Quantitative Economics 13 (2022), 1321–1360 1759-7331/20221321

Global trends in income inequality and income dynamics: New insights from GRID

Fatih Guvenen
Department of Economics, University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER

Luigi Pistaferri
Department of Economics, Stanford University, SIEPR, and NBER

Giovanni L. Violante
Department of Economics, Princeton University, CEPR, and NBER

The Global Repository of Income Dynamics (GRID) is a new open-access, cross-country database that contains a wide range of micro statistics on income inequality, dynamics, and mobility. It has four key characteristics: it is built on micro panel data drawn from administrative records; it fully exploits the longitudinal dimension of the underlying data sets; it offers granular descriptions of income inequality and income dynamics for finely defined subpopulations; and it is designed from the ground up with the goals of harmonization and cross-country comparability. This paper introduces the database and presents a set of global trends in income inequality and income dynamics across the 13 countries that are currently in GRID. Our results are based on the statistics created for GRID by the 13 country teams who also contributed to this special issue with individual articles.

Keywords. Administrative data, cross-country, database, granular, harmonized, inequality, longitudinal, mobility, volatility.

JEL classification. E24, J24, J31.

Fatih Guvenen: guvenen@umn.edu
Luigi Pistaferri: pista@stanford.edu
Giovanni L. Violante: violante@princeton.edu

GRID is a joint initiative between the Minnesota Economics Big Data Institute at the University of Minnesota, Princeton University, and Stanford University. GRID can be accessed at https://www.grid-database.org/. We owe special thanks to Serdar Ozkan (University of Toronto) and Sergio Salgado (Wharton School of the University of Pennsylvania), who wrote and regularly updated the master code used by all country teams to produce statistics. We are grateful to all country team members who supported the project from the very beginning and gave us vital feedback on many aspects of the empirical analysis. We also thank Francisco Bullano for outstanding research assistance on this paper, and Kjetil Storesletten and Chris Taber for having supported this special issue from its inception. Finally, we thank the institutions that provided funding for the GRID Project: the Data-Driven Social Science Initiative and Simpson Center at Princeton University, Stanford University, Heller-Hurwicz Economics Institute and the Minnesota Economics Big Data Institute (MEBDI) at University of Minnesota, Litigation Analytics, and the Washington Center for Equitable Growth. All the statistics available in GRID and used in this paper have been cleared for distribution by the relevant authorities in respective countries. The US statistics were released by the US Census Bureau on June 15, 2022, with clearance number CBDRB-FY22-283.

 https://doi.org/10.3982/QE2260

1. Introduction

In the last few decades, interest in the distribution of income has grown enormously in academia, policy circles, and popular media.[1] This trend is motivated by many concurring factors: its global nature (affecting countries as diverse in their institutions as the United States, Sweden, and China); the observation that higher inequality is paired, in some countries, with lower economic mobility; the rise of top income shares in some countries; the concern that extreme income disparity may distort the political process, and thus undermine the proper functioning of democracies; and finally, the belief that the key forces behind this transformation—technology and trade liberalization—have also generated widespread prosperity, and thus involve a complex trade-off between growth and inequality.[2]

Today, virtually every area of economics is contributing to the inequality debate. A rough gauge of the growing interest is the number of academic publications on the subject. Between 1980 and 1989, only 38 articles published in the top five economics journals included the word "inequality" in their abstracts. This number increased to 59 in the 1990–1999 period, 89 in 2000–2009, and 148 in 2010–2018.[3] For the conversation to progress in the right direction, we believe that economists need rich microdata that accurately represent the evolution of the income distribution in all of its many facets. A number of cross-country databases that document trends in world income inequality already exist—most notably, the World Inequality Database (WID), the World Income Inequality Database (WIID) at the United Nations University, the Luxembourg Income Study, and several other data sources hosted by the OECD. Most of these databases have two defining characteristics: they are cross-sectional, not longitudinal and, therefore, do not provide information on income dynamics or economic mobility; and second, they provide statistics for fairly aggregated demographic groups, with limited information on finely defined subpopulations.[4]

This special issue of *Quantitative Economics* introduces the Global Repository of Income Dynamics (GRID), a new open-access, cross-country database of harmonized micro statistics that overcomes these limitations.[5] The current (launch) version of GRID includes 13 countries: Argentina, Brazil, Canada, Denmark, France, Germany, Italy, Mexico, Norway, Spain, Sweden, the United Kingdom, and the United States. These countries were chosen not only for the availability of suitable administrative data but also because they represent a fairly broad spectrum of levels of development and institutions. GRID has been developed over the past 4 years with the participation of more

[1] Throughout the paper, we use the terms "earnings" and (labor) "income" interchangeably.

[2] The *IFS Deaton Review* offers a comprehensive and thorough discussion about how inequalities arise, which ones matter, why they matter, and how they should be addressed; see https://ifs.org.uk/inequality/directory/.

[3] Similarly, a Google Scholar search of the term "income inequality" turns up 12,900 articles between 1950 and 1989, 130,000 between 1990 and 2007, and 323,000 since 2008.

[4] The project most closely related to GRID is the OECD's LinkEED, which uses employer-employee matched panel data from 17 countries (see Criscuolo et al. (2020)). The LinkEED project produced a book on the role of firms in wage inequality (, 2021) but does not have a publicly available database.

[5] The GRID database can be accessed at https://www.grid-database.org/.

than 50 economists in 13 country teams. In addition to producing the GRID statistics, each team has also written an article about their respective countries, and together with the present paper, these articles make up this special issue.

The goal of the present paper is twofold. First, we provide a brief introduction to GRID and discuss its key features, as well as explore the ways in which it is similar to or different from existing databases. Second, we present a series of global trends in income inequality and income dynamics that we identified from a bird's-eye view of the GRID statistics across all countries. The global trends we present are by no means exhaustive. Rather, we present these stylized facts as examples of interesting regularities that can be discovered by the sort of cross-country comparisons that GRID easily allows. We begin with an overview of GRID, which is built on four pillars: longitudinal, administrative, granular, and harmonized.

Longitudinal

The first feature of GRID is its longitudinal (or panel) dimension, which enables researchers to study the dynamics of individual (labor) income over time, as opposed to static snapshots of distributions as cross-sectional inequality measures do. This distinction is crucial for any welfare analysis and for designing redistributive and social insurance programs.[6] A key design goal for GRID was that all statistics for every country would be computed from administrative *panel* data, which allows us to analyze the entire distribution of individual income *changes* (including the tails), document the nature of income risk that workers face (e.g., the size of individual income shocks, their persistence, and how they vary with the business cycle), and estimate the rank mobility of individuals within the income distribution both over the life cycle and, potentially, across successive generations.[7]

This wealth of additional statistics greatly enriches our understanding of the dynamics of income distributions. The usefulness for applied economists extends even beyond this point, because income dynamics are a key input into structural models used for quantitative analysis, policy counterfactuals, and welfare calculations.

Administrative

All country-level data sets come from administrative records (e.g., social security records and other government registers). Administrative data offers several advantages over survey data. First, by their nature, survey data suffer from sample attrition, measurement

[6]A classic example that illustrates the pitfall of cross-sectional snapshots is their inability to draw welfare inferences from a constant poverty rate. A constant 10% poverty rate across 2 subsequent years is compatible with 10% of the population being permanently poor or with the entire population facing a 10% chance every year of falling into poverty (as well as the more realistic spectrum of intermediate cases). With cross-sectional data, the two cases are indistinguishable; with panel data, one can follow the fortunes of people over time, and the distinction becomes immediate.

[7]In its current version, GRID has no intergenerational component; we plan to study the feasibility of adding that component in the future.

error, and lack of representativeness of the tails (especially at the top).[8] Moreover, because of their small size, they run into statistical power problems when trying to produce nonparametric analyses. These issues are not present with administrative data, which collect information on either the entire population or very large random samples that are tracked over time using government identifiers (such as social security numbers). This approach makes attrition mostly a nonissue. Moreover, measurement error is minimal since income data are not self-reported but instead reported by third parties (typically, the employer) and misreporting is subject to heavy penalties by government authorities. Sample sizes in the millions (and sometimes tens of millions) of observations per year allow granular analyses for detailed subpopulations, estimation of tail statistics (e.g., the top 0.1% share), measurement of higher-order moments of the data (e.g., skewness and kurtosis), and nonparametric representations. Of course, also administrative data have their own limitations. For example, in countries where the informal sector is large, they can miss a significant share of the population. For this reason, the articles in this special issue that study developing countries also contain comparisons to survey data where informal workers are better captured. A final consideration is that, for most researchers, access to administrative microdata may be prohibitively costly or infeasible. The vast menu of statistics on the income distribution and income dynamics available in the GRID database should substantially alleviate the need to access the underlying administrative microdata for many users.

Granular

GRID provides micro statistics on income inequality, income fluctuations, and mobility for finely defined subpopulations. The initial database that is currently online includes statistics by year, age, gender, and a measure of permanent income as a proxy for the skill level.[9] This last characteristic is the most detailed: statistics are available for every 2.5 percentile bin group as well as for the top 1% and 0.1% groups, for a total of 42 permanent income groups. Many studies document that income inequality, risk, and mobility vary substantially across demographic groups. Computing disaggregated statistics allows researchers to study the different dynamics of these groups and to separate the role of ex ante heterogeneity in types from that of ex post uncertainty in outcomes.

Harmonized

A primary goal of the GRID project is to produce statistics that are as comparable as possible across countries. Harmonization is an inherently challenging task, given the discrepancies in variable definitions and data collection methods in different countries. We spent a great deal of effort in harmonizing the output produced by each country:

[8] Households have become increasingly less likely to answer surveys, and when they do provide answers, they are less likely to be accurate (Meyer, Mok, and Sullivan (2015)). These threats to survey quality are present in many of the most important US data sets for social science research and government policy (National Research Council, 2013).

[9] Permanent income is defined as average individual income in the preceding 3 years.

all statistics for all countries are produced by one unique master code, which ensures that a long list of small, but potentially critical, steps are carried out the same way in each country. The code is minimally tailored to a specific country only when absolutely needed. Section 2 describes in more detail the common guidelines followed by each country for sample selection and variable construction. This harmonization effort means that adding additional countries to GRID in the future will be relatively easy.

Besides the present article, this special issue includes articles for each of these 13 countries. The articles all follow the same two-part template. The first part ("common core," usually Section 3 and possibly Section 4) is harmonized and reports the same set of statistics on: (i) the evolution of the cross-sectional distribution of income in levels, in order to paint a picture of the trends in income inequality at the bottom and top of the distribution; (ii) a detailed analysis of income volatility, asymmetry (skewness), and the importance of tail events (i.e., the probability of extreme drops or rises in earnings, reflected in kurtosis among other statistics) in learning about the nature of labor market risk workers face; and an estimate of intragenerational income rank mobility to assess the gap between current (or short-run) income inequality and permanent (or long-run) income inequality. These analyses are conducted for the whole population as well as separately by gender and other demographic characteristics.

In the second part (Section 4 or 5) of the papers, each team has leveraged its own data set to study a topic of special relevance for that country. Blanco, de Astarloa, Drenik, Moser, and Trupkin (2022) quantify nominal wage rigidity at times of low and high inflation in Argentina. Engbom, Gonzaga, Moser, and Olivieri (2022) combine administrative and survey data to study the wage dynamics of workers switching between the formal and informal sectors in Brazil. Bowlus, Gouin-Bonenfant, Liu, Lochner, and Park (2022) document the relationship between the earnings dynamics of workers and the size and growth of their employers in Canada. Leth-Petersen and Saeverud (2022) compare the dynamics of gross labor income with those of disposable income in Denmark. Kramarz, Nimier-David, and Delemotte (2022) examine the spatial dimension of inequality in France. Drechsel-Grau, Peichl, Schmieder, Schmid, Walz, and Wolter (2022b) compare the earnings dynamics of workers and entrepreneurs in Germany. Hoffmann, Malacrino, and Pistaferri (2022) investigate the role of structural labor market reforms in shaping earnings dynamics in Italy. Puggioni, Calderon, Zurita, Bujanda, Gonzalez, and Jaume (2022) study how time away from formal employment shapes future earnings in Mexico. Halvorsen, Ozkan, and Salgado (2022) analyze the intergenerational transmission of income dynamics in Norway. Arellano, Bonhomme, De Vera, Hospido, and Wei (2022) separate the predictable component from the uncertain component of earnings to obtain more accurate measures of individual income risk in Spain. Friedrich, Laun, and Meghir (2022) show the role of the generous but evolving welfare state in determining the earnings dynamics of immigrants and natives in Sweden. Bell, Bloom, and Blundell (2022) estimate the responsiveness of earnings and hours in the UK to firm-level shocks and aggregate shocks, including the Covid-19 recession. McKinney, Abowd, and Janicki (2022) study long-term average earnings differentials across workers of different races and ethnicities in the US. The remarkable breadth of the topics investigated illustrates the research potential of our database as well as its usefulness in many fields of economics.

We now provide a brief summary of the global trends in income inequality and income dynamics that we present in the rest of the paper.[10]

On cross-sectional income inequality, we document four global stylized empirical facts. First, GRID countries do not display any discernible global trend toward rising income inequality, despite the often-repeated assertions to that effect. In fact, inequality remains fairly stable in about half of the countries, with the rest evenly split between those experiencing a rise and those experiencing a decline in inequality (see Figure 2). Perhaps because of their rapid economic development or because of a declining incidence of informal labor, Latin American countries actually record declining income inequality. On the other side of the spectrum are continental European countries, where inequality rises, perhaps as a result of the effects of various waves of labor market reforms. Second, in countries where inequality rises (or declines) significantly, the right and left tails both widen (or shrink), whereas in countries with small changes in inequality, the two tails go in the opposite direction (one expands while the other shrinks). Third, in the vast majority of countries, income levels at the very top (top 1% and 0.1%) do not show very fast trend growth relative to historical growth rates in GDP per capita and in average wages. In countries where the top income *shares* have grown fast, this growth reflects the stagnation of earnings for the rest of the population, especially for those below the median, rather than accelerating growth at the top. Fourth, the gap between the dispersion in women's and men's earnings has closed in many countries: the convergence is toward gender equality in the levels of earnings inequality.

Turning to the distribution of income *growth*, we see a remarkably homogeneous picture across countries, with a few exceptions. In particular, in all countries, the density of income growth has a very large variance, peaks at the center, and has thick Pareto tails, resulting in very high kurtosis. In addition, the left tail is thicker than the right, giving rise to negative skewness for all countries except Italy and Mexico. Furthermore, in all countries, skewness comoves with the business cycle in a robustly procyclical fashion. These skewness fluctuations are driven by both the upper tail (income changes above the median) compressing and the lower tail expanding in recessions, and vice versa in expansions. As a result, idiosyncratic income risk is countercyclical and asymmetric over the business cycle.

The remarkable homogeneity across countries carries over to the patterns by permanent income and age.[11] In particular, the dispersion of income growth rates declines up to about the 80th to 95th percentiles of the permanent income income distribution, after

[10]The 2010 special issue of the *Review of Economic Dynamics*, "Cross-sectional Facts for Macroeconomists," is a relevant precedent, as well as a source of inspiration for us, because it also aimed at organizing in a coherent way stylized facts on income inequality across countries. It did, however, differ from our project in three key aspects: it was built on easily accessible survey data, it almost exclusively exploited the cross-sectional dimension of the data, and it did not aim at building a global database. See Krueger, Perri, Pistaferri, and Violante (2010) for an introduction to that special issue.

