Revisiting 'Income inequality and economic growth: a panel VAR approach'

Johannes Coetsee - 19491050[a]

[a] *Stellenbosch University*

Abstract

This paper attempts to replicate and extend a paper by Atems & Jones (2015), in which the authors attempt to model the contemporaneous effects of income inequality and economic growth of the United States at the state level, using a Panel Vector Autoregression (PVAR) approach. This approach is reproduced to ascertain whether Atems & Jones (2015) hold up to scrutiny given new data. Two models are therefore estimated. One with only the replicated sample of data, and one that includes new data from 2005-2018. We find their results largely robust to the additional data. Further extensions include the reporting of forecast error variance decompositions, orthogonalised impulse response functions, and re-estimations of the model with geographical subsamples.

Keywords: Panel vector autoregression, Income Inequality, Economic Growth

1. Introduction

Atems & Jones (2015) utilize a Panel Vector Autoregression (PVAR) approach to examine the dynamic effects of US state-level inequality on income per capita and vice versa. The current analysis attempts to replicate various parts of this paper by specifically emphasizing the methodological approach used by these authors, in an attempt to uncover whether their PVAR specification holds against further checks for robustness. As such, like the authors, this paper 1) tests the series for unit roots, 2) estimates the baseline model for annual data from 1930-2005, and 3) replicates the cumulative Impulse Response Functions (IRFs) for this baseline bivariate PVAR. Additionally, another model will then be estimated for the entire series of data from 1930-2018 in order to compare the cumulative IRFs for both models. This comparison is the primary objective for this paper. In order to add more nuance to their argument, however, we will also test whether certain subsamples of observational units drive their results, as well as report the Forecast Error Variance Decompositions (FEVDs) and Orthogonalized IRFs (OIRFs) for both estimated models.

The paper will be structured in the following manner: for completeness, a overview of the primary differences between the PVAR and VAR approaches is given, after which the contribution of Atems &

Email address: `19491050@sun.ac.za` (Johannes Coetsee - 19491050)

Jones (2015) is discussed. Section 2.2 looks at their methodological approach through a critical lens, thereby also informing which robustness checks are considered the important ones to include in the current analysis. Section 3 then reporta some descriptive statistics as well as the results of the various estimations. Section 4 concludes.

2. Literature Review

2.1. Panel Vector Autoregression

Panel Vector Autoregressions are, as the name suggests, a variation of the standard VAR approach applied to panel data. Panel data, in contrast to time series data, is comprised out of various cross-sectional units observed over time - in our case, states - meaning that a VAR approach to model interactions between endogenous variables need to account for the fact that the underlying structure of the model might differ across these units. Whereas both VARs and PVARs treat all variables in a given system as endogenous, the PVAR approach thus allows for unobserved individual heterogeneity between the different cross-sectional units of observation. In order to overcome this difficulty, a PVAR approach therefore imposes an additional restriction, one that attempts to ensure a homogeneous underlying structure between all of the units of observation.

However, Love & Zicchino (2006) note that this restriction is highly likely to be violated in practice, consequently requiring circumvention. To this end, Love & Zicchino (2006) suggest the introduction of fixed effects that allow individual heterogeneity in the levels of the variables. Crucially, these fixed effects are correlated with the regressors because of the necessary inclusion of dependent variable lags in the model (the ‘autoregressive’ aspect of VARs), meaning that mean-differencing - the standard method used to eliminate fixed effects - will bias the regression coefficients. To solve this problem, Arellano & Bover (1995) advocate for the usage of the ‘Helmert procedure’, where the means of only the future observations for each unit is removed. This procedure therefore transforms the variables in a way that preserves the orthogonality between the variables and the lagged regressors - an important requirement for isolating shocks to the system. This, in turn, allows for the usage of the lagged regressors as instruments whereby the coefficients of the systems can be estimated. Moreover, these orthogonal relationships provide the necessary moment conditions that allow for VAR estimation using Generalized Method of Moments (GMM), which is, as will be shown, the estimation method employed both in this paper and in Atems & Jones (2015).

