Will Inequality Continue to Rise? Forecasting Income Inequality in the United States

Marina Gindelsky

Bureau of Economic Analysis, U.S.

Discussant

Thesia I. Garner, Bureau of Labor Statistics, U.S.

Session 5 (Plenary)

Thursday, August 25, 2016

Starting in 1980….

- 1980 represents a turning point in inequality after which it “exploded" (Piketty 2014, Goldin and Katz 2008)
- Trends are different afterwards such that beginning in a previous period leads to worse forecasts
- Change in labor market conditions starting in 1980 due to skill-biased technological change (SBTC)
- Analysts do not expect future trends to resemble the prior period
- Swaps crosswalk only to 1975 too many changes in CPS survey methods

Motivation and Aim

- Efforts to explain trends in inequality-hot topic for policy
- Recent work by Piketty (2014) to predict future trends
- Role of skill-biased technological change—role in inequality trends
- Best model to predict trends?

➢ "This paper [attempts to] answers question by choosing models to forecast several inequality measures and providing short-term forecast."

Motivation

- Research Questions:
  - Will income inequality continue to rise in theshort-run?
  - Does this depend on other macro, human capital and labor market variables?
- Strong growth in income inequality in the United States since 1980.This growth has diered by group and by income share
- Contribution: Forecasting short-run inequality for the United States using microdata from Current Population Survey and potential determinants

To predict inequality, need answers to questions

- Most appropriate measure?
- Determinates?

Measure

- What to measure? Income
  - Individual earnings-wages, self-employment, and farm income
  - Household income-labor (70% of hh Y), capital income, capital income, government transfers
- How to measure? Dependent variable in regression model
  - Gini index
  - 90/10 income ratios
  - Income shares
- Where to measure using Gini?
  - Overall distribution
  - Top 1% income share
  - Top 0.1% income share

Determinants on Inequality Forecast

- Human capital attainment
- Labor Force structure
- Macroeconomic variables

Predictive Factors I: Human Capital Attainment Indicators

- % Population 25+ Years Who have Completed College
- % Female Population 25+ Years Who have Completed College
- % of Population 25+ Years Who have Completed High School
- % of Female Population 25+ Years Who have Completed High
- Skill Premium (College Wage/High School Wage)

Predictive Factors II: Labor Force Structure Indicators

- High-Skill Employment (Non-routine Cognitive)
- Middle-Skill Employment I (Routine Cognitive)
- Middle-Skill Employment II (Routine Manual)
- Low-Skill Employment (Non-routine Manual)
- Share of Services in GDP
- Labor Force Participation
- Female Labor Force Participation

❖ *"Skill" vars. defined as log number employed in group*

Predictive Factors III: Macroeconomic Indicators

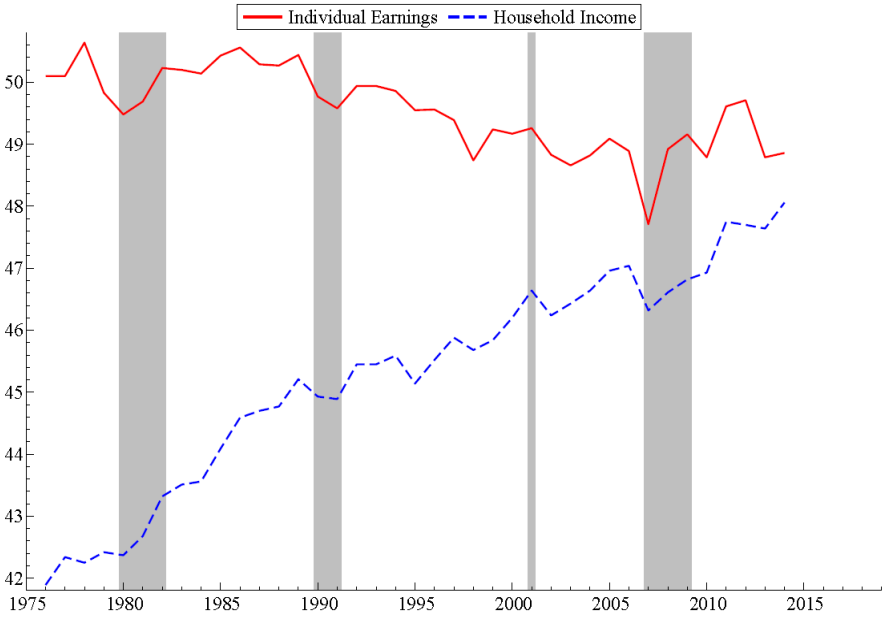
- Real GDP (*gdp*)
- Government Expenditure as a Share of GDP (*gov/gdp*)
- Inflation (*infl*)
- Unemployment (*unemp*)
- Male Unemployment (*m_unemp*)