[11]At a given age, permanent income is proxied by the average income of the individual in the preceding 3 years.

which point it rises sharply, forming a hockey-stick shape (see Figure 9). In fact, in most countries, the volatility of income growth is higher for the top 1%–2% group than for the bottom 10% group. Volatility also declines with age, especially between the early part of the life cycle and the middle part, a pattern that is common to almost all countries. The pattern for skewness is very similar to that for volatility, with skewness becoming more negative with permanent income up to about the 70th to 90th percentiles, then reverting after that to form a similar hockey-stick shape (see Figure 10). Finally, kurtosis displays the mirror image (upside down) of the hockey-stick shape described: rising for most of the permanent income range and then declining at the very top. The maximum kurtosis level reached is extremely high—as high as 30 to 40 for middle-age men around the 90th percentile of the permanent income distribution. For all three statistics, the only exceptions to the hockey-stick shape are Brazil, Mexico, and (partially) Italy, three countries that are somewhat affected by top coding, which may explain the distinct patterns at the top end.

The analysis of cross-sectional inequality summarizes properties of the income distribution. The study of individual income *changes* describes how workers' income evolves over time. Also of interest is understanding the extent to which these income changes reshuffle workers' relative positions within the distribution itself. For this purpose, each country team also computed measures of 5- and 10-year rank mobility. Overall, we uncover fairly sizable differences in the degree of income mobility across the countries in our database. Perhaps not surprisingly, Scandinavian countries feature the lowest degree of income persistence, and some of the Latin American countries, together with Italy and France, feature the highest. Intragenerational mobility is, in general, higher for women and younger workers. We discerned no significant time trend in life-cycle mobility, suggesting that the trends in cross-sectional earnings we documented have largely translated into similar trends for permanent earnings. Finally, we document the existence of a negative cross-country relationship between rank mobility and inequality, an intragenerational version of the so-called Great Gatsby curve (Durlauf and Seshadri (2018)).

The GRID project, as well as the papers in this special issue, build upon two separate and vast literatures on income inequality and income dynamics, respectively. This limited space precludes our doing justice to a thorough review of the work in this area. See Katz and Autor (1999) and Acemoglu and Autor (2011) for surveys of the income inequality literature, and Meghir and Pistaferri (2011) and Altonji, Hynsjö, and Vidangos (2022) for reviews of the income dynamics literature that goes back to the 1970s. The approach to income dynamics in GRID is most closely related to a recent strand of literature that emphasizes nonparametric approaches in measurement and modeling that allow for nonlinearities and nonnormalities in income dynamics. Starting with Geweke and Keane (2000) and followed by Bonhomme and Robin (2009), Guvenen, Ozkan, and Song (2014), and Arellano, Blundell, and Bonhomme (2017), among others, this literature emphasizes higher-order moments such as skewness and kurtosis as well as the nonlinear persistence of income shocks. The implications of these features for a wide range of economic questions, from taxation to monetary policy, asset pricing, and others, are increasingly being studied in recent work. As for socioeconomic mobility, Fields and Ok (1999) survey the literature and discuss the importance of accounting for mobility in analyses of income inequality. Finally, our paper draws some inspiration from the overview paper of the 2010 special issue of the *Review of Economic Dynamics* by Heathcote, Perri, and Violante (2010) as well as from the excellent papers that make up this special issue of *Quantitative Economics*.

The rest of the article is organized as follows. Section 2 gives an overview of the country data sets and the common variable definitions. The next sections discuss the stylized facts on, respectively, cross-sectional income inequality (Section 3), distribution of income growth (Sections 4 and 5), and life-cycle income mobility (Section 6) that altogether emerge from the 13 countries. Section 7 concludes.

2. Data sets, variables, and sample selection

Table 1 gives an overview of the key characteristics of each of the underlying databases used in GRID. All 13 data sets in our database are, as explained, of an administrative nature and assembled by government agencies.[12] The sample period covers at least 20 years for 10 of the countries, averaging 26 years over the 13 countries. Spain has the shortest sample period (14 years) and the UK the longest (45).

Income data are originally recorded at monthly to annual frequencies (with the exception of weekly data in the UK) and aggregated to an annual frequency when needed for calculating all GRID statistics. The data are not top-coded for 10 of the 13 countries. The exceptions are Brazil, which has a very high threshold of 120 times the minimum wage, and Italy and Mexico, which have somewhat lower thresholds that nevertheless bind for less than a few percent of the population. Germany has two data sets that are used jointly in GRID: the IAB from the Social Security administration and TPP from the tax authorities. The former has been used extensively in past research (e.g., Card, Heining, and Kline (2013), Song, Price, Guvenen, Bloom, and Von Wachter (2019)) but has fairly severe top coding (about 10% of the population), whereas the latter is nontop-coded but bottom-coded as a result of nonfiling. The Germany team synthetically combined these two data sets to obtain statistics that do not suffer from bottom coding or top coding. As for sample size, for 7 out of 13 countries (Brazil, Canada, Denmark, Mexico, Norway, Sweden, US), the data sets have nearly complete coverage of the relevant population, and the remaining data sets cover about 3% to 25% of the population with the exception of the UK, which as 1% coverage. The size of the final cross-sectional sample (defined below) varies from about 100,000 individuals for Argentina and the UK to 2 to 3 million individuals for the middle group of countries (Denmark, Norway, and Sweden) to as high as 25 million for Germany, 45 million for Brazil, and 95 million for the US.

[12] While the UK data set is technically a survey, it is a survey of firms (by the Office of National Statistics) rather than of households, so earnings data are not self-reported, and the data set has broad and consistent coverage over time.

Table 1. Summary characteristics of administrative data sets used in GRID.

	Argentina	Brazil	Canada	Denmark	France	Germany	Italy	Mexico	Norway	Spain	Sweden	UK	US
Data set name	RELS (Registered Employment Longitudinal Sample)	RAIS (Relação Anual de Informações Sociais)	CEEDD (Canadian Employer–Employee Dynamics Database)	*Several registry databases*	DADS (Déclaration Annuelle des Données Sociales)	IAB/TPP (Integrated Employment Biographies/ German Taxpayer Panel)	INPS (Istituto Nazionale della Previdenza Sociale) LoSal	IMSS (Instituto Mexicano del Seguro Social)	Inntekts-og formues-registeret	MCVL (Muestra Continua de Vidas Laborales)	LOUISE (Longitudinel databas kring utbildning, inkomst och sysselsä ttning)	ASHE (Annual Survey of Hours and Earnings)	LEHD (Longitudinal Employer-Household Dynamics)
Income record frequency	Monthly	Monthly	Annual	Annual	Job spell	IAB (Job spell)/TPP (annual)	Job spell	Monthly	Annual	Job spell	Annual	Weekly earnings	Quarterly
Top coding	No (pooled)	120 × MW	No	No	No	IAB (Yes), TPP (No)	Yes (645 euro daily threshold)	25 × MW	No	No	No	No	No
Bottom coding	No	No	No	No	No	IAB (No)/TPP (nonfilers)	No	1 × MW	No	No	No	No	No
Time span	1996–2015	1985–2018	1983–2016	1987–2016	1991–2016	2001–2016	1985–2016	2005–2019	1993–2017	2005–2018	1985–2016	1975–2020	1998–2019
Number of years	20	34	34	30	26	16	32	15	25	14	32	45	22
Available data set total size	3% random sample/ 130K to 230K	Population/ 40M	89–93% of all ages 25–55 Canadians	Population	4% random sample	IAB (10% R.S.)/TPP (25% R.S.)	6.6% R.S.	Population/ 17M to 26M	Population	4 % Random Sample	Population	1% Sample/140K to 180K	Population (excludes a few states)
CS sample† [min-max]	97.2K to 167.6K	15.7M to 45.9M	8 Mto 10.8M	1.8M to 1.9M	0.58 M to 1.34 M (2002 ×2)	IAB(23.1 to 24.9 M)/TPP(16 to 22.4 M)	700K	12.4M to 19.6M	1.96M to 2.0M	405K to 463K	2.95M to 3.2M	93.1K to 121K	82.1M to 95.6M

Note: All data are aggregated to an annual level to calculate GRID statistics. In each country, the databases are produced by the following agencies: RELS by the Ministry of Labor, Employment, and Social Security; RAIS by Ministério da Economia; CEEDD by Statistics Canada; Danish data registries by Statistics Denmark; DADS by INSEE; IAB by Institute for Employment Research and TPP by Research Data Centre of the Statistical Offices of the Federal States; INPS-LoSal by INPS (Istituto Nazionale della Previdenza Sociale); IMSS data by Instituto Mexicano del Seguro Social; Inntekts-og formuesregisteret by Statistics Norway; MCVL by Dirección General de Ordenación de la Seguridad Social; LOUISE by Statistics Sweden; ASHE by ONS (Office of National Statistics); and the LEHD by the US Census Bureau. †CS sample is the cross-sectional sample as defined in the text. R.S. stands for random sample.

Sample construction

To enhance harmonization and allow meaningful comparisons across countries in the project, we start by imposing three common restrictions. First, we focus on workers between 25 and 55 years old, a range within which most education choices are usually completed and after which workers tend to leave the labor force for retirement. Second, for most of the analysis we drop observations with earnings (defined next) below a threshold (call it $\underline{y}$) to avoid using records from workers without a meaningful attachment to the labor force or with very low earnings, which could skew log-based statistics. Specifically, we discard observations with earnings below what workers would earn if they were to work part-time for one quarter at the national minimum wage. For countries without a national minimum wage, we have used the US-specific threshold (in PPP terms). Third, for the countries where labor income is top-coded (Brazil, Italy, and Mexico) we use an imputation procedure.

Each team constructed three separate samples to be used for different parts of the analysis:[13]

1. The cross-sectional (CS) sample is the one used to compute cross-sectional inequality statistics. All individuals who satisfy the three criteria above are in this sample at date t. This sample is the most comprehensive and uses the longest possible time series available.

2. The longitudinal (LX) sample is used to study the distribution of earnings changes. It includes all individuals in the CS sample who, in addition, have 1-year and 5-year forward earnings changes.

3. The heterogeneity (H) sample is used to study variation across demographic groups defined by observable characteristics (such as age, gender, and permanent income). It includes all individuals in the LX sample for whom, in addition, a permanent earnings measure (see the definition below) can be constructed. For this sample, we only select the last 15–20 years available and always pool observations across years.[14]

Variable definitions

Our main variable of interest is annual individual labor earnings (i.e., market income from employment services) comprehensive, whenever possible, of bonuses, overtime pay, tips, commissions, and so on, earned from all jobs held during the calendar year but excluding self-employment income.[15] We asked country teams to construct several measures of earnings for worker i in year t:

[13] Additional documentation with greater detail and exact definitions of all the variables in the database is available on the GRID website: https://www.grid-database.org/documentation.

[14] Even when a very long panel is available for a country, using the full length is not ideal for analyses of heterogeneity, since economies evolve over time and, in some cases, change quite dramatically. For example, many countries have experienced rising female labor force participation, population aging, increasing immigration, skill-biased technical change, and so on, which makes more recent periods potentially more informative about today and the future.

[15] We exclude self-employment income because it is not available in some of the pilot countries.

1. Raw real earnings in levels, y_{it}, and logs, $\log(y_{it})$. Real earnings are computed from nominal earnings and a measure of CPI inflation for each country.

2. Residualized log earnings, ε_{it}. This measure is the residual from a regression of log real earnings on a full set of age dummies,[16] separately for each year and gender. It is intended to control for predictable changes in individual earnings (life cycle and business cycle effects).

3. Permanent earnings, P_{it-1}. They are defined as average earnings over the previous 3 years, $P_{it} = \sum_{s=t}^{t-2} y_{is}/3$, where y_{is} can include earnings below $\underline{y}$ for at most 1 year. The measure is intended to average over transitory income changes and proxy for skill levels.

4. 1-year change in residualized log earnings, g_{it}^1. It is the 1-year forward change in ε_{it}, defined as $g_{it}^1 = \Delta\varepsilon_{it} = \varepsilon_{it+1} - \varepsilon_{it}$, where earnings must be above $\underline{y}$ for both years.[17]

5. 5-year change in residualized log earnings, g_{it}^5. It is the 5-year forward change in ε_{it}, defined as $g_{it}^5 = \Delta\varepsilon_{it} = \varepsilon_{it+5} - \varepsilon_{it}$, where earnings must be above $\underline{y}$ for both years.

We now proceed to summarize the stylized facts that emerge from a systematic analysis of the statistics in the common core of the 13 country papers in this special issue.

3. Income inequality

In this section, we summarize some key trends in income inequality across the 13 countries that are part of the GRID project.

3.1 *Measuring inequality*

For men and women, separately as well as for the combined population, the papers in this volume report key percentiles (10th, 25th, 50th, 75th, 90th) of the cross-sectional distribution of various measures of income: log earnings, residualized log earnings, and permanent earnings. This was done for each year. We also asked the country teams to compute statistics that would give a more granular view of the top part of the distribution (the 95th, 99th, 99.9th, and 99.99th percentiles), as well as more traditional measures of inequality, such as the standard deviation and the Gini coefficient.[18]

As mentioned above, other publicly available databases contain information about trends in cross-sectional inequality for the countries we study in this special issue. Comparing inequality trends in GRID with those from other existing databases is thus instructive. Because of its ease of use, we focus on the WID and on a broader measure of inequality, the Gini coefficient.

[16] In the next iterations of the project, we plan to add controls by race, education, and geographical location when available.

[17] We take "leads" to avoid mechanical mean reversion when conditioning on permanent earnings. In other words, for a given year t, the statistics we are interested in are calculated using t-years forward, whereas the permanent earnings measure which the statistics are conditioned on is computed for years $t-1$ and earlier, avoiding any overlapping years.

[18] Interested researchers will find many more statistics on the GRID website.

Figure 1 shows that the trends in overall cross-sectional inequality for the countries in the GRID project are very similar to those produced from the WID database.[19] The exception is Brazil, where our administrative data show declining inequality and the WID source shows a slightly increasing pattern, albeit with considerable year-to-year volatility.[20] It is possible that coverage accounts for the differences: administrative data miss the informal sector, which should be represented in survey data, and survey data have other problems discussed above. On the other hand, trends for Argentina and Mexico, where informality is also high, show little discrepancy between GRID and WID.

Next, we turn to GRID to discuss the inequality trends in more detail—how they differ in different parts of the distribution and the extent to which they differ for men and women. To reduce the influence of compositional effects arising from life cycle, business cycle, or increased female participation, the analysis below uses summary statistics for the cross-sectional distribution of residualized log earnings, ε_{it}.

From information on the percentiles of the distribution, we can immediately obtain summary measures of cross-sectional inequality, such as the difference between the 90th and 10th percentiles of the (residualized) log income distributions ($P9010$ from now on). For example, a $P9010$ of 2.5 means that the individual at the 90th percentile of the earnings distribution makes 12.2 (i.e., $e^{2.5}$) times more than what the individual at the 10th percentile makes.[21] An increase in earnings dispersion may come from either the income-rich getting richer or from the income-poor getting poorer (at least in relative terms) or both. To get a sense of what drives trends in inequality, one can decompose $P9010$ into two components: the 90–50 percentile difference, or $P9050$ (which measures the gap between high-income workers and the median worker), and the 50–10 percentile difference, or $P5010$ (the gap between low-income workers and the median worker). This is a simple additive decomposition because $P9010 = P9050 - P5010$. For brevity, we will refer to $P9050$ as inequality "at the top" or "above the median" and to $P5010$ as inequality "at the bottom" or "below the median."