2.2. Inequality and Economic Growth (Atems & Jones, 2015)

Atems & Jones (2015) utilize a panel of annual state level income inequality data to consider the relationship between per capita income and income inequality using a panel VAR approach. This approach allows them to examine two things: first, the correlation between these variables, and second, the dynamic responses of both variables given shocks to income and inequality. The motivation for using a PVAR approach is that it captures more complexity than either standard VARs or traditional panel date models due to the allowance of dynamic effects, as well as allowing for unobserved heterogeneity across units of observation. Atems & Jones (2015) also report that it does well at fitting the data, whilst being parsimonious enough without making strong identifying assumptions (Atems & Jones, 2015; Love & Zicchino, 2006). Additionally, the inclusion of specifically state level panel data reduces possible measurement error due to greater homogeneity between states than, for instance, countries (Frank, 2009a).

Atems & Jones (2015)'s results are displayed using cumulative IRFs, which describe the response of one variable to the innovations in the other variable in the system.[1] They find that shocks to inequality has significant negative effects on the level of income per capita. They also find that the relation between income per capita and inequality varies over time, and is sensitive to specific subsamples of time. Overall, their contribution is novel in that it is the first study to employ a panel VAR approach to estimate the effect of inequality on income per capita and vice versa by utilising US state-level data. In order to gauge whether their analysis is sufficiently specified, a brief overview of the data and employed methodology is discussed below.

Data and Unit Root Testing

There are three datasets used by Atems & Jones (2015) relevant to our discussion. The first, data on state-level economic growth, is measured by the annual change in per capita real income for the 48 contiguous US states - plus District of Columbia - for the period 1930-2005.[2] The second series is state-level income inequality data (sourced by Frank (2009a)), for the same period. The measure of inequality, the Gini coefficient, is constructed using tax filing data.[3]. The third relevant series is the US CPI for all urban consumers, the measure of inflation used to convert nominal rates to real rates.

The first step to any VAR study is to conduct unit root tests on the series of data. Atems & Jones (2015) perform five different unit root tests - which include the Fisher-type Augmented Dickey-Fuller (ADF), Levin-Lin-Chu (LLC), IM-Pesaran-Shin (IPS), Harris-Tzavalis (HT) and Hadri (LM) tests - on demeaned data.[4] The Akaike Information Criterion (AIC) is used to select the appropriate lag length for the tests. Although the first four tests reject the null hypothesis of a unit root, the Hadri test, which tests the null hypothesis of no unit root, cannot be rejected. This is interpreted as being sufficient evidence to suggest that nonstationarity might be present at a 5% level in some of the series.

[1] Cumulative IRFs must be interpreted as capturing the effect on the levels of the variables, and not their growth rates.

[2] Hawaii and Alaska are therefore excluded.

[3] The usage of tax data is often considered problematic in that it excludes low-income earners, thereby introducing possibly misleading results. This possibility informs the authors' choice to check for robustness by using other inequality metrics.

[4] Levin *et al.* (2002) suggest to perform these tests on demeaned data, as it reduces the effects of dependence between cross-sectional units.

In order to overcome this, the authors first-difference both series of data, concluding that the PVAR should be estimated on the differenced series.

Methodology

The authors estimate a baseline structural bivariate VAR model of the growth rate of real income per capita and changes in the Gini index, whilst also implementing various robustness checks. These robustness checks include subsampling (structural break testing) on the time-period of analysis, as well as using four differing measures of inequality.[5]

Their reduced-form empirical specification is as follows:

$$Y_{it} = A(L)Y_{i,t-1} + \delta_i + \varphi_i + \varepsilon_{it} \quad \varepsilon_{it}\ N(0, \Sigma_i), \tag{1}$$

where $A(L)$ is the polynomial matrix of the lag operator L, δ_i is the unobservable time effects and φ_i is a vector of constant-over-time fixed effects across states. Y_{it} is equal to the vector of the growth rate of real income per capita (Δy_{it}) of state i in year t, and the change in the Gini coefficient of state i in year t (Δg_{it}), thereby equaling $[\Delta y_{it} \Delta g_{it}]'$. Further $\varepsilon_{i,t} = [\varepsilon_{i,t}^{\Delta y} \varepsilon_{i,t}^{\Delta g}]'$, which denotes the vector of errors.

It is necessary to impose further structure on equation (1) to uncover the underlying structural behaviour of shocks to the system, and therefore to make IRFs interpretable. These restrictions are often untestable and must be guided by economic theory. There are two restrictions imposed on this system. The first is necessitated by the requirement of orthogonality due to the structure of panel data, and is discussed in Section 2.1. The second restriction is guided by economic theory, and is concerned with the ordering of the variables - there cannot be contemporaneous effects of changes in the Gini coefficient on economic growth. By Cholesky identification, Atems & Jones (2015) argue that the Gini coefficient should be ordered second in the structural specification of the VAR. This argument is sound; the Gini coefficient is calculated using tax and income data, meaning that there will be contemporaneous effects of changes on income on the Gini coefficient. The Gini coefficient, however, has delayed effects on income, a fact that is established in the literature (Barro, 2008; Cingano, 2014; Frank, 2009b).