Data: Current Population Survey

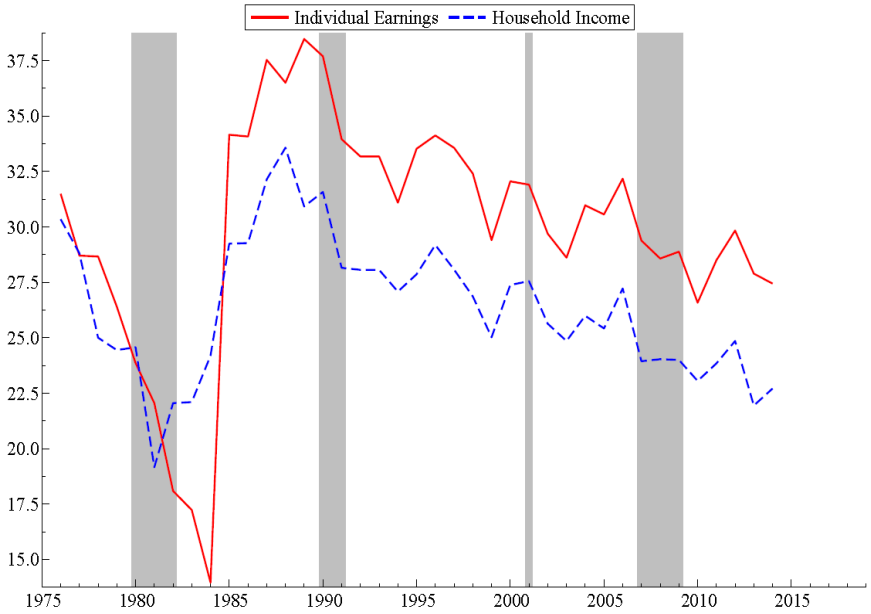
- Annual microdata in CPS 1975-2014 (public use)
- Structural breaks
  - 1993: Structural break due to survey instrument change
    - Asked specically about other sources of income
    - Allowed higher values for income reporting (internal topcode)
    - Weighting and imputation changes
    - Change in interview mode
  - Most of increase is not increase in income inequality but to structural break from 1992 to 1993
  - Others: 1985, 1987, 2007
- Top-coding - existence and consistency: used rank proximity swapping technique (all value greater >= topcode swapped with other values within a bounded interval to better represent internal data and allow for more accurate inequality calculations)