Two facts that are apparent from looking at the GRID countries as a whole are that trends in cross-sectional inequality are not homogeneous and that watershed events

[19]We use the P9010 statistic from GRID for the UK because the Gini coefficient was not authorized for disclosure at the time of this writing. Note that there are differences in the levels of the Gini coefficients because GRID and WID define the income variable differently (in GRID we use income from employment, while the concept of income used in WID is pre-tax personal income, which includes income from labor and capital, social insurance benefits minus the corresponding contributions and excluding other forms of redistribution). The population of reference is also different. In GRID, it is individuals aged 25-55; in WID, it is individuals over age 20. The base unit is the individual but resources are split equally within couples when a household income concept is available.). The goal here is simply to show that broad measures of cross-sectional income inequality are similar across the two databases.

[20]For Argentina, Brazil, and Mexico, WID statistics come from survey microdata harmonized by the Statistics Division of the United Nation's Economic Commission for Latin America and the Caribbean (ECLAC).

[21]One advantage of the $P9010$ as a summary measure of inequality is that it is robust to deviations from normality. Nonetheless, each country team was also asked to produce, alongside the $P9010$, the standard deviation of the log income distribution (σ), or more precisely, $\tilde{\sigma} = 2.56 \times \sigma$, which under normality coincides approximately with the $P9010$. We asked each team to produce a graph plotting $P9010$ and $\tilde{\sigma} = 2.56 \times \sigma$ against time. This way, readers are offered a direct visual view of the extent of deviations from normality.

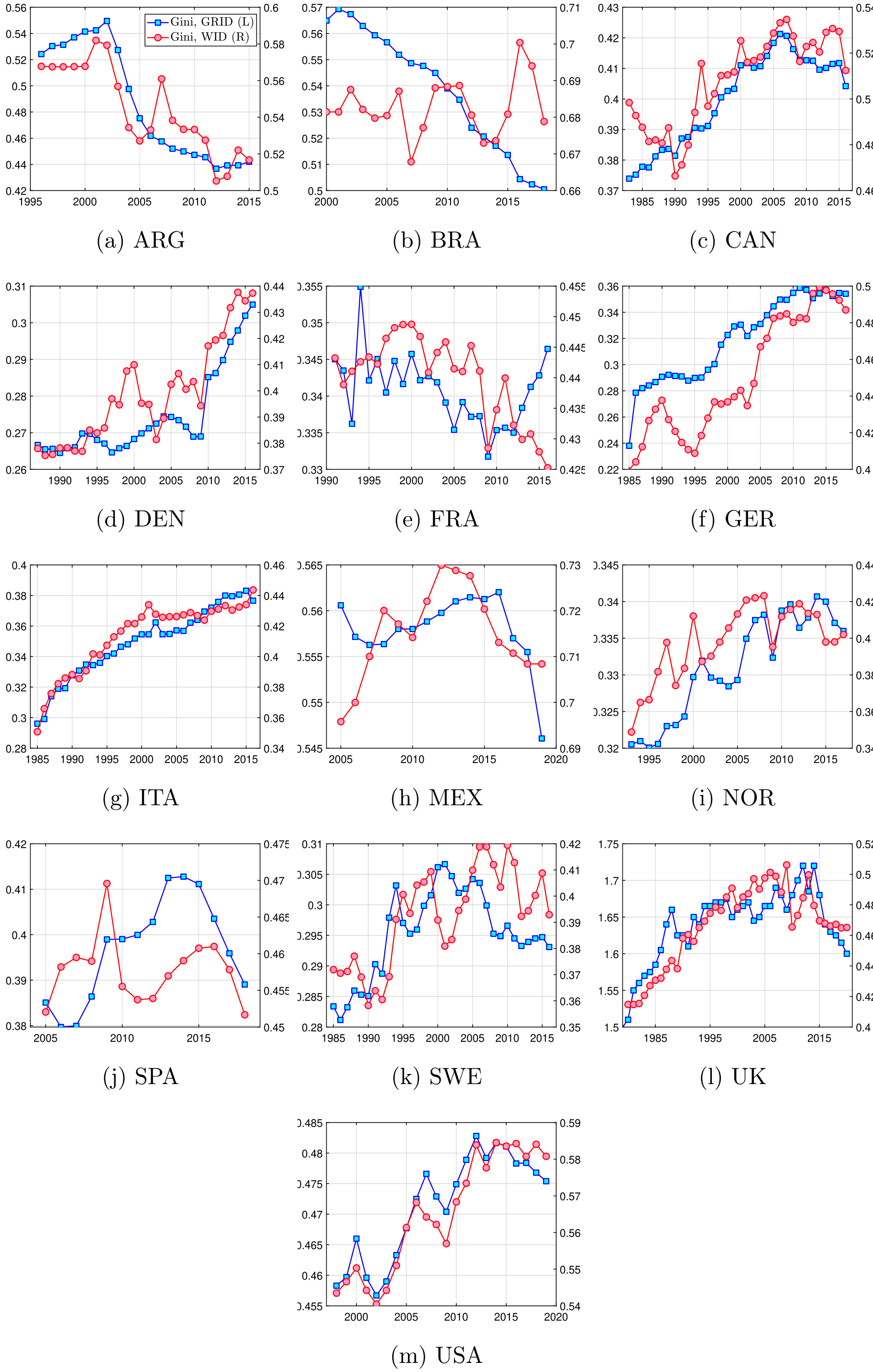


Figure 1. Comparing inequality trends in GRID and WID. Note: We use the P9010 statistic for the UK because the Gini coefficient was not authorized for disclosure in GRID. For Argentina, Brazil, and Mexico, the Gini coefficients in WID are calculated from survey data.

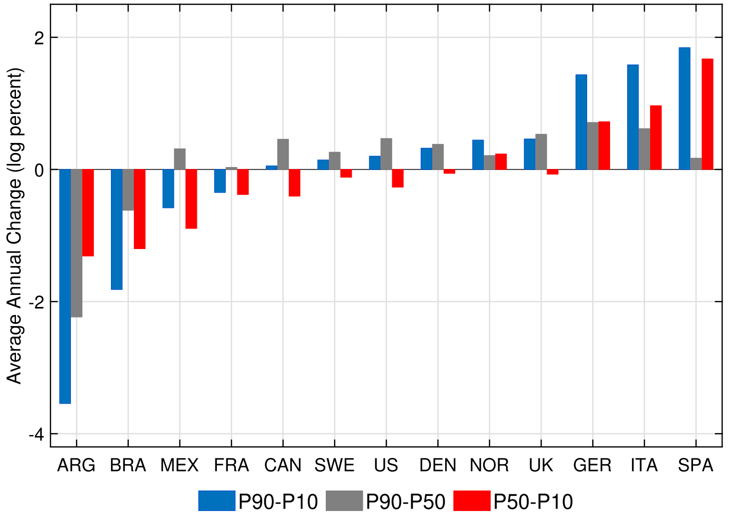


FIGURE 2. Trends in overall, top-end, and bottom-end income inequality for GRID countries.

such as deep recessions bring distinctive trend shifts. To get a visual representation of the broad trends, we conduct the following simple exercise. First, we run a regression of $P9010$ against a time trend and the growth in real GDP per capita (computed as $\Delta\log(\text{GDP per capita}_{c,t})$, separately for each country.[22] Controlling for the growth in GDP per capita allows us to separate the trend from the cycle. We then show how the coefficient on the time trend differs across the 13 countries (represented by the leftmost bars in Figure 2).[23]

Figure 2 shows that the countries that are part of the GRID project constitute a fairly heterogeneous group in terms of trends in overall inequality. However, it is possible to discern some interesting patterns that are common to geographical/language blocks.

Before delving into the details, recording the overall impression one gets from Figure 2 is useful: it is fairly symmetric around zero, with three countries showing falling inequality, three showing rising inequality, and the remaining seven countries ranked in the middle showing modest changes of either sign. So, the notion that income inequality is rising everywhere (and by large magnitudes) is not borne out in this fairly diverse set of countries. Looking more closely, we see that the three countries with declining inequality—Argentina, Brazil, and Mexico—are all in Latin America, although the level of inequality in these countries continues to be among the highest in the economies we

[22]The data for real GDP per capita are downloaded from the World Bank DataBank and are in constant local currency (except for Canada, where we use IMF data). One caveat from running country-specific regressions is that they are based on a few time-series observations, hence the differences in Figure 2 (below) may be noisy.

[23]For Germany, we use the longer 1985–2018 West Germany sample from the IAB. Moreover, because of confidentiality issues, the highest percentile that can be reported from the IAB is the 87.5th, so the 90th is obtained by a simple extrapolation procedure.

study.[24] The changes are in some cases substantial. In Argentina, over the 1996–2015 period, P9010 declines on average by more than 3 percentage points a year.

France and Canada experience modest changes in overall inequality, slightly declining in France and slightly increasing in Canada. However, in France inequality increases after the Great Recession whereas in Canada it declines. The Nordic countries (Denmark, Norway, and Sweden) all experience rising inequality of modest magnitudes, starting from extremely low bases. The Anglo-Saxon countries (the UK and the US) display a positive trend, although the rise in the US is fairly small. This finding may seem surprising at first, given the well-documented rise in income inequality in previous work. The reason is that the time span of the US data in GRID is from 1998 to 2019, whereas the large rise in US income inequality happened in the 1980s and 1990s.[25] Finally, the remaining continental European countries (Germany, Italy, and Spain) are those that experience the more robust upward inequality trends during the sample period.[26]

The other two bars in Figure 2 are informative about whether the trends in inequality are driven by changes at the top (the middle bar, displaying trends in the P9050 gap) or at the bottom (the rightmost bar, displaying trends in the P5010 gap). These trends are obtained the same way described above for the P9010. One striking finding is that in all countries (except for Brazil and Argentina), inequality above the median displays a positive trend: high-earners are pulling away from the median worker everywhere. In contrast, the P5010 trends are more heterogeneous, with half of the countries experiencing shrinking inequality at the bottom and half experiencing an increase, a trend that is especially significant in continental Europe. Indeed, in Spain almost all the rise in overall inequality is driven by a deterioration of conditions for the bottom percentiles. In the Anglo-Saxon countries, the rise in overall inequality is driven by the right tail with bottom-end inequality either flat (the UK) or slightly declining. A similar hollowing out of the middle, with increases in inequality at the top being accompanied by a fall in inequality at the bottom, is visible in France, Mexico, Canada, Denmark, and Sweden.

Top incomes All of the articles in this special issue report trends in top percentiles (including the 99th, 99.9th, and 99.99th percentiles if top coding allows). The behavior of these extreme fractiles of the income distribution has generated a large amount of interest. Here, we summarize how growth in these extreme percentiles compares with growth in the economy, as measured by per capita GDP. In all cases, we compute the relevant growth rates over the period for which country data are available. The left panel of Figure 3 plots the growth in the 99th percentile against GDP growth; the right panel repeats

[24] Some of this decline may be compositional, since—as noted above—administrative data from these three countries miss informal workers, and the extent of informality may have declined over time owing to the process of economic development.

[25] The bulk of the rise in US inequality in this earlier period happened in the upper tail, the P9050 gap, with the P5010 differential showing more "episodic" changes than a trend. See, for example, Autor, Katz, and Kearney (2008) for more details on this earlier period.

[26] In the case of Spain, the sharp decline at the end of the period, visible in Figure 1, is almost completely absorbed by the cyclical controls, so a strong positive trend remains. If instead one focuses on the difference between the first and last year of the sample (an alternative way of measuring broad trends), the rise in inequality is much more modest.

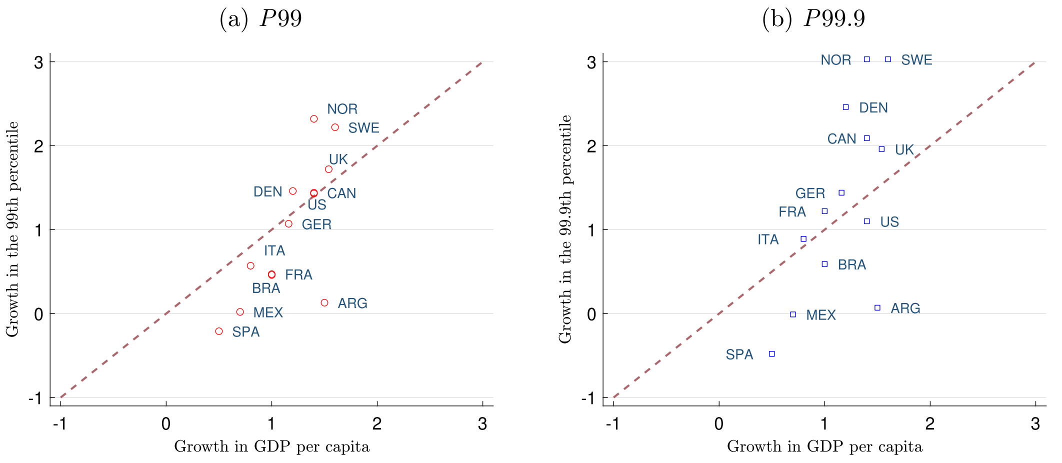


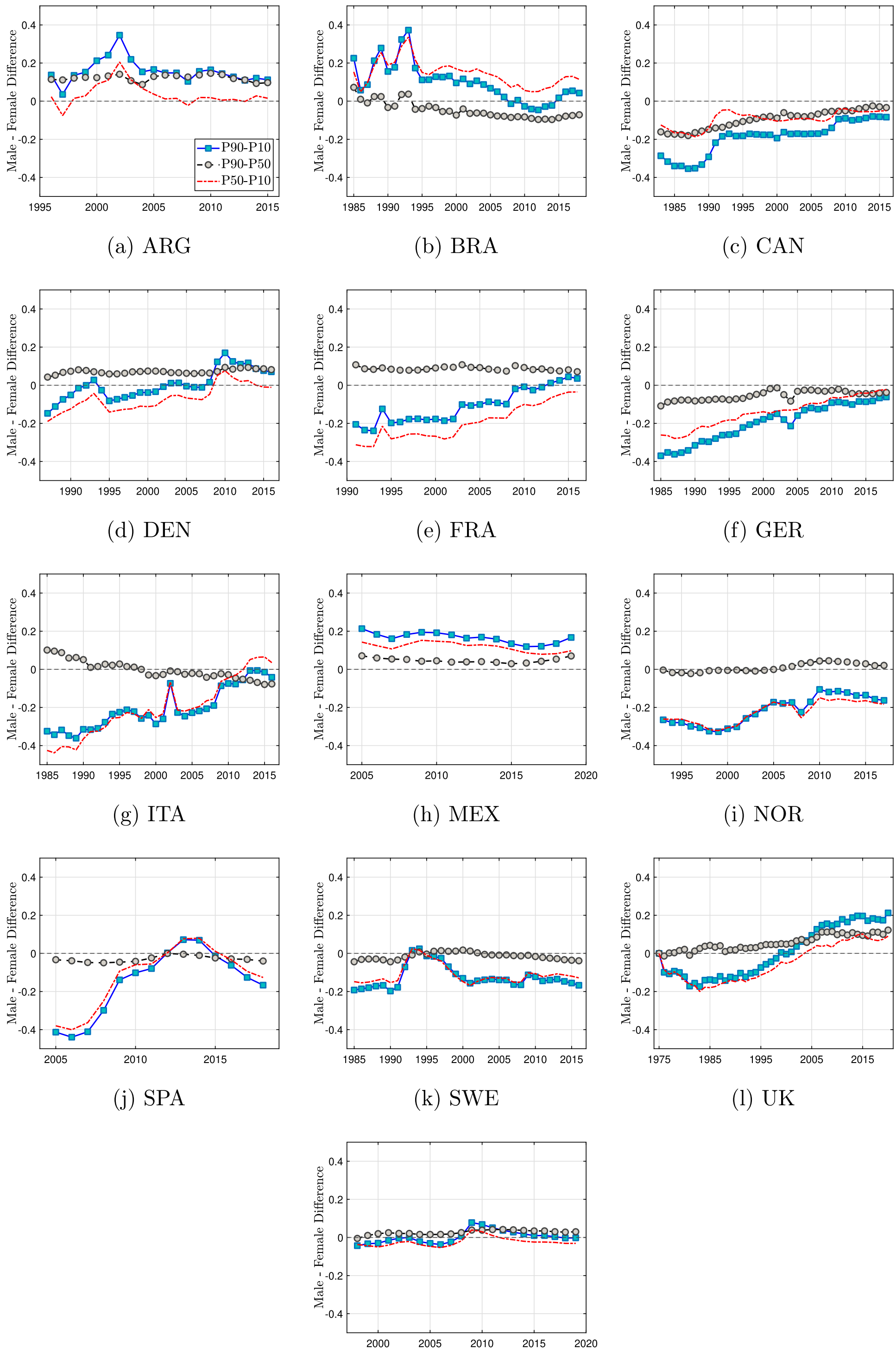
FIGURE 3. Growth in top percentile versus GDP growth.

