skewness and Crow-Siddiqui measure of kurtosis (Guvenen, Karahan, Ozkan, and Song, 2019, among others); and 5) reporting the mean absolute deviation (Arellano, Bonhomme, De Vera, Hospido, and Wei, 2021).

The first approach, based on mixed normal distribution, essentially assumes an exponentially decaying tail, which cannot characterize tail heaviness. The second one is based on nonparametric density estimation, which tends to behave poorly in the tails due to few extreme observations. The third approach based on sample moments may yield misleading information if the population moment does not exist due to heavy tails. In particular, the population kurtosis may not exist even if a researcher can always get a finite sample kurtosis. We indeed find that the population kurtosis, skewness, and even variance may not exist. The fourth one is robust to heavy tails: these non-extremal quantile-based measures are always well-defined even if the population moments do not exist. However, they fail to capture extreme income changes and tail events. Among other things, information on tail events is central to quantifying earnings risk under heavy-tailed distributions. Like the fourth one, the fifth approach based on the first moment alone is not informative about extreme income changes and tail events.

Given the above concerns about existing approaches, we aim for an alternative measure of tail heaviness to characterize the earnings risk conditional on individuals’ attributes. The conditional Pareto exponent is proposed as our robust measure of tail heaviness. We have the elegant characterization that the conditional kurtosis/skewness/variance is infinite if and only if the conditional Pareto exponent is smaller than the value of four/three/two. We propose methods of estimation and inference about the Pareto exponent in the conditional distribution of income risk given observed attributes, such as gender, age, and base-year income level. Unlike the aforementioned approaches, this new method can be used to characterize the tail heaviness robustly even if the rate of tail decay is slower than exponential ones.
and even if the population kurtosis, skewness, or variance is infinite.

To illustrate the empirical significance, we apply the proposed method to two data sets: the UK NESPD and the US PSID, with our focus on the former. NESPD is an employer-based survey data set on individual earnings in the UK and has been investigated very recently (cf. De Nardi, Fella, and Paz-Pardo 2021; Bell, Bloom, and Blundell 2021). We contribute to this new literature with the following three findings. First, we find that the distribution of the earnings changes (measured by the difference between the log earnings across two subsequent years) conditional on age, gender, and past earnings exhibits substantially heavy tails. The conditional Pareto exponent is significantly less than four for the majority of the subpopulations, raising the concern that the kurtosis may be infinite. Also, the conditional Pareto exponent is significantly less than three or even two for some subpopulations, and we thus reject the hypothesis of finite conditional skewness and even finite conditional variance. These results are robust across years, including the period of recession in 2007-2008 and the period of positive growth in 2015-2016 among others.

Second, we find remarkable patterns that the tail earnings risk is increasing over the life cycle. This is documented as that 40- and 50-year-old workers face higher earnings risk than 30-year-old workers. Third, we quantify the tail heaviness of the distribution of earnings changes for job-stayers and find that, at low earnings levels, conditional kurtosis does not exist, contrary to what we observe for the whole population. Excluding middle-aged men at top quantiles of earnings, we reject the hypothesis of finite conditional kurtosis for job-stayers at all earnings levels. In Appendix C, we also apply our method to the US PSID and obtain similar empirical findings on the tail heaviness of the conditional distribution of income risk.

The above findings differ from those in Arellano, Bonhomme, De Vera, Hospido, and Wei (2021), who examine the Spanish administrative data, and find that the income risk is inversely related to income and age, and the income risk inequality increases markedly in the recession. To understand the difference, we note that their measure of risk is based on the conditional mean absolute deviation, which is less influenced by extreme observations, while ours explicitly captures the conditional tail heaviness and hence focuses more on the tails. In this sense, our results compare more closely to the findings in Guvenen, Karahan, Ozkan, and Song (2019) and the
income literature inspired by their methodology. In particular, Guvenen, Karahan, Ozkan, and Song (2019) find in the US data set that the conditional variance of earnings risk is higher for younger workers at the bottom of the income distribution, and the kurtosis increases with age and with lagged income up to the top 5% of the income distribution where it sharply declines. Since the data exhibit substantially heavier tails than those of the normal distribution, these sample moments could be less informative if their population counterparts are indeed infinite.

Organization: The rest of the paper is organized as follows. Section 2 describes the data to be examined in this article and provides some descriptive analysis of the heavy-tail phenomena in the earnings growth risk. Section 3 formally introduces our econometric method, and Section 4 applies it to the UK data set. Section 5 concludes with some remarks. The Appendix collects asymptotic theories, computational details, Monte Carlo simulations, additional empirical results with the UK and US data sets, and mathematical proofs.

2 Data and Preview

2.1 New Earnings Survey Panel Data

We start with introducing the New Earnings Survey Panel Data (NESPD), an administrative data set at the individual level on UK earnings from the UK Social Security. It is an annual panel running from 1975 to 2016, and it surveys around 1% of the UK workforce. All employees whose National Insurance Number (NIN) ends in a given pair of digits are included in the survey. The NIN number is randomly issued to all UK residents at age 16 and kept constant throughout the lifetime of an individual. The NESPD is a survey directed to all employers whose employees qualify for the sample: the employers complete the questionnaire based on payroll records for their employees. As a result of being directed to the employer, NESPD has a low non-classical measurement error.

The survey reports the employees’ gender, age, and detailed work-related information: annual, weekly, and hourly earnings, hours of work, occupation, industry, working area, firms’ number of employers, and unionization. This information re-
lates to a specified week in April of each year: the data sample is taken on the first
day of April of each calendar year and concerns complete employee records only.
The NESPD contains complete information on the employees’ working life from the
first year they started working (or 1975 if later) until retirement age (or 2015 if
earlier), as long as the employer answered the questionnaire and the individual was
working with the last recorded employer in April.

Given that NESPD contains detailed information on earnings and a long panel
component, it has been used for a broad range of topics in the literature: among the
others, Goos and Manning (2007) have used NESPD to document job polarization
in the UK; Nickell and Quintini (2003) and Elsby, Shin, and Solon (2016) have
employed this study for analysis of wage rigidities; Adam, Phillips, and Roantree
(2019) provide an analysis of valid response rates in NESDP. However, the study of
income and earnings dynamics in the UK has not received lots of attention, with a
couple of notable exceptions in recent years: De Nardi et al. (2021) and Bell et al.
(2021) are two main references for earnings dynamics in the UK using NESPD for
the analysis.

The information in NESPD is of exceptionally high quality. However, it is
worth mentioning that not all workers are covered and there are non-negligible
non-response issues. In particular, data in NESPD are collected in a specified week
of April. As a result, NESPD might under-sample part-time workers if their weekly
earnings fall below the threshold for paying National Insurance and those that moved
jobs recently. Furthermore, the data set is unbalanced with possibly non-random
missing observations. This issue might be problematic for representativeness of those
at the bottom of the distribution and with unstable spells. We refer to Bell et al.
(2021) and De Nardi et al. (2021) for a deeper investigation of the data limitations
of NESPD.

Following the earnings literature, we further extract a subset of NESPD, which
correspond to employees with a strong attachment to the labor force. Specifically,
we follow De Nardi et al. (2021) to use the following selection criteria: we drop
the observations below 5% of the median earnings (roughly £1,300 a year), indivi-
duals whose total working hours exceed 80 hours per week, and individuals that
display negative values in earnings or hourly wages; we do not consider individuals
whose hours worked or weekly pay are missing. Earnings are deflated using CPI (2015=100). The earnings measure is the residual obtained by regressing the logarithm of earnings on year and age dummies. We extract a portion of data for a pair of years across the period of the great recession, 2007-2008. The above-described sample selection leaves 78,531 individuals.

For the purpose of comparisons, we also use the portion of data for the pair, 2015 and 2016, of the most recent survey years. Note that this period is associated with positive economic growth in the UK. The sample selection procedure described above leaves 95,906 individuals for this period. The results for other years are similar and hence postponed to Appendix B for readability.

Apart from age, gender, and income, we could further consider industry and occupation as heterogeneous attributes in the analysis. NESPD contains information on both industry and occupation with the Standard Industry Classification (SIC) and Standard Occupation Classification (SOC) codes.

We define a measure $Y$ as the difference between the log earnings in 2007 and the log earnings in 2008 (i.e., one-year earnings growth rate). Similarly, we also construct this variable for the period between 2015 and 2016. Figures 1 and 2 show the kernel density estimates of the measure $Y$ for the period 2007–2008 and the period 2015–2016, respectively. In each figure, the top and bottom panels show the densities for men and women, respectively. The figures illustrate clear departures from normality: each kernel density exhibits a large spike in the middle of the distribution sticking upward out of the reference normal density. Moreover, each kernel density has heavier tails compared to the reference normal density. These features of the estimated densities suggest that the actual distributions of $Y$ indeed have heavier tails than normal distributions, as documented in the previous literature.

