The GRID Project: An Overview and Some New Insights

Fatih Guvenen, Luigi Pistaferri, and Gianluca Violante

U of Minnesota, Stanford, Princeton

WCEG Webinar

April 19, 2023

The GRID Project: A Bird's Eye View

GRID: Global Repository of Income Dynamics

The GRID Project: A Bird's Eye View

GRID: Global Repository of Income Dynamics

Joint initiative: Princeton (Gianluca Violante), Stanford (Luigi Pistaferri), U of Minnesota & MEBDI (Fatih Guvenen)

The GRID Project: A Bird's Eye View

GRID: Global Repository of Income Dynamics

Joint initiative: Princeton (Gianluca Violante), Stanford (Luigi Pistaferri), U of Minnesota & MEBDI (Fatih Guvenen)

Goal: To build an open-access, cross-country

The GRID Project: A Bird's Eye View

GRID: Global Repository of Income Dynamics

Joint initiative: Princeton (Gianluca Violante), Stanford (Luigi Pistaferri), U of Minnesota & MEBDI (Fatih Guvenen)

Goal: To build an open-access, cross-country

1. harmonized database
2. of detailed micro statistics
3. on income inequality & income dynamics
4. based on panel data
5. from administrative records

Current State: GRID Version 1.0

- Project has been ongoing for the past 5 years.

Current State: GRID Version 1.0

- Project has been ongoing for the past 5 years.
- 13 countries: US, UK, CAN, FRA, ITA, SPA, GER, NOR, SWE, DEN, MEX, BRA, ARG.
- 54 economists in 13 country teams.
- 500,000 to 1,500,000 statistics per country.

Current State: GRID Version 1.0

- Project has been ongoing for the past 5 years.
- 13 countries: US, UK, CAN, FRA, ITA, SPA, GER, NOR, SWE, DEN, MEX, BRA, ARG.
- 54 economists in 13 country teams.
- 500,000 to 1,500,000 statistics per country.
- Database and user-friendly website recently launched: www.grid-database.org

Current State: GRID Version 1.0

- Project has been ongoing for the past 5 years.
- 13 countries: US, UK, CAN, FRA, ITA, SPA, GER, NOR, SWE, DEN, MEX, BRA, ARG.
- 54 economists in 13 country teams.
- 500,000 to 1,500,000 statistics per country.
- Database and user-friendly website recently launched: www.grid-database.org
- A Special Issue of *Quantitative Economics* with 13 papers written by country teams was published in November 2022.

QE QUANTITATIVE ECONOMICS
JOURNAL OF THE ECONOMETRIC SOCIETY

SPECIAL ISSUE ON GLOBAL INCOME DYNAMICS

Vol. 13, No. 4 — November, 2022

In Progress: GRID Version 2.0

- We are currently expanding GRID to about 25 more countries:

In Progress: GRID Version 2.0

- We are currently expanding GRID to about 25 more countries:
  - Europe: Austria, Belgium, Finland, Greece, Hungary, Iceland, Ireland, Israel, Netherlands, Poland, Portugal, Switzerland
  - Australasia: Australia, India, Japan, New Zealand, Singapore, South Korea, Taiwan
  - South America: Chile, Colombia, Costa Rica, Ecuador, Peru, Uruguay.

In Progress: GRID Version 2.0

- We are currently expanding GRID to about 25 more countries:
  - Europe: Austria, Belgium, Finland, Greece, Hungary, Iceland, Ireland, Israel, Netherlands, Poland, Portugal, Switzerland
  - Australasia: Australia, India, Japan, New Zealand, Singapore, South Korea, Taiwan
  - South America: Chile, Colombia, Costa Rica, Ecuador, Peru, Uruguay.
- Goal is to have 40 countries in GRID by Summer 2024.

Background: Data on Income Inequality

- Several harmonized cross-country databases (of statistics) on income inequality are available:
  - World Inequality Database (WID.world) spearheaded by the work of T. Atkinson, T. Piketty and E. Saez
  - World Income Inequality Database (WIID2) maintained at the United Nations University
  - OECD Income Distribution Database (IDD)
  - Luxembourg Income Study (LIS)

Why This Project?

- Existing cross-country databases are typically:
  - based on survey data
  - focused on cross-sectional statistics (e.g. inequality)
  - reported for broad demographic groups (age, gender, etc.)

Why This Project?

- Existing cross-country databases are typically:
  - based on survey data
  - focused on cross-sectional statistics (e.g. inequality)
  - reported for broad demographic groups (age, gender, etc.)
- GRID aims to extend the focus to:
  - Administrative data, now available for many countries
  - The study of income dynamics (mobility, income instability, etc.) and other long-run phenomena
  - Finely-defined subpopulations (examples later)

I. Administrative Data

- Relative to survey data, admin data have minimal measurement error and little or no attrition

I. Administrative Data

- Relative to survey data, admin data have minimal measurement error and little or no attrition
- Size: granular analyses of sub-populations and tail events

I. Administrative Data

- Relative to survey data, admin data have minimal measurement error and little or no attrition
- Size: granular analyses of sub-populations and tail events
  - Income changes characterized by high probability of extreme events:
    - Effect on welfare or risk premia substantial.

I. Administrative Data

- Relative to survey data, admin data have minimal measurement error and little or no attrition
- Size: granular analyses of sub-populations and tail events
  - Income changes characterized by high probability of extreme events:
    - Effect on welfare or risk premia substantial.
- Ability to link data across various sources in some countries (income, expenditures, health, education)

I. Administrative Data

- Relative to survey data, admin data have minimal measurement error and little or no attrition
- Size: granular analyses of sub-populations and tail events
  - Income changes characterized by high probability of extreme events:
    - Effect on welfare or risk premia substantial.
- Ability to link data across various sources in some countries (income, expenditures, health, education)
- Caveat: In countries with large informal sectors (BRA, MEX, ARG), coverage limited to formal sector
  - Teams for these countries asked to validate findings using survey data; ongoing work

II. Panel Dimension

- Cross-sectional data (inequality):
  - Provides static snapshots of a distribution, but misses dynamics
  - Two societies can have similar levels of inequality, but very different chances of mobility

II. Panel Dimension

- Cross-sectional data (inequality):
  - Provides static snapshots of a distribution, but misses dynamics
  - Two societies can have similar levels of inequality, but very different chances of mobility
- Panel data:
  - Allows study of several dimensions important for welfare:
    - Income Dynamics (volatility, skewness, persistence, etc.)
    - Mobility (intra- and inter-generationally)
    - Tail behavior (probability of extreme wage cuts vs. wage hikes)
    - Inequality in long-run ("permanent") income

III. Heterogeneity Dimension

- Income risk, inequality, and mobility patterns vary significantly:
  - by birth year, gender, education, employer, permanent income, geographical location, etc.

III. Heterogeneity Dimension

- Income risk, inequality, and mobility patterns vary significantly:
  - by birth year, gender, education, employer, permanent income, geographical location, etc.
- In GRID, many statistics are reported for groups defined by
  - the interaction of gender x cohort x age group x "permanent" income.