the exercise for the 99.9th percentile. In both panels, the dashed line is the 45-degree line. Hence, in countries that are displayed above the line, top percentiles grow at a faster rate than the overall economy and vice versa in countries below the line. The visual impression from the figure is that for most countries, the growth in the 99th percentile has been at or below GDP growth. The exceptions are the Nordic countries and the UK. The evidence for the 99.9th percentile is that there are more countries with faster growth at the very top than in the rest of the economy. However, the deviations from uniform growth are not large, again with the exception of the Nordic countries. In the United States, top fractiles grow at the same rate as or even less than GDP per capita.[27]

Gender gap in earnings inequality Are these broad trends similar for men and women? To investigate this question, in Figure 4 we compute the gender gap between the P9010 differentials of men's and women's earnings distributions. A positive value means that the distribution of earnings for men is more disperse than that for women. At the beginning of the sample periods for which we have data, only the Latin American countries exhibit more dispersion among men than among women. In all other countries, moving from the 10th to the 90th percentile of the women's earnings distribution entails a larger growth of earnings relative to what we see in the men's distribution.[28] Perhaps more interesting, all countries exhibit a process of "convergence," whereby earnings inequality

[27] The time span of the US data in GRID is 1998 to 2019, which leaves out the 1980s and most of the 1990s when overall income inequality rose much faster. This raises the question of whether top percentiles may have grown much faster during that time than what we find in this analysis. Using data from the US Social Security Administration earnings records, Guvenen, Kaplan, and Song (2020) report that, between 1981 and 1998, the top 1% and 0.1% thresholds grew at 1.1% and 2.1% per year, respectively, compared with a GDP per capita growth rate of 1.7% (calculated from Figure 1).

[28] This is true in most cases because the gender gap at the bottom is larger than it is at the top. That is, there is less of a pay difference between executives of different genders than there is between service workers of different genders. Note that since the GRID database includes information about average (and median) earnings by gender and year, researchers may study the evolution of the gender gap over time and across countries.



(a) ARG (b) BRA (c) CAN (d) DEN (e) FRA (f) GER (g) ITA (h) MEX (i) NOR (j) SPA (k) SWE (l) UK (m) USA

FIGURE 4. Gender differences in inequality trends.

among women becomes closer to that of men by the end of the sample period. Figure 4 also makes clear that this pattern is primarily driven by the convergence in the lower tail (P5010), with the upper tail gap remaining stable in most countries.

4. Income dynamics: Variation over time

In this section, we discuss the salient properties of the income growth distribution. Whereas the distribution of income levels is informative about the cross-sectional dispersion in income or income inequality, the distribution of income growth tells us about something quite different: how income evolves for the same individuals over time. As such, it is closely related to income risk or uncertainty and is often equated to the latter under assumptions commonly made in the literature.[29] Thus, not only do income levels and income growth correspond to different concepts, but their properties can also be vastly different—and indeed that will be the case, as we see in this section. We begin with the density of annual growth in log income, which turns out to succinctly summarize many of the key features we will discuss in the rest of this section and the one that follows.

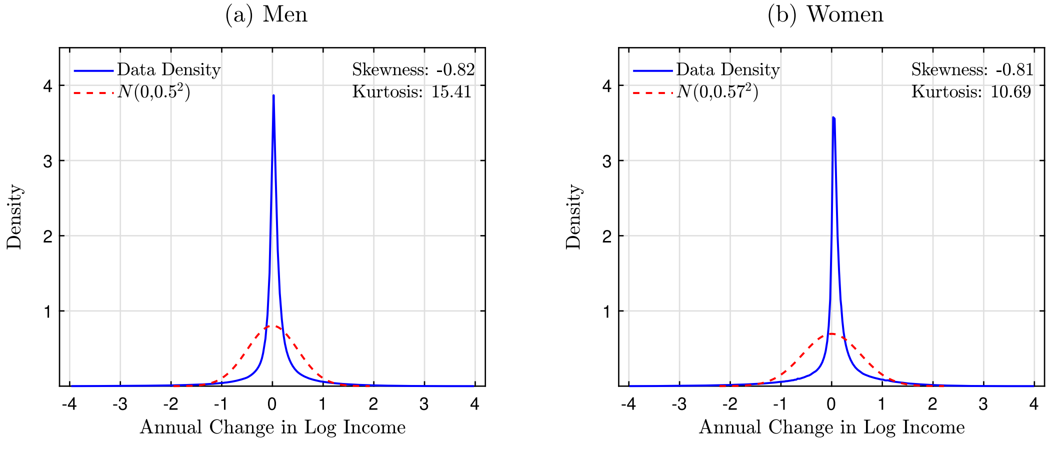


Figure 5. Density of annual change in log income by gender.

[29] Specifically, if income follows a random walk plus a purely transitory shock—the widely used permanent plus transitory model—income growth reflects the accumulated innovations to the random walk plus the mean reversion of transitory shocks. In this setting, longer horizon changes in income increasingly reflect permanent income risk. One issue with this interpretation is the assumption that the econometrician and the individual have access to the same information set, which may not be the case. To disentangle the individual's information set from that of the econometrician, some papers have used survey expectations of individual future income (e.g., see Pistaferri (2001) and Manski (2004) and the references therein) whereas others have used individual's economic choices to infer the same thing (see, e.g., Cunha, Heckman, and Navarro (2005), Guvenen (2007), and others).

4.1 *Inspecting the density*

Figure 5 plots the density of annual log income changes for men and women in Canada (the basic features we describe here hold for all the countries in GRID). Throughout the next two sections, income growth refers to the residualized measure defined in Section 2 (g_{it}^1) unless otherwise noted. Before even looking at the shape, notice that the standard deviation of 1-year log changes is 0.51 for men and 0.59 for women, which indicates a remarkably high level of volatility. If the data were Gaussian, it would imply that the average worker in Canada faces a typical income shock between 50% and 60%. Although large, 0.50 is in fact the average value in the sample of countries, which ranges from 0.38 for Germany to 0.66 for Mexico.

Earnings growth, however, is not close to being Gaussian. That is the second main takeaway from the figure, which superimposes a Gaussian density with the same standard deviation on top of the empirical density. Relative to the Gaussian density, the data have a very sharp peak in the middle, indicating a much larger probability of small income changes; very thin shoulders, indicating a much lower probability of middling shocks (around $\pm\sigma$); and longer tails or extreme shocks. The peakedness and long tails are reflected in a very high kurtosis in excess of 15 for men and 10 for women, relative to 3 for a Gaussian.

The sharp peak compresses the scale and makes it difficult to see some other important features of the density. Therefore, in Figure 6 we plot the *log* of the density for all GRID countries, which reveals two more key features that are harder to see when plotting the empirical density. All panels are chosen to have the same x- and y-axis limits for ease of comparability across countries. First, the tails of the distribution of income growth are very long and close to linear for the vast majority of countries. For comparison, the Gaussian log density with the same variance that is superimposed has tails that fall very quickly. The near-linear shape of the log density (which can be seen by the good fit of the linear regression line approximately beyond $\pm 3\sigma$) highlights an important feature: that income growth has a double-Pareto tail distribution. Furthermore, the tails are asymmetric, with the left tail thicker than the right tail, and much more so in some countries. To quantify this, we can look at the estimated slopes of the linear regression fit for the right and left tails, respectively, reported in Table 2. The left tail is thicker than the right tail in every country, and except for Italy and Mexico, the gap is sizable, exceeding 0.8 for every country except for Latin American countries (Argentina, Brazil, Mexico) and Italy.[30]

Comparing across countries, a few remarks are in order. First, there is no one-to-one mapping between the thickness of the tails and the standard deviation of income change. For example, the correlation between the average of the two tail indices for each country and the standard deviation is only -0.42. This is because the tail index measures

[30]We should note that Italy and Mexico are the only countries in GRID with top coding in income records: in Italy, the limit is 645 euros per day (or 161,250 euros for a full-time job with 50 weeks a year, and in Mexico, it is 25 times the legal minimum wage (see Hoffmann, Malacrino, and Pistaferri (2022)). These thresholds bind for a small but nonnegligible fraction of workers, which is likely to be playing some role in the more symmetrical tails found here. Below, we will see other examples where these two countries exhibit slightly different patterns in the tails, which gives further support to this conjecture.

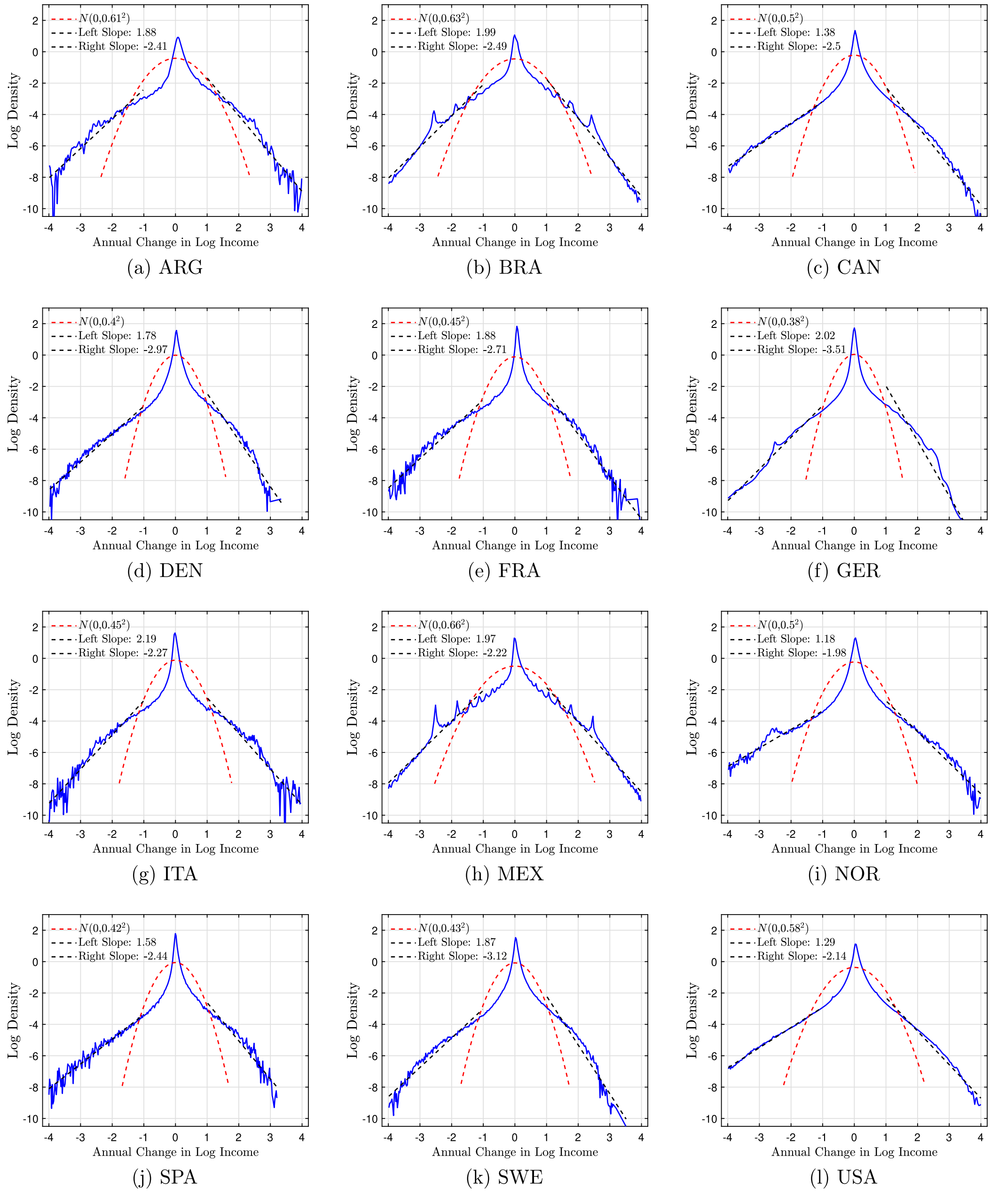


FIGURE 6. Density of annual change in log income (men).

the *slope* of the density in the tails only, so a density with a thicker tail may still have a smaller standard deviation if the *level* of density at the point we start to measure is low. In other words, the tail index can be interpreted as measuring the likelihood of a very large shock relative to a middling shock without reference to the likelihood of the former. This point is often overlooked in the discussion of Pareto tail indexes.

TABLE 2. Pareto tail indices of the distribution of log annual income change.

	Left tail	Right tail	Left − |Right|	$\sigma(\Delta Y)$
Argentina	1.88	−2.41	−0.53	0.61
Brazil	1.99	−2.49	−0.50	0.63
Canada	1.38	−2.50	−0.98	0.50
Denmark	1.78	−2.97	−1.19	0.40
France	1.88	−2.71	−0.83	0.45
Germany	2.02	−3.51	−1.42	0.38
Italy	2.19	−2.27	−0.09	0.45
Mexico	1.97	−2.22	−0.26	0.66
Norway	1.18	−1.98	−0.80	0.50
Spain	1.58	−2.44	−0.86	0.42
Sweden	1.87	−3.12	−1.50	0.43
US	1.29	−2.14	−0.96	0.57

Note: The numbers in the first two rows report the slope of the linear regression fit to the tails over $[-4, -1]$ and $[1, 4]$, respectively, which are the tail indices of the Pareto tails. The UK is omitted from this table because information on the tail index is not available.