Individual and Household: All and Top 1%: Gini

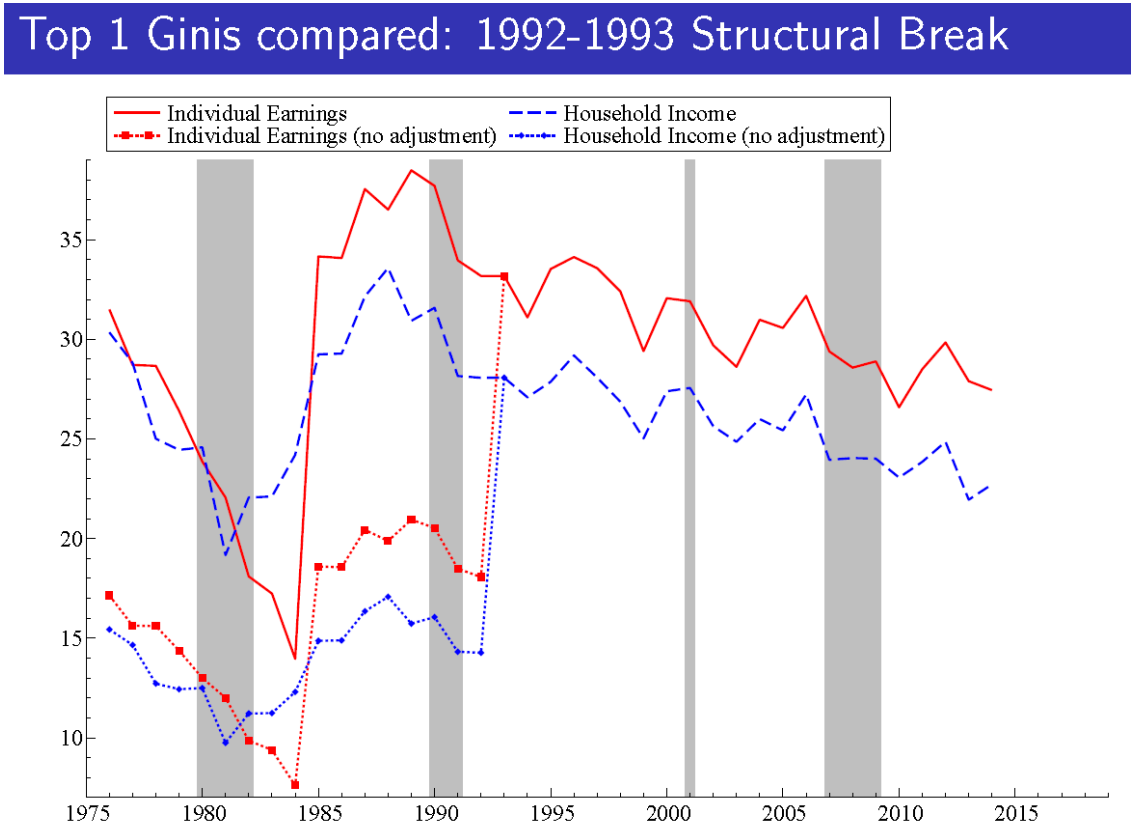
Ginis compared: Individuals & Households



Top 1 Ginis compared: Individuals & Households



Income Data Adjusted for Break



Model Selection

- Dickey-Fuller test: all series stationary after first differencing, except labor force participation and female labor force participation which were second differenced
- Model in differences, converted back into levels
- Max 4 lags
- Standard autoregressive models were chosen
- General-to-Specific modeling approach[1], with Impulse Indicator
- Saturation (Impulse Indicator Saturation) at 1% level
- Forecast comparison in pseudo-out-of-sample periods with Diebold-Mariano tests and White Reality Check
- Cautions: Correlation of indicators and overfitting

