Economic Growth Analysis When Balanced Growth Paths May Be Time Varying

Mountford, Andrew

19 August 2022

Economic Growth Analysis When Balanced Growth Paths May Be Time Varying*

ANDREW MOUNTFORD

Department of Economics, Royal Holloway, University of London, Egham, TW20 OEX, U.K.

(email: A.Mountford@rhul.ac.uk)

December 2023

Abstract

The determinants of an economy's growth path for income per head may vary over time. In this paper we apply unobserved components analysis to an otherwise standard panel model of economic growth dynamics so that an economy's long run relative income per head can change at any point of time. We apply this model to data for US states for 1929-2021 and the world economy for 1970-2019. In both datasets an economy's initial relative income per head is a good predictor of its long run relative income per head. Relatively poor economies on average remain relatively poor.

WORD COUNT 5031

JEL classification numbers: C11, O40

Keywords: Bayesian Econometrics, Economic Growth, State Space Models

*Many thanks to Fabio Canova, Etsuro Shioji and Jonathan Wadsworth for very helpful comments on an earlier draft. All errors are my own.

I Introduction

Economic growth theory describes how an economy's income per head on its balanced growth path may be determined by, *inter alia*, its saving propensity, human capital accumulation, population growth, institutional quality and economic policy. All these factors may be changing over time. Even without the anchor of economic theory, it is intuitive that an economy's long run path is determined by factors that are likely to change. In this paper we therefore add unobserved components analysis to an otherwise standard empirical model of economic growth dynamics so that an economy's growth path can change at any point in time. We estimate the model using two publicly available datasets: US Bureau of Economic Analysis (BEA) data for personal income per head for the 48 contiguous US states for 1929-2019, and the Penn World Table data for GDP per head in the world economy for 1970-2019. We find in both datasets that an economy's initial relative income per head is a good predictor of its long run relative income per head, with little mobility in long run relative growth paths. These results provide support for 'The Poor Stay Poor' hypothesis of Canova and Marcet (1995).

A fundamental question in macroeconomics is whether one should think of the world as made up of economies that are slowly converging to the same balanced growth path or whether one should think that economies are converging to their own individual balanced growth paths. Examples of the former view include Barro and Sala-i-Martin (1991) on the convergence of US states, and recently Patel, Sandefur and Subramanian (2021) on convergence in the world economy. In the latter view, economies' growth dynamics are described by economy specific parameters and so panel estimation methods are used. Examples of this approach include Canova and Marcet (1995), Caselli, Esquivel and Lefort (1996) and Shioji (2004) and more recently by Acemoglu et al. (2019) and Acemoglu and Molina (2022).

This paper contributes to the literature in three ways. Firstly, it introduces a time varying individual economy specific parameter into this analysis. As well as being a more reasonable description of reality, this modelling approach has the advantage of providing a general form for modelling changes in growth paths. The literature has tended to focus on one specific growth determinant at a time. Important examples include Acemoglu et al (2019), Cerra and Saxena (2008) and Wacziarg and Welch (2008) who, respectively, have shown how changes to democracy, financial and political stability and trade openness can affect an economy's long run growth path. These effects are not mutually exclusive and can change in any economy at any time. Therefore, rather than look at these determinants one by one, this paper provides a more general form which nests all - observed and unobserved - potential effects into one model, albeit in reduced form, so that an economy's balanced growth path is able to move persistently at any point in time.

Secondly, the empirical growth literature has typically estimated models where the autoregressive coefficients, and so the speed of convergence to the steady state, are restricted to be the same across economies, with Canova and Marcet (1995) and Andrés et al (2004) important exceptions to this. While, as Mankiw, Romer and Weil (1992) argue, some determinants of the speed of convergence, such as production technology, may be determined at a global level, this is not true for all determinants. We therefore also estimate a hierarchical model where, in addition to time varying individual economy parameters, the autoregressive coefficients in the panel regression are able to differ across economies while remaining related to each other by belonging to a common population distribution. As well as being intuitive and linking better to economic growth theory, the reduced level of aggregation in this model may also reduce bias in the panel estimates, see Canova (2023).

Finally, the paper is able to address the argument of Shioji (2004) that US states' income per head data is more consistent with a slow convergence of states to the same balanced growth path than with states converging to their own individual balanced growth paths. Shioji (2004) showed that panel models produced parameter estimates that implied a relatively fast rate of convergence to the long run balanced growth path and this seemed inconsistent with the large distance of many states' initial conditions from their long run balanced growth paths. This paper addresses this issue by allowing the long run balanced growth path to change over time so that, for example, an economy could initially be close to its initial balanced growth path but far away from its ultimate balanced growth path.

This paper applies Bayesian state space analysis to the empirical economic growth literature. Recently Startz (2020) and Imam and Temple (2023) have used state space methods to model economies' growth paths as switching between different growth states in a Markov process. Bayesian macroeconomic state space analysis has also been used to, *inter alia*, decompose time series such as GDP and inflation into a trend and cyclical component, see e.g. Canova (2007) and Chan et al. (2019) for examples of this large literature. This paper applies Bayesian macroeconomic state space analysis to decompose an economy's change in relative income per head into a long run growth path component which can change through time and transitory components around this path.

We estimate two variations of our model: (i) A baseline model where each individual economy's intercept term follows an independent local level model but with the same convergence coefficients across economies. This model is designed to be as close to the panel model structure used by the literature as possible while allowing time varying individual economy terms. (ii) A hierarchical model where the autoregressive parameters may also vary across economies but where they are drawn from a common population distribution.

The estimation process for both models follows the Gibbs sampling procedures set out in Chan et al. (2019) and is applied to two important datasets from the literature. The US states dataset is of contiguous, free trading, democratic economies operating under free interstate capital and labour mobility. The conditions for convergence in this dataset are therefore as good as could reasonably be expected. The Penn World Table dataset allows analysis of the dynamics of output per head across countries in the world economy which is of intrinsic interest.

The paper is organised as follows. Section II describes the estimated models and their relationship to the empirical growth literature. Sections III and IV describe the results from applying the models to data on US states and the world economy respectively. Section V discusses the choice of priors and the performance of the models in out-of-sample prediction exercises. Section VI concludes. The Gibbs sampling algorithms used in the estimation and some further results are presented in the Appendix.

II Time varying balanced growth paths

There is a large literature analysing economic growth dynamics using dynamic panel models, see for example Chen et al (2019). These models have the form

$$Y_{i,t} = \alpha_i + \gamma_t + \sum_{l=1}^{L} \beta_l Y_{i,t-l} + \delta D_{i,t} + e_{i,t} \tag{1}$$

where $e_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$, α_i are individual fixed effects, γ_t are time effects, $Y_{i,t}$ is the log of income per head in economy i at time t and Y_{t-l} is its l'th lag. $D_{i,t}$ is an indicator variable for a significant change. Notable examples of this estimated model include Acemoglu *et al* (2019) where $D_{i,t}$ is an indicator for democracy, Wacziarg and Welch (2008) where $D_{i,t}$ is an indicator for trade openness and Cerra and Saxena (2008) where $D_{i,t}$ is an indicator for financial instability.

The model can also be estimated in terms of relative income per head, where the time effects cancel out, so that the estimated equation becomes,

$$Y^*_{i,t} = \alpha_i + \sum_{l=1}^{L} \beta_l Y^*_{i,t-l} + \delta D_{i,t} + e_{i,t} \tag{2}$$

where $Y^*_{i,t}$ is the deviation of the log of the income per head in economy i at time t, from the average across all economies at time t. The relative income per head form of the model was used by Canova and Marcet (1995) and Shioji (2004) and we use this form in our estimation below.

Unobserved components model

The contribution of this paper is to estimate the dynamics of income per head where the country specific intercept term, now denoted $\alpha_{i,t}$, is free to move persistently at any point in time, rather than being tied to particular policies or institutional variables.

We estimate two variations of the model. Our baseline model has the intercept term for each economy following an independent local levels model together with a transitory cyclical term, as in Chan et al (2019), but where the autoregressive β coefficients are the same for each economy and estimated using data from all economies as in equation (2). This model is designed to be as close as possible to that used by the literature except for time varying intercept terms. Our second model is a hierarchical model where individual economies' β coefficients can vary but are related to the population distribution.

We choose diffuse priors for most parameters but not for the priors for the local level and transitory terms. These parameters significantly impact on the volatility of the long run growth path and so one's choice of prior for these parameters will reflect the degree of variation one thinks is reasonable for a long run growth path. We discuss the implications of the choice of priors briefly below and at greater length in Section V.

The baseline model

The baseline model varies equation (2) by treating each economy's intercept term as a standard local levels model and allowing for a transitory shock to each economy's growth process, $c_{i,t}$, so that

$$\begin{aligned} Y_i^* &= \alpha_{i,t} + \sum_{l=1}^{L} \beta_l Y_{i,t-l}^* + c_{i,t} + e_{i,t} \ \forall i \\ c_{i,t} &= \phi_{i,1} c_{i,t-1} + \phi_{i,2} c_{i,t-2} + u_{i,t} \ \forall i \\ \alpha_{i,t} &= \alpha_{i,t-1} + \nu_{i,t} \ \forall i \end{aligned} \tag{3}$$

where $e_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$, $u_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{c,i}^2)$ and $\nu_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \omega_i^2)$. The local levels model specifies that α_i follows a random walk with a Normally distributed error term. This is appropriate for our case as the change in the intercept in the literature is also persistent, i.e. the change in $D_{i,t}$ from 0 to 1 in equation (2).[1] The $c_{i,t}$, process is also unobserved and provides economy level parameterization of potentially cyclical transitory deviations around the balanced growth path.