Second, the two Anglo-Saxon countries, the US and Canada, are surprisingly similar to each other in terms of the thickness of each tail (1.33 and −2.31 for Canada versus 1.38 and −2.34 for the US) as well as in the overall dispersion (0.50 versus 0.56). Third, two of the Nordic countries, Denmark and Sweden, together with Germany, rank at the other end of the spectrum, with the lowest overall standard deviation of income growth and thinnest right tails, indicating a smaller chance of large upward income swings in these countries relative to others. Another interesting exception is Norway, which has the thickest tails, both right and left, in the sample (1.18 and −1.98). This compares to 1.78 and 1.87 for Denmark and Sweden for the right tail and −3.37 and −2.97 for the left tail. Although we are not aware of an explanation for why this is the case for Norway (given the low income inequality and the similarities in labor market institutions to other Scandinavian countries), this fact certainly seems worth further investigation. Finally, the Latin American and Southern European countries are between these two extremes.

Taking stock Our first look at the density reveals four key properties of the income growth distribution. The distribution has: (i) very high dispersion, (ii) very high excess kurtosis, (iii) thick double Pareto tails, and (iv) negative skewness, especially out in the tails as seen from a significantly thicker left tail relative to the right. These four properties confirm that the earlier results documented by Guvenen, Karahan, Ozkan, and Song (2021) for the US are robust across a broad cross-section of countries.[31]

[31] Incidentally, the US team's findings are very similar both qualitatively and quantitatively to those from that paper, despite using data from the LEHD rather than the Social Security Administration. For example, McKinney, Abowd, and Janicki (2022) estimate left and right tail indices of 1.29 and −2.14 for 2010 compared with 1.40 and −2.18 in Guvenen, Karahan, Ozkan, and Song (2021) for 1997–1998.

4.2 *Long-term trends in idiosyncratic income risk*

We begin by investigating whether there are any trends in idiosyncratic income risk. This question is of obvious interest given the importance of idiosyncratic risk for individual decisions and welfare, and consequently, for social insurance and government policy, among others issues. Since the 1990s, the conventional wisdom among economists has been that idiosyncratic risk increased substantially since the 1970s, a conclusion from empirical analyses of survey-based panel data sets showing rising income volatility. Following the seminal work of Gottschalk and Moffitt (1994) and Moffitt and Gottschalk (1995), a long list of papers that analyze US survey data confirmed their finding and found evidence of a continued rise in volatility all the way to the 2010s.[32]

Against this backdrop, several recent papers studied US administrative data from the Social Security Administration on earnings histories and reached the opposite conclusion: income volatility at both the short and long horizon has been either flat (Congressional Budget Office (2007)) or declining (Sabelhaus and Song (2010) and Bloom, Guvenen, Pistaferri, Sabelhaus, Salgado, and Song (2017)) since the early 1980s. GRID provides an ideal opportunity to not only revisit this question for the US but also examine possible trends in a wide cross-section of countries.

Figure 7 plots the standard deviation of annual income growth for both men (line with squares) and women (line with circles). Most lines are fairly flat, with a few countries (e.g., Argentina and Brazil) showing clear declining trends and a few countries (e.g., Italy, Norway, and Sweden) showing a rising trend, more so for men than for women. Second, among Anglo-Saxon countries, the trend is flat or slightly declining for Canada and the UK and strongly declining for the US. Recall that the GRID data source for the US is the Longitudinal Employer Household Dynamics (LEHD) programs from the US Census Bureau (see Table 1), not the SSA as in the studies cited above. Hence, this constitutes independent evidence on the flat/declining income volatility trend for the United States. Other countries, such as Denmark, France, and Mexico, show an overall flat pattern, indicating no specific trends in income volatility. Finally, Spain shows a cyclical rise in volatility during the Great Recession and its aftermath, but volatility falls back to its initial level for men and is lower for women by the end of the period.

To summarize, Figure 7 paints a somewhat mixed picture. with volatility flat for about half of the countries, declining for some countries and rising for others. It does not provide any evidence of a widespread rise in volatility or income risk around the world. This conclusion echoes our findings above for income inequality, which also showed a mixed picture, with trends in inequality being more idiosyncratic and country specific than reflecting a global rise in inequality. As we will see in a moment, this is not always the case, and for a number of empirical questions we study next, global trends that are far clearer are observed in the vast majority of countries.

4.3 *Business cycle variation in idiosyncratic income risk*

How does idiosyncratic income risk change over the business cycle? Do any robust patterns hold across this broad set of countries? Or does the answer depend on the labor

[32] See Dynan, Elmendorf, and Sichel (2012) for a review of these papers and Moffitt and Zhang (2018) for an update.

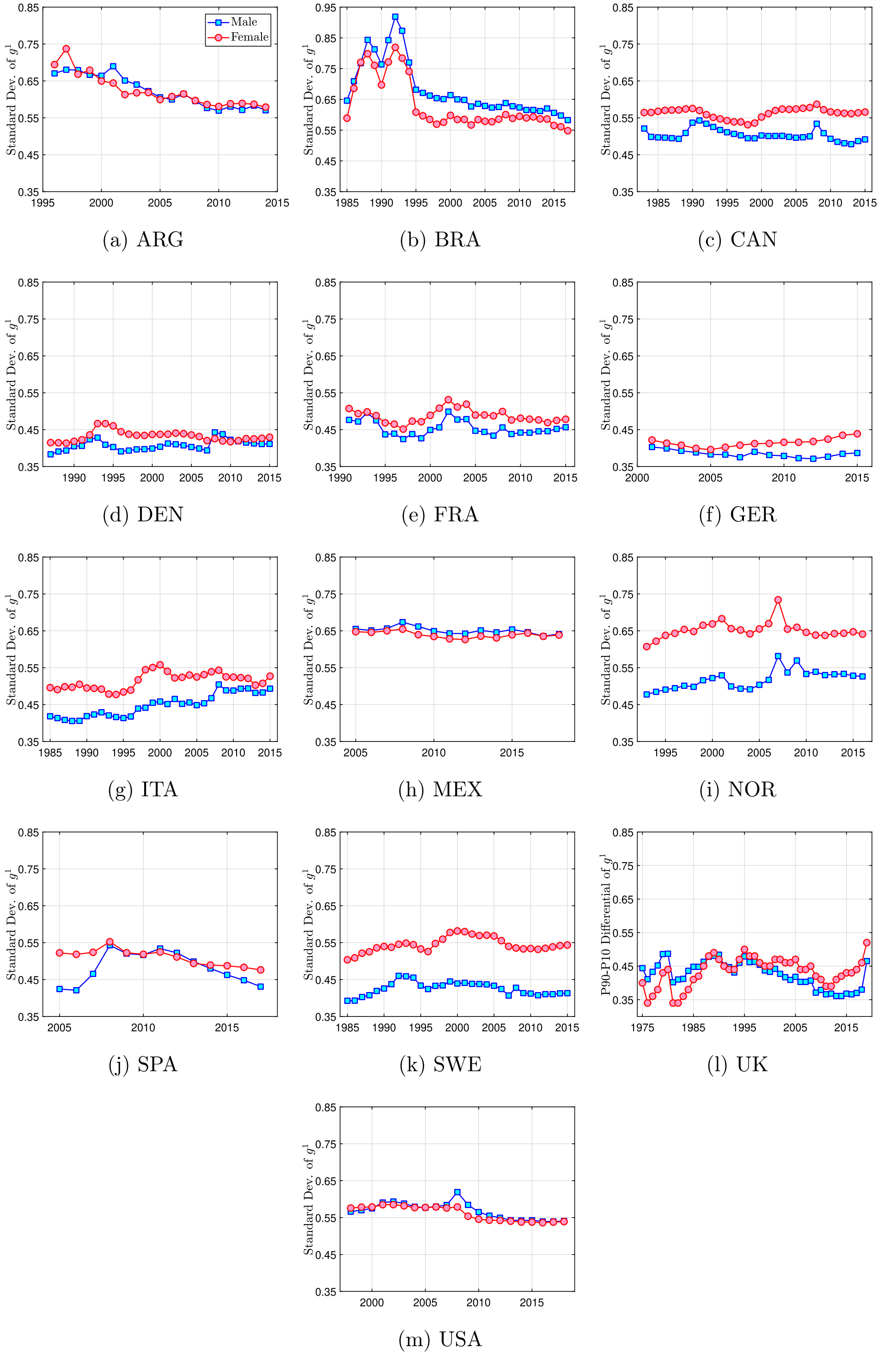


FIGURE 7. Trends in income volatility (men and women). Note: The figure shows the P9010 differential for the UK because the standard deviation is not available.

market and other institutions of each country? The answers to these questions are critical for many macroeconomic and policy design features that account for individual heterogeneity and incomplete insurance.

To answer these questions, we begin by examining how different moments of the income growth distribution vary with the business cycle. Following earlier work documenting the strong procyclicality of the skewness of income changes, in Figure 8, we plot the Kelley skewness measure of the 1-year log income change for men (line with squares, left axis) together with the annual growth rate of GDP per capita (dashed line, right axis) over time.[33] The latter is a natural indicator of the business cycle, so the comovement between the two series would give a direct visual indication of the cyclicality of the statistic plotted. As seen in the figure, the two lines comove to a remarkable extent in almost every country, especially during deep recessions, showing that the skewness of income change is strongly procyclical in all countries in GRID. The pattern for women looks qualitatively very similar, with a somewhat smaller amplitude of fluctuations in skewness, so the analogous figures are included in the Online Appendix (Guvenen, Pistaferri, and Violante (2022)) for brevity.

Regarding the magnitudes, are the procyclical fluctuations large? For most countries, the answer is yes. The advantage of the Kelley statistic is that it can easily be mapped into the relative sizes of P9050 and P5010 (or the upper and lower tails) in P9010 of the shock distribution. For example, in Argentina, Kelley skewness went from -0.29 in 2001 to $+0.32$ in 2003, implying that the share of P9010 of the income growth distribution accounted for by the upper and lower tails flipped from 1/3 and 2/3 to 2/3 and 1/3 in a short span of 2 years.[34] (Notice that this comparison controls for the rise in median income growth.) This is a major reversal of the income shock distribution in a very short period. Of course, for Argentina this period was preceded by a large decline in skewness coinciding with the deep 2001 recession, illustrating the procyclical nature of skewed income risk.

Similar or larger swings happened in several other countries (e.g., Spain, Italy, Denmark, and the US) and during different episodes, such as the severe recession in Europe during the early 1990s as well as the Great Recession (with skewness falling from 0.24 to -0.49 in 3 years in Spain and from 0.22 to -0.37 in Italy in 2 years). To sum up, a major manifestation of changes in income risk between expansions and recessions is in the skewness of the income change or shock distribution. These procyclical swings are large and synchronized with the business cycle.

[33] The *Kelley measure* is defined as

$$\mathcal{S}_k = \frac{(P90 - P50) - (P50 - P10)}{P90 - P10},$$

or the share of overall dispersion in income change ("shock") in the upper tail minus the share in the lower tail. A negative Kelley value indicates that the lower tail (income falls) makes up a larger fraction of the dispersion than the upper tail (income rises). Relative to the third standardized moment, the Kelley measure has the advantage of being robust to outliers, and its magnitude has a clearer interpretation as we will discuss in a moment. That said, we have replicated the results presented in this section with standardized moments and found them to be robust except where noted (see Online Appendix A).

[34] More concretely, P9050 was $0.5 - 0.29/2 = 0.35\%$ of P9010 in 2001 and $0.5 + 0.32/2 = 0.66\%$ in 2003.

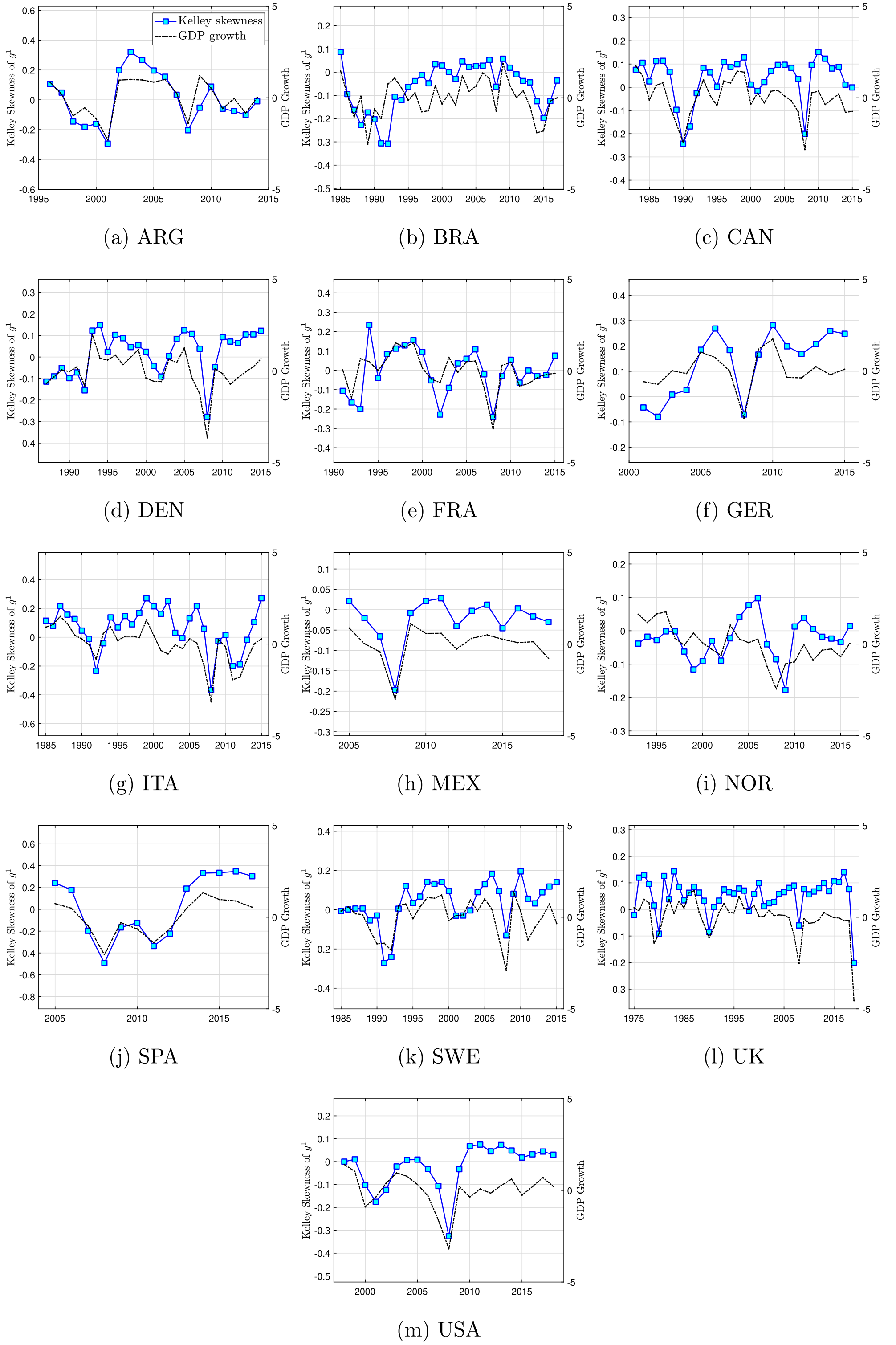


FIGURE 8. (Kelley) Skewness fluctuations over the business cycle (men).