III. Heterogeneity Dimension

- Income risk, inequality, and mobility patterns vary significantly:
  - by birth year, gender, education, employer, permanent income, geographical location, etc.
- In GRID, many statistics are reported for groups defined by
  - the interaction of gender x cohort x age group x "permanent" income.
  - Permanent income based on past 3-year average income. 40 quantile bins plus more at the top.

Detailed Statistics

- List of statistics inspired by recent empirical and theoretical work

Detailed Statistics

- List of statistics inspired by recent empirical and theoretical work
- Emphasis on providing flexible quantile-based statistics to allow computation of alternative measures:
  - Inequality: Top shares, Gini, Standard deviation, P90-P10, etc.
  - Volatility: P90-P10 for income changes, etc.
  - Skewness: 3rd centralized moment, Kelley's skewness, etc.

Detailed Statistics

- List of statistics inspired by recent empirical and theoretical work
- Emphasis on providing flexible quantile-based statistics to allow computation of alternative measures:
  - Inequality: Top shares, Gini, Standard deviation, P90-P10, etc.
  - Volatility: P90-P10 for income changes, etc.
  - Skewness: 3rd centralized moment, Kelley's skewness, etc.
- Statistics calculated for 4 different income measures: raw income, log income, residualized income, "permanent" income.

Overview of Available Statistics

- Income levels (static):
  - Time-series statistics (up to 35+ years for some countries)
  - Measures of inequality for annual and "permanent" income
  - Empirical densities

Overview of Available Statistics

- Income levels (static):
  - Time-series statistics (up to 35+ years for some countries)
  - Measures of inequality for annual and "permanent" income
  - Empirical densities
- Income changes (dynamics):
  - 1-year and 5-year income changes (transitory and persistent changes)
  - Log change and arc-percent change (to allow extensive margin—zero income)
  - Moments on dispersion, skewness, and kurtosis (tails)
  - Short- and Long-run mobility statistics
  - All available for very finely defined groups (e.g., Canadian women of age 30 in 2005)

Overview of Available Statistics

- Income levels (static):
  - Time-series statistics (up to 35+ years for some countries)
  - Measures of inequality for annual and "permanent" income
  - Empirical densities
- Income changes (dynamics):
  - 1-year and 5-year income changes (transitory and persistent changes)
  - Log change and arc-percent change (to allow extensive margin—zero income)
  - Moments on dispersion, skewness, and kurtosis (tails)
  - Short- and Long-run mobility statistics
  - All available for very finely defined groups (e.g., Canadian women of age 30 in 2005)
- A quick look at the GRID website: https://data.grid-database.org

IV. Harmonization

- Harmonization has been a key focus from the beginning.

IV. Harmonization

- Harmonization has been a key focus from the beginning.
- On the data side:

IV. Harmonization

- Harmonization has been a key focus from the beginning.
- On the data side:
  - Comparable income definitions: before-tax, annual, wage/salary income (exclude self employment income), including overtime, bonuses, vacation pay, and other items.

IV. Harmonization

- Harmonization has been a key focus from the beginning.
- On the data side:
  - Comparable income definitions: before-tax, annual, wage/salary income (exclude self employment income), including overtime, bonuses, vacation pay, and other items.
  - Same sample selection criteria
  - Similar time periods: For heterogeneity statistics, all countries use 20 years of data, going back from last available year.

IV. Harmonization

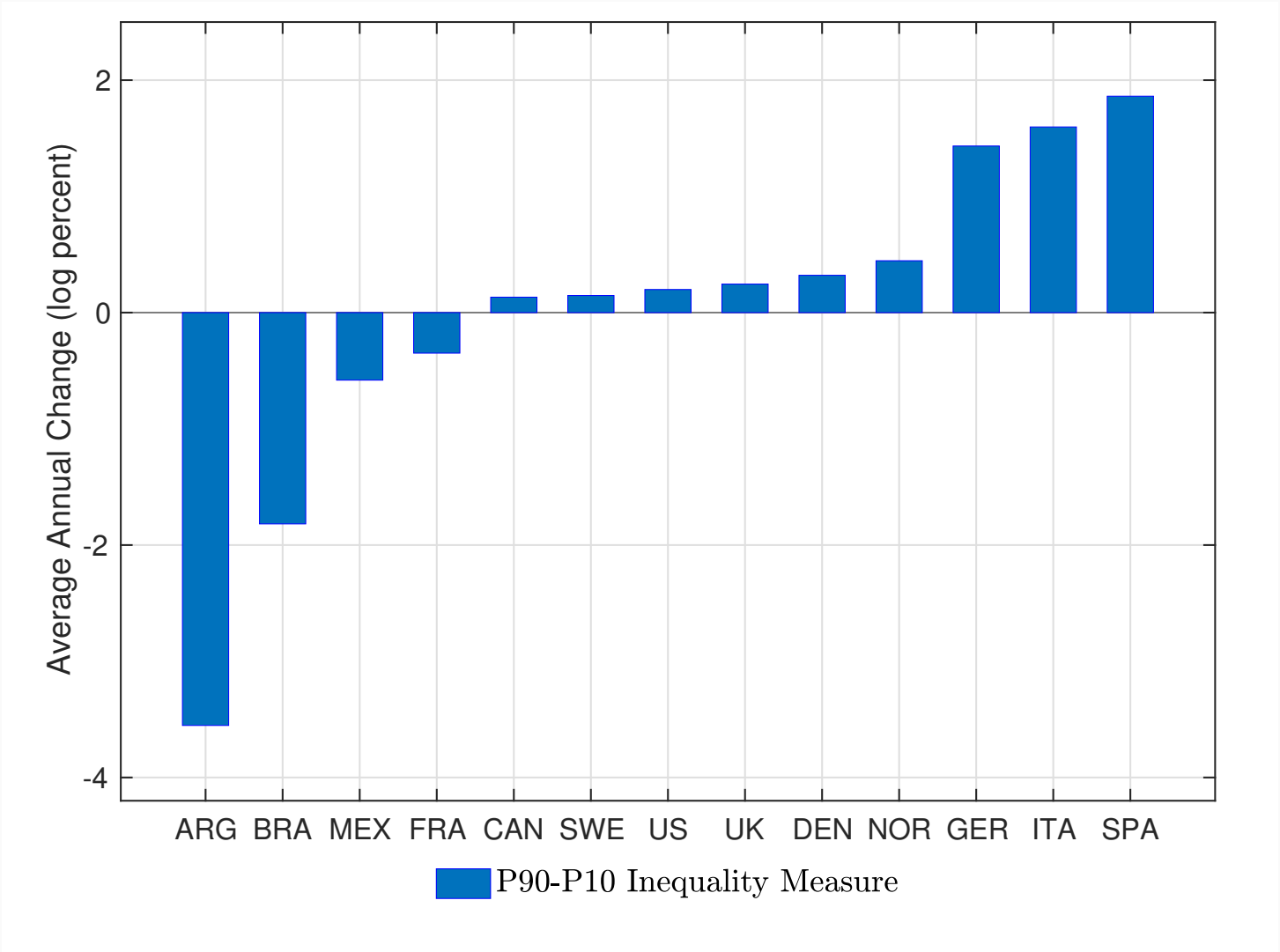
- Harmonization has been a key focus from the beginning.
- On the data side:
  - Comparable income definitions: before-tax, annual, wage/salary income (exclude self employment income), including overtime, bonuses, vacation pay, and other items.
  - Same sample selection criteria
  - Similar time periods: For heterogeneity statistics, all countries use 20 years of data, going back from last available year.
- But most importantly:
  - Single master code to generate all statistics from very first step of cleaning raw data to the very end results.