Therefore, incorporating these restrictions to the model and transforming the variables according to the Helmert procedure, one can specify the final transformed equation as taking the form:

[5] These are the Gini coefficient, the Relative Mean Deviation, the Theil Entropy Index, and the income share of the top income decile and top percentile of the state population. All measures are sourced from Frank (2009a).

$$\tilde{Y}_{it} = A(L)\tilde{Y}_{i,t-1} + \tilde{\varepsilon}_{i,t} \tag{2}$$

In addition to the specification above, Atems & Jones (2015) identify the cumulative IRF's using a Cholesky decomposition of the covariance matrix of the residuals, and decide on an arbitrary number of lags, arguing that four lags is sufficient to capture the system dynamics. The cumulative IRFs are given for a one standard deviation shock, with 5% confidence bands generated by Monte Carlo simulation methods.

Overview

The VAR model by Atems & Jones (2015) seems to hold up to initial scrutiny. It is difference-stationary and accounts for the paneled aspects of the data. Furthermore, the ordering restriction is economically sound, and the data seems to be of good quality. Moreover, their initial specification holds up to robustness checks with respect to different inequality metrics and structural break tests. The authors also insulate themselves against possible criticism when it comes to omitted variable bias by including another variable, human capital, and re-estimating the VAR as a trivariate model. However, their metric of 'human capital' - average years of schooling per state - seems to be fairly simplistic. Arguments can be made for a more nuanced metric that includes labour market information.[6]

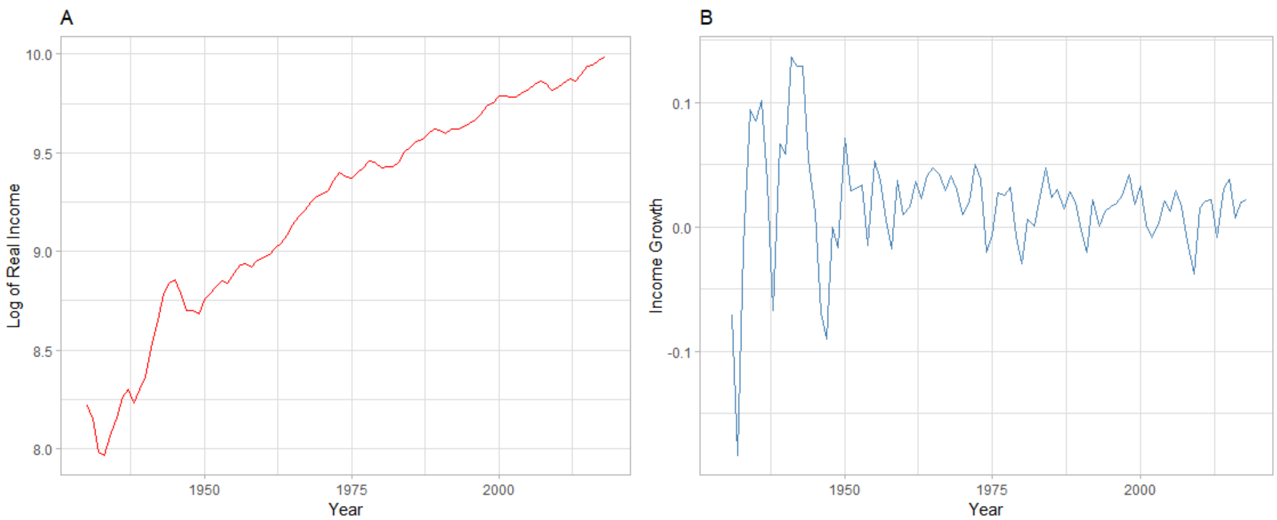
Given these factors, several extensions to this study can nonetheless be identified. The first is to extend the model to include new data from 2005 to 2018, which, at the time of writing, was not available to the authors of the original paper. Another extension to consider would be the inclusion of state subsamples delineated according to geographical regions. The intuition for the latter is to ascertain whether some regions drive the results found by Atems & Jones (2015), which would add more nuance to their analysis. For the purposes of this study, emphasis will be placed on extending the model to include the new data, and to include these regional subsamples. Additionally, as another way of summarizing and visualizing the VAR results, it might be useful to include forecast error variance decompositions (FEVDs). This method displays how much variation between the endogenous variables is due to which shock - Gini or Income - over time. The subsequent section endeavors to attempt the three extensions mentioned above.