To more precisely quantify the extent of cyclicality in different moments, we adopt a simple regression framework. In particular, we regress a given moment m on a constant, a linear time trend, and log GDP per capita growth—$\Delta(\log \text{GDP}_t) \equiv \log(\text{GDP}_{t+1}) - \log(\text{GDP}_t)$,[35]

$$m(\Delta y_t) = \alpha + \gamma t + \beta^m \times \Delta(\log \text{GDP}_t) + u_t, \tag{1}$$

for each country and separately for men and women. We normalize the GDP per capital growth to have unit standard deviation, which makes the estimated $\beta' s$ easier to compare across countries. Table 3 reports the parameter estimates of β^m (multiplied by 100), which measures the cyclical sensitivity of moment m. A significant and positive β^m indicates a procyclical moment and *vice versa* for a negative coefficient.

The top panel contains the estimates for Kelley skewness, which confirms the visually evident strong procyclicality of skewness for men and shows the same for women: β^m is statistically significant with t-statistics that range from 4 to 17 for men and 2 to 14 for women.[36] The magnitudes are smaller for women, with the lowest coefficients found in Nordic countries with large public sectors that heavily employ female workers. This finding is consistent with Busch, Domeij, Guvenen, and Madeira (2022), who find that the industry of employment is a key determinant of how cyclical skewed income risk is for a worker. The magnitude of the sensitivity is large, considering that the Kelley statistic is bounded between zero and one. For example, a coefficient of 15.0 for Argentina indicates that a two standard deviation swing (which would be typical when going from a normal expansion to a recession), implies a 0.30 drop in the Kelley skewness of income changes.

Clearly, a change in skewness can be driven by a change in the right tail, left tail, or both. To investigate which tail drives the procyclical fluctuations, we run the same cyclicality regressions for P9050 and P5010 separately (next two panels of Table 3). The results suggest that, with one or two exceptions, the right tail is strongly procyclical and the left tail strongly countercyclical for all countries. The magnitudes of the coefficients are comparable, with those on the left tail often slightly larger than those on the right.

Finally, the dispersion (P9010) and kurtosis show more mixed patterns. For men, there is evidence of countercyclical volatility, while for women the pattern is less clear and noisier. Kurtosis does not show a clear cyclical pattern, with coefficients sometimes positive and sometimes negative, and t-statistics indicating significance at the 5% level for only a few countries. Overall, this evidence appears too noisy to be economically informative.

Taking stock The skewness of income changes (a proxy for income risk) varies significantly from expansions to recessions in all countries represented in GRID. In particular, income shocks become more negatively skewed in recessions, with the probability

[35] Other indicators of business cycles are possible, but previous research has found them to deliver similar results (see, e.g., Guvenen, Ozkan, and Song (2014)).

[36] The standard errors used to construct the t-statistics are Newey–West corrected using a lag length of 3. However, unlike the rest of our results that rely on a very large number of observations, the cyclicality regressions are based on a relatively short time series, which should be kept in mind when reading the statements about significance in this section.

TABLE 3. Cyclicality of the moments of 1-year log income change.

	ARG	BRA	CAN	DEN	FRA	GER	ITA	MEX	NOR	SPA	SWE	UK	USA
							Kelley Skewness						
Males	15.0	5.6	8.5	8.0	8.8	7.8	14.8	5.2	3.7	27.3	7.1	5.2	7.3
	(6.0)	(4.6)	(9.7)	(12.3)	(9.3)	(5.9)	(10.0)	(15.8)	(4.7)	(16.7)	(4.8)	(6.9)	(7.7)
Females	8.5	4.7	3.4	2.0	3.9	3.2	6.3	4.5	2.1	12.3	1.8	2.8	3.8
	(4.5)	(3.8)	(7.7)	(2.0)	(5.2)	(3.5)	(5.0)	(8.1)	(2.6)	(14.0)	(2.6)	(4.2)	(8.2)
							P9050: Upper Tail						
Males	6.7	2.7	2.4	1.8	1.7	1.4	3.0	2.6	0.5	5.9	1.8	0.5	2.3
	(4.7)	(4.2)	(8.4)	(5.4)	(12.7)	(8.0)	(6.1)	(14.6)	(1.4)	(8.3)	(5.5)	(3.4)	(8.7)
Females	4.2	1.7	0.8	1.5	0.6	0.4	2.3	1.9	−0.9	4.5	1.5	0.1	1.6
	(3.0)	(2.8)	(3.7)	(2.6)	(2.9)	(1.9)	(2.2)	(3.5)	(−1.2)	(8.9)	(3.7)	(0.5)	(8.4)
							P5010: Lower Tail						
Males	−9.6	−4.6	−4.2	−2.5	−2.9	−1.9	−5.6	−4.3	−1.9	−13.4	−2.5	−1.8	−5.0
	(−6.7)	(−3.2)	(−7.9)	(−5.6)	(−7.3)	(−4.6)	(−11.1)	(−15.1)	(−6.4)	(−12.3)	(−3.4)	(−4.4)	(−5.9)
Females	−4.7	−3.7	−2.7	0.2	−2.4	−1.2	−3.0	−3.6	−3.4	−6.0	−0.1	−1.0	−2.0
	(−5.3)	(−2.9)	(−8.2)	(0.3)	(−4.1)	(−2.7)	(−4.9)	(−19.1)	(−4.5)	(−16.1)	(−0.2)	(−2.7)	(−5.6)
							P9010: Volatility						
Males	−2.9	−1.9	−1.8	−0.7	−1.2	−0.5	−2.6	−1.7	−1.3	−7.5	−0.7	−1.3	−2.7
	(−4.9)	(−1.3)	(−4.2)	(−1.1)	(−2.8)	(−1.2)	(−3.6)	(−9.6)	(−2.9)	(−4.3)	(−1.1)	(−2.5)	(−4.0)
Females	−0.5	−2.0	−1.9	1.7	−1.8	−0.9	−0.7	−1.7	−4.3	−1.5	1.4	−1.0	−0.4
	(−0.4)	(−1.4)	(−6.6)	(1.3)	(−2.6)	(−1.5)	(−0.5)	(−4.3)	(−3.8)	(−2.3)	(1.7)	(−1.9)	(−1.1)
							Crow Kurtosis						
Males	−0.73	0.06	0.01	0.21	0.31	0.04	0.58	0.21	−0.15	−1.17	0.32	0.35	0.01
	(−1.9)	(0.3)	(0.1)	(0.9)	(1.5)	(0.2)	(2.0)	(1.6)	(−1.3)	(−2.5)	(2.5)	(3.4)	(0.1)
Females	−0.85	0.09	0.08	0.49	0.31	0.10	0.54	−0.02	0.30	−0.70	0.15	0.48	−0.20
	(−2.4)	(0.3)	(0.5)	(1.8)	(3.2)	(1.3)	(2.9)	(−0.1)	(2.4)	(−2.5)	(1.1)	(4.3)	(−2.1)

Note: Each cell reports the cyclical sensitivity coefficient, β^m, in a regression of statistic m on log annual GDP change plus a constant and a time trend (equation (1)). Except for the Crow kurtosis, the reported coefficient is multiplied by 100 for ease of interpretation. The numbers in parentheses are the t-statistics computed using Newey–West standard errors with 3 lags.

of large negative tail shocks rising and the likelihood of large positive shocks falling in recessions. The opposite happens in expansions, which see a rise in the likelihood of large positive shocks and a decline in the likelihood of large negative shocks. As for overall dispersion, while the estimates indicate countercyclical variation, the magnitudes of fluctuations are relatively modest for most countries. Finally, we can identify no robust cyclical patterns in kurtosis.

5. Income dynamics: Cross-sectional and life-cycle variation

In this section, we discuss how the higher-order moments vary with permanent income status and age in different countries and the extent to which these patterns also vary by gender. The main set of statistics we focus on are the standardized 5-year change in income moments; however, other than some exceptions we mention, the patterns we discuss here also hold true for 1-year changes as well as for the quantile-based statistics we focused on in the previous section. The reason for the switch is twofold. First, while 1-year changes were convenient in the previous discussion to more clearly see the alignment of skewness with the business cycle, 5-year changes put more emphasis on the persistent component of income fluctuations, which are more important for economic decisions. Second, we use standardized moments to show that these results are robust to different measures used. For each fact, we first describe the general patterns and then discuss exceptions by country (group), measure, and gender, if any.

Volatility

We begin with volatility, the second moment. Figure 9 plots the standard deviation of 5-year log income growth. Three facts are immediately noticeable. First, in all countries, the volatility of future income fluctuations declines significantly with the permanent income level of the worker (regardless of age). In the majority of countries, the decline in volatility slows down when permanent income rises above the median, giving the graph a convex shape. The exceptions to this pattern are Spain and the Latin American countries, for which the pattern is more linear than convex. Notice that, except for Spain, the countries in this group have some of the highest levels of volatility in GRID.

Second, income volatility rises sharply at the very top of the permanent income distribution. Indeed, in most countries. the degree of volatility for workers in the top 1% group is as much as —or even more than—volatility for workers at the very bottom of the distribution.[37] The exceptions to this pattern are Brazil, Italy, and Mexico. But recall that earnings records for Italy and Mexico are top-coded, which could be tempering the large positive shocks for workers at the very top, thus dampening measured volatility. Overall, it appears that—across countries—workers at the very top (1%–2%) of the permanent income distribution have some of the highest income volatility in the population.

[37] Because the sample excludes workers whose income in $t+5$ is less than $1/3$ of the minimum threshold, this statement excludes workers who receive long-term nonemployment shocks that last longer than the calendar year. However, this group is fairly small in many countries.

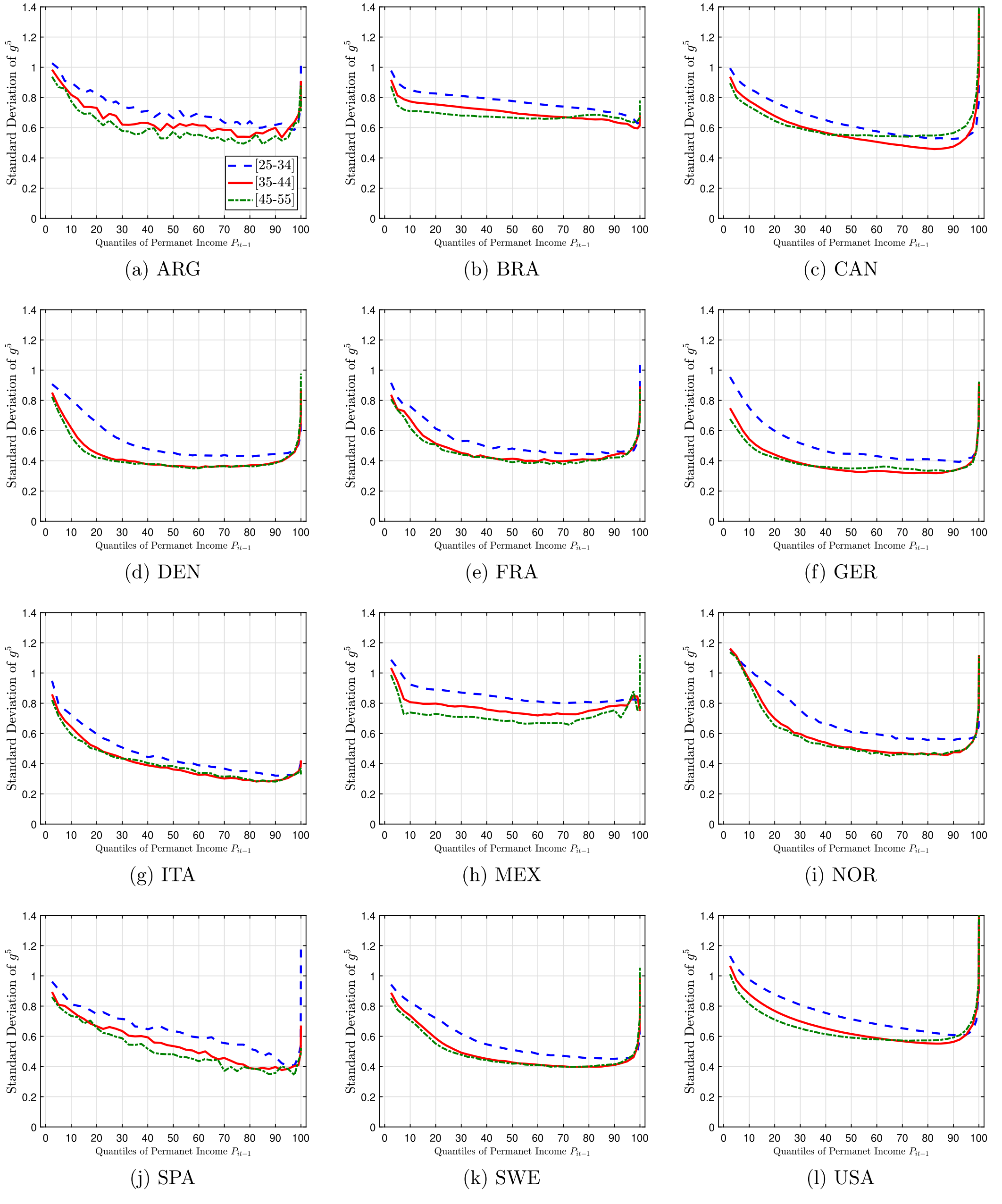


FIGURE 9. Standard deviation of 5-year income growth, by permanent income and age (men).

Third, volatility declines with age, most notably between the first and second decades of the prime working years (from ages 25–34 to 35–44). The pattern between the second and third decades is less clear, with volatility further declining in the Latin American countries and Spain and staying largely the same for the remaining countries.

To sum up, volatility is significantly lower for higher-income groups with the important exception being the very top earners, who face very high income volatility. Volatility also declines with age, although the variation across income levels is larger than across age groups. The patterns described here also hold for the P9010 differential; see Online Appendix for the counterpart of Figure 9.

Skewness

Turning to skewness, the patterns we see in Figure 10 share important similarities with those just discussed for volatility in Figure 9. In particular, in all countries, workers with high permanent income face more negatively skewed income shocks, and the (negative) slope of this relationship is steep in all countries except in Latin America. The steepness also increases with age. For example, in many countries—such as Canada, Denmark, Germany, Italy, Norway, and Spain—for middle-age workers, skewness ranges from almost zero at the lowest permanent income levels to -3 or -4 at the 80th to 90th percentiles. Furthermore, as with volatility, the negative relationship between skewness and income reverts (with the exception of Latin American countries), and skewness becomes increasingly less negative for the top 10% to 20% of the permanent income distribution. For Brazil and Mexico, skewness becomes more negative at the very top end, which may be a result of top coding.[38]

Kurtosis

Finally, Figure 11 plots the kurtosis coefficient. Again, clear regularities manifest themselves across all countries. First, at low levels of permanent income, kurtosis is low, in fact close to a Gaussian (which is marked as zero in the figure showing excess kurtosis). As income rises, kurtosis rises substantially and in a monotonic fashion, reaching levels as high as 30 to 40 for some countries (Denmark, France, Germany, Italy, Norway, Spain, and the US). To put these values into context, note that a distribution with an excess kurtosis of 5 to 10 is generally considered to be highly leptokurtic. At the very top, the pattern reverses again, and kurtosis falls for the very top earners. Second, kurtosis rises with age as well, and the gap is quite large—doubling in magnitude from the youngest and oldest groups—for some countries. These patterns hold for all countries with the exception of Brazil and Mexico, for which both levels and the rise of kurtosis are muted relative to other countries, and there is no decline in kurtosis at the very top end.

