Revisiting ‘Income inequality and economic growth: a panel VAR approach’

Johannes Coetsee - 19491050[a]

[a] *Stellenbosch University*

Abstract

This paper attempts to replicate and critique 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.

Keywords: Panel vector autoregression, Income Inequality, Economic Growth

1. Introduction

This paper attempts to replicate various parts of the paper by Atems & Jones (2015). It will specifically emphasise the methodological approach used by these authors in an attempt to uncover whether their PVAR approach holds against further checks for robustness. As such, this paper will, like the authors, test the series for unit roots, whilst attempting to replicate the cumulative IRFs for the baseline bivariate PVAR. Furthermore, we will test whether certain subsamples of observational units drive their results, something the authors did not consider.

The paper will be structured in the following manner: for completeness, a overview of the primary differences between the PVAR and VAR approaches will be given, after which the contribution of Atems & Jones (2015) will be 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.

2. Literature Review

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

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 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, namely, that the underlying structure is the same for all of the units of analysis.

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).

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

Atems & Jones (2015) utilise 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. Their results are displayed using cumulative Impulse Response Functions (IRFs), which describe the response of one variable to the innovations in the other variable in the system. 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. In order to gauge whether their analysis is sufficiently specified, a brief overview of the data and employed methodology is discussed below.

Data

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 (listed in the Appendix) for the period 1930-2005. The second series is state-level income inequality data (sourced by Frank (2009a)), for the same period. Their measure of inequality, the Gini coefficient, is constructed using tax filing data.[1]

Unit Root Testing

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 Augmented Dickey-Fuller (ADF), Levin-Lin-Chu (LLC), IM-Pesaran-Shin (IPS), Harris-Tzavalis (HT) and Hadri (LM) tests - on demeaned data.[2] After using the Akaike Information Criterion (AIC) to choose the appropriate lag length of 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 significance level of 5% in some of the series. In order to respond to this, the authors conclude that first-differencing is the appropriate response. As such, the PVAR is estimated on the differenced series of the data.

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 three other measures of inequality.[3]

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

[1]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.

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

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

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 equalling $[\Delta y_{it} \Delta g_{it}]'$. Further $\varepsilon_{i,t} = [\varepsilon_{i,t}^{\Delta y} \varepsilon_{i,t}^{\Delta g}]'$, which denotes the vector of errors for each series.

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 employing a Cholesky decomposition, 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 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).

Critiques

3. Replication

- Descriptive Statistics
- Unit Root tests
- ACF's? Checking for White Noise Errors
- IRF's
- State Subsample
- Adding new data (three years, 2015-2018)

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.

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.

Appendix

- Table of 48 states and the state sample splits