We also inspect the sample moments of the distribution of one-year earnings changes conditional on previous earnings. Figure 3 shows the standard deviation, the Kelly’s measure of skewness, and the Crow-Siddiqui measure of kurtosis, respectively, of earnings changes by previous earnings, for men (left three figures) and women (right three figures). The Kelly’s measure of skewness is defined as $(Q_{0.9}Y - Q_{0.5}Y) - (Q_{0.5}Y - Q_{0.1}Y)/(Q_{0.9}Y - Q_{0.1}Y)$ and the Crow-Siddiqui measure of kurtosis as $(Q_{0.75}Y - Q_{0.25}Y)/(Q_{0.75}Y - Q_{0.25}Y)$, where $Q_{\tau}Y$ denotes the $\tau$-quantile of
the distribution of \( Y \). As in [De Nardi et al. (2021)](#), we find that the Crow-Siddiqui measure of kurtosis is higher than the reference value of 2.91 for the normal distribution, and the Kelly’s measure of skewness deviates from zero, ranging from positive values for low earnings groups to negative values for higher values of previous earnings. However, as mentioned in the introduction, these quantile-based measures do not take into account the very tail observations that are above, say \( Q_{0.975}Y \). These observations are indeed informative about the tail feature of the earnings risk and inequality. In comparison, our proposed method takes all samples into consideration and is explicitly designed to characterize the tail heaviness.

In the rest of the paper, we analyze the heaviness of the tails of the conditional distributions of \( Y \) given the earnings level and age in the base year (i.e., the base year is 2007 for the period 2007–2008 and it is 2015 for the period 2015–2016). Specifically, the heaviness of the tails is quantified by the conditional Pareto exponent, and we propose an econometric method of estimation and inference about this measure of tail heaviness. Its value, in particular, informs whether the \( r \)-th conditional moments of the income growth exist for \( r = 2, 3, 4 \), and so on. If the conditional Pareto exponent, \( \alpha(x_0) \), given \( X = x_0 \) is less than \( r \), then the conditional distribution of \( Y \) given \( X = x_0 \) does not have a finite \( r \)-th moment. Section 3 introduces a method of estimation and inference, and Section 4 presents the empirical results. It turns out that we reject the hypotheses of finite conditional kurtosis, finite conditional skewness, and even finite conditional standard deviation for some subpopulations.

### 3 Measurement of Conditional Tail Risk

We now introduce our proposed measure of conditional tail risk, and present method of estimation and inference. Formal theoretical justifications are relegated to the appendix. Let \( Y \) denote the variable of interest and let \( X \) denote a vector of individual characteristics, such as the income level and age in the base year. Assume that the sample is i.i.d. so that the joint distribution \( F_{Y,X}(\cdot, \cdot) \) of \((Y,X)\) is unique. In this section, we present an econometric method to study tail features (e.g., existence of the variance, skewness, kurtosis, and so on) of the conditional distribution \( F_{Y|X=x_0} \) of \( Y \) given \( X = x_0 \) for pre-specified values \( x_0 \) of the individual characteristics \( X \).
Figure 1: Kernel density estimates (black line) of $Y$, which is defined as a one-year change in log earnings, in 2007 in the NESPD. The top (respectively, bottom) panel shows the density of men (respectively, women). Also shown in gray dashed lines are the normal density fit to data. Number of individuals: 38,955 men (39,576 women).
Figure 2: Kernel density estimates (black line) of $Y$, which is defined as a one-year change in log earnings, in 2015 in the NESPD. The top (respectively, bottom) panel shows the density of men (respectively, women). Also shown in gray dashed lines are the normal density fit to data. Number of individuals: 45,985 men (49,921 women).
Figure 3: Moments of one-year log earnings changes by previous earnings, for men (left) and women (right) based on the NESPD.
Consider a rectangular array \( \{(Y_{ij}, X_{ij}) : i \in \{1, \cdots, I\}, j \in \{1, \cdots, J\}\} \). Such a data structure can be constructed by randomly splitting a cross-sectional data set of size \( N \) into an \( I \times J \) array such that \( I \cdot J \approx N \), which is what we implement in Section 4. See Appendix A.3 for how to choose \( I \) and \( J \) in practice under such a construction of an array.

Assume that the conditional distribution \( F_{Y|X=x_0}(\cdot) \) is regularly varying at infinity, that is,
\[
1 - F_{Y|X=x_0}(y) \propto y^{-\alpha(x_0)} \text{ as } y \to \infty,
\]
where \( \alpha(x_0) > 0 \) denotes the \( x_0 \)-conditional Pareto exponent that characterizes the tail heaviness of the conditional distribution of \( Y \) given \( X = x_0 \). We note that this Pareto-type tail condition is mild and satisfied by many commonly used distributions, such as Student-t, F, and gamma distributions. See Appendix A.1 for more discussions and primitive conditions.

With the parameter \( \alpha(x_0) \), we can characterize the existence of the conditional moments as follows. For any \( r \in \mathbb{R}^+ \),
\[
\mathbb{E}[Y_{ij}^r | X_{ij} = x_0] < \infty \text{ if } \alpha(x_0) > r \text{ and } \mathbb{E}[Y_{ij}^r | X_{ij} = x_0] = \infty \text{ if } \alpha(x_0) < r.
\]
Thus, the test of finite conditional \( r \)-th moment can be represented by the competing hypotheses
\[
H_0 : \alpha(x_0) > r \text{ against } H_1 : \alpha(x_0) \leq r. \tag{1}
\]
In particular, we set \( r = 2, 3, \) and \( 4 \) for tests of the existence of the standard deviation, skewness, and kurtosis, respectively.

To construct a feasible test for (1), we need to obtain a random sample from the conditional distribution \( F_{Y|X=x_0}(\cdot) \). This would be straightforward if \( X \) is discrete. However, random sampling from such a conditional distribution is infeasible when \( X \) includes at least one continuous random variable, such as the income level, which we use in our data analysis. To overcome this issue, we propose the following procedure to extract the local sample. Let \( ||\cdot|| \) denote the Euclidean norm.

1. For each \( i \), find the nearest neighbor (NN) of \( \{X_{ij}\}_{j=1}^J \) to \( x_0 \) and the induced \( Y_{ij} \) associated with the NN, that is \( Y_{ij} = Y_{ij}^* \) where \( ||X_{ij}^* - x_0|| = \)
\[
\min_{j \in \{1, \ldots, J\}} \|X_{ij} - x_0\|. \text{ Denote the induced } Y \text{ by } \{Y_i, [x_0]\}_{i=1}^J.
\]

2. Sort \(\{Y_i, [x_0]\}_{i=1}^J\) descendingly as \(\{Y_1, [x_0] \geq Y_2, [x_0] \geq \ldots \geq Y_(\alpha), [x_0]\}\). Collect the largest \(K + 1\) of them as

\[
Y(x_0) = \{Y_1, [x_0], Y_2, [x_0], \ldots, Y_(K+1), [x_0]\}.
\]

3. Use \(Y(x_0)\) to estimate the Pareto exponent \(\alpha(x_0)\) by the formula in (2) below.

We postpone until Appendix A.1 formal discussions of conditions under which this procedure works and why it works in theory. Here, we instead discuss its intuition. First, for each \(i\), we select the NN among \(\{X_{ij}\}_{j=1}^J\) to \(x_0\). When \(J\) (as well as \(I\)) is large enough, such a NN gets close enough to \(x_0\) and therefore its induced value \(Y_i, [x_0]\) behaves as if it were generated from \(F_{Y|X=x_0}\). Second, given the i.i.d. assumption, the subsample \(\{Y_i, [x_0]\}_{i=1}^I\) collected from all individuals serves as a random sample \(\text{approximately}\) generated from \(F_{Y|X=x_0}\). We can thus use the extreme order statistics of \(\{Y_i, [x_0]\}_{i=1}^I\) to estimate the Pareto exponent \(\alpha(x_0)\) by using one of the many existing methods. In particular, if we apply the estimator of Hill (1975) to this NN sample \(\{Y_i, [x_0]\}_{i=1}^I\), then we get an estimator

\[
\hat{\alpha}(x_0) = \left[\frac{1}{K} \sum_{k=1}^K \{\log(Y_k, [x_0]) - \log \left(Y_(K+1), [x_0]\right)\}\right]^{-1},
\]

of the conditional Pareto exponent \(\alpha(x_0)\). See Appendix A.3 for how to choose \(K\) in practice.