[38]These patterns look somewhat different when we switch from the third moment to Kelley skewness (see Figure A.2 in the Online Appendix). First, the bottom of the U-shape for Kelley skewness is reached at a much lower permanent income level than in Figure 10. The difference implies that tail shocks (outside of the P90 and P10 bounds that Kelley focuses on) that are negatively skewed are increasingly more important at higher permanent income percentiles, driving the skewness coefficient more negative as income rises. The opposite seems to happen at the very top end: Kelley skewness turns more negative for the top income group, as opposed to the rise for the skewness coefficient, suggesting that tail shocks are more positively skewed for these workers than shocks within the P10 and P90 bounds.

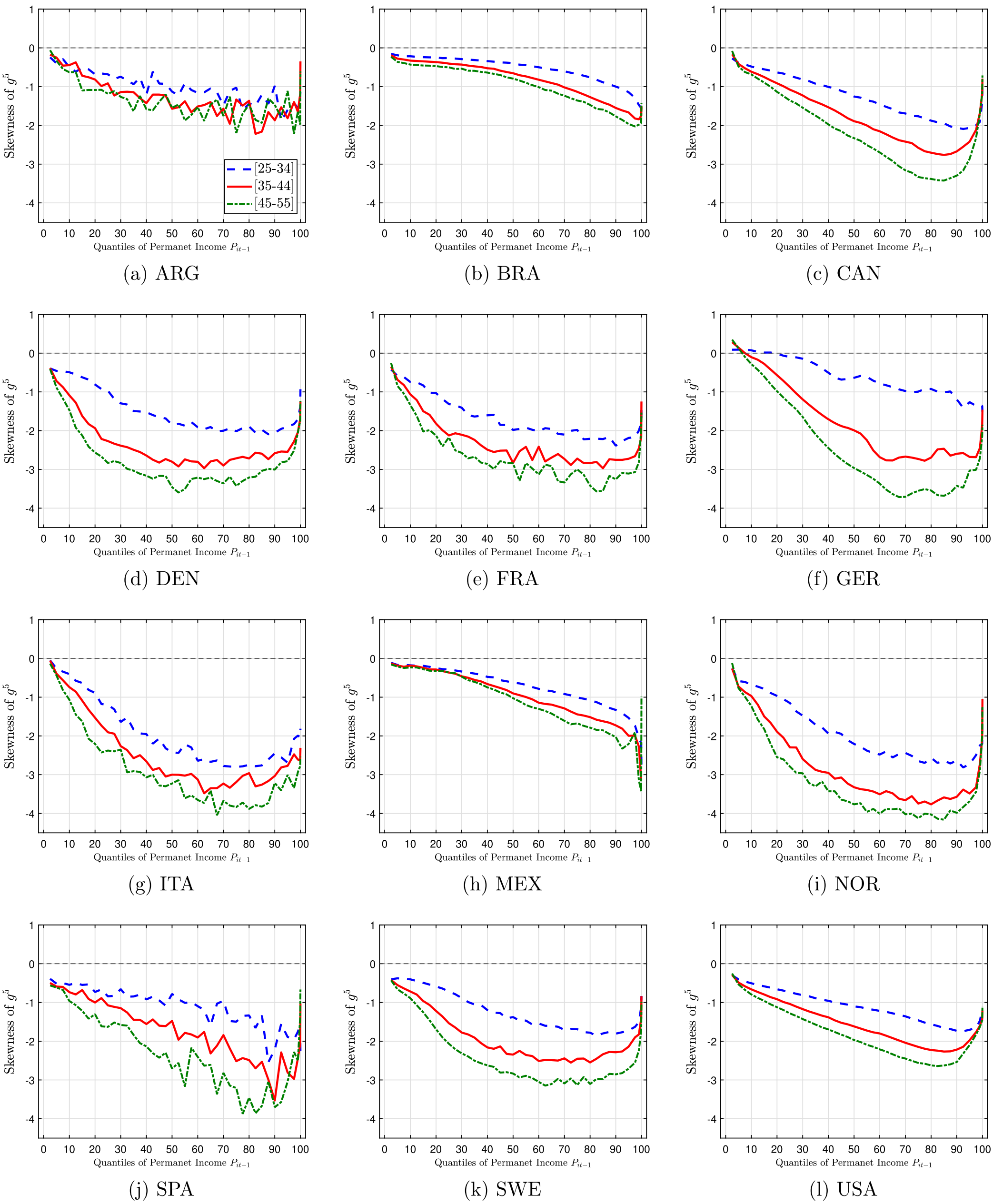


FIGURE 10. Skewness of 5-year income growth, by permanent income and age (men).

6. Rank mobility over the life cycle

The statistics we discussed in the previous sections characterize in great detail the distribution of individual income levels and changes. We provided ample evidence that workers are subject to idiosyncratic shocks of various sizes, signs, and persistence. Because these income fluctuations are largely uncorrelated across workers, individual

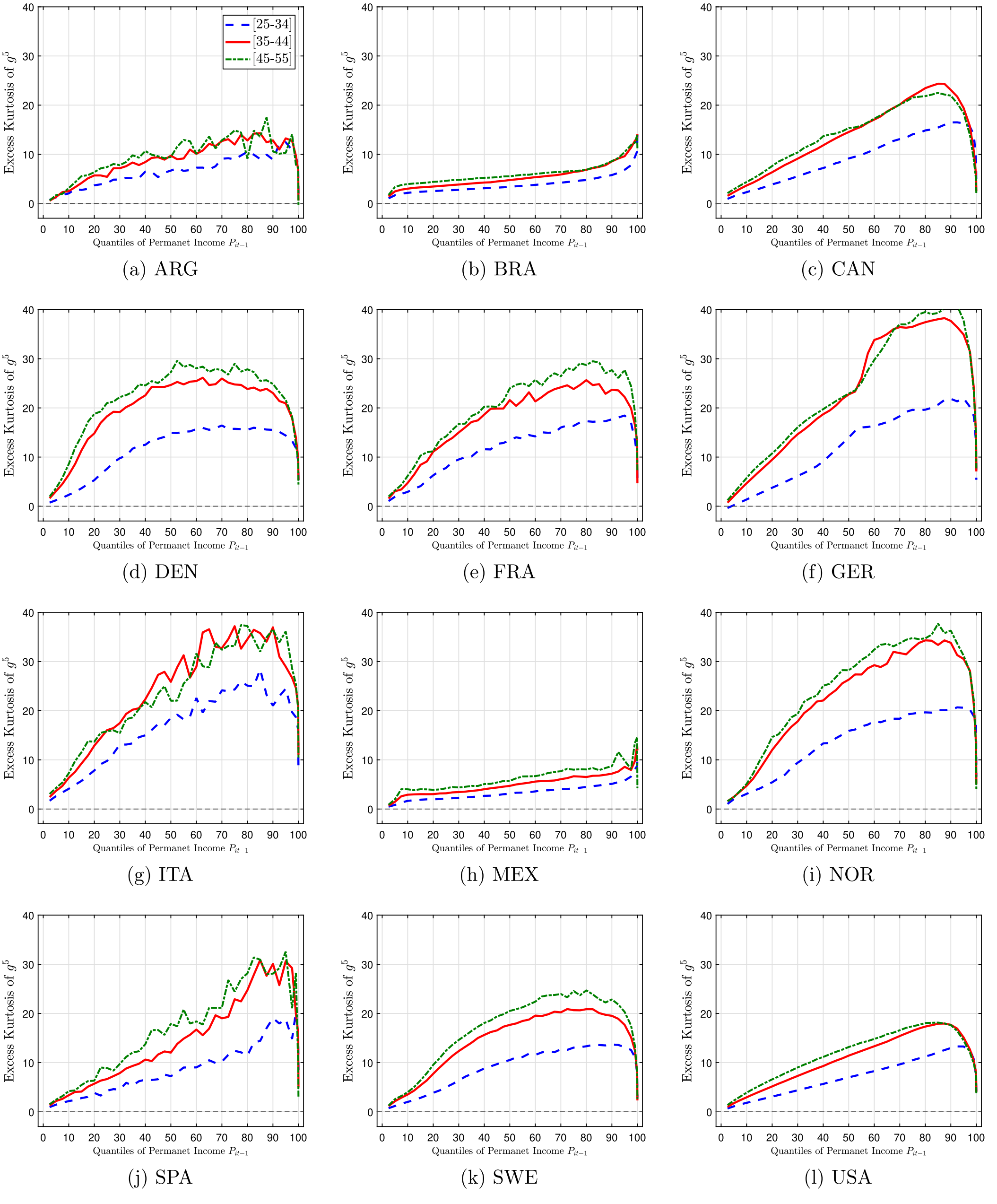


FIGURE 11. Kurtosis of 5-year income growth by permanent income (men).

ranks within the cross-sectional income distribution change over time. In this section, we investigate the extent of this positional mobility over the life cycle across countries. The extent of intragenerational mobility of earnings is important because high levels of individual mobility imply less cross-sectional inequality in *long-run earnings* for any given level of cross-sectional inequality in *annual earnings*. Policy prescriptions and welfare calculations are greatly affected by the relation between current and permanent income inequality. Fields and Ok (1999) discuss the importance of accounting for mobility in analyses of income inequality and compare key mobility indicators.

We start by ranking an individual i based on the individual's permanent income in period t.[39] Let R_{it} denote this rank (percentile). We then compute, for each percentile of the permanent earnings distribution at time t, the average rank for all individuals in that percentile 5 or 10 years later (i.e., at date $t+k$, with $k=5$ and 10). The individual articles in this issue report plots of this rank-rank relation for each country, and for various sample periods and subpopulations. Table 4 summarizes these results by providing two types of statistics for 5-year mobility over the period 1997–2007, which is common to all countries.

First, we compute the rank-rank slope (RRS), that is, the coefficient β of the following regression:

$$R_{it+5} = \alpha + \beta R_{it} + \varepsilon_{it}. \tag{2}$$

This indicator, also common in the literature on intergenerational mobility (Chetty, Hendren, Kline, and Saez (2014)), measures rank persistence over the life cycle. In addition, we calculate a set of mobility indicators conditional on various initial positions within the distribution. The first is an index of mobility from the bottom, or absolute upward mobility (AUM), that is, the expected rank at $t+5$ for individuals who are below the median at time t,

$$\text{AUM} = \mathbb{E}[R_{it+5} | R_{it} \leq 50].$$

The second is an index of mobility from the top, or absolute downward mobility (ADM),

$$\text{ADM} = \mathbb{E}[R_{it+5} | R_{it} > 50].$$

In a world without any rank persistence, $\beta = 0$, and the indicators of bottom and top mobility would all be equal to 50 (the median rank). In a world with maximum persistence, $\beta = 1$, and ranks would perpetuate: AUM would equal 25 and ADM would equal 75. Finally, we fully exploit the large sample size of our data sets to learn about mobility at the very top of the earnings distribution and estimate an indicator for those in the top 1% (M99):

$$M99 = \mathbb{E}[R_{it+5} | R_{it} > 99].$$

Rank persistence

All RRS in Table 4 are systematically below one, meaning that there is some degree of mean reversion in ranks over a 5-year horizon: those who start with a low position in

[39] Using an estimate of permanent income averages very transitory shocks and mitigates concerns about a mechanical reversion of incomes to the mean. We use a slightly modified version of permanent income relative to the previous sections. We keep in the sample individuals with earnings above the minimum threshold for at least 1 year (it was 2 in the previous definition). We want a broad representation of the cross-sectional distribution, and hence, also include workers with a weak attachment to the labor market.

Table 4. Key statistics on income mobility.

	Pooled, 1997–2007				AUM Pooled, 1997–2007				AUM		
	RRS	AUM	ADM	M99	Males	Females	25–34	45–55	1987	1997	2007
Argentina	0.71	33.3	69.2	94.7	32.9	34.3	35.2	31.0	–	33.5	32.9
Brazil	0.87	28.9	73.6	98.7	29.2	28.3	30.4	27.2	29.3	28.8	29.4
Canada	0.74	32.2	70.1	93.0	31.9	32.5	34.7	30.5	32.8	32.3	32.0
Denmark	0.68	33.8	68.5	92.3	33.3	34.3	38.2	30.6	33.1	33.8	33.2
France	0.83	29.8	72.6	97.4	29.3	30.4	32.3	28.1	–	29.2	29.7
Germany	0.76	30.1	68.7	91.7	29.8	31.8	34.7	28.2	–	30.5	30.1
Italy	0.86	29.0	73.4	98.2	28.3	30.2	31.6	27.0	28.8	28.8	29.1
Mexico	0.82	30.0	72.4	96.9	30.2	29.6	31.5	28.3	–	–	30.0
Norway	0.70	33.6	68.8	93.2	32.2	34.9	37.4	30.5	–	34.1	33.4
Spain	0.79	31.1	71.4	96.4	31.0	31.3	33.9	28.4	–	–	31.1
Sweden	0.67	34.5	68.0	93.4	32.6	36.5	38.8	30.9	34.2	35.0	33.9
UK	0.87	28.6	71.5	97.1	28.7	28.4	30.1	26.2	–	28.9	28.2
US	0.75	31.7	70.5	92.8	31.4	32.1	33.8	32.1	–	32.0	31.4

Note: RRS: rank-rank slope; AUM: expected rank at $t+5$ conditional on being below the median at time t; ADM: expected rank at $t+5$ conditional on being above the median at time t; M99: expected rank at $t+5$ conditional on being in the top 1% at time t; Pooled, 1997–2007: calculated using all the available years between 1997 and 2007. In the calculations of the last three columns, if 1987, 1997, or 2007 are not available for a particular country, the closest year in the sample is chosen.

the national income distribution gain ranks, and those who start with a high position lose ranks over time. RRS are, however, much closer to one than to zero, suggesting that there is significant 5-year persistence in ranks. The data present sizable variation across countries. The countries displaying the largest degree of intragenerational mobility are, unsurprisingly, the Nordic trio (Denmark, Norway, and Sweden). Argentina also shows high levels of life cycle mobility, in sharp contrast to the other two Latin American countries, Brazil and Mexico, which display a high degree of persistence. Among the most developed countries, UK, Italy, and France have very low levels of 5-year mobility.

It is also of interest to compare intra to intergenerational mobility across countries. Kenedi and Sirugue (2021, Table 2) collect recent estimates of the RRS of the relationship between the expected income of children and the income of parents for most countries in our sample (Brazil, Canada, Denmark, France, Germany, Italy, Norway, Sweden, UK, and the US).[40] The cross-country correlation between these estimates of intergenerational persistence and our estimates of life-cycle persistence in Table 4 is 0.69, indicating that there might be strong common forces that jointly determine the two. These simple calculations are only suggestive: long longitudinal samples that enable linking subsequent generations would allow a full investigation of this relationship.

Bottom and top mobility

The expected rank 5 years ahead for an individual below the median varies between 29 in Brazil and 35 in Sweden, whereas for an individual starting above the median rank

[40] The intergenerational RRS from Brazil is obtained from Britto, Fonseca, Pinotti, Sampaio, and Warwar (2022), and the one for the UK from Rohenkohl (2019).

varies between 68 in Sweden and 74 in Brazil. While these cross-country differences (6 percentiles between the most and least mobile countries) seem small at first sight, recall that they are generated only over the course of 5 years. Using the linear relation in (2) together with a simple first-order Markovian model, one obtains that over the course of 20 years, or half a worker career, this gap would amplify to 10 percentiles.[41]

At the very top, ranks are even more persistent; only in Canada and the Scandinavian countries does an individual who is in the top 1% rank below the top 5% 5 years later. In Brazil and Italy, persistence in the top 1% is nearly full.