[6] For instance, an inclusion of labour force participation rates, unemployment figures or a variable measuring average work experience might be suitable.

3. Replication and Extensions

3.1. Descriptive Statistics

In order to prove that the replicated study uses the same series as the original, as well as to visualize the additional data from 2005-2018, descriptive statistics on the US income and inequality series are plotted. Figure 3.1 below displays the per capita real income and per capita real income growth for the United States as a whole, averaged at the state level. These series closely follow the same plots provided by Atems & Jones (2015).



A) United States average of real income per capita and B) real income per capita growth at the state level, 1930–2018. Author's own calculations.

Figure 3.1

In terms of inequality metrics, Figure 3.2 gives the state-average inequality measures from 1930 to 2018, which, again, replicates the same plot in Atems & Jones (2015). However, it is important to note that the current analysis will emphasize only the Gini-coefficient (displayed in green), as the robustness checks performed by Atems & Jones (2015) with respect to different measures were deemed adequate for our purposes.

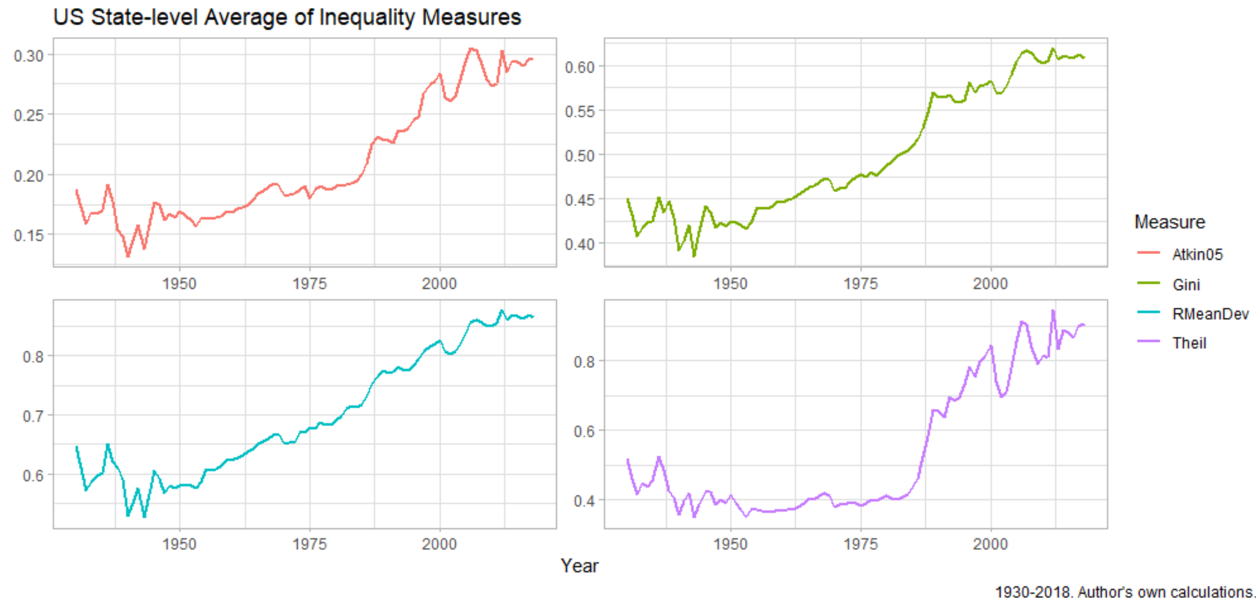


Figure 3.2

3.2. Unit Root Tests

We employ three tests for unit roots[7], and one explicitly for stationarity[8]. Like Atems & Jones (2015), these tests are performed on demeaned data, and the panel data is balanced. Furthermore, all tests are performed for both the intercept and with a trend for all series, whilst the new data from 2005 to 2018 is also included. The test results mirror the findings of Atems & Jones (2015), and are displayed in the Appendix, section 4.1. Although the three unit root tests indicate stationarity in levels for both income per capita and inequality, the Hadri test indicates that the null of no unit roots cannot be rejected for the series of both variables. After first differencing, the same process is repeated, and all tests indicate stationarity in the growth rates of the two variables.