We show in Appendix A.1 that this estimator asymptotically follows the normal distribution as

\[
\sqrt{K} (\hat{\alpha}(x_0) - \alpha(x_0)) \xrightarrow{d} \mathcal{N} \left(0, \alpha(x_0)^2\right)
\]

under suitable conditions. Moreover, for any \(x_0 \neq x_1\), we also have the joint asymptotic normality:

\[
\sqrt{K} \begin{pmatrix} \hat{\alpha}(x_0) - \alpha(x_0) \\ \hat{\alpha}(x_1) - \alpha(x_1) \end{pmatrix} \xrightarrow{d} \mathcal{N} \begin{pmatrix} 0, & \begin{pmatrix} \alpha(x_0)^2 & 0 \\ 0 & \alpha(x_1)^2 \end{pmatrix} \end{pmatrix}.
\]

We provide a formal statement of these results as Proposition 1 in Appendix A.1.
Given (3) and (4), we can construct the standard t and F tests for our hypothesis testing problem (1). Specifically, we reject the (one-sided) null hypothesis at the 5% nominal level if
\[
\frac{\hat{\alpha}(x_0) - r}{\sqrt{K}\hat{\alpha}(x_0)} < \Phi^{-1}(0.05),
\] where \(\Phi^{-1}(\cdot)\) denotes the quantile function of the standard normal distribution and \(r = 2, 3,\) and \(4\). Moreover, we reject the null hypothesis that \(\alpha(x_0) = \alpha(x_1)\) (against the two-sided alternative) at the 5% nominal level if
\[
\frac{|\hat{\alpha}(x_0) - \hat{\alpha}(x_1)|}{\sqrt{K}\sqrt{\hat{\alpha}(x_0)^2 + \hat{\alpha}(x_1)^2}} > \Phi^{-1}(1 - 0.025).
\] Appendix A.2 contains a simulation study, which supports our theoretical results.

We end this section with discussions about the related econometrics and statistics literature. To estimate the unconditional Pareto exponent of some underlying distribution, researchers have developed numerous methods, including the popular and widely used Hill’s (1975) estimator \(^2\) and those proposed by Smith (1987) and Gabaix and Ibragimov (2011). We refer to de Haan and Ferreira (2007) for a comprehensive review. In recent work, Beare and Toda (2017) show that the cumulative sum of Markov multiplicative processes generates an approximate Pareto tail. This result supports our theoretical conditions and fits our empirical setup.

Unlike the unconditional Pareto exponent of, say \(Y\), estimating its conditional counterpart given some other variable \(X\) has received less attention. This is technically more challenging when \(X\) contains continuous random variables as one could not obtain a random sample from the conditional distribution of interest, say \(F_{Y|X=x_0}\) for some pre-specified value \(x_0\). Wang and Tsai (2009) and Wang and Li (2013) impose some parametric assumptions such that the Pareto exponent is a single index function of \(X\). Gardes and Girard (2008) and Gardes, Guillou, and Schorgen (2012) develop fully nonparametric methods based on local smoothing. Estimation and inference based on these local smoothing methods sensitively rely on the bandwidth parameter unlike our proposed method based on nearest neighbors. Furthermore,

---

\(^2\)This method fits large values in the sample to the Pareto distribution and implements the maximum likelihood estimation.
these existing methods underperform compared to ours in terms of bias and standard deviation, as demonstrated through simulation studies presented in Appendix A.2. For these reasons, we suggest using our novel method proposed above.

Finally, in a related paper, Trapani (2016) proposes a randomized testing procedure to test the finiteness of moments of a random variable. Unlike his proposal, we test the existence of moments of the conditional distribution of a random variable given some other variables. As previously stated, allowing for conditional Pareto exponents poses several challenges. Moreover, our analysis is not limited to inference but also allows for estimation of the conditional Pareto exponent. Thereby, we can propose a novel measure of earnings risk based on the estimate of the tail index. As such, our proposal is suitable for studying the tail features of the conditional distribution of income growth and relates these to the results in the income literature.

4 Results

4.1 Benchmark Sample

Applying the method introduced in Section 3 (and also in more detail in Appendix A) to NESPD described in Section 2.1, we analyze the conditional tail risk of earnings of adult individuals in the United Kingdom. We define $Y$ by the absolute difference of the log earnings in 2007 and the log earnings in 2008 for our baseline analysis. For the conditioning variables $X$, we include the quantile of earnings level and the age of the individual in the base year (2007), following Guvenen et al. (2019). With this setting, we study the conditional Pareto exponent $\alpha(x_0)$ for each point $x_0$ of earnings levels from $\{0.05, 0.10, \cdots, 0.90, 0.95\}$ (in quantile) and ages from $\{30, 40, 50\}$ for each of men and women.

Figure 4 illustrates the estimates of the conditional Pareto exponents $\alpha(x_0)$ (in black lines) along with the upper bounds of their one-sided 95% confidence intervals (in gray lines) for 30-, 40- and 50-year-old individuals. The left (respectively, right) panel shows the results for men (respectively, women) in each figure, which suggests different patterns of heterogeneous earnings risk across age, base-year earnings, and
gender groups.

We describe the empirical findings in the order of age and gender. For 30-year-old men (the top left panel in Figure 4), the conditional Pareto exponents (in point estimates) range from 1.4 to 7.0, and the upper bounds of the one-sided 95% confidence intervals range from 2.3 to 11.4. Given any quantile of earnings received in 2007, with the exception of the very top and bottom quantiles, the conditional Pareto exponent is significantly less than four, implying that the conditional kurtosis of earnings growth does not exist for these middle-earnings groups of young men. Overall, the kurtosis barely exists for most of the base-year earnings levels even if we fail to reject the hypothesis of finite kurtosis. Furthermore, given that earnings received in 2007 were between the 45th and the 75th percentile, the Pareto exponent is significantly less than three, implying that even the conditional skewness does not exist. For 30-year-old women (the top right panel in Figure 4), the conditional Pareto exponents (in point estimates) range from 1.5 to 6.1, and the upper bounds of the one-sided 95% confidence intervals range from 2.5 to 10.0. For this subpopulation, the conditional Pareto exponent is significantly less than four for high- and middle-high-earnings groups of women, those above the 60th percentile of the earnings distribution in 2007, implying that the conditional kurtosis does not exist. Furthermore, given that earnings received in 2007 were at or above the 70th percentile, the Pareto exponent is significantly less than three, implying that even the conditional skewness does not exist. Comparing the results between men and women at age 30, we observe that women were more vulnerable to earnings risk than men at the top quantiles of the base-year income level, above the 75th percentile.

For 40-year-old men (the middle left panel in Figure 4), the conditional Pareto exponents (in point estimates) range from 1.2 to 9.1, and the upper bounds of the one-sided 95% confidence intervals range from 1.9 to 14.9. Remarkably, the earnings risks of 40-year-old men are higher than those of 30-year-old men. For this age group of men, we reject the hypothesis of finite kurtosis at any level of base-year earnings, apart from the bottom quantiles, below the 15th percentile. Furthermore, given that earnings received in 2007 were between the 20th and the 35th percentile and above or at median excluding the very top quantile (between the 75th and 80th percentile), the Pareto exponent is significantly less than three (two), implying that even the
Figure 4: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2007–2008. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show the results for 30-, 40- and 50-year-old individuals.
conditional skewness (respectively, standard deviation) does not exist. For 40-year-old women (the middle right panel in Figure 4), the conditional Pareto exponents (in point estimates) range from 1.4 to 6.4, and the upper bounds of the one-sided 95% confidence intervals range from 2.4 to 11.0. For this age group of women, we reject the hypothesis of finite kurtosis at any earnings level above the 40th percentile. Furthermore, given that earnings received in 2007 were at or above the 65th percentile, the Pareto exponent is significantly less than three, implying that even the conditional skewness does not exist. Comparing the results between men and women at age 40, we observe that men are almost as vulnerable to earnings risk as women for this middle age group, with some difference at the bottom quantiles.

For 50-year-old men (the bottom left panel in Figure 4), the conditional Pareto exponents (in point estimates) range from 1.1 to 4.8, and the upper bounds of the one-sided 95% confidence intervals range from 1.8 to 8.0. For this age group of men, we reject the hypothesis of finite kurtosis and finite skewness at any level of base-year earnings, apart from the bottom quantiles, below the 20th percentile. Moreover, we reject the hypothesis of finite standard deviation at the 45th percentile and between the 65th and the 75th percentile. For 50-year-old women (the bottom right panel in Figure 4), the estimated conditional Pareto exponents range from 1.1 to 15.0, and the upper bounds of the one-sided 95% confidence intervals range from 1.9 to 25.7. For this age group of women, we reject the hypothesis of finite kurtosis at any level of base-year earnings, apart from the bottom quantiles, below the 30th percentile. Moreover, we reject the hypothesis of finite skewness (standard deviation) at any level of base-year earnings above the 40th percentile (between the 55th and the 70th percentile).