New Insights from GRID: Inequality

Question 1: Is Income Inequality Rising Globally?

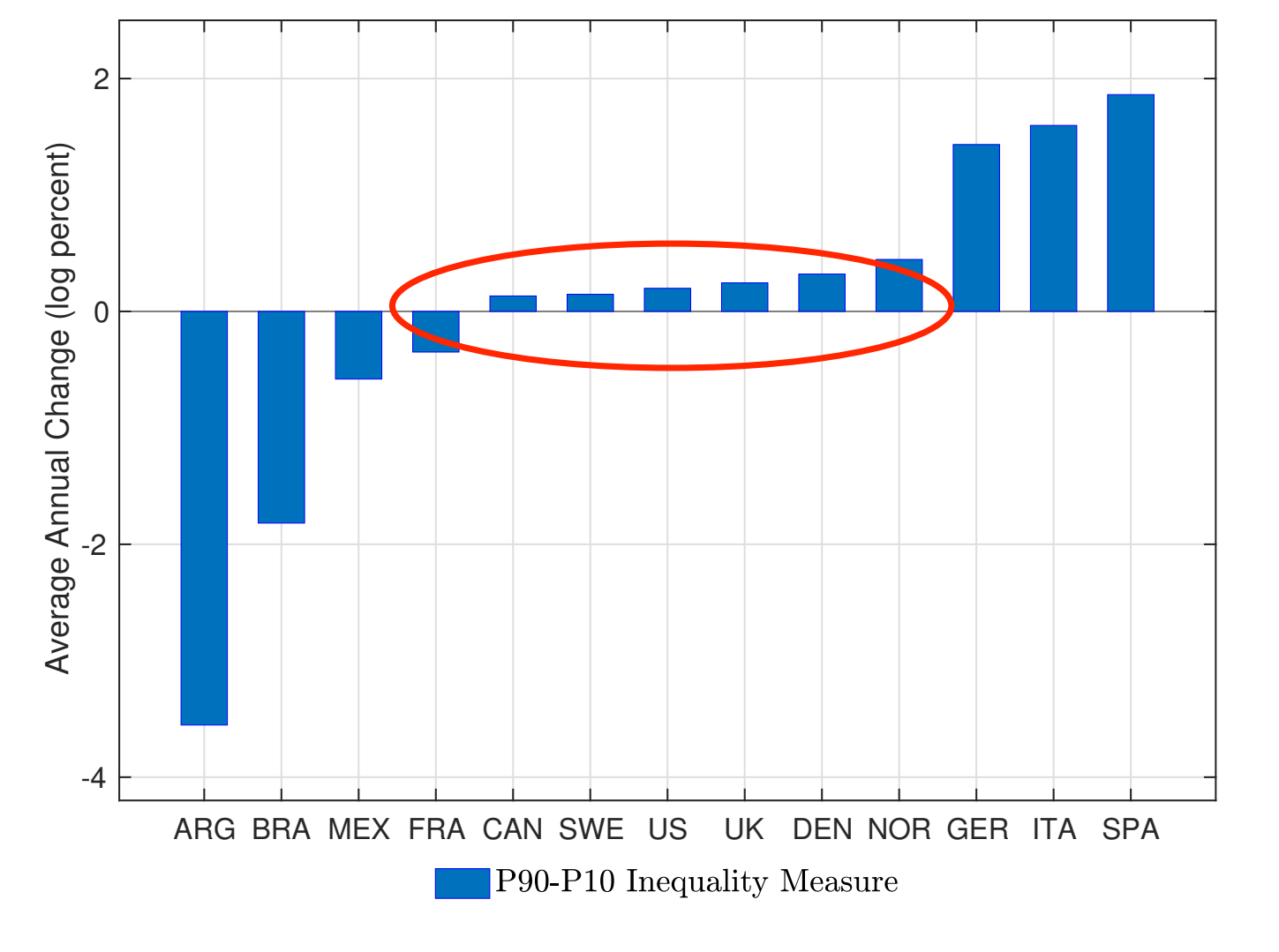
Question 1: Is Income Inequality Rising Globally?

Figure 1: Trends in Overall Income Inequality for GRID Countries



Question 1: Is Income Inequality Rising Globally?

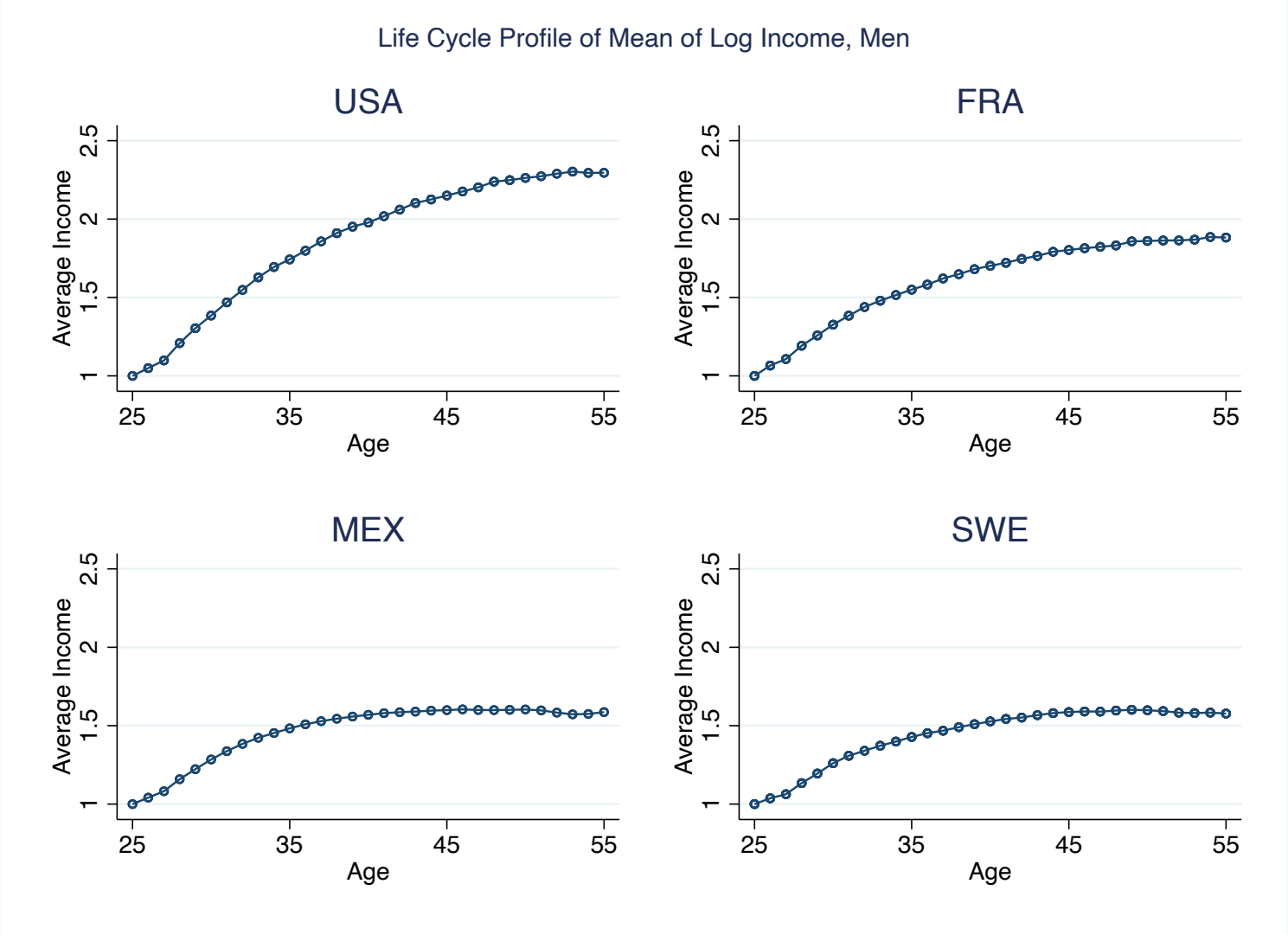
Figure 1: Trends in Overall Income Inequality for GRID Countries