Subpopulations

Bottom and top mobility are, in general, larger for women than for men. This result could be caused by the fact that women, especially in the initial phase of their career, enter and exit the labor force more often than men, especially around childbirth. Brazil and Mexico are the exceptions: in these two countries, women display more rank persistence. In every country in our sample, young workers exhibit larger upward mobility than older ones: as workers age, their position in the income distribution becomes stickier. This finding clearly reflects the rising profile of earnings over the life cycle, but it is not the whole story because downward mobility at the top of the distribution is also higher for the young. This result is in line with what we observed for cross-sectional volatility, that is, the dispersion of earnings shocks declines with age. Note that in Nordic countries the gap in 5-year upward mobility between workers ages 25–34 and 45–55 is substantial, around 7–8 percentiles, and much bigger than in every other country.

Time trends

The last section of Table 4 examines whether any changes in the degree of intragenerational mobility have occurred over time. We detect no clear difference between 1987 and 2007 in any of our countries. As a result, the trends in cross-sectional inequality in current earnings that we documented in Section 3 have not been offset by stronger mobility, and thus, have largely translated into similar trends of inequality in permanent earnings.

Great Gatsby curve

The empirical literature on intergenerational mobility has documented the existence of a negative statistical relation with various measures of cross-sectional inequality, the so-called Great Gatsby curve (Durlauf and Seshadri (2018)): more unequal countries are also less intergenerationally mobile. Does this relation also hold with respect to intragenerational mobility? Figure 12 shows that in our sample, inequality, as measured by the Gini coefficient, is indeed positively correlated with our estimates of income rank persistence over 5 years. The correlation is, however, not particularly strong because of countries like Italy and the UK, which appear to have low Gini but high persistence, and Argentina which displays the opposite pattern.

[41] From equation (2), it is clear that $\alpha = (1-\beta)50$. One can then recursively apply the relation $P_{t+5} = (1-\beta)50 + \beta P_t$ to obtain the N-step (or $N \times 5$ years ahead) expected rank.

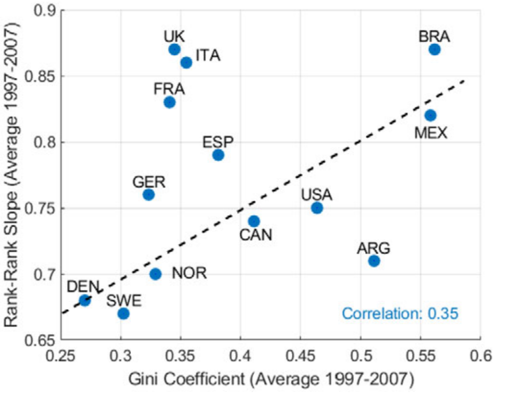


FIGURE 12. Great Gatsby curve for countries in our sample.

7. Conclusions

This paper offered an overview of GRID, the Global Repository of Income Dynamics, a new open-access, cross-country database containing a wealth of micro statistics on earnings inequality, the distribution of earnings changes, and earnings rank mobility over the life cycle. The database has four key characteristics: it fully exploits the longitudinal dimension of the administrative data sets, where—unlike survey data—attrition is not a result of nonresponse or refusal but rather a reaction to rare life events (death, international migration, or—for some countries—a transition out of dependent employment); it is built on administrative microdata (from Social Security records or other government registries); it cuts the data along various demographic traits of the population, making the database granular in nature; and it is maximally harmonized across countries. In the paper, we also presented key global trends that emerge from the analysis of the 13 GRID countries.

Our plan going forward is to extend the database in at least three ways. First, we will populate the database with many additional countries. Our flexible master code makes it straightforward to generate thousands of micro statistics from virtually any administrative data set on earnings. At least a dozen of additional countries have expressed interest and will soon enter the database. Second, whenever possible, we will add other definitions of income. For example, many countries have also information on self-employment income.[42] In addition, for all those countries that have tax authorities as a data source, it is an easy matter to construct accurate measures of disposable income as well.[43] Finally, perhaps the biggest advantage of using administrative data is the ability for these data to be linked to other data sources. Whenever possible, we plan to make these data linkages and incorporate into the database additional information on workers' spouses, employers, government taxes and transfers, and other critical statistics.

[42] Drechsel-Grau, Peichl, Schmieder, Schmid, Walz, and Wolter (2022a) analyzes how different earnings dynamics are for workers and entrepreneurs in Germany.

[43] Leth-Petersen and Sæverud (2022) present a comparison of pre-government and post-government labor income for Denmark.

While still at the initial stage, we expect the GRID project to develop quickly into one of the leading cross-country databases that researchers, policy-makers, journalists, and the public can freely access to investigate many questions related to the themes of income inequality, income risk, and economic mobility.

References

(2021), *The Role of Firms in Wage Inequality: Policy Lessons From a Large Scale Cross-Country Study*. OECD Publishing, Paris. [1322]

Acemoglu, Daron and David H. Autor (2011), “Skills, tasks and technologies: Implications for employment and earnings.” *Handbook of Labor Economics*, 4, 1043–1171. [1327]

Altonji, Joseph, Disa Hynsjö, and Ivan Vidangos (2022), “Individual earnings and family income: Dynamics and distribution.” Working paper 30095, National Bureau of Economic Research. [1327]

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme (2017), “Earnings and consumption dynamics: A nonlinear panel data framework.” *Econometrica*, 85 (3), 693–734. [1327]

Arellano, Manuel, Stephane Bonhomme, Micole De Vera, Laura Hospido, and Siqi Wei (2022), “Income risk inequality: Evidence from Spanish administrative records.” *Quantitative Economics*, 95. [1325]

Autor, David, Lawrence Katz, and Melissa S. Kearney (2008), “Trends in U.S. wage inequality: Revising the revisionists.” *Review of Economics and Statistics*, 90 (2), 300–323. [1335]

Bell, Brian, Nicholas Bloom, and Jack Blundell (2022), “Income dynamics in the UK and the impact of the Covid-19 recession.” *Quantitative Economics*, 41. [1325]

Blanco, Julio, Bernardo Diaz de Astarloa, Andres Drenik, Christian Moser, and Danilo Trupkin (2022), “The evolution of the earnings distribution in a volatile economy: Evidence from Argentina.” Quantitative Economics. [1325]

Bloom, Nicholas, Fatih Guvenen, Luigi Pistaferri, John Sabelhaus, Sergio Salgado, and Jae Song (2017), *The Great Micro Moderation.* Technical Report. University of Minnesota. [1342]

Bonhomme, Stéphane and Jean-Marc Robin (2009), “Assessing the equalizing force of mobility using short panels: France, 1990–2000.” *Review of Economic Studies*, 76 (1), 63–92. [1327]

Bowlus, Audra, Emilien Gouin-Bonenfant, Huju Liu, Lance Lochner, and Youngmin Park (2022), “Four decades of Canadian earnings inequality and dynamics across workers and firms.” *Quantitative Economics*, 44. [1325]

Britto, Diogo, Alexandre Fonseca, Paolo Pinotti, Breno Sampaio, and Lucas Warwar (2022), *Intergenerational Mobility in the Land of Inequality.* Technical Report. Bocconi University. [1354]

Busch, Christopher, David Domeij, Fatih Guvenen, and Rocio Madeira (2022), "Skewed idiosyncratic income risk over the business cycle: Sources and insurance." *American Economic Journal: Macroeconomics*, 14 (2), 207–242. [1346]

Card, David, Jörg Heining, and Patrick Kline (2013), "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly Journal of Economics*, 128 (3), 967–1015. [1328]

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014), "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129 (4), 1553–1623. [1353]

Congressional Budget Office (2007), "Trends in earnings variability over the past 20 years." Technical Report, Congressional Budget Office. [1342]

Criscuolo, Chiara, Alexander Hijzen, Cyrille Schwellnus, Wen-Hao Chen, Richard Fabling, Priscilla Fialho, Katarzyna Grabska, Ryo Kambayashi, Timo Leidecker, Oskar Nordstrom Skans, et al. (2020), "Workforce composition, productivity and pay: The role of firms in wage inequality." [1322]

Cunha, Flavio, James Heckman, and Salvador Navarro (2005), "Separating uncertainty from heterogeneity in life cycle earnings." *Oxford Economic Papers*, 57 (2), 191–261. [1338]

Drechsel-Grau, Moritz, Andreas Peichl, Johannes Schmieder, Kai D. Schmid, Hannes Walz, and Stefanie Wolter (2022b), "Inequality and income dynamics in Germany." *Quantitative Economics*, 41. [1325]

Drechsel-Grau, Moritz, Andreas Peichl, Johannes F. Schmieder, Kai D. Schmid, Hannes Walz, and Stefanie Wolter (2022a), "Inequality and income dynamics in Germany." Quantitative Economics. [1356]

Durlauf, Steven N. and Ananth Seshadri (2018), "Understanding the great Gatsby curve." *NBER Macroeconomics Annual*, 32 (1), 333–393. [1327, 1355]

Dynan, Karen E., Douglas W. Elmendorf, and Daniel E. Sichel (2012), "The evolution of household income volatility." Available at SSRN: http://ssrn.com/abstract=2138990. [1342]

Engbom, Niklas, Gustavo Gonzaga, Christian Moser, and Roberta Olivieri (2022), "Earnings inequality and dynamics in the presence of informality: The case of Brazil." *Quantitative Economics*, 71. [1325]

Fields, Gary S. and Efe A. Ok (1999), "The measurement of income mobility: An introduction to the literature." In *Handbook of Income Inequality Measurement*, 557–598. [1327, 1328, 1353]

Friedrich, Benjamin, Lisa Laun, and Costas Meghir (2022), "Earnings dynamics of immigrants and natives in Sweden 1985–2016." *Quantitative Economics*, 96. [1325]

Geweke, John and Michael Keane (2000), "An empirical analysis of earnings dynamics among men in the PSID: 1968–1989." *Journal of Econometrics*, 96, 293–356. [1327]

Gottschalk, Peter and Robert Moffitt (1994), "The growth of earnings instability in the U.S. labor market." *Brookings Papers on Economic Activity*, 25, 217–272. [1342]

Guvenen, Fatih (2007), "Learning your earning: Are labor income shocks really very persistent?" *American Economic Review*, 97 (3), 687–712. [1338]

Guvenen, Fatih, Greg Kaplan, and Jae Song (2020), "The glass ceiling and the paper floor: Changing gender composition of top earners since the 1980s." *NBER Macroeconomics Annual 2020*, 35, 309–373. [1336]

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021), "What do data on millions of U.S. workers say about labor income risk?" *Econometrica*, 89 (5), 2303–2339. [1341]

Guvenen, Fatih, Serdar Ozkan, and Jae Song (2014), "The nature of countercyclical income risk." *Journal of Political Economy*, 122 (3), 621–660. [1327, 1346]

Guvenen, Fatih, Luigi Pistaferri, and Giovanni L. Violante (2022), "Supplement to 'Global trends in income inequality and income dynamics: New insights from GRID'." *Quantitative Economics Supplemental Material*, 13, https://doi.org/10.3982/QE2260. [1344]

Halvorsen, Elin, Serdar Ozkan, and Sergio Salgado (2022), "Earnings dynamics and its intergenerational transmission: Evidence from Norway." *Quantitative Economics*, 115. [1325]

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante (2010), "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006." *Review of Economic Dynamics*, 13 (1), 15–51. [1328]

Hoffmann, Eran B., Davide Malacrino, and Luigi Pistaferri (2022), "Earnings dynamics and labor market reforms: The Italian case." *Quantitative Economics*, 50. [1325, 1339]

Katz, Lawrence F. and David H. Autor (1999), "Changes in the wage structure and earnings inequality." In *Handbook of Labor Economics* (Orley Ashenfelter and David Card, eds.). [1327]

Kenedi, Gustave and Louis Sirugue (2021), *The Anatomy of Intergenerational Income Mobility in France and Its Spatial Variations.* Technical Report, Sciences Po Economics Discussion Paper 2021-09. [1354]

Kramarz, Francis, Elio Nimier-David, and Thomas Delemotte (2022), "Inequality and earnings dynamics in France: National policies and local consequences." *Quantitative Economics*, 84. [1325]

Krueger, Dirk, Fabrizio Perri, Luigi Pistaferri, and Giovanni L. Violante (2010), "Cross-sectional facts for macroeconomists." *Review of Economic dynamics*, 13 (1), 1–14. [1326]

Leth-Petersen, Soren and Johan Saeverud (2022), "Inequality and dynamics of earnings and disposable income in Denmark 1987–2016." *Quantitative Economics*, 61. [1325]

Leth-Petersen, Søren and Johan Sæverud (2022), "Inequality and dynamics of earnings and disposable income in Denmark 1987–2016." Quantitative Economics. [1356]

Manski, Charles (2004), "Measuring expectations." *Econometrica*, 72 (5), 1329–1376. [1338]

McKinney, Kevin L., John M. Abowd, and Hubert P. Janicki (2022), "U.S. long-term earnings outcomes by sex, race, ethnicity, and place of birth." *Quantitative Economics*, 77. [1325, 1341]

Meghir, Costas and Luigi Pistaferri (2011), "Earnings, Consumption, and Life Cycle Choices." In *Handbook of Labor Economics*, Vol. 4B (Orley Ashenfelter and David Card, eds.), 773–854, Elsevier. [1327]

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan (2015), "Household surveys in crisis." *Journal of Economic Perspectives*, 29 (4), 199–226. [1324]

Moffitt, Robert and Sisi Zhang (2018), "Income volatility and the PSID: Past research and new results." *AEA Papers and Proceedings*, 108, 277–280. [1342]

Moffitt, Robert A. and Peter Gottschalk (1995), "Trends in the variances of permanent and transitory earnings in the U.S. and their relation to earnings mobility." Boston College Working Papers in Economics 444. [1342]

National Research Council (2013), *Nonresponse in Social Science Surveys: A Research Agenda*. The National Academies Press, Washington, DC. [1324]

Pistaferri, Luigi (2001), "Superior information, income shocks and the permanent income hypothesis." *Review of Economics and Statistics*, 83 (3), 465–476. [1338]

Puggioni, Daniela, Mariana Calderon, Alfonso Cebreros Zurita, Leon Fernandez Bujanda, Jose Antonio Gonzalez, and David Jaume (2022), "Inequality, income dynamics, and worker transitions: The case of Mexico." Quantitative Economics. [1325]

Rohenkohl, Bertha (2019), "Intergenerational income mobility in the UK: New evidence using the BHPS and understanding society." *The Sheffield Economic Research Paper Series (SERPS)*, 2019, 2019017. [1354]

Sabelhaus, John and Jae Song (2010), "The great moderation in micro labor earnings." *Journal of Monetary Economics*, 57, 391–403. [1342]

Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter (2019), "Firming up inequality." *Quarterly Journal of Economics*, 134 (1), 1–50. [1328]

Co-editor Stephane Bonhomme handled this manuscript.

Manuscript received 28 September, 2022; final version accepted 6 October, 2022; available online 6 October, 2022.