

# Timely Estimates of Household Income Growth and Income Inequality

Gary **Koop**<sup>1</sup>, Stuart **McIntyre**<sup>1</sup>,  
James **Mitchell**<sup>2</sup> and Ping **Wu**<sup>1</sup>

<sup>1</sup>University of Strathclyde    <sup>2</sup>Federal Reserve Bank of Cleveland

*ESCoE Conference on Economic Measurement 2023*

The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System. All authors are research associates at ESCoE

# Nowcasting and Mixed Frequency Econometrics

- Timeliness of economic statistics is central to their usefulness
- Statistical agencies such as ONS release data with a delay and often at lower frequency than policymakers wish
- E.g. regional GVA data annual prior to 2012 (now quarterly) and 6 months delay (longer delay in earlier data)
- Mixed frequency time series methods can be used to produce quickly released, high frequency nowcasts (flash estimates)
- Done in our previous work for ESCoE
- Regional output growth in the United Kingdom: More timely and higher frequency estimates from 1970, *Journal of Applied Econometrics*, 2020 by Koop, McIntyre, Mitchell and Poon

# Nowcasting Using MF-VARs

- Earlier nowcasting work used Mixed Frequency Vector Autoregressions (MF-VARs): flexible models which do not impose much structure on the problem and let the data speak
- The MF-VAR is a VAR using  $y_t$  as the dependent variables:

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + u_t \quad (1)$$

- In conventional VAR  $y_t$  is vector of macroeconomic variables, in MF-VAR some elements of  $y_t$  are unobserved
- E.g. classic setup: monthly GDP growth (unobserved) with other elements being monthly macro variables (e.g. unemployment, industrial production).
- MF-VAR is a state space model with equation linking monthly GDP growth to (observed) quarterly GDP growth as additional equation completing the model
- Bayesian methods for state space models can be used (machine learning priors to handle issue that these models are big)

# Combining the Micro and the Macro

- Existing nowcasting work focuses on macro aggregates (e.g. GDP growth)
- But macro aggregates only tell you so much
- Upsurge of in interest in using micro (individual level) data in macro models
- Inequality: need entire distribution of income (not just aggregate GDP)
- Resilience: how will policy intervention/macro shock effect poorest 10% of society?
- Econometric challenge: How to build micro data into macro models like the MF-VAR?
- This is what our present research is about

# Combining the Micro and the Macro

- Our goal is to produce timely, quarterly nowcasts of the income distribution
- Each time UK quarterly GDP data (and other macro aggregates) released produce new nowcasts of income distribution
- Initially at the UK level, future work at regional level

# The Econometrics of Combining Micro and the Macro

- Hot topic in modern macroeconomics
- Older approaches would simply convert micro into macro data and model in a VAR
- E.g. use micro income data to produce Gini coefficient or quantiles of income distribution and include as variables in a VAR
- Mumtaz and Theophilopoulou (2017). The impact of monetary policy on inequality in the UK. An empirical analysis. European Economic Review.
- Recent work uses functional VARs (i.e. some "variables" in VAR are not variables but probability density functions like the income distribution)
- E.g. Chang and Schorfheide (2022). On the effects of monetary policy shocks on earnings and consumption heterogeneity. Centre for Economic Policy Research Paper
- Interesting to use these approaches and extend to mixed frequency (future research agenda)

# The Econometrics of Combining Micro and the Macro

- In this paper propose a new approach
- Let  $x_{it}$  be micro variable (income) of individual  $i$  at time  $t$
- $x_t$  be the vector containing income for all individuals
- Our idea: MF-VAR with  $y_t$  (macro) and  $x_t$  (micro) being dependent variables
- Such an approach retains rich granular information in the panel
- Not just summarizing them through (e.g.) Gini coefficient
- Nor try to approximate their p.d.f. (e.g. using spline based methods) like functional VAR
- MF-VARs are flexible, familiar set of tools that have had great success with nowcasting

# Problems with Our Idea

- Two problems with our idea
- We don't have a panel, rather we have a repeated cross-section
- $N_{it}$  is big (roughly 15000 per year in the data we use)
- Our idea leads to enormous VAR
- Second problem partly solved with recent developments in large Bayesian VARs (we can now handle hundreds of variables)
- Machine learning priors, restrictions (e.g. individual A's equation depends on individual A characteristics and UK variables, but not on individual B variables)
- But we have one solution which addresses both of these problems....