Question 2: How Much Does Average Income Grow with Age?

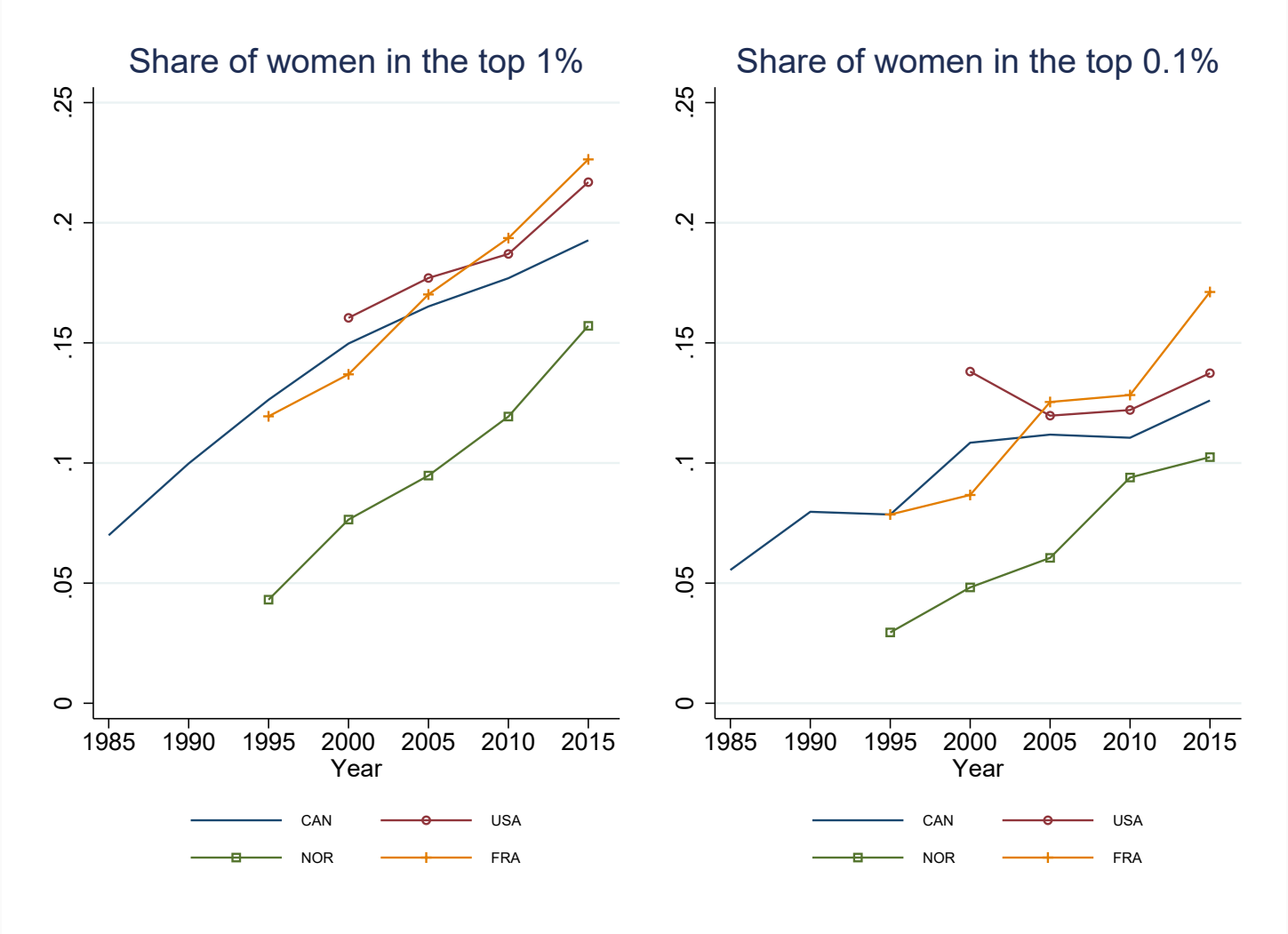
Question 2: How Much Does Average Income Grow with Age?

Figure 3: Average Life Cycle Profile of Income, Men



Question 3: What Fraction of Top Earners are Women?

Question 3: What Fraction of Top Earners are Women?



New Insights from GRID: Dynamics of Income

Question 4: Has Income Instability Increased Over Time?

- Answer crucial for many questions and policy design.

Rising income uncertainty in the US became conventional wisdom.