We estimate this model using Bayesian methods set out in Chan et al (2019). The prior for the β parameters, $\underline{\beta} \equiv [\beta_1 \dots \beta_l]'$ is distributed, $\mathcal{N}(\beta_0, S_\beta)$, where $\mathcal{N}$ denotes the Normal distribution. The prior for the ϕ parameters, $\underline{\phi} \equiv [\phi_1, \phi_2]'$ is distributed, $\mathcal{N}(0, S_\phi)$ and where $\underline{\phi}$ is restricted to ensure stationarity. The prior for the initial value of α, denoted α_0 in each economy is also normal with distribution $\mathcal{N}(a_0, b_0)$. The prior for σ^2 is distributed so that $\sigma^{-2} \sim \mathcal{G}(\nu_\sigma, S_\sigma)$ where $\mathcal{G}$ denotes the gamma distribution.[2] We set $\nu_\sigma = 2.5$ and $S_\sigma = 4 \times 10^{-4}$. The value of β_0 is taken from a pooled OLS regression of all economies with a common intercept term and we set $S_\beta = 10^8 I_p$ and $S_\phi = 10^8 I_2$ which implies a diffuse prior. The parameters for the prior of α_0 is also chosen so that the prior is diffuse with $a_0 = 0$ and $b_0 = 10000$.

The priors for ω_i^2 and $\sigma_{i,c}^2$ are very important as they significantly impact the variation in the path of $\alpha_{i,t}$ and so the cross-sectional variation in the balanced growth path. We use the same priors for each economy in the panel and assume the priors for ω_i^2 and $\sigma_{i,c}^2$ are distributed so that $\omega_i^{-2} \sim \mathcal{G}(\nu_\omega, S_\omega^{-1})$ and $\sigma_{i,c}^{-2} \sim \mathcal{G}(\nu_c, S_c^{-1})$ respectively, where we set

[1] The normality assumption is standard and is a good fit for the US regional dataset where relative log income per head is approximately Normal in the cross section.

[2] We are using the following form for the Gamma Distribution. A random variable y, where $y \sim \mathcal{G}(\nu_\omega, S_\omega)$, has a gamma distribution with the following form

$$f_\gamma(y \mid \nu, S) = \begin{cases} \frac{S^{-\nu}}{\Gamma(\nu)} y^{\nu-1} e^{-\frac{y}{S}} & \nu, S, > 0 \qquad y > 0, \\ 0 & \text{otherwise} \end{cases}$$

This implies that y has a mean of νS and y^{-1} has a mean of $\frac{1}{S(\nu-1)}$.

$\nu_\omega = \nu_c = 3$. In Section V below, we discuss the performance of the model under different prior specifications, including their forecasting performance. Intuitively, if ω_i^2 is restricted to be very small then the estimated balanced growth path will not be able to explain or predict much of the variation in $Y_{i,t}^*$ over time and the model will resemble the panel model of equation (2) with fixed effects but without the $D_{i,t}$ indicator variable which we are interested in. Conversely, one also does not want to set S_ω to be so high that the balanced growth path is too variable, as this would conflict with the notion of a long run balanced growth path which intuitively should be slow moving. We therefore choose intermediate values of S_ω and S_c and set $S_\omega = S_c = 1$ to illustrate the model in Sections III and IV below. We discuss the choice of priors further including their forecasting performance in Section V.

Given the $\underline{\beta}$ coefficients, for each economy, i, the model described by equation (3) is a local levels model with an autoregressive transitory component as in Chan et al (2019), where the dependent variable is $\widehat{Y_{i,t}}$ where $\widehat{Y_{i,t}} \equiv Y_{i,t}^* - \sum_l^L \beta_l Y_{i,t-l}^*$. Similarly given vectors $\underline{\alpha_i} \equiv [\alpha_1, \ldots, \alpha_T]$ and $\underline{c_i} \equiv [c_1, \ldots, c_T]$ the model is a simple linear regression model of the form $\tilde{Y}_t = \sum_l^L \beta_l Y_{t-l}^*$ where $\tilde{Y}_t$ is the stacked vector $Y_{i,t}^* - \alpha_{i,t} - c_{i,t}$, which can be estimated using standard methods, The model overall can be estimated with a Gibbs sampler which is described in the Appendix.

The hierarchical model

The baseline model assumes that the autoregressive $\underline{\beta}$ coefficients are the same for all economies. This implies that the speed of convergence to the steady state is the same across economies. In our second model we follow Canova and Marcet (1995) and relax this assumption. Our second model is a hierarchical model where in addition to time varying intercept terms, an individual economy's autoregressive coefficients, now denoted $\underline{\beta_i}$, are drawn from a population distribution so that each individual economy's coefficients can vary from each other to the extent allowed by the variance of the population distribution.[3] This model is described as follows,

$$
\begin{aligned}
Y_{i,t}^* &= \alpha_{i,t} + \sum_{l=1}^{L} \beta_{l,i} Y_{i,t-l}^* + c_{i,t} + e_{i,t} \ \forall i \\
c_{i,t} &= \phi_{i,1} c_{i,t-1} + \phi_{i,2} c_{i,t-2} + u_{i,t} \ \forall i \\
\alpha_{i,t} &= \alpha_{i,t-1} + \nu_{i,t} \ \forall i
\end{aligned}
\tag{4}
$$

where now $\underline{\beta_i} \sim \mathcal{N}(\underline{\overline{\beta_l}}, \overline{\Sigma}_{\overline{\beta}^2}) \quad \forall i$, and where $\underline{\overline{\beta_l}}$ and $\overline{\Sigma}_{\overline{\beta}^2}$ are the mean and variance of the population distribution respectively. The prior for the mean of the population distribution is distributed normally, i.e. $\underline{\overline{\beta_l}} \sim \mathcal{N}(\psi, C)$, where ψ is taken from a pooled OLS regression of all economies with a common intercept term and where we set $C = I_p \times 10^8$ which is diffuse. The prior for the variance $\overline{\Sigma}_{\overline{\beta}^2}$ is distributed with a Inverse Wishart density function so that $(\overline{\Sigma}_{\overline{\beta}^2})^{-1} \sim \mathcal{W}([\rho R]^{-1}, \rho))$, where $\mathcal{W}$ denotes the Wishart distribution. We set $R = I_p \times 10^{-2}$ and $\rho = 100$. These priors are tight around $\underline{\overline{\beta_l}}$ but still allow for variation across economies as

[3]The differences between this model and Canova and Marcet (1995) are that this model includes time varying parameters, a full hierarchical model for the autoregressive parameters and uses different datasets.

we describe below. The other priors are set as in the baseline model with the same parameter values.

This model can also be estimated using Gibbs sampling. Given the $\underline{\beta_i}$ coefficients, for each economy, i, the model described by equation (4) is a local levels model with an autoregressive transitory component as in Chan et al (2019), where the dependent variable is $\widehat{Y_{i,t}}$ where $\widehat{Y_{i,t}} \equiv Y^*_{i,t} - \sum_l^L \beta_l Y^*_{i,t-l}$. Similarly given vectors $\underline{\alpha_i} \equiv [\alpha_1, \ldots, \alpha_T]$ and $\underline{c_i} \equiv [c_1, \ldots, c_T]$ for each economy i, the model is a hierarchical linear regression model with dependent variable $\hat{Y_{i,t}} = Y^*_{i,t} - \alpha_{i,t} - c_{i,t}$ and so given the population parameters, for each economy i the $\underline{\beta_i}$ can be estimated using standard methods. Given the estimated parameters for each individual economy the population parameters $\underline{\beta_i}$, $\overline{\Sigma}_{\overline{\beta}^2}$ and the inverse variance $(\frac{1}{\sigma^2})$ can be shown to be distributed, Normal, Inverse Wishart, and Gamma respectively, as described in Chan *et al* (2019).

III US states growth dynamics 1929-2019

We estimate the models described in section II using data for personal income per head for the 48 contiguous US states for 1929-2021. We follow Shioji (2004) in using the log of income per head for each state and taking its deviation from the average across states in each year. This data is publicly available from the BEA.[4]

Shioji (2004) used data for 1929-2001 and analysed the relationship between the initial level of relative income per head and the long run relative income per head, $Y^{*,LR}$. We define $Y_i^{*,LR}$ for economy i in the baseline model by

$$Y_i^{*,LR} = \frac{\alpha_{i,t}}{1 - \sum_{l=1}^{L} \beta_l} \tag{5}$$

In the hierarchical model the β_l become $\beta_{i,l}$. In Shioji's model there is no term $c_{i,t}$ and the economy specific intercept is fixed so that $\alpha_{i,t} = \alpha_i \ \forall t$.

Shioji (2004) found a strong positive relationship between initial relative income per head and long run relative income per head, $Y_i^{*,LR}$. In Figure 1 we plot this relationship for different periods in the evolution of $\alpha_{i,t}$ using the posterior means from the baseline model of equation (3) and data for 1929-2021. The first panel, Figure 1a, plots this relationship in 1933. This relationship is similar to Shioji (2004) with the variables being strongly correlated but where the estimated relationship is away from the 45° line.

The key difference between Shioji (2004) and the baseline model of equation (3) is that the latter allows $\alpha_{i,t}$ to move over time so that states' long run balanced growth paths can evolve. Thus potentially a state's initial conditions may have been close to its initial balanced growth path but far away from its present day balanced growth path. The evolution of this relationship is shown in the other panels of Figure 1. These plot the relationship for the years 1953, 1973, and 1993, so that e.g. Figure 1b plots long run relative income calculated using

[4] It is available via the website https://www.bea.gov/data/income-saving/personal-income-by-state. We use the variable personal income, which is Table SAINC4, line code 30. This is a nominal variable but since relative income is the variable used in the analysis, an inflation adjustment would cancel out.