4.2 Period of Positive Growth

While our main focus has been on the period of great recession between 2007 and 2008, we next look at the period of positive growth, between 2015 and 2016. Figure 5 illustrates the results for earnings growth between 2015 and 2016 for 30-, 40- and 50-year-old individuals. These results share similar qualitative patterns to those reported in Figure 4. First, 30-year-old men at high quantiles of base-year earnings
enjoy less earnings risk. Second, 40-year-old men have an overall higher earnings risk than 30-year-old men. Lastly, and most remarkably, for men the earnings risk is not necessarily lower in the period 2015–2016 than in the period 2007–2008, even though the former period enjoyed a positive GDP growth (2.2% in 2015 and 1.8% in 2016) and the latter period suffered from negative GDP growth (0.7% in 2007 and −4.4% in 2008). Women, instead, are less vulnerable to earnings risk in the period 2015–2016 than in the period 2007–2008. Overall, there are more similarities than differences despite the contrast between a recession and a positive growth in the UK economy.

In summary, we find that: 1) the kurtosis, skewness, and even standard deviation may not exist for the conditional distribution of earnings growth given certain attributes (age, gender, and earnings); 2) 40- and 50-year-old workers have overall higher earnings risk than 30-year-old workers and we thus document that tail earnings risk is increasing over the life cycle; and 3) these patterns appear both in the period 2007–2008 of great recession and the period 2015–2016 of a positive growth for men, while there are some differences for women. Finally, we remark that, although we use the two periods, 2007–2008 and 2015–2016, to draw the above conclusion, we also present additional empirical results for other periods between 2005 and 2016 in Appendix B to demonstrate a robustness. In Appendix C, we also apply our proposed method to the US PSID and obtain similar empirical findings on the tail heaviness of the conditional distribution of income risk.

4.3 Job Stayers

We also experiment with a different set of selection criteria: we restrict the sample to workers who did not change their jobs in the last twelve months. We refer to them as job stayers. Figures 6 and 7 illustrate estimates of the conditional Pareto exponents \( \alpha(x_0) \) (in black lines) along with the upper bounds of their one-sided 95% confidence intervals (in gray lines) for the two pairs of years: 2007–2008 and 2015–2016, with the restricted sample of job stayers. Figures 6 and 7 show that, for almost all attributes, conditional kurtosis does not exist at any level of baseline earnings.
Figure 5: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2015–2016. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
This finding is in line with the US evidence documented by Guvenen et al. (2019) that job stayers experience different earnings dynamics and, in particular, face earnings innovations that are more leptokurtic, especially at the bottom of the earnings distribution.

More specifically, we find that, at low earnings levels, conditional kurtosis does not exist, in contrast to what we observe for the whole population. Excluding middle-aged men at top quantiles of earnings, we reject the hypothesis of finite conditional kurtosis for job-stayers at all earnings levels.

5 Concluding Remarks

Heavy tails of income and earnings risk distributions are costly in economies with risk-averse agents. Despite this common knowledge, the tail heaviness has been arguably less investigated in the literature on earnings and income dynamics. In this paper, we consider the conditional Pareto exponent as a novel robust measure of conditional earnings risk given observed attributes of individuals and propose a method of estimation and inference about this measure.

Applying the proposed method to the UK NESPD and, in Appendix, to the US PSID, we obtain the following findings. First, the population kurtosis, skewness, and even standard deviation might be infinite for the conditional distribution of income growth given certain attributes, such as age, gender, and lagged income, and, hence, their sample counterparts might be less informative. Second, 40- and 50-year-old workers have overall higher earnings risk than 30-year-old workers, thus earnings risk increases over the life cycle. Third, conditional earnings risk is higher for job stayers, in particular, at the bottom of the earnings distribution. Fourth, these patterns appear both in the period of great recession and the period of a positive growth, while there are differences as well, especially for women.

To the best of our knowledge, this is the first work to shed light on tail features of heavy-tailed distributions of earnings and income growth with a formal econometric method. While we focused on two specific data sets for empirical analysis, we hope that our proposed method will spur further empirical research with other data sets and for the study of other economies.
Figure 6: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2007–2008, for job-stayers. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 32,411 men (32,509 women).
Figure 7: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2015–2016, for job-stayers. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 38,680 men (41,393 women).
Appendix

Appendix A provides additional details about the econometric method introduced in Section 3. Appendix B provides additional empirical results with the UK NESPD. Appendix C provides the results with the US PSID.

A Additional Details about Section 3

Appendix A.1 presents a formal statement of the asymptotic normality results, (3) and (4). Appendix A.2 presents simulation studies. Appendix A.3 suggests a choice of $I$, $J$, and $K$ and a finite sample adjustment in practice.

A.1 Econometric Theory

In this appendix section, we provide a formal statement of the asymptotic normality results, (3) and (4). We consider the following conditions to establish these results.

Condition 1

1. $(Y_{ij}, X_{ij}^\top)$ is i.i.d. across $i$ and $j$. In addition, $f_X(x)$ is uniformly continuously differentiable and bounded away from zero in an open ball centered at $x_0$.

2. $1 - F_{Y|X=x}(y) = c(x)y^{-\alpha(x)} (1 + d(x)y^{-\gamma(x)} + r(x,y))$ uniformly as $y \to \infty$ where $c(\cdot) > 0$ and $d(\cdot)$ are continuously differentiable functions and uniformly bounded between 0 and $\infty$, $\alpha(\cdot) > 0$ and $\gamma(\cdot) > 0$ are continuously differentiable functions and $r(x,y)$ is continuously differentiable with bounded derivatives w.r.t. both $x$ and $y$ and satisfies $\limsup_{y \to \infty} \sup_{x \in B_\delta(x_0)} |r(x,y)/y^{-\gamma(x)}| = 0$ for some open ball $B_\delta(x_0)$ centered at $x_0$ with radius $\delta > 0$.

3. $K = o\left(\frac{n^{2\gamma(x_0)} / (\alpha(x_0) + 2\gamma(x_0))}{\lambda}\right) \to 0$ as $n \to \infty$.

4. $I \to \infty$ and $\lim_{I \to \infty} J/I^\lambda = C \in (0, \infty)$ for some $C$ and $\lambda \in (0, \infty)$. In addition, $\lambda/2 \geq \gamma(x_0)/(\alpha(x_0) + 2\gamma(x_0))$.

Condition 1.1 requires the data to be independent across $i$ and $j$. Condition 1.2 assumes that the target conditional distribution satisfies a second-order Pareto
tail approximation. The parameter $\gamma(x)$ governs the preciseness of the Pareto tail approximation, and the approximation error $r(x, y)$ is asymptotically negligible. The unconditional version (i.e., without $x$) of this condition has been commonly imposed in the statistics literature. See de Haan and Ferreira (2007) for a comprehensive review.

Condition 1.3 specifies the choice of $K$, a tuning parameter that characterizes the number of larger order statistics to be considered as stemming from the tail. Such a choice is to eliminate the asymptotic bias. If we choose $K$ such that $K \times n^{-2\gamma(x_0)/(\alpha(x_0)+2\gamma(x_0))} \to \mu(x_0)$ for some constant $\mu(x_0) \in \mathbb{R}$, then the asymptotic distribution in Proposition 1 involves one additional item $-\mu(x_0)\xi(x_0)$. To avoid such bias, we can select a smaller order $K$ as in Condition 1.3 such that $\mu(x_0)$ becomes zero asymptotically. This is close in spirit to choosing the undersmoothing bandwidth in the standard kernel estimation. Condition 1.4 requires a large $I$ and a large $J$. Note that $J$ can be substantially smaller than $I$ when $\lambda$ is less than one.

The following proposition characterizes the asymptotic behavior of $\hat{\alpha}(x_0)$ given in (2) under Condition 1.