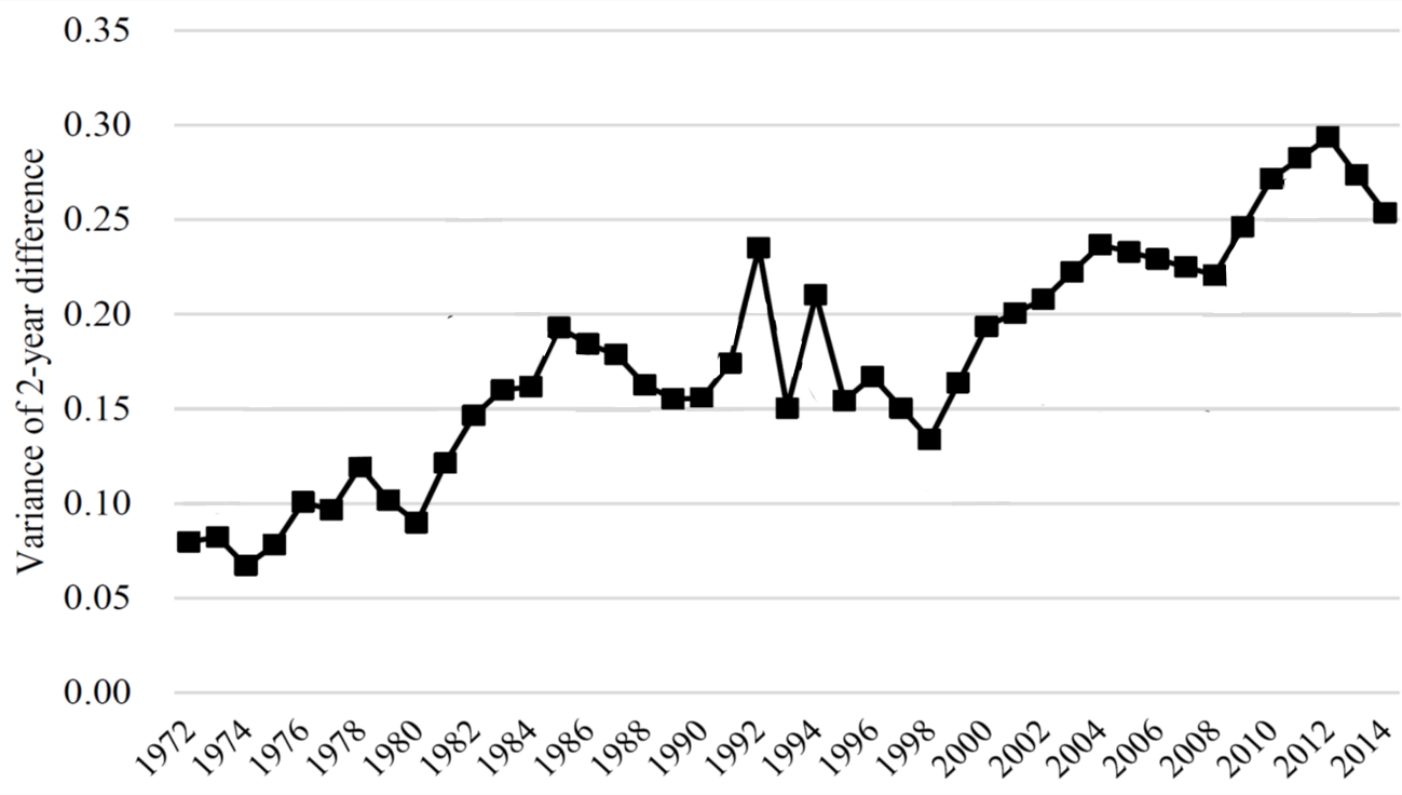
Question 4: Has Income Instability Increased Over Time?

- Answer crucial for many questions and policy design.
- Seminal paper by Moffitt and Gottschalk (1994) found:
  - Income volatility increased substantially in survey data from 1968 to 1988.
  - 30+ papers confirmed the result and extended to 2010s.

Rising income uncertainty in the US became conventional wisdom.

US Income Volatility in Survey Data Up

Figure 4: Variance of Income Growth (2-year)



Moffitt and Zhang (AEA P&P, 2018)

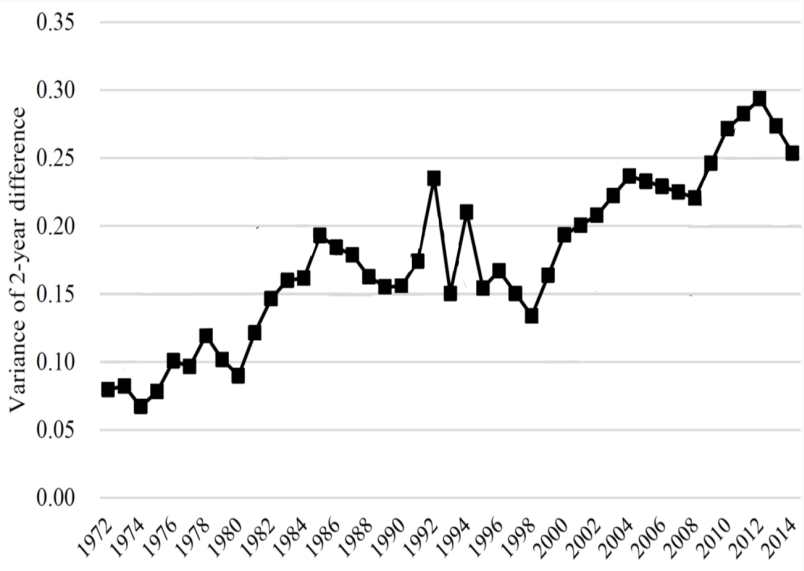
- Variance of income growth in PSID nearly tripled from 1970s to 2010s.

Has Income Instability Increased Over Time?

- Answer crucial for many questions and policy design.
- Seminal paper by Moffitt and Gottschalk (1994):
  - Income volatility increased substantially in PSID from 1968 to 1988.
  - 30+ papers confirmed the result and extended to 2010s.
- That income uncertainty is higher today has become conventional wisdom.

US Income Volatility Trending UP

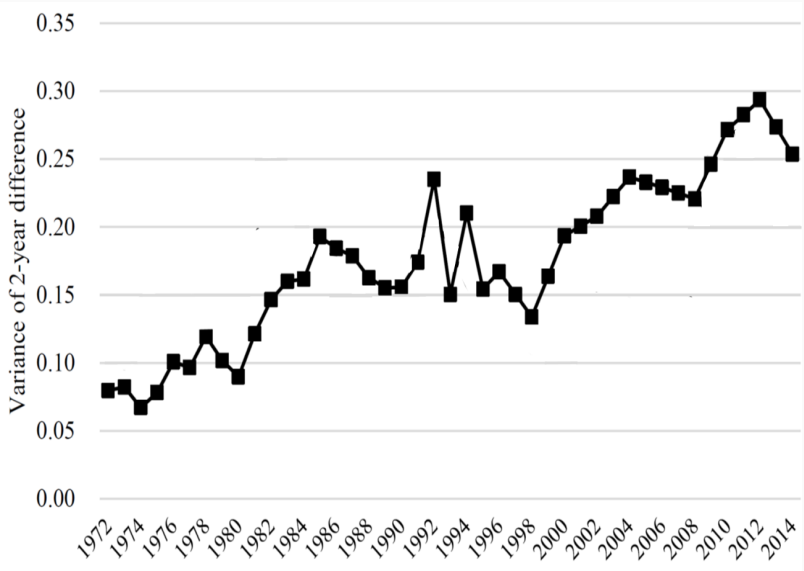
(a) Survey (PSID)



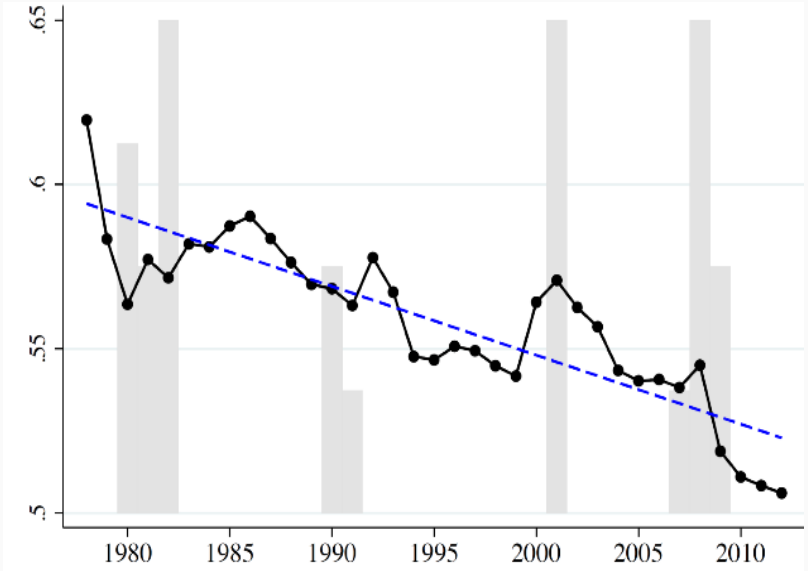
Left: Moffitt and Zhang (AEA P&P, 2018)

US Income Volatility Trending Up? or Down?