TABLE 1

Income mobility of US states by quartile 1929-2021

	Quartile in 2021			
Quartile in 1929	**First**	**Second**	**Third**	**Fourth**
First	7	3	2	0
Second	4	2	5	1
Third	0	6	3	3
Fourth	1	1	2	8

Notes: The mobility in income per head in US states from 1929-2021

$\alpha_{i,1953}$ against relative income in 1953. These figures show that over time the relationship between the long run balanced growth and initial income tends towards the 45° line. Indeed, Figure 1b shows this relationship is mostly established after only 20 years.

An alternative way of showing the stability of relative balanced growth paths over time is to directly plot the evolution of each states' balanced growth paths. This is done in Figure 2 for the baseline model and Figure 3 for the hierarchical model. Figure 2a shows that in the baseline model from 1933 to around the mid-1970's there was noticeable convergence in the distribution of income per head - i.e. σ convergence. However, since the mid-1970's the variance of the distribution has remained fairly constant. A noticeable feature of Figure 2 is that the balanced growth paths of most states have been largely stable and slow moving. Individual states' paths do cross with their neighbours in the distribution but tend to stay in the same part of the distribution. This is consistent with the concept of a balanced growth path. There are however exceptions to this, notably North Dakota which Figure 2a highlights together with the more stable balanced growth paths for New York and South Carolina.

This stable pattern is also consistent with the raw data on income mobility presented in Table 1. In terms of upward mobility, Table 1 shows that over the 93 years of the sample only one state has moved from the bottom quartile to the top quartile. This state is North Dakota which was ranked 41st in the raw data in 1929 and 12th in 2021 and is a state with a small population and whose income per head is highly influenced by the price of oil and grain. Its unusual degree of volatility is evident in Figure 2a and also in Figure 1c where it is a noticeable outlier with its relative income level in 1973 much higher than its long run growth path. This datapoint is coincident with a large temporary rise in the price of wheat which, intuitively, the estimated long run growth path does not respond greatly to. In terms of downward mobility, no state has moved from the top quartile to the bottom quartile and only two states that were in the first quartile in 1929 dropped to the third quartile in 2021.

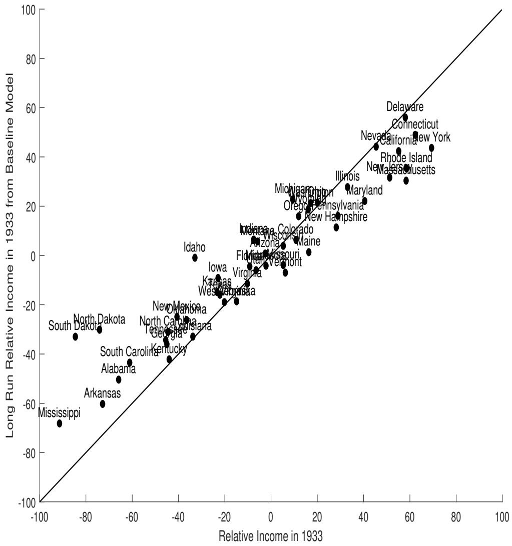
These were Nevada which was ranked 9th in 1929 and 26th in 2021 and Michigan that was ranked 10th in 1929 and 32nd in 2021. Thus change does happen but, as Figure 2a shows, this occurs against a backdrop of stability.

Figure 2b plots the 84th and 16th percentiles of the posterior distribution for the balanced growth path of the baseline model for New York and South Carolina whose balanced growth paths belong to the first and fourth quartiles, together the mean for each state. These confidence intervals can clearly distinguish between the top and bottom of the distribution and are consistent with the data described in Table 1 which shows very little mobility between the first and fourth quartiles.

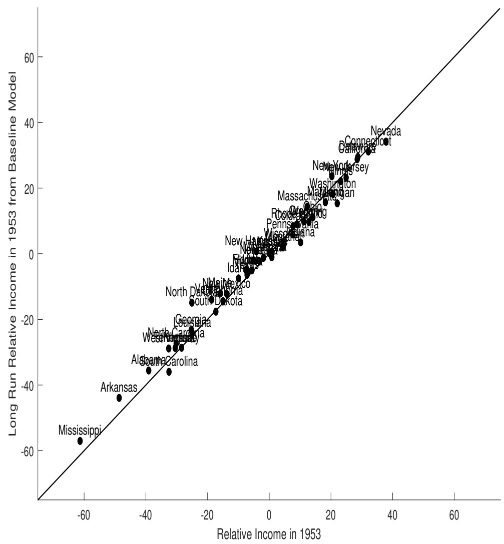
The estimated balanced growth paths for the hierarchical model are plotted in Figure 3 and are similar to those of the baseline model in Figure 2. The tight prior on the population β variance effectively restricts the β_i coefficients to be quite close to the population β, although there is still variability of the β_i across economies. The value of the sum of the mean population β's is 0.71 while that for the sum of the individual states' mean β_i coefficients varies from 0.48 to 0.85. At the moderate level of variation implied by the priors $S_\omega = S_c = 1$ which we have used in this section and Section IV, there is little difference in the performance of the two models. However, in general, the models do differ as their forecasting performance, discussed below in Section V demonstrates.

Figure 1: The evolution of balanced growth paths and initial income, US states 1933-1993

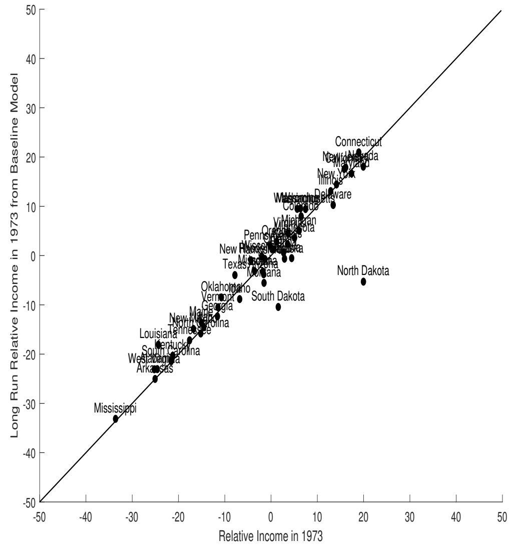
(a) 1933



(b) 1953



(c) 1973



(d) 1993

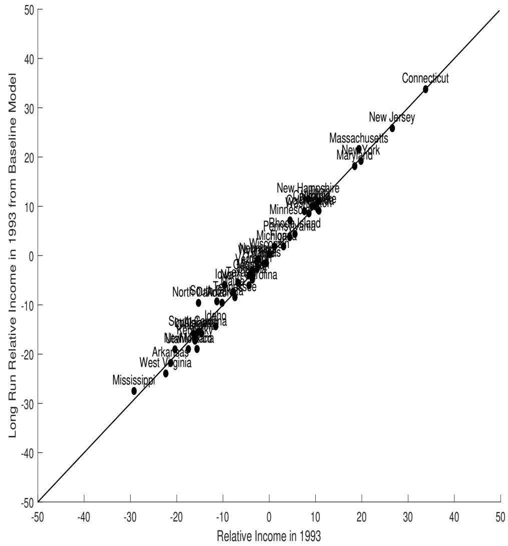
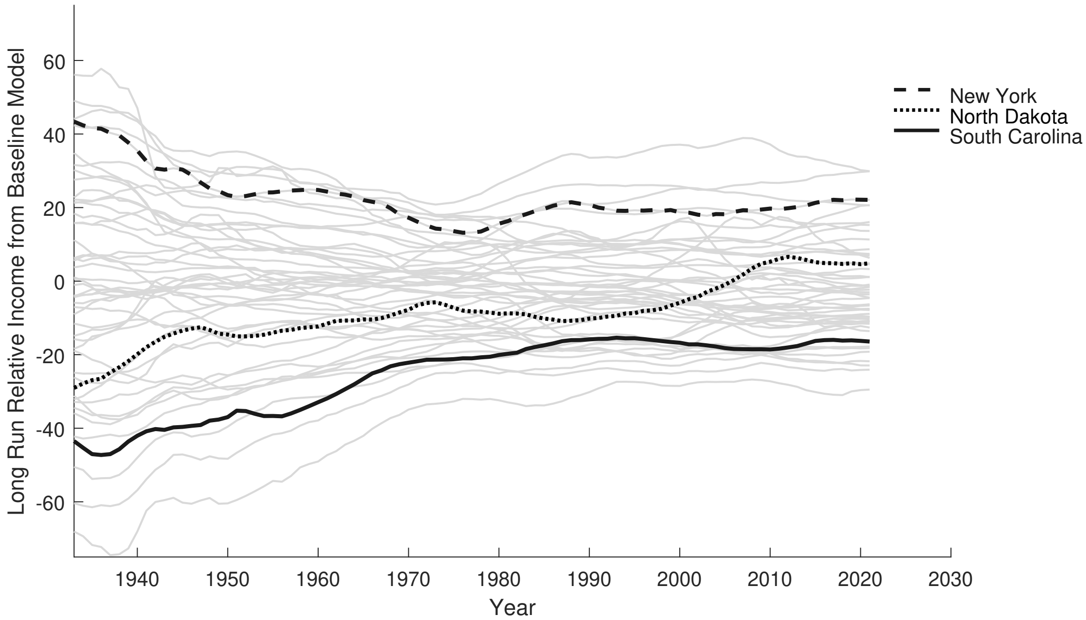
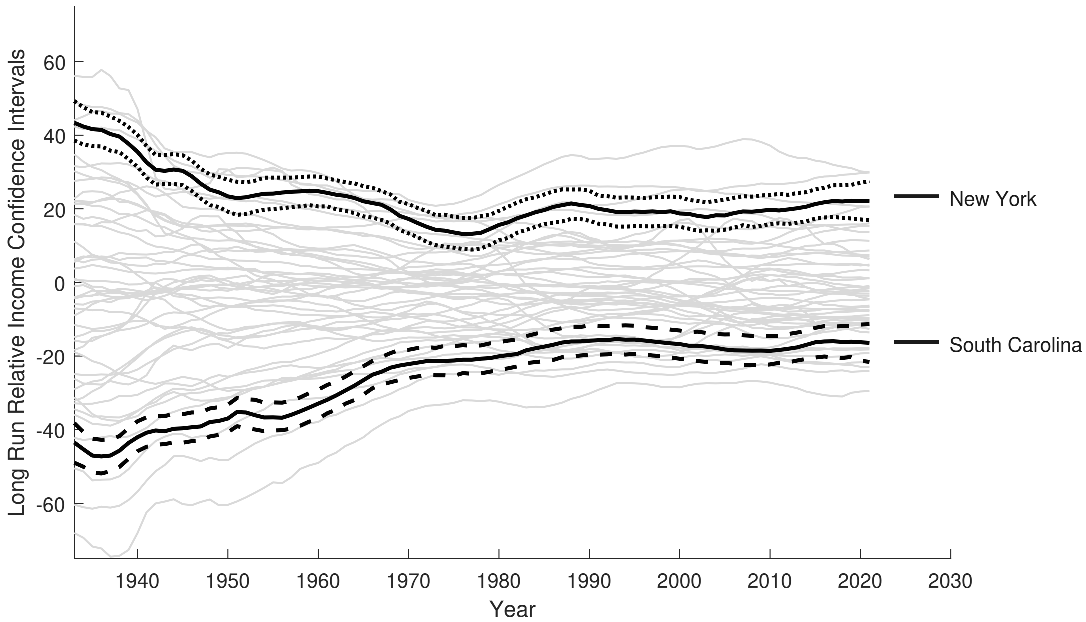