❖ *[1]Autometrics (OxMetrics) used to select best model*

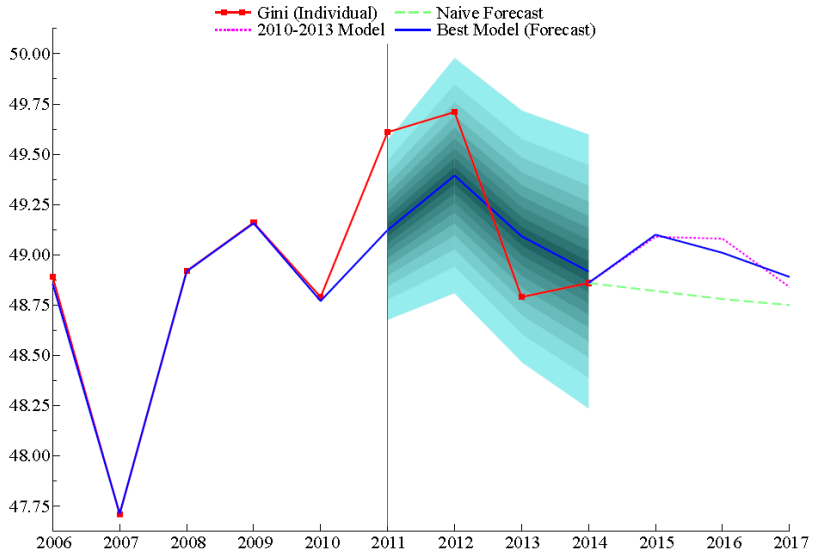
Results

- Best predictors
  - Human capital attainment
  - Labor force structure
- Model selection not robust but yields robust forecasts
- Out of sample forecasts differ between models by <6% for all variables and <2% for 4/8 measures
- Top 1% is projected to rise, while share of top 0.1% predicted to fall - consistent with inequality within top 1% falling
- Overall (Gini) individual inequality constant while household inequality rises

Gini: Individual Earnings

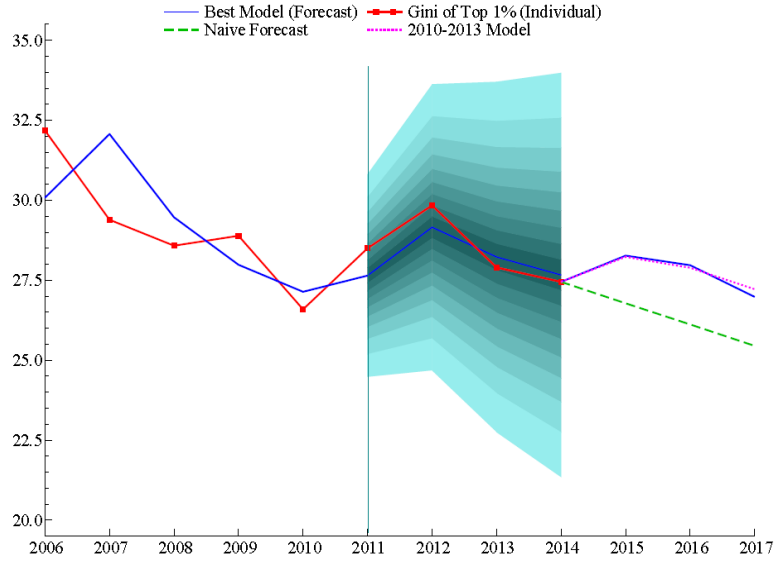
Gini: Individual Earnings

$$\widehat{\triangle Gini_t} = -0.04 - 1.97d2007 + 1.05d2008 - 0.15 \triangle Gini_{t-1} - 0.37 \triangle Gini_{t-2}$$
$$-09 \triangle mskill2_{t-1} - 3.55 \triangle mskill2_{t-2} + \epsilon_t$$



Gini of Top 1%: Individual Earnings

$$\widehat{\triangle Top1Gini_t} = -0.69 + 19d1985 + 208 \triangle \frac{mskill1}{lskill}_{t-1} - 258 \triangle \frac{mskill1}{lskill}_{t-2} - 185 \triangle \frac{mskill1}{lskill}_{t-3}$$

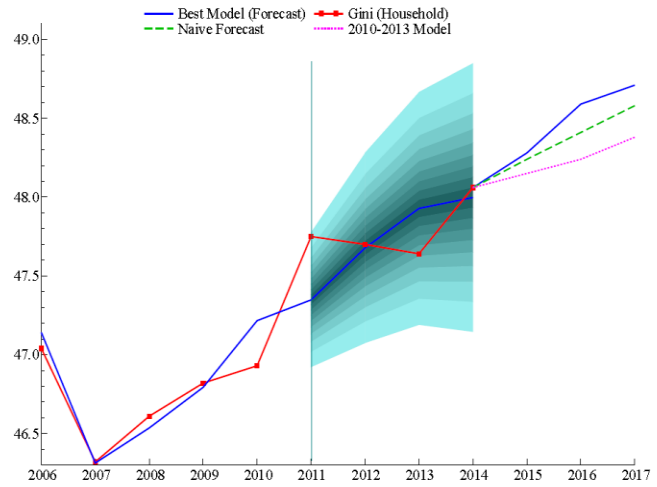


Pseudo out-of-sample forecast (2011-14) vs True out-of-sample forecast (2015-2017)