(a) Survey (PSID)



(b) Administrative (SSA Data)

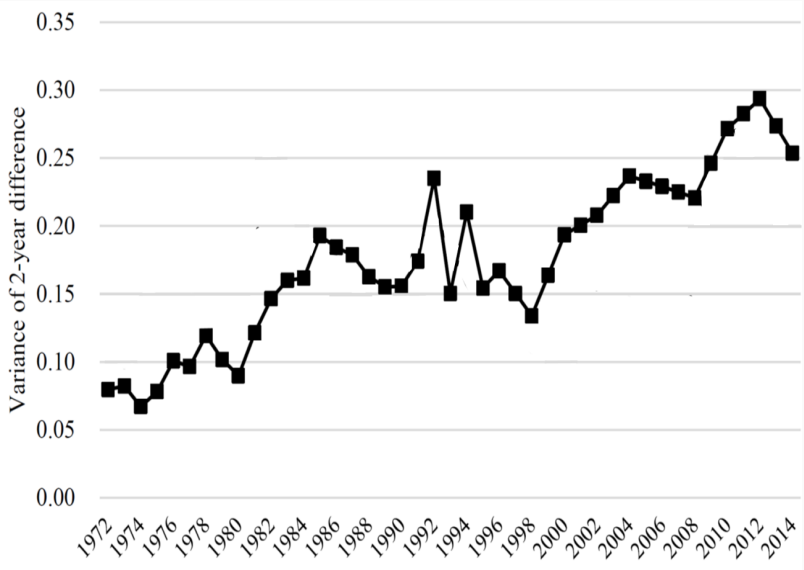


Left: Moffitt and Zhang (AEA P&P, 2018)

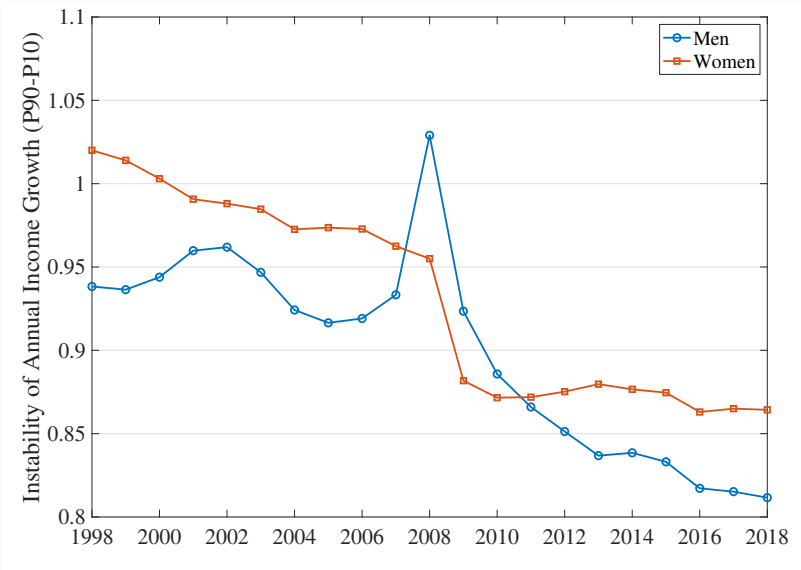
Right: Guvenen, Ozkan, Song (JPE, 2014)

US Income Volatility Trending Up? or Down?

(a) Survey (PSID)



(b) LEHD Data for GRID project

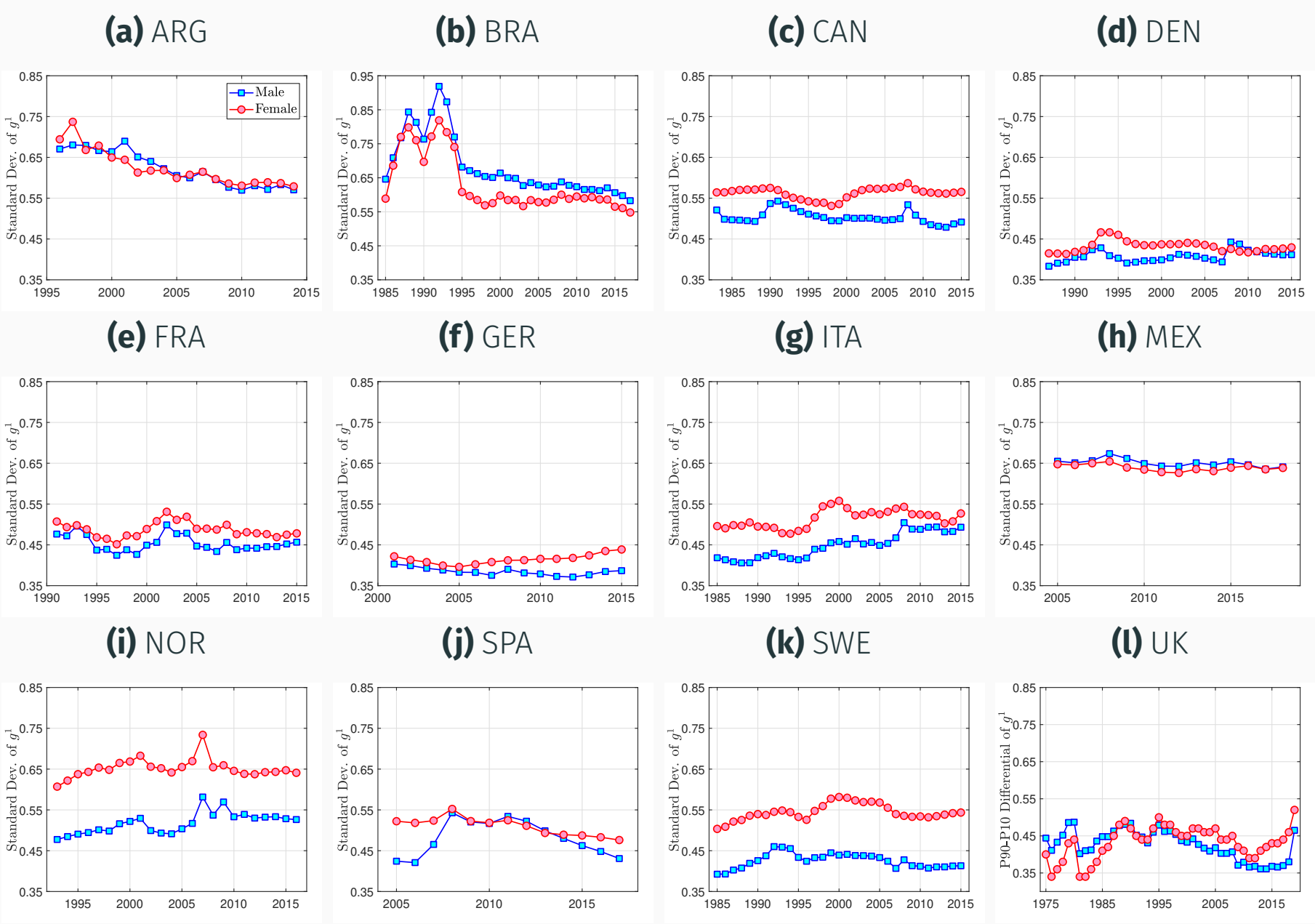


Left: Moffitt and Zhang (AEA P&P, 2018)

Right: GRID Project

How About Other Countries? Evidence from GRID

How About Other Countries? Evidence from GRID



Taking Stock

- Overall, income volatility has been
  - declining in several countries (Brazil, Argentina, USA, UK after mid-1980s)
  - relatively flat in some others (Canada, Denmark, France, Germany, Mexico, Spain, Sweden), and
  - increasing in a few (men in Italy, Norway, and women in UK pre-1990s).