Figure 1 shows US states' initial relative income per head against the estimated long run growth path level of relative income per head for years 1933, 1953, 1973 and 1993. The estimates are the posterior means from the baseline model with four lags using data from 1929-2021 and with prior values of $S_\omega = 1$ and $S_c = 1$. The scale in the figures, and subsequent figures, is multiplied by a factor of 100 so that, e.g. the x axis shows 100× the deviation of log income per capita from the sample mean.

Figure 2: Evolution of balanced growth paths, US states 1933-1993 - Baseline Model



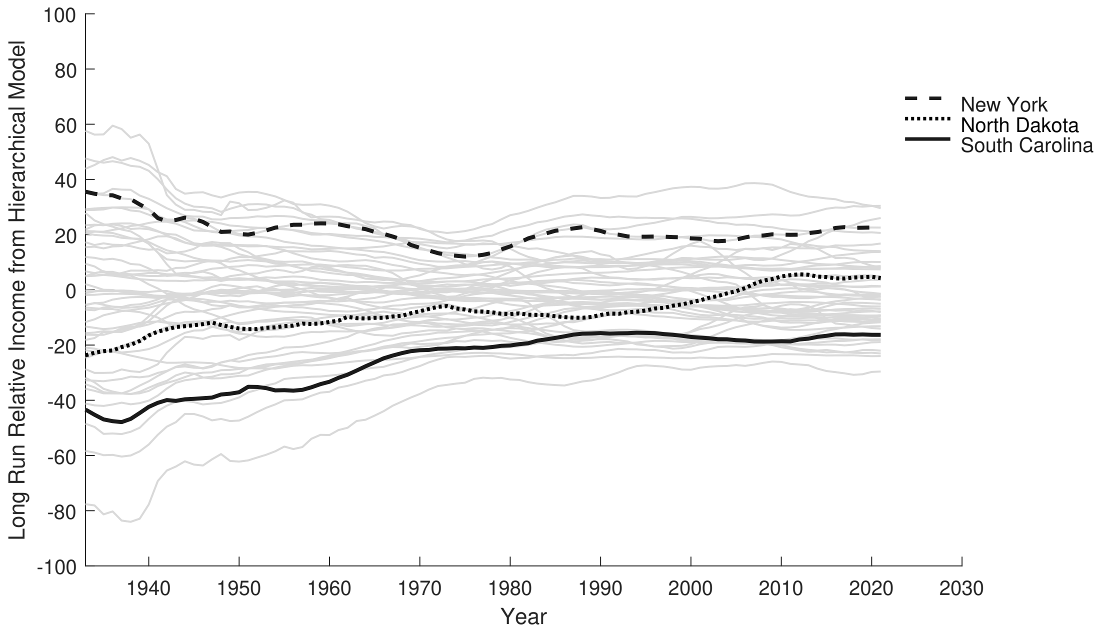
(a) Balanced growth paths in US states 1933-2021 highlighting selected states



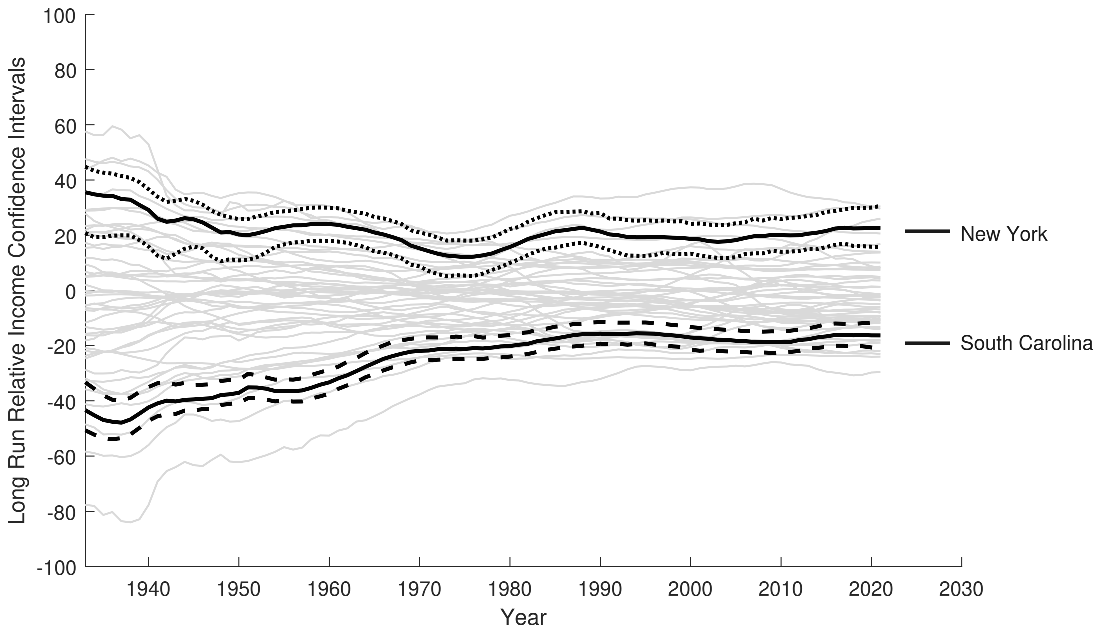
(b) Balanced growth paths in US states 1933-2021 highlighting confidence bands of two states

Figure 2 shows the evolution of the estimates of the balanced growth path for US states from the baseline model with four lags using data from 1929-2021 and with prior values of $S_\omega = 1$ and $S_c = 1$. Panel a) plots the posterior means for all states in light grey and highlights the path of three states. Figure (B1a) in the Appendix is a colour version of panel a) which highlights regions rather than individual states. Panel b) plots the 84th and 16th percentiles of the posterior distribution for two states together with the posterior means for all states in light grey, as discussed in the text.

Figure 3: Evolution of balanced growth paths, US states 1933-1993 - Hierarchical Model



(a) Balanced growth paths in US states 1933-2021 highlighting selected states



(b) Balanced growth paths in US states 1933-2021 highlighting confidence bands of two states

Figure 3 shows the evolution of the estimates of the balanced growth path for US states from the hierarchical model with four lags using data from 1929-2021 and with prior values of $S_\omega = 1$ and $S_c = 1$. Panel a) plots the posterior means for all states in light grey and highlights the path of three states. Figure (B1b) in the Appendix is a colour version of panel a) which highlights regions rather than individual states. Panel b) plots the 84th and 16th percentiles of the posterior distribution for two states together with the posterior means for all states in light grey, as discussed in the text.

IV World growth dynamics 1970-2019

In this section we carry out the same analysis as in section III but for GDP per head in the world economy for 1970-2019 using the Penn World Table 10.0 dataset. We follow section III and use the log of GDP per head and take its deviation from the average across countries in each year.[5] Following, for example, Patel *et al* (2021), we exclude oil producing economies and small economies which leaves a dataset of 123 countries.[6] As we discuss below, the results of the analysis for the world economy share many of the characteristics of that for the US states above, most notably that an economy's initial level of income per head is a strong predictor of its long run balanced growth path and that there is little relative mobility.

The relationship between an economy's initial conditions in 1974 and its estimated balanced growth path in 1974 is shown in Figure 4. The first panel, Figure 4a, plots this relationship for the baseline model. In contrast to the US states dataset described in Section III, the relationship for the world economy fits the 45° line well from the beginning of the dataset, albeit it with significant variation. Figure 4b plots the same relationship for the hierarchical model, which is very similar. In both models therefore the balanced growth paths do not need to move over time in order to fit the 45° line. This is consistent with the plots of the balanced growth paths in Figure 5 for the baseline model and Figure 6 for the hierarchical model, which show little evidence of any convergence in the distribution of relative GDP per head.