3.3. Results of PVAR Estimation

The baseline PVAR estimated in this section will follow the same methodology described in Atems & Jones (2015), using equation 2. The models are estimated using GMM for four lags, and a Cholesky decomposition is used to identify the structural error terms. Impulse response functions are generated accordingly. Our results are displayed using cumulative orthogonalized IRFs (COIRFs), which, as mentioned earlier, should be interpreted as long run responses to a permanent shock in the level of

[7] The Fischer-type ADF, LLC and IPS tests. The HT test was excluded as it was not available on the statistical program this study employed (the '*plm*' package from Croissant *et al.* (2020)).
[8] The Hadri LM test.

the series.[9]. COIRFs were computed manually from the OIRFs given in Appendix 4.2, which have been included for completeness, and as an additional extension.[10]

To test for robustness given new data, we estimate two PVARs. The first includes only the sample of the annual series from 1930 to 2005, and is thus an exact replica of the baseline PVAR by Atems & Jones (2015). The second model applies the PVAR methodology to the full dataset, including new data. The intuition behind this approach is comparative - the baseline model can be considered robust to new data if the two models display similar results. As an additional visualization tool, the FEVDs for both models are also reported.

Figure 3.3 displays the replicated COIRFs for the baseline model for the 1930-2005 sample.[11] It seems largely similar to the results obtained by Atems & Jones (2015), but with some noticeable differences, the most pertinent of which is the response to Gini given its own shock. Whilst both indicate a decrease over 5 years for the initial shock, the replicated COIRF stabilizes at 2.5 percent compared to the 1.25 percent in Atems & Jones (2015). In terms of the other results, however, our replicated study is similar. The permanent income shock to income has the expected result of increasing the level of income per capita permanently, with the greatest response coming at around year four post-shock. The response to income given a permanent shock to inequality shows a clear permanent decrease in the level of income per capita. An income shock, however, has only a slight positive effect on the level of inequality, essentially remaining near zero, whilst a deviation for inequality increases the level of inequality permanently, converging to a level below the initial standard deviation shock value.

Now turning towards incorporating the new data, Figure 3.4 displays the COIRFs for the full sample from 1930-2018. As is evident, the signs of the permanent changes in levels are the same for all four IRFs. However, and notably, the level of inequality given a shock to income is no longer insignificantly near zero. Instead, this deviation increases the level of inequality quite substantially above its initial value, and is the primary result obtained by this analysis. This can perhaps be explained by the high levels of inequality shown in Figure 3.2 for the years 2005-2018. If this result holds, it might indicate that the model by Atems & Jones (2015) is sample-specific with respect to inequality, and that periods of sustained higher inequality might react differently to permanent unexpected increases in income per capita.

[9]All estimation of PVARs, and computation of OIRFs, were done using a statistical package in R, '*panelvar*', developed recently by Sigmund & Ferstl (2019).

[10]They are displayed and briefly interpreted in the appendix as the focus of our findings is on cumulative IRFs.

[11]All confidence intervals are calculated at the 95% level and are generated by Monte Carlo simulation based on 100 draws. The low number of draws were chosen due to computational limitations.