**Proposition 1** Suppose that Condition 1 holds. Then
\[ \sqrt{K} (\hat{\alpha}(x_0) - \alpha(x_0)) \xrightarrow{d} \mathcal{N}(0, \alpha(x_0)^2) . \]
Moreover, for any $x_0 \neq x_1$, $\hat{\alpha}(x_0)$ and $\hat{\alpha}(x_1)$ are asymptotically independent, or equivalently,
\[ \sqrt{K} \left( \frac{\hat{\alpha}(x_0) - \alpha(x_0)}{\hat{\alpha}(x_1) - \alpha(x_1)} \right) \xrightarrow{d} \mathcal{N} \left( 0, \begin{pmatrix} \alpha(x_0)^2 & 0 \\ 0 & \alpha(x_1)^2 \end{pmatrix} \right) . \]

**Proof.** For convenience of writing, we also introduce the notation $\xi(x_0) = 1/\alpha(x_0)$, which is referred to as the tail index. We estimate $\hat{\xi}(x_0)$ by the reciprocal of $\hat{\alpha}(x_0)$, that is,
\[ \hat{\xi}(x_0) = \frac{1}{K} \sum_{k=1}^{K} \{ \log(Y_{(k),[x_0]}) - \log(Y_{(K+1),[x_0]}) \} . \]
Using similar lines of arguments to the proof of Theorem 2 in Sasaki and Wang (2021), we have that
\[ \sqrt{K} \left( \xi(x_s) - \xi(x_s) \right) \xrightarrow{d} \mathcal{N} \left( 0, \xi(x_s)^2 \right) \]
for $s = 1, 0$. In particular, we use $\alpha(x, 0)$ and $\gamma(x, 0)$ to denote their $\gamma(x, 0)$ and $\tilde{\gamma}(x, 0)$, respectively. Besides, their asymptotic bias term $\mu_H$ becomes zero by our Condition 1.3. Thus, the asymptotic distribution of $\hat{\alpha}(x, s)$ for $s = 1, 0$ is derived by the delta method.

Now, it remains to show the asymptotic independence between $\hat{\xi}(x, 1)$ and $\hat{\xi}(x, 0)$ for any $x \neq x_0$. Given their joint asymptotic Gaussianity, it suffices to show that

$$K \cdot \text{Cov} \left[ \hat{\xi}(x, 1) - \xi(x, 1), \hat{\xi}(x, 0) - \xi(x, 0) \right] \to 0,$$

which is proved as follows.

First, given the i.i.d. condition across $i$, we have that

$$\text{Cov} \left[ \log(Y_{i,1}, [x_0]), \log(Y_{i,2}, [x_1]) \right] = 0 \text{ for any } i_1 \neq i_2. \quad (8)$$

Moreover, given the i.i.d. condition across $j$, the standard argument in the induced order statistics literature (e.g., Lemma 3.1 in Bhattacharya (1984)) yields that, for any $i$, conditional on $X_i \equiv \{X_{i,1}, ..., X_{i,j} \}$, the induced order statistics $\{Y_{i,1}, ..., Y_{i,j} \}$ are independent. Denote $X_{i,[x]}$ as the NN of $\{X_{i,j} \}_{j=1}^{J}$ to $x$. Therefore as long as $X_{i,[x_0]} \neq X_{i,[x_1]}$, we have that

$$\text{Cov} \left[ \log(Y_{i,[x_0]}), \log(Y_{i,[x_1]}), \right] = \mathbb{E} \left[ \text{Cov} \left[ \log(Y_{i,[x_0]}), \log(Y_{i,[x_1]}) \right] \middle| X_i \right] = \mathbb{E} \left[ \text{Cov} \left[ \log(Y_{i,j_0}), \log(Y_{i,j_1}) \right] \middle| X_{i,j_0} = X_{i,[x_0]}, X_{i,j_1} = X_{i,[x_1]} \right] = 0 \quad (9)$$

by the law of iterated expectations. Since $X_{i,j}$ contains at least one continuous component, Lemma 1 in Sasaki and Wang (2021) implies that $||X_{i,[x_0]} - x_0|| = o_a.s.(1)$ and $||X_{i,[x_1]} - x_1|| = o_a.s.(1)$ as $J \to \infty$. This yields that, for any $i$,

$$\mathbb{P} \left[ X_{i,[x_0]} = X_{i,[x_1]} \right] = o(1) \text{ as } J \to \infty. \quad (10)$$

Now, recall that $\hat{\xi}(x, 1) = K^{-1} \sum_{i=1}^{K} \{ \log(Y_{i,[x_1]} - \log(Y_{i,[x_0]} \})$, which involves $\{Y_{i,[x_1]}, ..., Y_{i,[x_1]} \}$ for some indices $i_1, ..., i_{K+1} \in \{1, ..., I \}$. Symmetrically, we have that $\hat{\xi}(x, 0) = K^{-1} \sum_{i=1}^{K} \{ \log(Y_{i,[x_0]} - \log(Y_{i,[x_0]} \})$, which involves
For some indices $i_1', \ldots, i_{K+1}' \in \{1, \ldots, I\}$. By (8), the only potentially non-zero components in $\text{Cov}[\hat{\xi}(x_1), \hat{\xi}(x_0)]$ will be $\text{Cov}[\log(Y_{i_1'[x_1]}, \log(Y_{i_1'[x_0]})]$ for $i \in \{i_1, \ldots, i_{K+1}\} \cap \{i_1', \ldots, i_{K+1}'\}$. However, these terms are still zero by (9) if $X_{i_1[x_0]} \neq X_{i_1'[x_1]}$, which happens with probability approaching one by (10). This completes a proof of the proposition.

### A.2 Simulation Analysis

In this section, we present simulation studies of finite sample performance of the proposed estimation and inference method. In each iteration, we randomly generate an $I \times J$ array $\{Y_{ij}, X_{ij}\}$ according to the following two designs. Design 1: $Y_{ij} \sim \text{Pareto}(\alpha(X_{ij}))$ where $\alpha(x) = 1 + 10x$ and $X_{ij} \sim \text{Uniform}(0, 1)$. Design 2: $Y_{ij} \sim \text{Pareto}(\alpha(X_{ij}))$ where $\alpha(x) = 10 \times (x^2 - x + 1)$ and $X_{ij} \sim \text{Uniform}(0, 1)$. We estimate $\alpha(x_0)$ by $\hat{\alpha}(x_0)$ for each of $x_0 \in \{0.1, 0.2, \ldots, 0.8, 0.9\}$. Each set of Monte Carlo simulations consists of 1000 iterations of this process.

Tables 1 and 2 report the results across various values of $I = J \in \{500, 1000\}$ and $K \in \{10, 20\}$ under Design 1 and Design 2, respectively. Displayed statistics are the bias (Bias), standard deviation (SD), root mean square error (RMSE), and 95% coverage frequency (95% Cover). Observe that the RMSE shrinks as $K$ increases. Furthermore, the 95% coverage frequencies are close to the nominal probability. These results demonstrate desired finite sample performance of the proposed method of estimation and inference.

For comparison, we also implement the fully nonparametric method proposed by Gardes and Girard (2008). In particular, given a random sample of $(Y_i, X_i)$ of size $N$ and a pre-determined bandwidth $b$, we follow Gardes and Girard (2008) to select all the observations that satisfy $|X_i - x_0| \leq b$, where $x_0$ again denotes the conditional value. Given this selected subsample, we sort the $Y_i$’s and construct the standard Hill estimator based on the largest $K + 1$ order statistics as in (2). Since there is no theoretical justification of the choice of $b$, we implement various values of $b \in \{0.01, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2\}$ for sensitivity check. We generate data from Design 1 described above and set $x_0 = 0.5$. The sample size is $N = I \times J \in \{500^2, 1000^2\}$. Table 3 presents the results with 1000 iterations. We
| $I = J = 500$ | $x_0$ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
|---|---|---|---|---|---|---|---|---|---|
| $K = 10$ | Bias | 0.221 | 0.304 | 0.497 | 0.594 | 0.614 | 0.712 | 0.979 | 0.920 | 1.069 |
| | SD | 0.759 | 1.125 | 1.476 | 2.043 | 2.178 | 2.812 | 3.066 | 3.608 | 4.002 |
| | RMSE | 0.791 | 1.166 | 1.558 | 2.128 | 2.263 | 2.901 | 3.219 | 3.724 | 4.142 |
| | 95% Cover | 0.948 | 0.957 | 0.963 | 0.952 | 0.963 | 0.952 | 0.956 | 0.950 | 0.949 |
| $I = J = 500$ | $x_0$ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| $K = 20$ | Bias | 0.121 | 0.158 | 0.207 | 0.273 | 0.262 | 0.370 | 0.519 | 0.451 | 0.519 |
| | SD | 0.512 | 0.722 | 0.951 | 1.287 | 1.448 | 1.773 | 2.126 | 2.141 | 2.568 |
| | RMSE | 0.526 | 0.740 | 0.973 | 1.316 | 1.472 | 1.811 | 2.189 | 2.188 | 2.620 |
| | 95% Cover | 0.949 | 0.966 | 0.956 | 0.937 | 0.959 | 0.947 | 0.945 | 0.954 | 0.943 |
| $I = J = 1000$ | $x_0$ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| $K = 10$ | Bias | 0.244 | 0.333 | 0.580 | 0.574 | 0.643 | 0.861 | 0.813 | 0.981 | 1.205 |
| | SD | 0.815 | 1.233 | 1.589 | 1.996 | 2.357 | 2.780 | 2.947 | 3.287 | 3.809 |
| | RMSE | 0.851 | 1.277 | 1.692 | 2.077 | 2.443 | 2.910 | 3.057 | 3.431 | 3.995 |
| | 95% Cover | 0.957 | 0.951 | 0.963 | 0.953 | 0.956 | 0.947 | 0.954 | 0.962 | 0.965 |
| $I = J = 1000$ | $x_0$ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| $K = 20$ | Bias | 0.122 | 0.148 | 0.272 | 0.254 | 0.315 | 0.432 | 0.418 | 0.459 | 0.471 |
| | SD | 0.508 | 0.767 | 1.007 | 1.219 | 1.453 | 1.760 | 2.020 | 2.183 | 2.537 |
| | RMSE | 0.523 | 0.781 | 1.043 | 1.245 | 1.487 | 1.812 | 2.063 | 2.231 | 2.580 |
| | 95% Cover | 0.944 | 0.950 | 0.957 | 0.947 | 0.958 | 0.946 | 0.953 | 0.965 | 0.947 |