Gini: Household Income

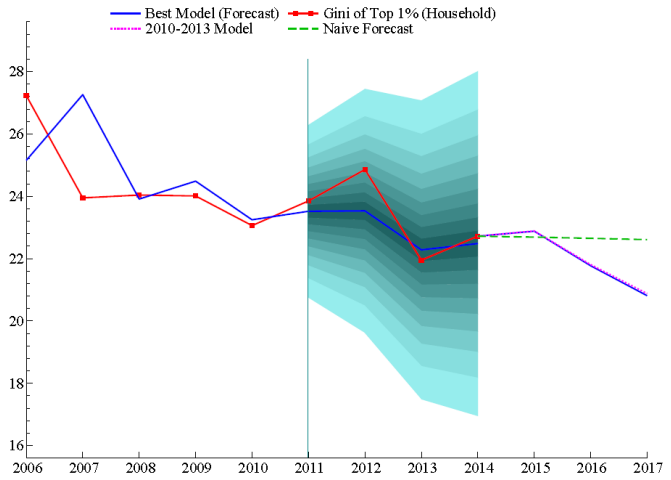
Gini: Household Income

$$\widehat{\triangle Gini}_t = 0.01 - 0.91 d2007 + 6.10 \triangle serv/gdp_{t-2} - 38.90 \triangle \frac{mskill1}{lskill}_{t-3} + 0.28 \triangle col\_fem_{t-3} + \epsilon_t$$



Gini of Top 1%: Household Income

$$\triangle \widehat{Top1Gini}_t = -0.73 - 5.78 d1981 + 3.83 d1985 + 36.89 \triangle lskill_{t-2} - 1.23 \triangle \triangle fem\_lfpr_{t-2} - 0.27 \triangle \triangle fem\_lfpr_{t-3} + 0.63 \triangle \triangle fem\_lfpr_{t-4} + \epsilon_t$$



Also results plotted for change in shares

Results: Model Selection

- Best predictors
  - Indicators of human capital attainment
  - Labor force structure - no one variable is super predictive
- Model selection in a General-to-Specific modeling approach is not robust but yields robust forecasts
- Sensitive to lag length and variable inclusion

Results: Forecasts

- 2010-2013 models forecast in the same direction as 2011-2014 models
- Out of sample forecasts differ between models by <6% for all variables and <2% for 4/8 measures
- Best models not not statistically different from naive models
- Overall Gini individual earnings inequality constant while household inequality increases steady-Inequality among individual and households continues to show different patterns
- Share of top 1% rising but inequality within top 1% falling

Comments-1

- Concern with forecast since income used appears to subject to topcodes (Burkhuser et al. 2011)—concern swapping adequate
  - Try analysis with internal data
  - Try analysis with imputes from SCF at top or tax data
- Are these data appropriate/adequate for measuring top 1% much less 0.1%?

- Issue of taxes, realized capital gains, definitions of transfers (especially at top given different compensation packages for executives)
- Farm income versus income from rental properties (one individual, one household)

- Use of equivalence scales? Major changes in households composition over this time period

Comments-2

- Determinants are macro—what about micro? You have the data…
  - Different determinants for individuals and households over the income distribution
- Macroeconomic econometrics– what about micro?

Contact Information

Thesia I. Garner, Ph.D.

Supervisory Research Economist
Division of Price and Index Number Research, BLS
U.S. Department of Labor, Washington, DC

202-691-6576
Garner.Thesia@bls.gov