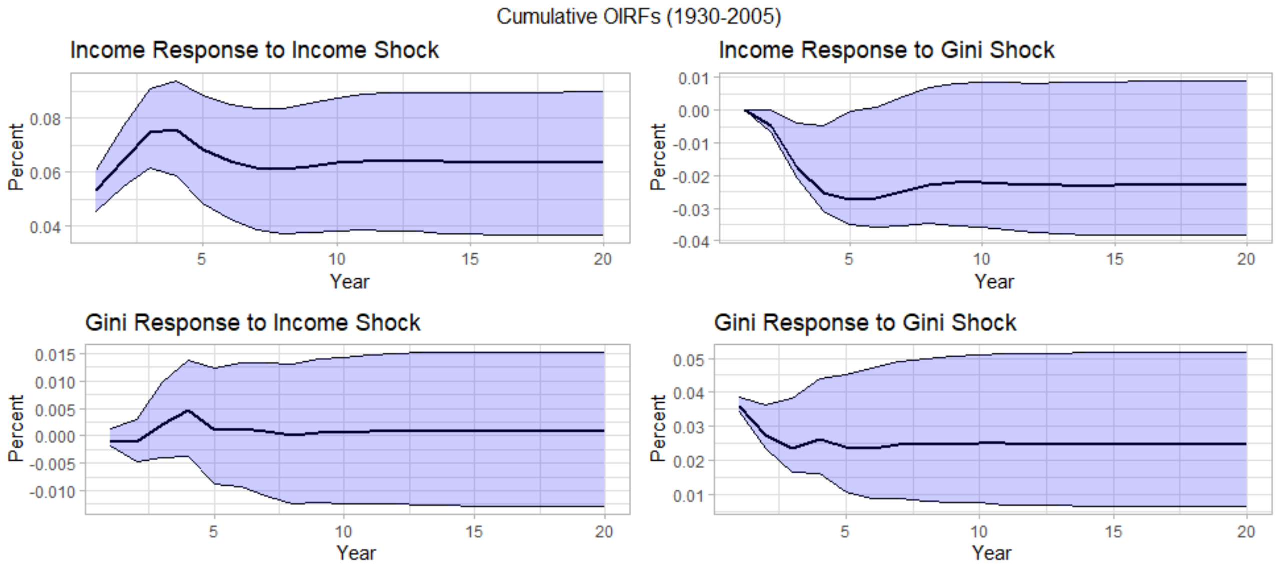


Figure 3.3

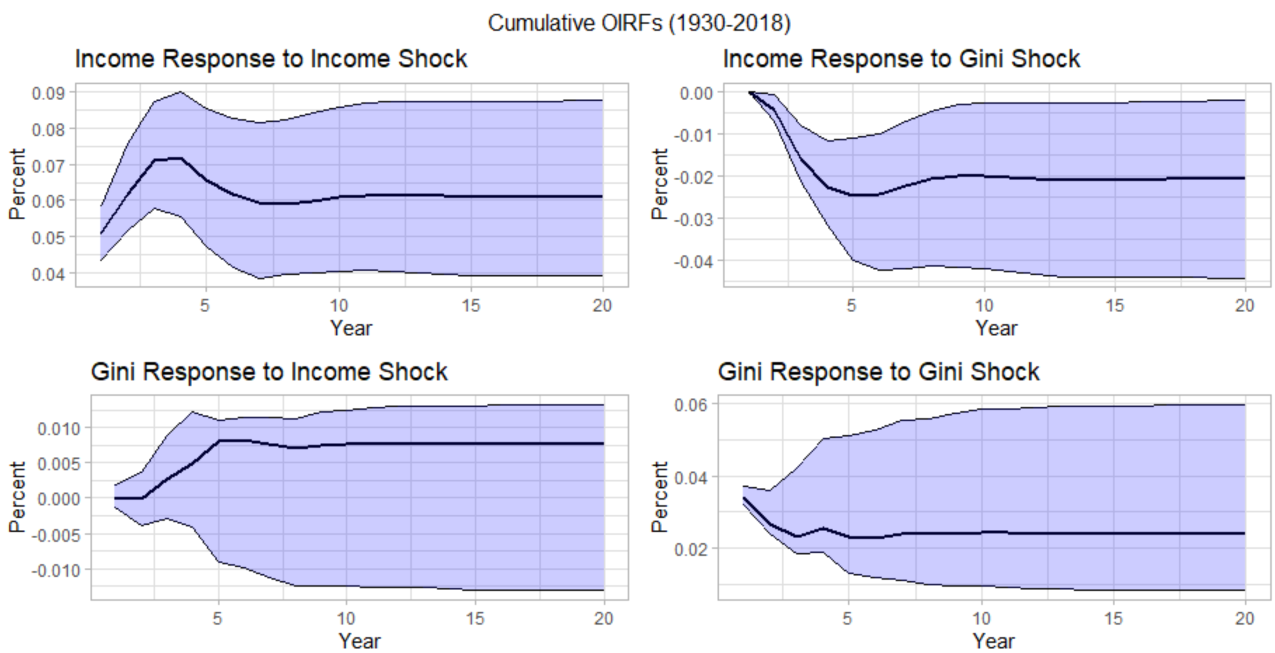


Figure 3.4

Figure 3.5 presents the results of the forecast error variance decompositions for both estimated models. As is evident, they are very similar, indicating that our model is robust to different sample periods. For the full-sample model, the variance of Income Growth at each forecast horizon is explained more by its own shock than for the replicated (smaller) sample. Conversely, Income Growth variance is thus also explained less by the shock to inequality. Similarly, Gini Growth variance