Table 1: Simulation results under Design 1.
| I = J = 500 | x₀ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| K = 10 | Bias | 1.020 | 0.921 | 0.989 | 0.967 | 0.872 | 0.702 | 0.856 | 0.986 | 1.000 |
| | SD | 3.483 | 3.164 | 3.060 | 2.895 | 2.957 | 2.922 | 3.083 | 3.332 | 3.689 |
| | RMSE | 3.629 | 3.295 | 3.216 | 3.052 | 3.083 | 3.005 | 3.200 | 3.475 | 3.822 |
| | 95% Cover | 0.957 | 0.960 | 0.963 | 0.955 | 0.948 | 0.950 | 0.951 | 0.965 | 0.953 |
| I = J = 500 | x₀ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| K = 20 | Bias | 0.523 | 0.408 | 0.448 | 0.514 | 0.375 | 0.335 | 0.397 | 0.576 | 0.527 |
| | SD | 2.266 | 2.026 | 1.979 | 1.881 | 1.822 | 1.784 | 1.887 | 2.105 | 2.252 |
| | RMSE | 2.325 | 2.066 | 2.030 | 1.950 | 1.860 | 1.815 | 1.929 | 2.183 | 2.313 |
| | 95% Cover | 0.958 | 0.951 | 0.936 | 0.948 | 0.956 | 0.947 | 0.953 | 0.953 | 0.953 |
| I = J = 1000 | x₀ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| K = 10 | Bias | 0.843 | 0.933 | 0.944 | 0.765 | 0.725 | 0.824 | 0.790 | 0.750 | 1.131 |
| | SD | 3.409 | 3.349 | 3.015 | 2.795 | 2.890 | 3.063 | 2.891 | 3.344 | 3.586 |
| | RMSE | 3.512 | 3.476 | 3.160 | 2.897 | 2.979 | 3.172 | 2.997 | 3.428 | 3.760 |
| | 95% Cover | 0.951 | 0.961 | 0.970 | 0.956 | 0.952 | 0.950 | 0.965 | 0.951 | 0.951 |
| I = J = 1000 | x₀ | 0.100 | 0.200 | 0.300 | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 |
| K = 20 | Bias | 0.474 | 0.401 | 0.446 | 0.359 | 0.339 | 0.413 | 0.440 | 0.324 | 0.495 |
| | SD | 2.256 | 2.130 | 2.069 | 1.831 | 1.910 | 1.962 | 1.913 | 2.033 | 2.299 |
| | RMSE | 2.306 | 2.168 | 2.116 | 1.866 | 1.940 | 2.005 | 1.963 | 2.059 | 2.352 |
| | 95% Cover | 0.953 | 0.948 | 0.944 | 0.947 | 0.947 | 0.948 | 0.960 | 0.947 | 0.941 |

Table 2: Simulation results under Design 2.
summarize the findings as follows. First, the choice of bandwidth has a substantial
effect on the performance of this nonparametric estimator. The bias is small only
when the bandwidth is within a narrow window, which theoretically depends on the
unknown higher-order parameters. Second, compared with Table 1, this nonpara-
metric estimator has larger bias and standard deviations than our proposed method
when the bandwidth is small enough to guarantee the correct coverage.

A.3 Choice of $I$, $J$, and $K$

In this appendix section, we present a rule of choosing $I$, $J$, and $K$ in practice
when a random sample of size $N$ is available. For simplicity, we consider the Pareto
distribution family as a reference. In this case, we have $\gamma(x_0) = \infty$ in Condition 1
and hence $\lambda \geq 1$. Since $K$ is the effective sample size in the asymptotic normality
result and $K$ in turn increasing in $I$, the goal is to choose $\lambda \geq 1$ such that $I$ is large.
By Condition 4, such a choice is $\lambda = 1$. We therefore suggest to set $I = J = \sqrt{N}$.

Once $I$ has been chosen, then we can choose $K$ by adapting the diagnostic
method proposed by Guillou and Hall (2001). We consider the case of estimating
$\alpha(x_0)$ here. Define $Z_i = i \log(Y_{(i)},[x_0]/Y_{(i+1)},[x_0])$ for $i = 1, \ldots, I$. Suppose that $Y_{i,[x_0]}
$ is exactly Pareto distributed with exponent $\alpha(x_0) = 1/\xi(x_0)$. Then, $Z_i$ should be
i.i.d. with the exponential distribution and satisfies $E[Z_i] = \xi(x_0)$ and $\text{VAR}[Z_{m+j}] = \xi(x_0)^2$. For any given $K$ and any antisymmetric weights $\{w_i\}_{i=1}^K$ such that $w_i = -w_{K-i+1}$ and $\sum_{i=1}^K w_i = 0$, the weighted average statistic $U_K \equiv \sum_{i=1}^K w_i Z_i$ should
have zero mean and variance $\sum_{i=1}^K w_i^2 \xi(x_0)^2$. Therefore, we can construct

$$T_K \equiv \left( \sum_{i=1}^K w_i^2 \right)^{-1/2} \hat{\xi}(x_0)^{-1} U_K,$$

which has zero mean and unit variance if $Y_{i,[x_0]}$ is exactly Pareto and if $\hat{\xi}(x_0) = \xi(x_0)$.

Now under Condition 1, $\{Z_i\}_{i=1}^K$ should be approximately independent and exponentially
distributed. Then the above properties of $T_K$ hold asymptotically, following
a similar argument to that in Guillou and Hall (2001). When such an approxima-
tion performs well for a certain $K$, we expect the fluctuation of $T_K^2$ to be small.
| $N = 500^2$ | $b$ |
|-------------|-----|
| $K = 10$    |     |
| Bias        | 0.627 0.562 0.479 0.428 0.296 -0.064 -0.360 -0.549 -0.994 |
| SD          | 2.349 2.238 2.213 2.186 2.153 2.068 2.058 1.952 1.849 |
| RMSE        | 2.430 2.307 2.263 2.227 2.172 2.068 2.088 2.026 2.098 |
| 95% Cover   | 0.956 0.947 0.946 0.947 0.933 0.902 0.867 0.856 0.763 |

| $N = 500^2$ | $b$ |
|-------------|-----|
| $K = 20$    |     |
| Bias        | 0.284 0.334 0.118 0.086 -0.106 -0.383 -0.654 -0.916 -1.269 |
| SD          | 1.482 1.549 1.453 1.501 1.444 1.359 1.196 1.184 1.086 |
| RMSE        | 1.509 1.584 1.457 1.503 1.448 1.411 1.362 1.496 1.670 |
| 95% Cover   | 0.952 0.950 0.937 0.937 0.927 0.881 0.852 0.770 0.656 |

| $N = 1000^2$ | $b$ |
|-------------|-----|
| $K = 10$    |     |
| Bias        | 0.842 0.501 0.457 0.294 0.088 -0.238 -0.565 -0.804 -1.024 |
| SD          | 2.488 2.283 2.249 2.200 2.128 2.112 2.037 1.842 1.780 |
| RMSE        | 2.625 2.336 2.294 2.219 2.129 2.124 2.113 2.009 2.053 |
| 95% Cover   | 0.966 0.951 0.950 0.933 0.921 0.891 0.828 0.806 0.758 |

| $N = 1000^2$ | $b$ |
|-------------|-----|
| $K = 20$    |     |
| Bias        | 0.360 0.350 0.143 0.022 -0.219 -0.449 -0.713 -1.015 -1.360 |
| SD          | 1.490 1.507 1.410 1.428 1.447 1.346 1.321 1.170 1.089 |
| RMSE        | 1.532 1.546 1.416 1.428 1.463 1.418 1.501 1.548 1.742 |
| 95% Cover   | 0.961 0.958 0.947 0.930 0.892 0.869 0.807 0.747 0.631 |