Taking Stock

- Overall, income volatility has been
  - declining in several countries (Brazil, Argentina, USA, UK after mid-1980s)
  - relatively flat in some others (Canada, Denmark, France, Germany, Mexico, Spain, Sweden), and
  - increasing in a few (men in Italy, Norway, and women in UK pre-1990s).
- We do not see any broad-based increase in income instability as suggested by US survey data.

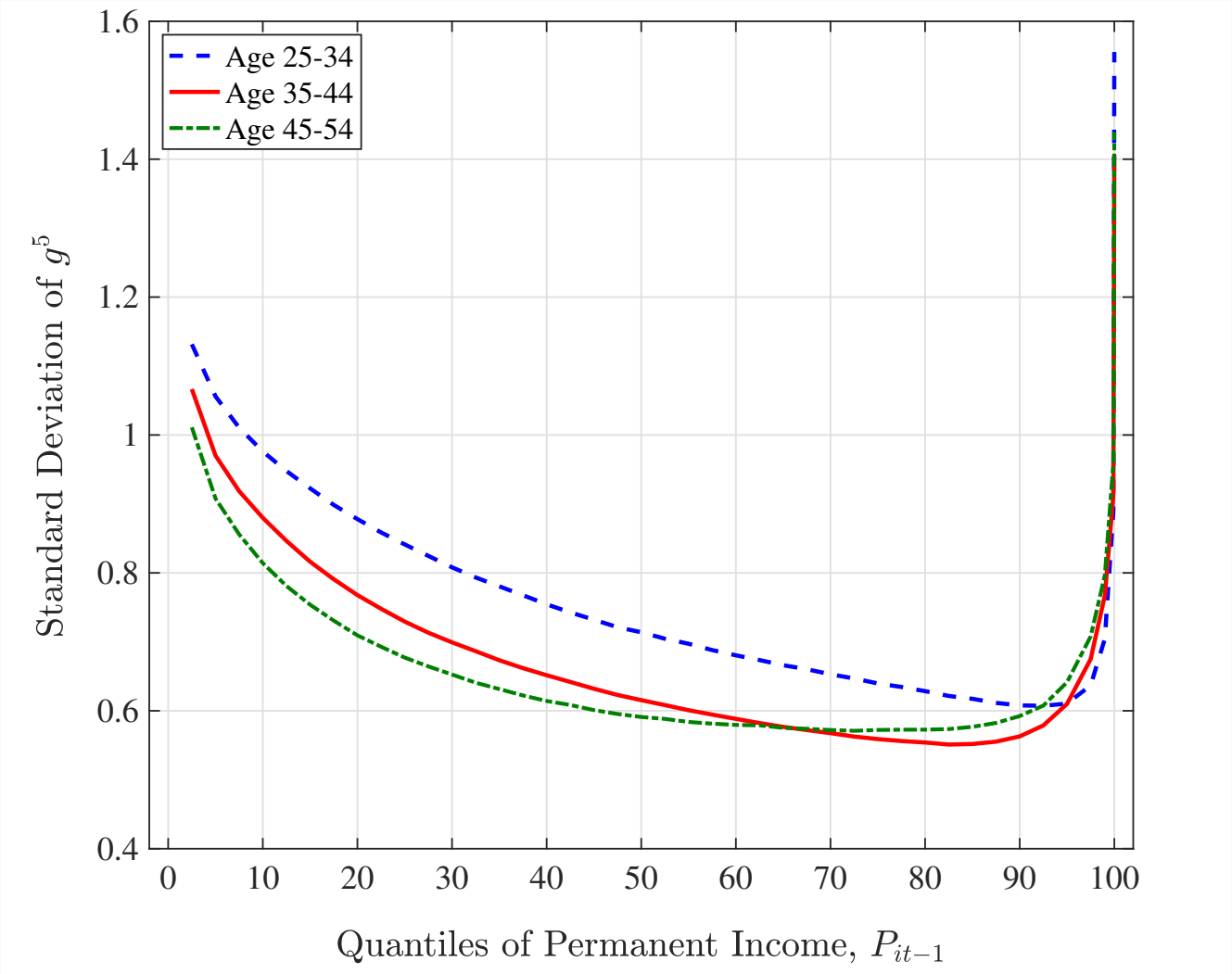
Taking Stock

- Overall, income volatility has been
  - declining in several countries (Brazil, Argentina, USA, UK after mid-1980s)
  - relatively flat in some others (Canada, Denmark, France, Germany, Mexico, Spain, Sweden), and
  - increasing in a few (men in Italy, Norway, and women in UK pre-1990s).
- We do not see any broad-based increase in income instability as suggested by US survey data.
- GRID contains very granular statistics that allows digging deeper into this question.

Question 5: How Does Income Volatility Vary by Income Level?