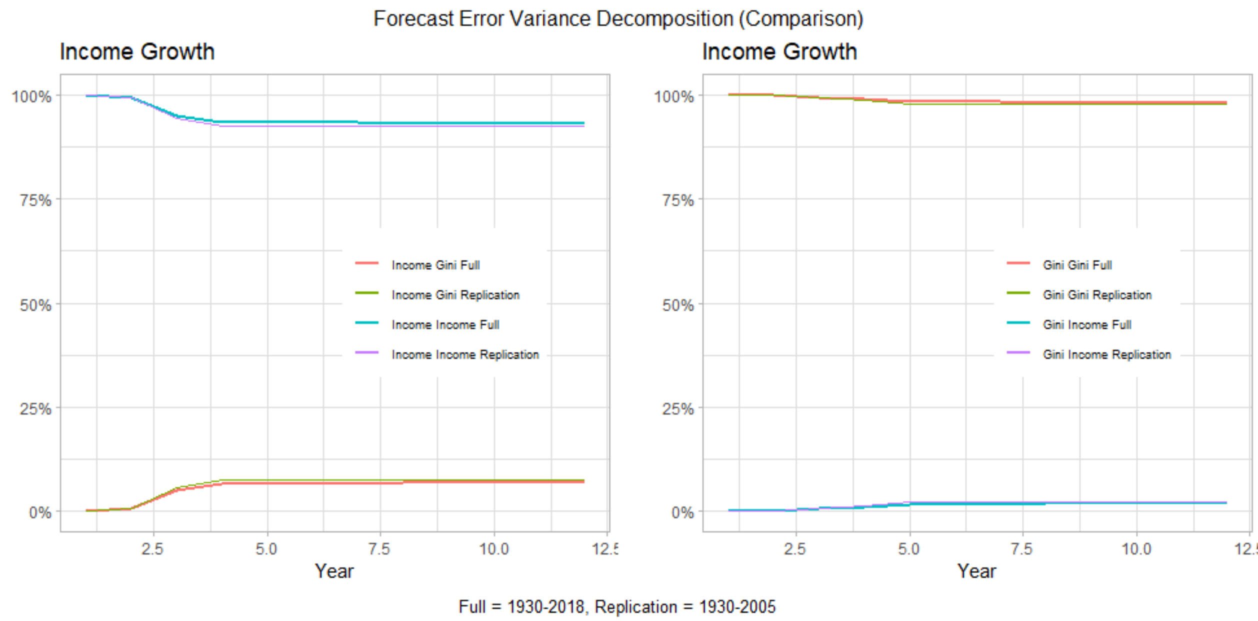


Figure 3.5

3.4. Regional Subsamples

In this section, we re-estimate the model for each regional subsample of the full data. These regions, of which there are eight, are defined according to geographical area in the US, and are the following:

4. Conclusion

References

Arellano, M. & Bover, O. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*. 68(1):29–51.

Atems, B. & Jones, J. 2015. Income inequality and economic growth: A panel var approach. *Empirical Economics*. 48(4):1541–1561.

Barro, R.J. 2008. *Inequality and growth revisited*. ADB Working paper series on regional economic integration.

Cingano, F. 2014. Trends in income inequality and its impact on economic growth.

Croissant, Y., Millo, G., Tappe, K., Toomet, O., Kleiber, C., Zeileis, A., Henningsen, A., Andronic, L., et al. 2020. Package "plm". *Choice*. 139(1):227–240.

Frank, M.W. 2009a. Inequality and growth in the united states: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*. 47(1):55–68.

Frank, M.W. 2009b. Income inequality, human capital, and income growth: Evidence from a state-level var analysis. *Atlantic Economic Journal*. 37(2):173–185.

Levin, A., Lin, C.-F. & Chu, C.-S.J. 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of econometrics*. 108(1):1–24.

Love, I. & Zicchino, L. 2006. Financial development and dynamic investment behavior: Evidence from panel var. *The Quarterly Review of Economics and Finance*. 46(2):190–210.

Sigmund, M. & Ferstl, R. 2019. Panel vector autoregression in r with the package panelvar. *The Quarterly Review of Economics and Finance*.