Table 3: Simulation results of the nonparametric estimator under Design 1.
Accordingly, we define the following criteria based on a moving average of $T^2_K$:

$$C_K = \left((2l + 1)^{-1} \sum_{j=-l}^{l} T^2_{K+j}\right)^{1/2},$$

where $l$ equals the integer part of $K/2$. When $K$ is too large relative to $I$, the Pareto approximation incurs a larger bias, and hence $C_K$ exceeds one by a larger magnitude. To obtain an implementable rule, we follow Guillou and Hall (2001) to use $w_i = \text{sgn} \left( K - 2i + 1 \right) |k - 2i + 1|$ and propose to choose the smallest $K$ that satisfies $C_t > 1$ for all $t \geq K$, that is,

$$K^* = \min_{1 \leq K \leq I} \{ K : C_t > 1 \text{ for all } t \geq K \}. \quad (11)$$

We close the section by presenting a method of finite-sample adjustments to incorporate the uncertainty induced by the random splitting of data. We follow a procedure suggested in Section 3.4 of Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and Robins (2018). Suppose that we obtain $S$ estimates $\{\hat{\alpha}_s(x_0)\}_{s=1}^S$ by (2) with $S$ times of random splitting. For point estimation, we use $\bar{\alpha}_S(x_0) = \text{median}\{\hat{\alpha}_s(x_0)\}_{s=1}^S$. For variance estimators, we use

$$\hat{\sigma}^2_S(x_0) = \text{median}\{\hat{\alpha}_s^2(x_0) + (\hat{\alpha}_s(x_0) - \bar{\alpha}_S(x_0))^2\}_{s=1}^S$$

by accounting for the variation introduced by random splitting. We use $S = 1000$ in our empirical analyses.

**B Additional Results with NESPD**

In Section 4.1 in the main text, we focus on the period 2007–2008 of the great recession and the period 2015–2016 for conservation of space. In this appendix section, we repeat the econometric analysis for other periods: 2005–2006, 2006–2007, 2008–2009, 2009–2010, 2010–2011, 2011–2012, 2012–2013, and 2013–2014 to demonstrate the robustness of the observed patterns reported in the main text. Figures 8, 9, 10, 11, 12, 13, 14, 15, and 16 illustrate the estimates of the conditional Pareto exponents $\alpha(x_0)$ (in black lines) along with the upper bounds of their one-sided 95% confidence intervals (in gray lines). In each figure, the left (respectively,
right column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. The findings are similar to those in Section 4.

C Panel Study of Income Dynamics

C.1 Data

We also apply our methodology to another data set, the US Panel Study of Income Dynamics (PSID), which is one of the most commonly used data sets for empirical research on income dynamics. The portion of data for the pair, 2007 and 2009, of years across the period of the great recession is extracted for our use. We select the subsample of both men and women whose total taxable incomes were recorded and non-zero in both 2007 and 2009 and were aged between 25 and 64 in 2007. This sample selection leaves 271 individuals. For the purpose of comparisons, we also use the portion of data for the pair, 2017 and 2019, of the most recent survey years at the time of our writing of this paper. Note that this period is associated with positive economic growth in the United States. The sample selection procedure described above leaves 389 individuals for this period.

We define a measure \( Y \) as the absolute difference of the log income in 2007 and the log income in 2009 (i.e., two-year income growth rate). Similarly, we also construct this variable for the period between 2017 and 2019. Figures 17 and 18 display kernel density estimates of the income measure \( Y \) for the period 2007–2009 and the period 2017–2019, respectively. In each figure, the left and right panels show the densities for men and women, respectively. Also shown in gray dashed lines are the normal density plots fit to data. Observe that each kernel density exhibits a large spike in the middle of the distribution sticking upward out of the reference normal density. Moreover, each kernel density has heavier tails compared to the reference normal density. These features of the estimated densities evidence that the actual distributions of \( Y \) indeed have heavier tails than normal distributions, as documented in the previous literature. That said, as emphasized in the introductory section, nonparametric density plots cannot informatively demonstrate evidence of
Figure 8: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2005–2006. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 48,542 men (47,851 women).
Figure 9: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2006–2007. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 39,903 men (39,546 women).
Figure 10: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2008–2009. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 39,715 men (40,449 women).
Figure 11: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2009–2010. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 48,368 men (50,372 women).
Figure 12: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2010–2011. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 49,367 men (51,672 women).
Figure 13: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2011–2012. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 49,720 men (52,874 women).
Figure 14: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2012–2013. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 48,728 men (51,799 women).
Figure 15: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2013–2014. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 49,455 men (52,586 women).
Figure 16: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the NESPD in the period 2014–2015. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals. Number of individuals: 48,520 men (52,109 women).
C.2 Results

Applying the method introduced in Section 3 (and also Appendix A) to the US Panel Study of Income Dynamics (PSID) described in Appendix C.1, we analyze conditional tail risk of income of adult individuals in the United States. We define $Y$ by the absolute difference of the log income in 2007 and the log income in 2009 for our baseline analysis. For the conditioning variables $X$, we include the quantile of income level and the age of the individual in the base year (2007), following Guvenen et al. (2019). With this setting, we study the conditional Pareto exponent $\alpha(x_0)$ for each point $x_0$ of income levels from $\{0.05, 0.10, \cdots, 0.90, 0.95\}$ (in quantile) and ages from $\{30, 40, 50\}$ for each of men and women.

Figure 19 illustrates the estimates of the conditional Pareto exponents $\alpha(x_0)$ (in black lines) along with the upper bounds of their one-sided 95% confidence intervals (in gray lines) for 30-, 40- and 50-year-old individuals, respectively. The top (respectively, bottom) panel shows results for men (respectively, women) in heavy tails.

We analyze the heaviness of the tails of the conditional distributions of $Y$ given the income level and age in the base year (i.e., the base year is 2007 for the period 2007–2009 and it is 2017 for the period 2017–2019).
Figure 18: Kernel density estimates (black line) of the risk measure $Y$ in 2017 in the PSID. The right (respectively, left) panel show the density of men (respectively, women). Also shown in gray dashed lines are the normal density fit to data.

For 30-year-old men (the top left panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 1.7 to 2.9, and the upper bounds of the one-sided 95% confidence intervals range from 3.1 to 5.3. Given that income received in 2007 was at or below the median, the conditional Pareto exponent is significantly less than four, implying that the conditional kurtosis of income growth does not exist for these lower-income groups of young men. The same conclusion also applies to the very top quantiles (top 5 percent). Overall, the kurtosis barely exists for most of the base-year income levels even if we fail to reject the hypothesis of finite kurtosis.

For 30-year-old women (the top right panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 0.7 to 2.9, and the upper bounds of the one-sided 95% confidence intervals range from 1.4 to 5.4. For this subpopulation, the conditional Pareto exponent is significantly less than four except for the very bottom quantiles (bottom 5 percent), implying that the conditional kurtosis does not exist. Furthermore, given that income received in 2007 was at or above the 15th (respectively, 40th) percentile, the Pareto exponent is significantly less than three (respectively, two), implying that even the conditional skewness (respectively, standard deviation) does not exist. Comparing the results between men and women...
Figure 19: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2007–2009. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
at age 30, we observe that women were more vulnerable to income risk than men, except at the very bottom quantiles of the base year income level.

For 40-year-old men (the middle left panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 1.2 to 1.6, and the upper bounds of the one-sided 95% confidence intervals range from 2.1 to 2.9. Remarkably, the income risk of 40-year-old men are higher than those of 30-year-old men. For this age group of men, we reject the hypothesis of finite kurtosis at any level of base-year income. For 40-year-old women (the middle right panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 0.9 to 2.0, and the upper bounds of the one-sided 95% confidence intervals range from 1.7 to 3.9. For this age group of women, we reject the hypothesis of finite kurtosis at any income level. Furthermore, given that income received in 2007 was at or above the 50th (respectively, 75th) percentile, the Pareto exponent is significantly less than three (respectively, two), implying that even the conditional skewness (respectively, standard deviation) does not exist. Comparing the results between men and women at age 30, observe that men are almost as vulnerable to income risk as women for this middle age group.