The estimated balanced growth paths for the world economy for the baseline model in Figure 5 are similar to those for US states in Figure 2 in that there is a great deal of stability. Figure 5a highlights the balanced growth paths for the United States, Cameroon and China which are all smooth and quite stable although with an upward trend for China and a downward trend for Cameroon. As in the US states dataset the confidence intervals for the baseline model Figure 5b can clearly distinguish between the top and bottom of the distribution. The confidence bands appear very narrow but this is due to the scaling of the Figure 5b which is much wider than in Figure 2b.

The estimated balanced growth paths for the hierarchical model are plotted in Figure 6 and are similar to those of the baseline model plotted in Figure 5 at the moderate level of variation implied by the priors $S_\omega = S_c = 1$ which we have used in this section. As with the US states dataset, there is variability of the β_i across economies. The value of the sum of the mean population β's is 0.88 while that for the sum of the individual states' mean β_i coefficients varies from 0.68 to 0.99. In general the two models do differ in their performance

[5] This data is publicly available via the website https://www.rug.nl/ggdc/productivity/pwt/. We follow e.g. Patel *et al* (2021) and use the variables 'rdpe' and 'pop' to calculate real GDP per head. 'rgdpe' is the Expenditure-side real GDP at chained purchasing power parity and 'pop' is the population. We note this is an output measure rather than the income measure used in Section III but in each case we are following the datasets used in the literature so that we cleanly demonstrate the implications of our different modelling approach.

[6] The oil producing countries are the same as in Patel *et al* (2021), and small countries are those with a population less than 100,000 in 1970.

TABLE 2

Income mobility in world economy by quartile 1970-2019

	Quartile in 2019			
Quartile in 1970	**First**	**Second**	**Third**	**Fourth**
First	25	5	0	0
Second	4	19	8	0
Third	1	5	12	13
Fourth	0	2	11	18

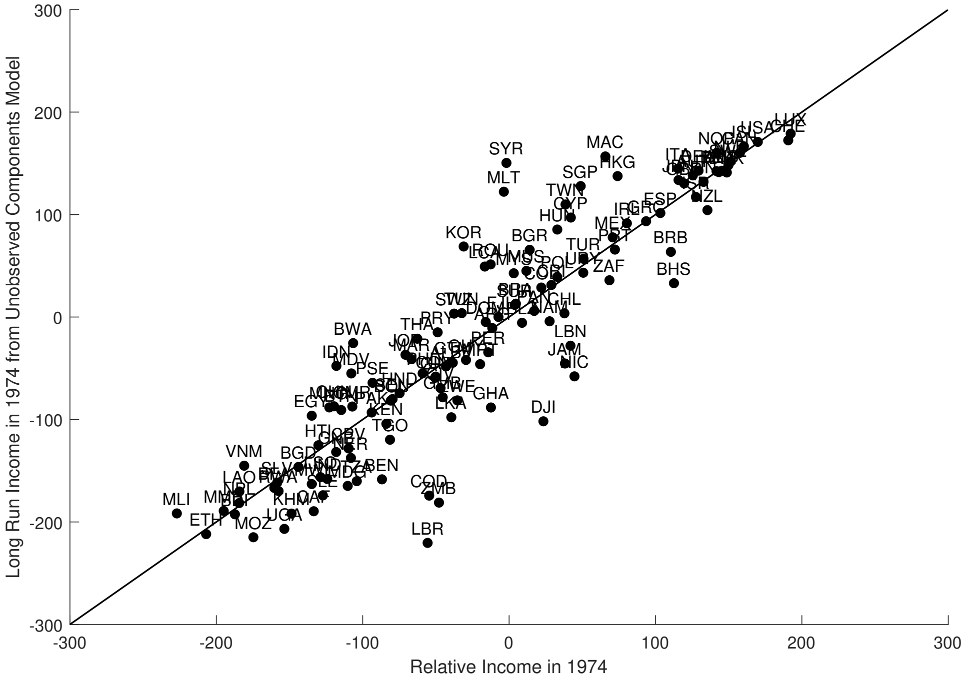
Notes: The mobility of GDP per head in the world economy dataset from 1970-2019

as we discuss below in Section V.

The stability of the balanced growth paths is again consistent with the raw data on income mobility presented in Table 2. This shows that over the 50 years of the sample no country has moved from the top quartile to the bottom half of the distribution and no country moved from the bottom quartile to the top quartile. One country moved from the third quartile in 1970 to the first quartile in 2019. This country is the Republic of Korea which was ranked 64th in 1970 and 27th in 2019. Change is possible, notable examples being China that rose from a ranking of 101st in 1970 to 63 in 2019 and India from 103rd in 1970 to 83rd in 2019. However, as with the US states dataset, this change is occurring in a broadly stable setting.

Figure 4: Balanced growth paths and initial income in the world economy

(a) Baseline model



(b) Hierarchical model

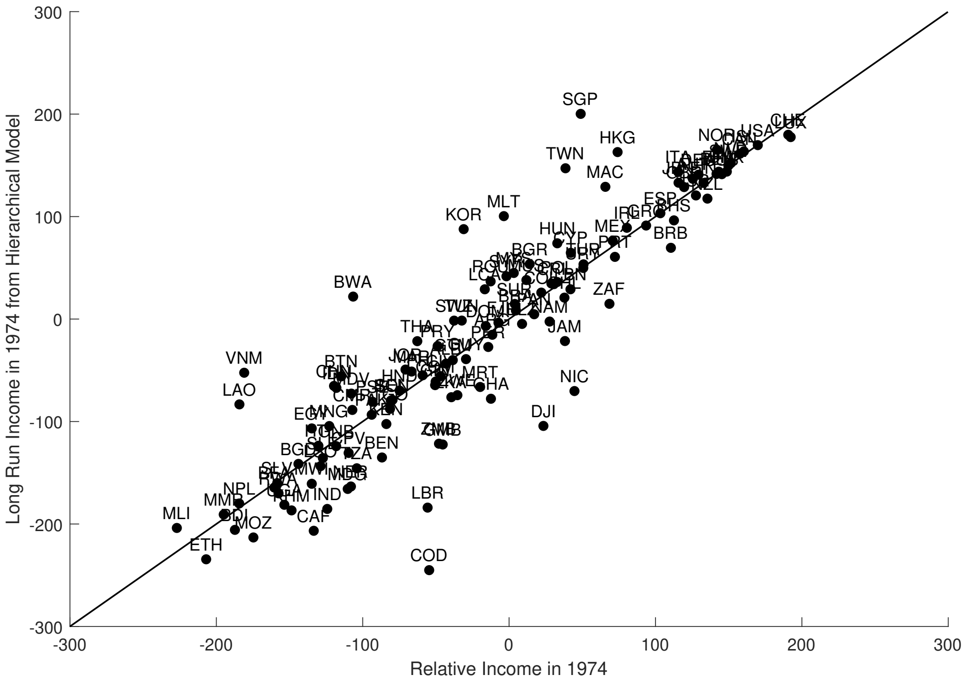
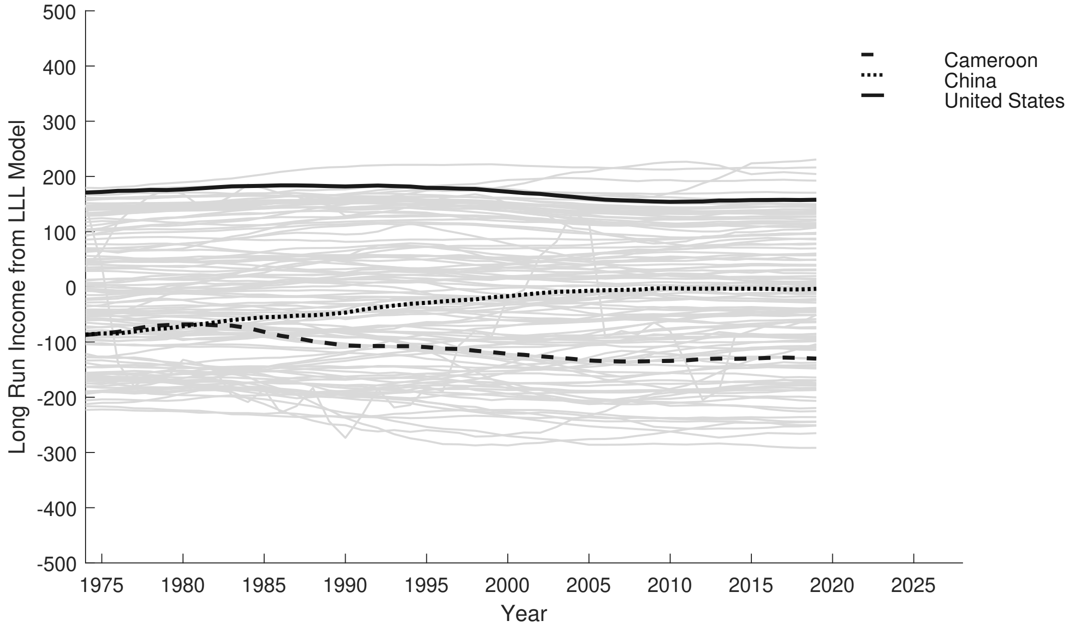
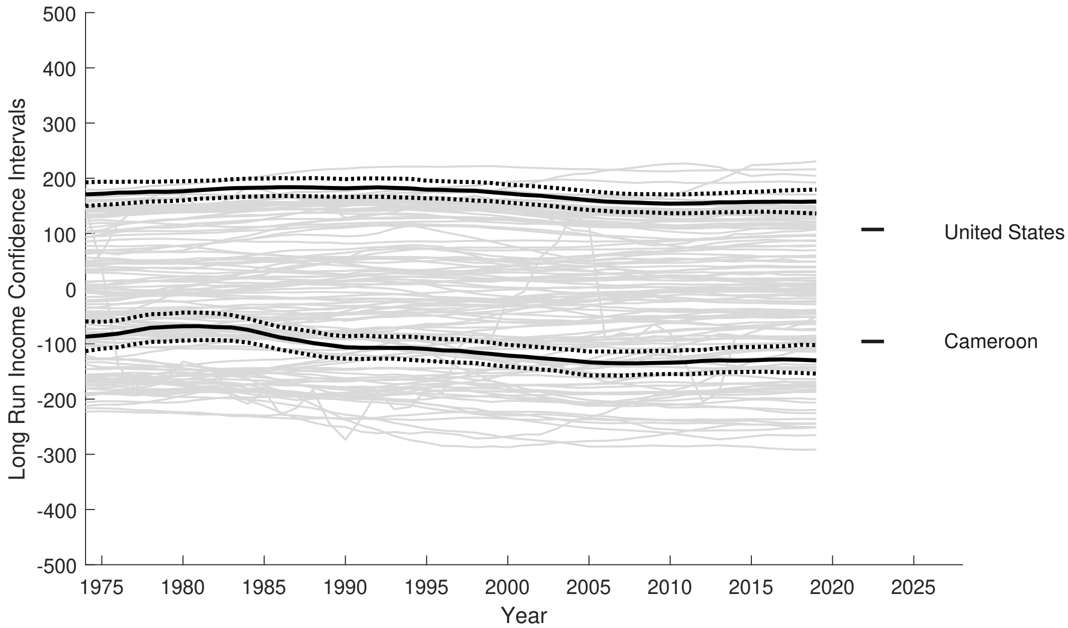