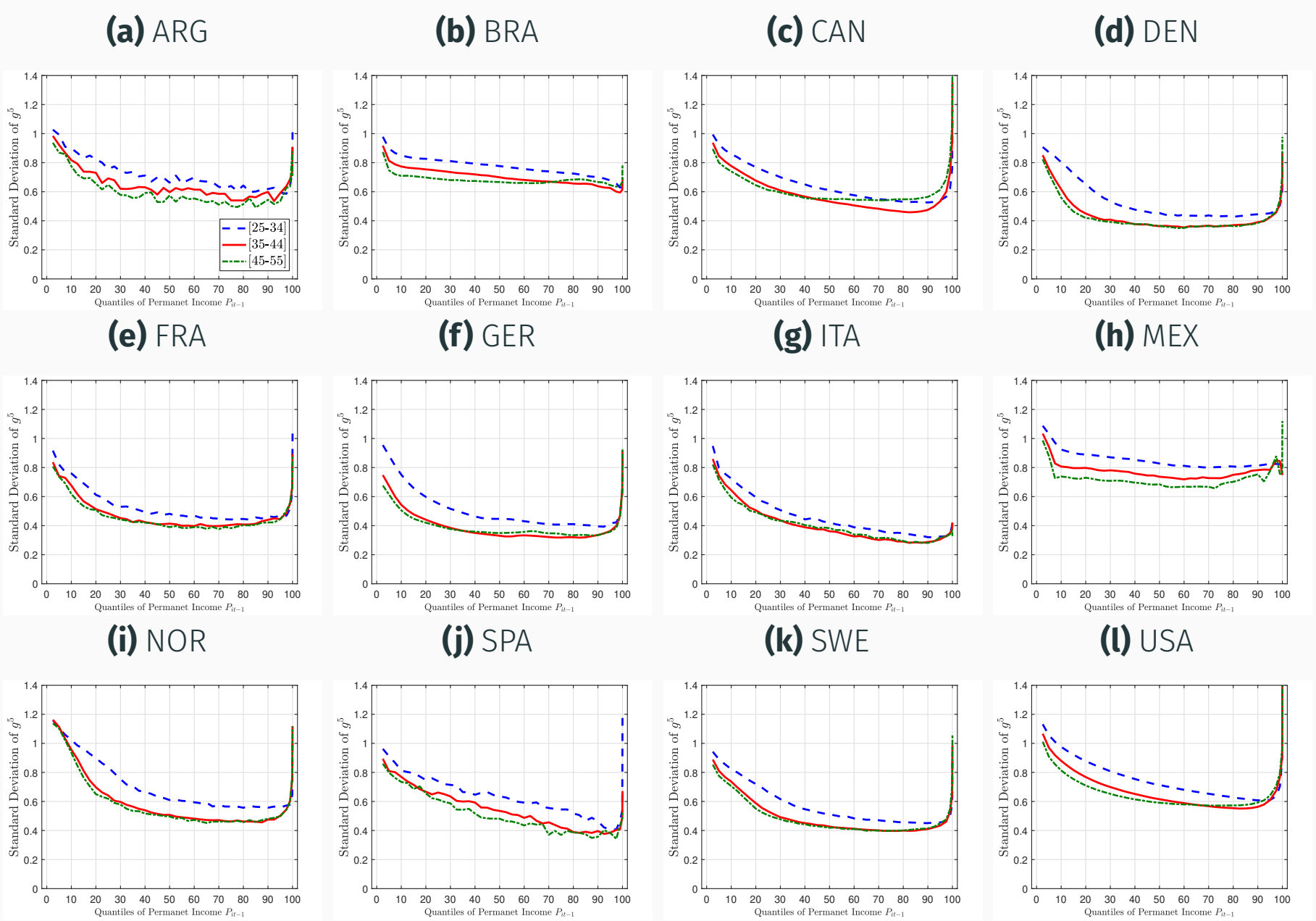
Question 5: How Does Income Volatility Vary by Income Level?

Figure 5: USA



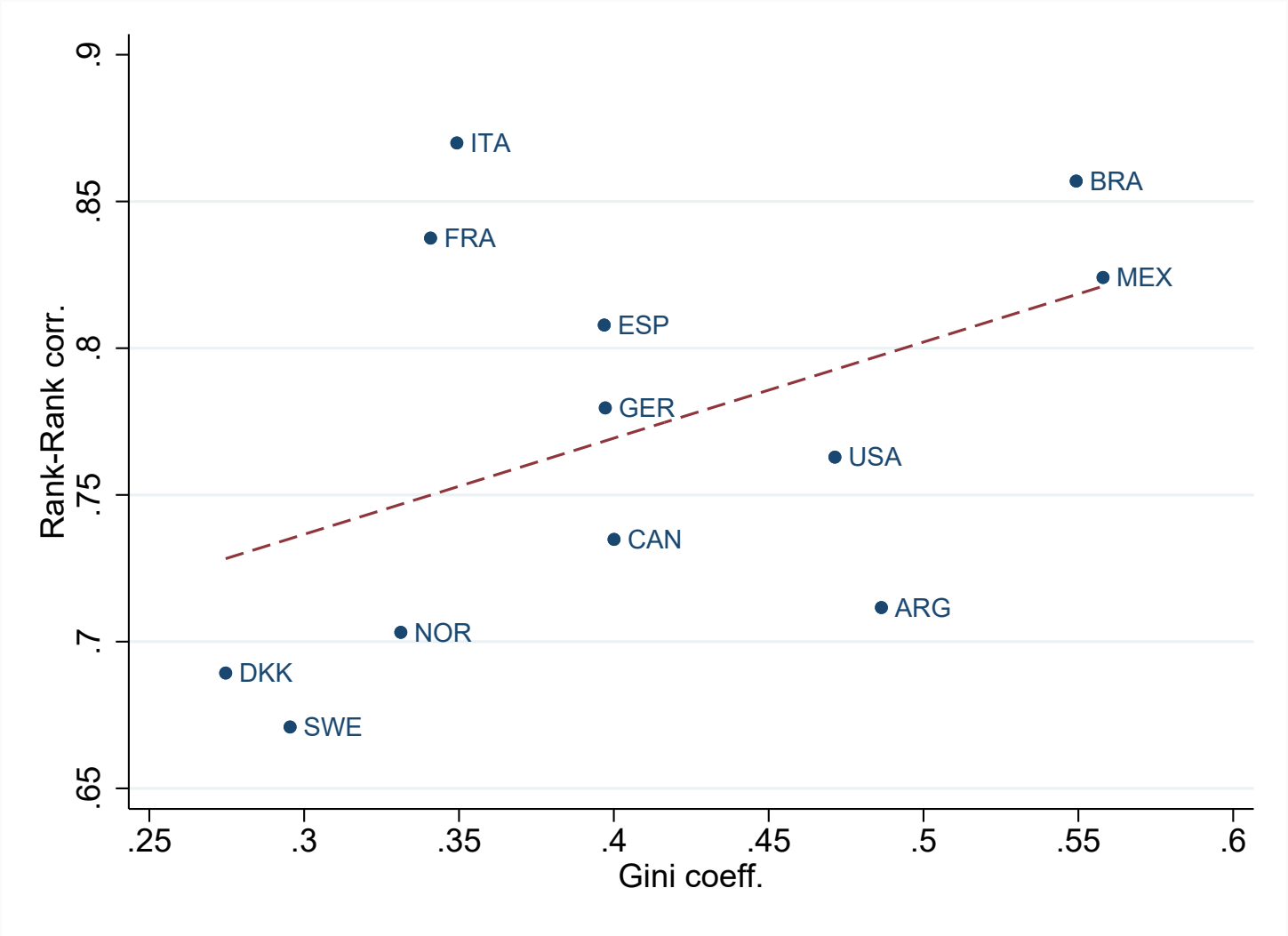
Question 5: How About Other Countries?

Question 5: How About Other Countries?



Question 6: Great Gatsby Curve

Question 6: Great Gatsby Curve



Conclusions

- GRID is an open-access, cross-country database of statistics on income inequality and income dynamics.
- GRID Version 2.0 will triple the number of countries in the database.
- We have future plans to expand the set of variables included in GRID.

Conclusions

- GRID is an open-access, cross-country database of statistics on income inequality and income dynamics.
- GRID Version 2.0 will triple the number of countries in the database.
- We have future plans to expand the set of variables included in GRID.
- We welcome all feedback and suggestions. Drop us a line at `support@grid-database.org`
- Follow us on Twitter for future updates: `@griddatabase`
- Thanks again to WCEG for giving us this opportunity to talk about GRID!

Thanks!