Appendix

4.1. Unit Root Test Results

Gini Levels (Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)

z = 4.3123, p-value = 1

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)

Wtbar = 9.1986, p-value = 1

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts)

chisq = 21.603, df = 98, p-value = 1

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts) (Heterosked. Consistent)

z = 360.4, p-value < 2.2e-16

alternative hypothesis: at least one series has a unit root

Gini Levels (Trend and Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

z = -10.034, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

Wtbar = -10.363, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

chisq = 300.73, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked. Consistent)

z = 144.43, p-value < 2.2e-16

alternative hypothesis: at least one series has a unit root

Income Levels (Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)

z = 4.9448, p-value = 1

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)

Wtbar = 14.295, p-value = 1

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts)

chisq = 5.0546, df = 98, p-value = 1

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts) (Heterosked. Consistent)

z = 402.2, p-value < 2.2e-16

alternative hypothesis: at least one series has a unit root

Income Levels (Trend and Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

z = -11.854, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

Wtbar = -11.073, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

chisq = 325.46, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked. Consistent)

z = 101.25, p-value < 2.2e-16

alternative hypothesis: at least one series has a unit root

Gini Growth (Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)

z = -53.409, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)

Wtbar = -53.122, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts)

chisq = 3030.6, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts) (Heterosked. Consistent)

z = -2.2375, p-value = 0.9874

alternative hypothesis: at least one series has a unit root

Gini Growth (Trend and Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

z = -55.653, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

Wtbar = -52.68, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

chisq = 3061.6, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked. Consistent)

z = 0.60399, p-value = 0.2729

alternative hypothesis: at least one series has a unit root

Income Growth (Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)

z = -51.387, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)

Wtbar = -51.238, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts)

chisq = 2835.8, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts) (Heterosked. Consistent)

z = -2.1457, p-value = 0.9841

alternative hypothesis: at least one series has a unit root

Income Growth (Trend and Intercept)

Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

z = -55.634, p-value < 2.2e-16

alternative hypothesis: stationarity

Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

Wtbar = -52.376, p-value < 2.2e-16

alternative hypothesis: stationarity

Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

chisq = 2912.5, df = 98, p-value < 2.2e-16

alternative hypothesis: stationarity

Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked. Consistent)

z = -1.0224, p-value = 0.8467

alternative hypothesis: at least one series has a unit root

4.2. Full Sample OIRFs

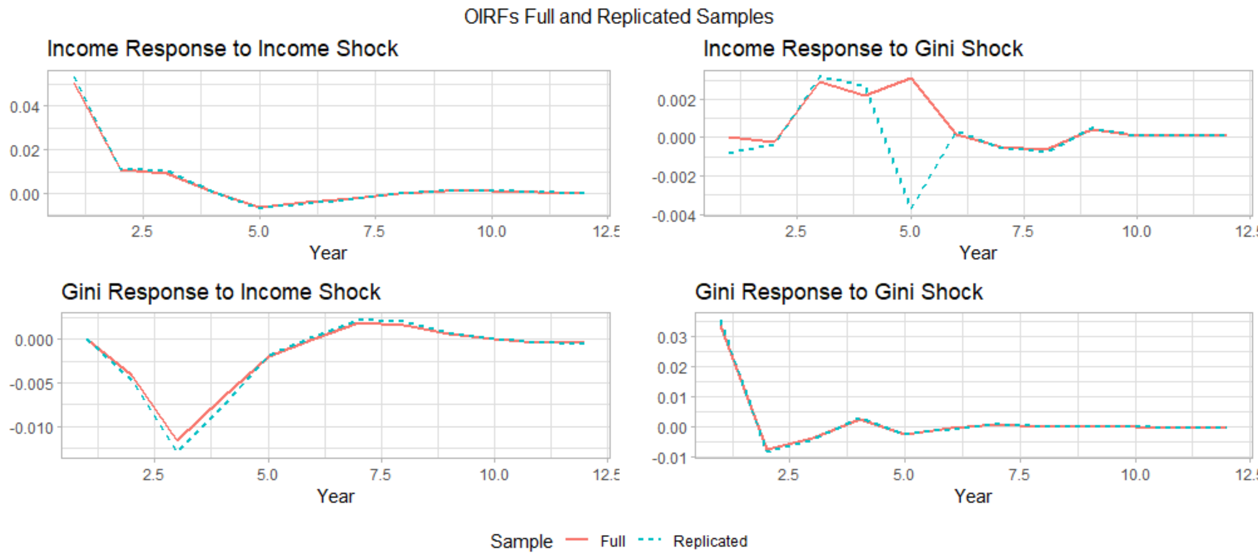


Figure 4.1