Figure 4 shows countries' initial relative GDP per head plotted against estimated long run balanced growth path relative GDP for 1974. Panel a) show the results for the baseline model and Panel b) shows those for the hierarchical model. The estimates are the posterior means and both models have four lags and use data from 1970-2019. The labels are three letter ISO codes.

Figure 5: Evolution of balanced growth paths, world economy 1974-2019 - Baseline model



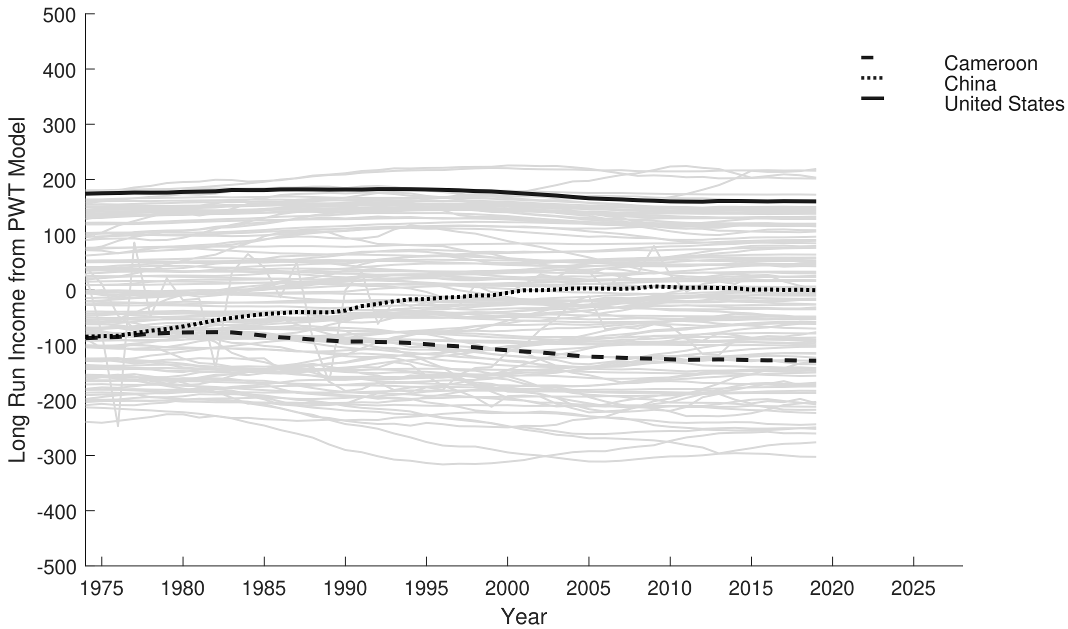
(a) Balanced growth paths in the world economy highlighting selected countries



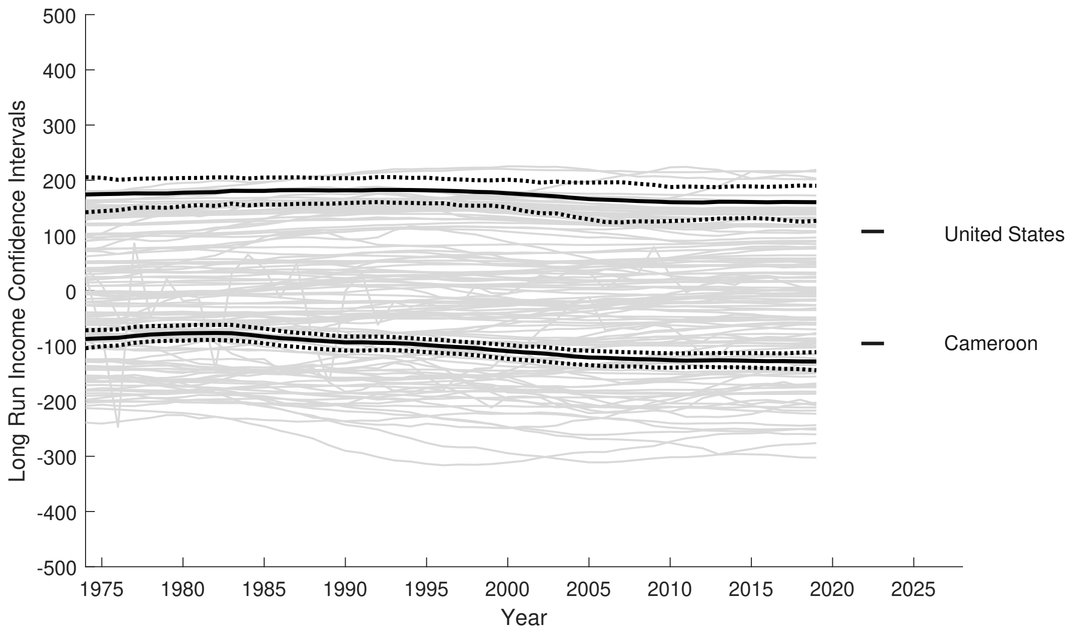
(b) Balanced growth paths in the world economy highlighting confidence bands for two countries

Figure 5 shows the evolution of the estimates of the balanced growth path for the world economy dataset from the baseline model with four lags using data from 1970-2019 and with prior values of $S_\omega = 1$ and $S_c = 1$. Panel a) plots the posterior means for all countries in light grey and highlights the path of three countries. Figure (B2a) in the Appendix is a colour version of panel a) which highlights continents rather than individual countries. Panel b) plots the 84th and 16th percentiles of the posterior distribution for two countries together with the posterior means for all states in light grey, as discussed in the text.

Figure 6: Evolution of balanced growth paths, world economy 1974-2019 - Hierarchical model



(a) Balanced growth paths in the world economy highlighting selected countries



(b) Balanced growth paths in the world economy highlighting confidence bands for two countries

Figure 6 shows the evolution of the estimates of the balanced growth path for the world economy dataset from the hierarchical model with four lags using data from 1970-2019 and with prior values of $S_\omega = 1$ and $S_c = 1$. Panel a) plots the posterior means for all countries in light grey and highlights the path of three countries. Figure (B2b) in the Appendix is a colour version of panel a) which highlights continents rather than individual countries. Panel b) plots the 84th and 16th percentiles of the posterior distribution for two countries together with the posterior means for all states in light grey, as discussed in the text.

V Choice of prior and the fit of the model

Section II gave two motivations for the choice of priors: (i) that the concept of a balanced growth path implies a path which is not very volatile and (ii) that the balanced growth path should help predict the future path of relative income per head. In this section we discuss these two motivations further. We first describe the implications of the choice of prior for the shape of the estimated balanced growth path. We then show that models whose priors imply a near constant balanced growth path perform poorly in terms of out-of-sample prediction compared to models whose estimated growth path is able to evolve significantly over time. Furthermore, in the US states dataset there is also some evidence that a moderate degree of variability performs better at longer forecast horizons, although for shorter horizons and for the world economy dataset, a better forecast performance is often associated with a higher level of volatility. Taken together these results give support for modelling growth paths as time varying as proposed by this paper. In addition for the US states dataset a balanced growth path interpretation of the results seems reasonable, whereas for the world economy dataset the long run growth paths seem too volatile for this interpretation.

The choice of prior and estimated balanced growth path

In Figure 7 we plot the estimated balanced growth paths of the baseline model for the US states dataset when we set a very low prior value, 10^{-5}, for S_ω (and where S_c was set to one). In this case the estimated balanced growth paths are essentially flat and the model approaches that of the fixed effects panel regression model of Shioji (2004). As discussed above we don't think this is a reasonable choice of prior because there are good reasons for believing that an economy's balanced growth path can change significantly through time. Furthermore we show below that the forecasting performance of models with very low prior levels of S_ω is not good. However, for a balanced growth path interpretation it is also not reasonable to set S_ω to be very high as a very volatile balanced growth path does not correspond to our understanding of what a balanced growth path is. For the world economy dataset we find that models with high levels S_ω perform well in out of sample forecasts compared to more moderate levels of S_ω and so for this dataset the balanced growth path interpretation is not appropriate.