# MF-VARs with Pseudo-individuals

- Panel data econometricians sometimes turn repeated cross sections into pseudo panels (and then use standard panel data econometric methods using their pseudo panels)
- We do something similar
- Construct "pseudo individuals" based on observed individual characteristics
- E.g. region, age, gender, etc.
- $x_{it}$  is now average income for all observations with a particular set of characteristics
- $i$  indexes a "pseudo individual"
- Build an MF-VAR which jointly models  $y_t$  (macro) with  $x_t$  (micro)
- This can be used to provide predictions/nowcasts for each pseudo-individual
- These can be used to produce nowcasts of income distribution

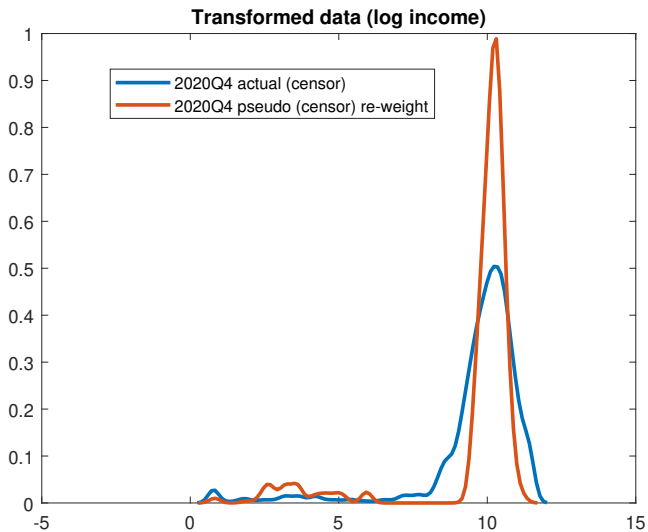
# Preliminary Empirical Results: Data

- Quarterly data 1975Q1 through 2020Q4 (no mixed frequency yet)
- Micro data: income for roughly 15000 individuals each year
- Data from Family Expenditure Survey, Expenditure and Food Survey and Living Costs and Food Survey.
- Use to construct panel of 144 (6 ages, 2 genders and 12 regions) pseudo individuals
- Five macro variables (GDP, average weekly earnings, CPI, exchange rate, claimant count)
- VAR involving 149 variables (fairly large, but manageable)

# Income Distribution Using Pseudo Data

- What Does the Income Distribution Based on Pseudo-individuals Look Like?
- Next figure compares income distribution using micro data to income distribution constructed using our pseudo-individuals
- Illustrates for one time period; 2020Q4

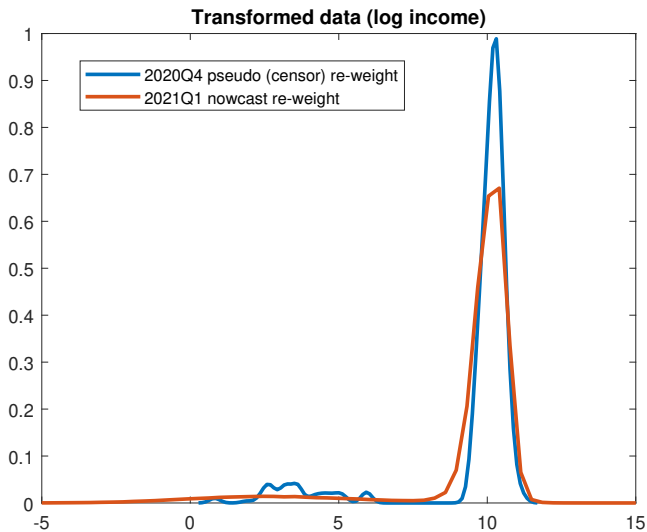
# Income Distributions: Actual Data and Pseudo Data



# Nowcasting Using Pseudo Data

- Next figure plots nowcast for 2021Q1 using data through 2020Q4
- Compares to actual data for 2020Q4
- Updating can be seen
- Nowcasts such as the one in this figure future goal of our analysis (updated each time new release of GDP data)

# Income Distributions: Pseudo Data and Nowcast 2021Q1



# Summary

- Our goal is nowcasting — not of variables like GDP — but of *distributions*
- MF-VARs have been a power tool for nowcasting, how to adapt to nowcasting distributions?
- Construct distributions based on pseudo individuals
- Model macro and pseudo individual data jointly in an MF-VAR
- Preliminary results suggests this strategy is feasible and produces sensible results