For 50-year-old men (the bottom left panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 1.3 to 3.1, and the upper bounds of the one-sided 95% confidence intervals range from 2.5 to 5.8. For this age group of men, the conditional Pareto exponents are relatively high, and we fail to reject the hypothesis of finite conditional kurtosis except at the very bottom quantiles (bottom 5 percent) of base-year income. For 50-year-old women (the bottom right panel in Figure 19), the conditional Pareto exponents (in point estimates) range from 1.0 to 1.5, and the upper bounds of the one-sided 95% confidence intervals range from 1.7 to 2.7. For this age group of women, we reject the hypothesis of finite conditional skewness at any income level. Furthermore, given that income received in 2007 was at or below the median, the Pareto exponent is significantly less than three, implying that even the conditional skewness does not exist.

While our main focus has been on the period of great recession between 2007 and 2009, we next look at the period of positive growth, between 2017 and 2019, ten years later than the baseline period. Figure 20 illustrates the results for income growth between 2017 and 2019 for 30-, 40- and 50-year-old individuals, respectively.
Figure 20: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2017–2019. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
These results share similar qualitative patterns to those reported in Figures 19 as follows. First, 30-year-old men at high quantiles of base-year income enjoy less income risk. Second, 30-year-old women suffer from high income risk except at the bottom quantiles of base-year income. Third, 40-year-old men have overall higher income risk than 30-year-old men. Lastly, and most remarkably, the income risk is not necessarily lower in the period 2017–2019 than in the period 2007–2009, even though the former period enjoyed a positive GDP growth (5.2% in two years) and the latter period suffered from negative GDP growth (−2.7% in two years). With all these similarities, there are differences as well – especially in the graphs that do not resemble across the period between 2017 and 2019 and the period between 2007 and 2009 for 40-year-old women or 50-year-old men. Overall, there are more similarities than differences despite the contrast between a recession and a positive growth in the US economy.

We repeat the econometric analysis for the remaining periods, 2009–2011, 2011–2013, 2013–2015 and 2015–2017 to demonstrate the robustness of the observed patterns reported above. Figures 21, 22, 23 and 24 illustrate estimates of the conditional Pareto exponents $\alpha(x_0)$ (in black lines) along with the upper bounds of their one-sided 95% confidence intervals (in gray lines) for the periods 2009–2011, 2011–2013, 2013–2015 and 2015–2017. In each figure, the left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.

For the periods 2007–2009 and 2017–2019 we found that: 1) the kurtosis, skewness, and even standard deviation may not exist for the conditional distribution of income growth given certain attributes (age, gender, and income); 2) younger women are more vulnerable to income risk than younger men; 3) middle-aged men are almost as vulnerable as middle-aged women; and 4) these patterns appear both in the period 2007–2009 of great recession and the period 2017–2019 of a positive growth, while there are differences as well. The first and second points robustly hold in the remaining periods, 2009–2011, 2011–2013, 2013–2015 and 2015–2017. The third point also continues to hold for the periods, 2009–2011 (to a less extent), 2011–2013 and 2013–2015. Thus, the fourth point largely extends to the remaining periods, 2009–2011, 2011–2013, 2013–2015 and 2015–2017.
Figure 21: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2009–2011. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
Figure 22: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2011–2013. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
Figure 23: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2013–2015. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
Figure 24: Estimates (black lines) and the one-sided 95% confidence intervals (gray lines) of the Pareto exponents $\alpha(x_0)$ of the conditional tail risk for men (left) and women (right) based on the PSID in the period 2015–2017. The left (respectively, right) column shows the results for men (respectively, women). The top, middle, and bottom panels show results for 30-, 40- and 50-year-old individuals.
References

Stuart Adam, David Phillips, and Barra Roantree. 35 years of reforms: A panel analysis of the incidence of, and employee and employer responses to, social security contributions in the uk. *Journal of Public Economics*, 171:29–50, 2019.

Manuel Arellano, Richard Blundell, and Stéphane Bonhomme. Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica*, 85(3): 693–734, 2017.

Manuel Arellano, Stéphane Bonhomme, Micole De Vera, Laura Hospido, and Sizi Wei. Income risk inequality: evidence from spanish administrative records. Working Paper, 2021.

Brendan K Beare and Alexis Akira Toda. Determination of pareto exponents in economic models driven by markov multiplicative processes. *arXiv preprint arXiv:1712.01431*, 2017.

Brian D Bell, Nicholas Bloom, and Jack Blundell. This time is not so different: Income dynamics during the covid-19 recession. 2021.

P. K. Bhattacharya. *Induced Order Statistics: Theory and Applications*, chapter 18, pages 383–403. Elsevier Sicence Publishers, 1984.

Stéphane Bonhomme and Jean-Marc Robin. Assessing the equalizing force of mobility using short panels: France, 1990–2000. *The Review of Economic Studies*, 76(1):63–92, 2009.

Stéphane Bonhomme and Jean-Marc Robin. Generalized non-parametric deconvolution with an application to earnings dynamics. *The Review of Economic Studies*, 77(2):491–533, 2010.

Irene Botosaru and Yuya Sasaki. Nonparametric heteroskedasticity in persistent panel processes: An application to earnings dynamics. *Journal of Econometrics*, 203(2):283–296, 2018.
Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018. doi: 10.1111/ectj.12097.

Laurens de Haan and Ana Ferreira. *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. Springer, NY, 2007.

Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz-Pardo. Wage risk and government and spousal insurance. Technical report, National Bureau of Economic Research, 2021.

Michael WL Elsby, Donggyun Shin, and Gary Solon. Wage adjustment in the great recession and other downturns: Evidence from the united states and great britain. *Journal of Labor Economics*, 34(S1):S249–S291, 2016.

Xavier Gabaix and Rustam Ibragimov. Rank$^{-1/2}$: A simple way to improve the OLS estimation of tail exponents. *Journal of Business and Economic Statistics*, 29(1):24–39, January 2011. doi: 10.1198/jbes.2009.06157.

Laurent Gardes and Stéphane Girard. A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99:2368–2388, 2008.

Laurent Gardes, Armelle Guillou, and Antoine Schorgen. Estimating the conditional tail index by integrating a kernel conditional quantile estimator. *Journal of Statistical Planning and Inference*, 142:1586–1598, 2012.

John Geweke and Michael Keane. An empirical analysis of earnings dynamics among men in the psid: 1968–1989. *Journal of econometrics*, 96(2):293–356, 2000.

Mikhail Golosov, Maxim Troshkin, and Aleh Tsyvinski. Redistribution and social insurance. *American Economic Review*, 106(2):359–86, 2016.

Maarten Goos and Alan Manning. Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics*, 89(1):118–133, 2007.
Armelle Guillou and Peter Hall. A diagnostic for selecting the threshold in extreme value analysis. *Journal of Royal Statistic Society, Series B*, 63:293–305, 2001.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of us workers reveal about life-cycle earnings dynamics? Federal Reserve Bank of New York Staff Report, 2019.

Bruce M. Hill. A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174, September 1975. doi: 10.1214/aos/1176343247.

Joel L Horowitz and Marianthi Markatou. Semiparametric estimation of regression models for panel data. *The Review of Economic Studies*, 63(1):145–168, 1996.

Yingyao Hu, Robert Moffitt, and Yuya Sasaki. Semiparametric estimation of the canonical permanent-transitory model of earnings dynamics. *Quantitative Economics*, 10(4):1495–1536, 2019.

Robert Moffitt and Sisi Zhang. Income volatility and the psid: Past research and new results. In *AEA Papers and Proceedings*, volume 108, pages 277–80, 2018.

Stephen Nickell and Glenda Quintini. Nominal wage rigidity and the rate of inflation. *The Economic Journal*, 113(490):762–781, 2003.

Yuya Sasaki and Yulong Wang. Fixed-\(k\) inference for conditional extremal quantiles. *Forthcoming in Journal of Business and Economic Statistics*, 2021.

Richard L. Smith. Estimating tails of probability distributions. *Annals of Statistics*, 15(3):1174–1207, September 1987. doi: 10.1214/aos/1176350499.

Lorenzo Trapani. Testing for (in) finite moments. *Journal of Econometrics*, 191(1): 57–68, 2016.

Hansheng Wang and Chi-Ling Tsai. Tail index regression. *Journal of the American Statistical Association*, 104:1233–1240, 2009.
Huixia Judy Wang and Deyuan Li. Estimation of extreme conditional quantiles through power transformation. *Journal of the American Statistical Association*, 108(503):1062–1074, 2013. doi: 10.1080/01621459.2013.820134.

**Data Citation:** Office for National Statistics. (2017). New Earnings Survey Panel Dataset, 1975-2016: Secure Access. [data collection]. 7th Edition. UK Data Service. SN: 6706, [http://doi.org/10.5255/UKDA-SN-6706-7](http://doi.org/10.5255/UKDA-SN-6706-7)