Choice of prior and prediction

We perform an out-of-sample forecasting exercise for both models using both datasets over a grid of different values for S_ω and S_c ranging from extremely low ($S_i = 10^{-5},\ i = \omega, c$) to extremely high ($S_i = 10^5,\ i = \omega, c$).[7] In the exercise we use the last 25 observations as the out-of-sample data with forecasts made at the one, five and ten periods ahead forecast horizon. Forecasts are made using the posterior mean of the parameter estimates using 20,000

[7] This forecasting exercise is included to illustrate the value of a model that allows balanced growth paths to change over time. It is far from an exhaustive search for the best performing prior specification which would be well beyond the scope of this study.

Figure 7: The US states estimated balanced Growth Paths with $S_\omega = 0.00001$

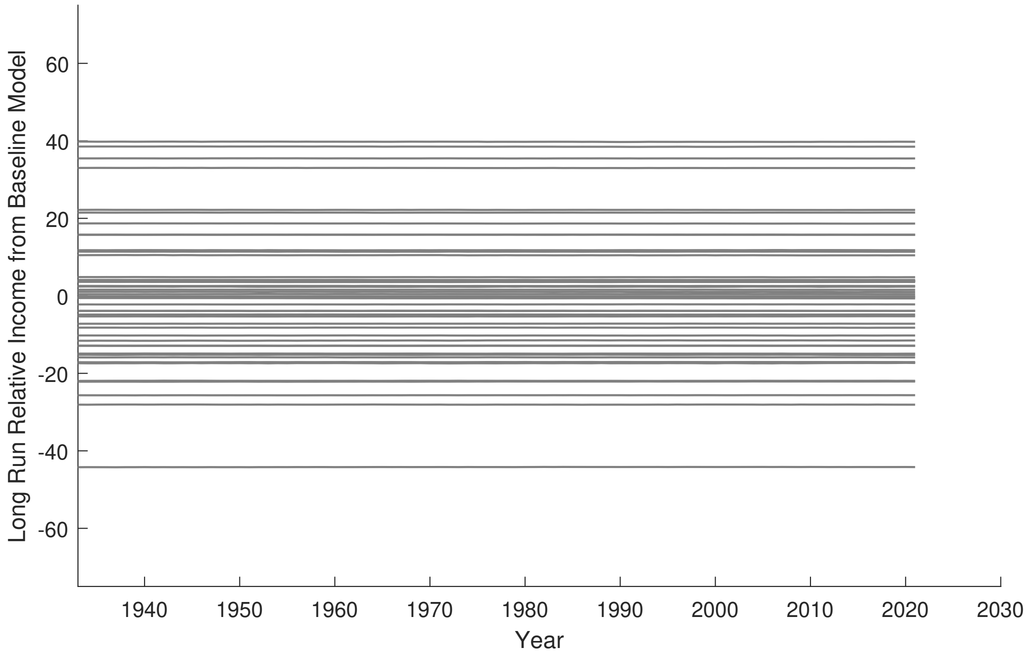


Figure 7 shows the balanced growth path for the 48 contiguous US states from 1933-2021 using the baseline model with $S_\omega = 0.00001$. The estimates are the posterior means.

draws from the posterior. Forecasts are made recursively so that parameters are re-estimated using new information as the out-of-sample roles forward, i.e. the root mean squared forecast error (RMSFE) is given by

$$\text{RMSFE} = \left[\frac{1}{N} \sum_{j=1}^{N} \sum_{t=\tau_0}^{T-h} \frac{(y_{j,t+h} - E(y_{j,t+h}|\text{Data}_t])^2}{T - h - \tau_0 + 1} \right]^{\frac{1}{2}} \tag{6}$$

where $h = \{1, 5, 10\}$.

We focus on the ten periods ahead forecast, $h = 10$, as this most corresponds to a long run forecast which is the primary interest of the paper. The shorter horizon forecasts will be influenced more by the transitory process $c_{i,t}$ and these results are placed in the Appendix.

Table 3 displays the ten year ahead forecasts for the US states dataset. The first row of the first block displays the RMSFE for the baseline model when S_ω is very small so that there is very little movement in the states' balanced growth paths over time, as in Figure 7. This row has a relatively poor forecasting performance compared to other rows and this illustrates the benefit of our modelling approach that allows the balanced growth path to move through time. In line with the intuition given above, the best performing row, displayed in bold, is at an intermediate level of variation, where of $S_\omega = 10^{-1}$. However the next best performance is where $S_\omega = 10^5$ and the one after that $S_\omega = 10^3$. Thus we can only conclude that the moderate levels of S_ω perform better than very low levels, and comparably to high and very levels of S_ω. The second block in Table 3 displays the results for the hierarchical model. Here the benefit of a moderate value of S_ω is much more apparent, as moving down each column there is a 'U' shape for forecast performance. The lowest RMSFE is where $S_\omega = 10$ $S_c = 100$ but its performance is close to that adjacent cells to the left, above and below. The forecasting performance of the first and last rows are now both relatively poor.

Table 3 therefore gives some support to hypothesis that intermediate values for S_ω are beneficial for forecasting long run balanced growth paths. However, this evidence is far from definitive. This forecasting exercise is not intended to be a rigorous search for the best performing prior. The grid is very sparse and while for the hierarchical model there does appear to be a local minimum at intermediate levels of S_ω for most columns, there is also evidence in the baseline model for beneficial forecasting performance of very high levels of S_ω. This also true for shorter forecast horizons in Tables A1 and A3 in the Appendix. However, across both models and at all forecast horizons the first row, with a very low level of S_ω, performs relatively poorly. These results therefore do give evidence for the benefit of modelling growth paths as time varying which is the primary purpose of the exercise.

The results of the ten year ahead forecasts for the world economy are displayed in Table 4 and in Tables A2 and A4 in the Appendix. These results are similar to those for the US states dataset in that for both the baseline and the hierarchical model and at all forecast horizons the first row, with a very low level of S_ω, is one of the worst performing rows. This again provides support for modelling growth paths as time varying. However, there is a much less evidence of the benefits of intermediate levels of S_ω in this dataset although there is some evidence of local minima at intermediate levels in some columns of Table 4. As stated above, the clear benefit of high levels of volatility for ω in the baseline model, is not consistent with a balanced growth interpretation of the data.

TABLE 3

Ten years ahead forecast performances (RMSFE): US states dataset

S_c	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^{2}	10^{3}	10^{5}
S_ω									
Baseline Model									
10^{-5}	8.715	7.202	7.620	7.416	9.406	8.484	15.834	11.276	15.785
10^{-3}	7.639	7.608	7.504	7.738	7.769	7.350	7.245	8.447	28.172
10^{-2}	7.316	7.166	7.515	7.401	7.097	6.709	6.580	7.294	50.397
10^{-1}	**7.189**	**6.999**	**6.963**	**7.077**	**7.069**	**6.589**	**6.354**	**6.688**	**44.651**
1	7.286	7.198	7.392	7.326	7.029	6.880	6.566	6.685	76.619
10	7.063	7.122	7.236	6.957	7.061	7.019	6.721	7.453	61.896
10^{2}	6.617	6.679	6.611	6.649	6.578	6.606	6.637	7.856	49.807
10^{3}	6.537	6.522	6.572	6.520	6.525	6.536	6.578	8.532	59.769
10^{5}	6.511	6.538	6.548	6.534	6.530	6.523	6.576	7.831	194.492
Hierarchical Model									
10^{-5}	7.649	7.528	7.698	7.622	7.787	8.017	9.636	10.584	26.154
10^{-3}	6.760	6.733	6.744	6.727	6.900	7.089	7.425	8.993	27.260
10^{-2}	6.571	6.579	6.586	6.560	6.766	6.865	7.020	8.138	34.886
10^{-1}	6.502	6.493	6.502	6.491	6.463	6.465	6.574	7.622	44.688
1	6.495	6.485	6.493	6.497	6.496	6.430	6.444	8.032	44.395
10	**6.496**	**6.500**	**6.496**	**6.489**	**6.495**	**6.477**	**6.424**	**8.021**	**43.533**
10^{2}	6.502	6.508	6.495	6.499	6.499	6.498	6.471	7.664	54.005
10^{3}	6.515	6.519	6.515	6.527	6.535	6.515	6.579	7.898	72.555
10^{5}	9.161	10.499	11.723	8.393	8.409	9.742	10.573	20.236	156.810

Notes: Table 3 reports the Root-Mean-Squared Forecast Errors for ten period ahead out-of-sample forecast, for the Regional US dataset. The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.

TABLE 4

Ten years ahead forecast performances (RMSFE): PWT Dataset

$\boldsymbol{S_c}$	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^{2}	10^{3}	10^{5}
$\boldsymbol{S_\omega}$									
Baseline Model									
10^{-5}	48.53	37.25	37.52	39.40	37.76	37.59	36.62	33.76	51.64
10^{-3}	45.03	45.81	45.15	45.38	46.48	44.82	41.49	40.62	79.72
10^{-2}	43.70	43.21	43.71	43.37	43.46	43.00	40.01	36.93	95.74
10^{-1}	41.63	41.98	42.24	41.80	42.18	41.01	38.20	32.64	123.52
1	40.78	41.08	40.44	40.05	41.36	39.99	37.81	30.69	114.08
10	39.04	38.96	38.92	38.64	38.55	38.07	37.91	31.72	97.20
10^{2}	31.27	31.57	30.99	31.25	31.79	31.89	33.46	32.25	87.77
10^{3}	**29.81**	**29.68**	**29.84**	**29.72**	**29.85**	**29.87**	**30.43**	**31.68**	**100.39**
10^{5}	30.04	30.13	29.71	29.87	30.00	29.88	30.87	34.71	308.85
Hierarchical Model									
10^{-5}	37.79	38.19	38.08	38.21	37.42	38.78	38.23	35.42	86.47
10^{-3}	37.64	38.03	37.37	37.58	37.55	38.15	36.87	33.49	85.97
10^{-2}	38.11	39.27	38.41	38.30	39.48	37.57	39.34	38.19	94.03
10^{-1}	37.21	37.35	36.48	36.78	36.74	35.78	39.31	38.22	117.32
1	32.96	33.08	33.23	33.31	32.64	34.19	37.28	36.79	128.66
10	**33.54**	**32.50**	**33.45**	**34.15**	**32.46**	**32.90**	**33.22**	**35.81**	**122.05**
10^{2}	33.30	34.52	34.20	34.39	34.88	34.35	34.75	35.85	116.61
10^{3}	38.73	36.76	38.20	38.10	38.63	37.89	38.15	41.59	154.26
10^{5}	51.33	54.35	55.72	54.89	54.93	52.94	54.81	58.82	373.42

Notes: Table 4 reports the Root-Mean-Squared Forecast Errors for ten period ahead out-of-sample forecast. The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.

VI Conclusion

Are economies converging to the same balanced growth path or their own individual growth paths? In this paper we have added unobserved components analysis to an otherwise standard empirical model of economic growth dynamics, so that an economy's growth path can change at any point in time. We applied this model to data on income per head from US states and the world economy. Although the empirical model allows growth paths to change through time, in both datasets there is very little evidence of convergence in the last 50 years. This result is most striking for the US states dataset. US states are free trading, democratic and peaceful, and operate under free interstate capital and labour mobility. The conditions for convergence are thus as good as can be expected. Yet there is little evidence of convergence over the last half century. 'The Poor' stay relatively poor.

If one accepts this paper's analysis, the natural next question to ask is 'Why then are US states converging to different balanced growth paths?'. This paper does not attempt to point to a specific determinant. The literature has provided many, possibly correlated and not mutually exclusive, potential candidates. The issue is highly multi-dimensional. In this context looking at the big picture, via a flexible reduced form analysis of income per head dynamics, is informative. The results show that an economy's initial level of relative income per head appears to be a good summary measure of what is important for long run relative growth.

References

Acemoglu, D., Naidu, S., Restrepo, P. and Robinson, J. (2019), 'Democracy Does Cause Growth', *Journal of Political Economy*, Vol. 127, pp. 47-100.

Acemoglu, D and Molina, C. (2022), 'Comment', in Editors: Eichenbaum, M., Hurst. E, and Parker, J. (eds) *NBER Macroeconomics Annual*, Vol. 36, pp. 425-442.

Andrés, J., Boscá, J. and Doménech, R. (2004), 'Convergence in the OECD: Transitional Dynamics or Narrowing Steady-State Differences?', *Economic Inquiry*, Vol. 42, pp. 141-149.

Barro, R. and Sala-i-Martin, X. (1991), 'Convergence Across states and Regions', *Brooking Papers on Economic Activity*, pp. 107-182.

Caselli, F., Esquivel, G. and Lefort, F. (1996), 'Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics', *Journal of Economic Growth*, Vol. 1, pp. 363:389.

Canova, F. (2007), 'Methods for Applied Macroeconomic Research', Princeton University Press.

Canova, F. (2022), 'Should we trust cross-sectional multiplier estimates?', , *Journal of Applied Econometrics* Forthcoming.

Canova, F. and Marcet, A. (1995), 'The Poor Stay Poor: Non-convergence Across Countries and Regions', Discussion Paper No. 1265, CEPR.

Cerra, V. and Saxena, S. (2008), 'Growth Dynamics: The Myth of Economic Recovery', *American Economic Review*, Vol. 98, pp. 439-57.

Chan,J., Koop, G. and Potter S. (2013), 'A New Model of Trend Inflation', *Journal of Business & Economic Statistics*, Vol. 31, pp. 94-106.

Chan, J., Koop, G., Poirier, D. and Tobias, J. (2019), *Bayesian Econometric Methods, Second Edition*, Cambridge University Press.

Chen, S., Chernozhukov, V. and Fernández-Val, I. (2019), 'Mastering Panel Metrics: Causal Impact of Democracy on Growth', *AEA Papers and Proceedings*, Vol. 109, pp. 77-82.

Imam, P. and Temple, J. (2023), 'Political Institutions and Output Collapses', Working Paper, Vol 2023, No. 036, International Monetary Fund.

Koop, G (2003), *Bayesian Econometrics*, John Wiley and Sons Ltd.

Kremer, M., Willis, J. and You Y. (2022), 'Converging to Convergence', in Editors: Eichenbaum, M., Hurst. E, and Parker, J. (eds) *NBER Macroeconomics Annual*, Vol. 36, pp. 337-412.

Mankiw, G., Romer, D. and Weil, D. (1992), 'A Contribution to the Empirics of Economic Growth', *The Quarterly Journal of Economics*, Vol. 107, pp. 407-437.

Patel, D., Sandefur, J. and Subramanian, A. (2021), 'The New Era of Unconditional Convergence', *Journal of Development Economics*, Vol. 152

Shioji, E. (2004), 'Initial Values and Income Convergence: Do "The Poor Stay Poor"?', *The Review of Economics and Statistics*, Vol. 86, pp. 444-446.

Startz, Richard, (2020), The next hundred years of growth and convergence, Journal of Applied Econometrics, 35: 99-113

Wacziarg, R. and Welch, K. (2008), 'Trade Liberalization and Growth: New Evidence', *The World Bank Economic Review*, Vol 22, pp 187-231.

Appendix A

Estimation Algorithms

Draws from the posterior for all the models are made iteratively via a Gibbs sampling algorithm which draws in sequence from the conditional posterior distributions described below. We follow Chan *et al* (2019) closely here, including the notation. Further detail and derivations can also be found at this reference.

Model 1: The Baseline Model

The baseline model of equation (3) with T time periods and N individual economies has the set of estimated parameters, Γ, where $\Gamma = [\{\underline{\alpha_i}\}_{i=1}^N, \{\omega_i^2\}_{i=1}^N, \{\underline{c_i}\}_{i=1}^N, \{\omega_{c,i}^2\}_{i=1}^N \{\alpha_{0,i}\}_{i=1}^N, \{(\phi_{1,i}, \phi_{2,i})\}_{i=1}^N, \underline{\beta}, \sigma^2]$ where $\underline{\alpha_i} \equiv [\alpha_1, \ldots, \alpha_T]'$, $\underline{c_i} \equiv [c_1, \ldots, c_T]'$ and $\{\psi_i\}_{i=1}^N$ denotes the set of coefficients ψ_i for all economies, $i = 1 : N$.

Denoting all parameters other than ψ by $\Gamma_{-\psi}$, then the conditional posterior distribution for parameter ψ given all the other parameters and the data, Y can be written $p(\psi \mid \Gamma_{-\phi}, Y)$. Using this notation, the Gibbs sampler algorithm can be described as follows:-

Choose starting values for $\underline{\beta}$, ω^2, σ^2,$\underline{c_i}$ and $\alpha_{0,i}$, and also the number of draws, n^{draw}. Then cycle through draws from the condition posterior distributions described in (i)-(viii) below, n^{draw} times, saving the draws after discarding an initial number, n^{burnin}. We choose $n^{\text{draw}} = 2 \times 10^4$ and $n^{\text{burnin}} = 10^4$ in all our estimations. Draws are thinned by a factor of 10 leaving 1000 draws for calculations.

(i) Draw from $p(\underline{\alpha_i} \mid \Gamma_{-\underline{\alpha_i}}, \underline{Y_i})$, for each economy i, separately. This has the distribution $\mathcal{N}(\underline{\hat{\alpha}_i}, V_\alpha)$ where

$$\underline{\hat{\alpha}_i} \quad = \quad V_\alpha \left(\frac{\alpha_{0,i}}{\omega_i^2} H' H 1_T + \frac{1}{\sigma^2} \widehat{Y_i^\alpha} \right) \quad V_\alpha = \left[\frac{1}{\omega^2} H' H + \frac{1}{\sigma^2} I_T \right]^{-1}$$

where $\widehat{Y_{i,t}^\alpha} \equiv Y_{i,t} - \sum_l^L \beta_l Y_{i,t-l} - c_{i,t}$ and where $\widehat{Y_i} = [\widehat{Y_{i,1}^\alpha}, \widehat{Y_{i,2}^\alpha}, \ldots, \widehat{Y_{i,N}^\alpha}]'$, 1_T is a $T \times 1$ vector of ones and H is a $T \times T$ matrix with 1 on the diagonal, and -1 below the diagonal and zeros elsewhere.

(ii) Draw from $p(\underline{c_i} \mid \Gamma_{-\underline{c_i}}, \underline{Y_i})$, for each economy i, separately. This has the distribution $\mathcal{N}(\underline{\hat{c}_i}, V_c)$ where

$$\underline{\hat{c}_i} \quad = \quad V_c(\frac{1}{\sigma^2} \widehat{Y_i^c}) \quad V_c = \left[\frac{1}{\sigma_c^2} H_\phi' H_\phi + \frac{1}{\sigma^2} I_T \right]^{-1}$$

where $\widehat{Y_{i,t}^c} \equiv Y_{i,t} - \sum_l^L \beta_l Y_{i,t-l} - \alpha$, 1_T is a $T \times 1$ vector of ones and H_ϕ is a $T \times T$ matrix with 1 on the diagonal, and $-\phi_{1,i}$ and $-\phi_{2,i}$ on the two rows immediately to the left of the diagonal -space permitting, see Chan et al (2019).

(iii) Draw from $p(\underline{\phi_i} \mid \Gamma_{-\underline{\phi_i}}, \underline{Y_i})$, for each economy i, separately. This has the distribution $\mathcal{N}(\underline{\hat{\phi}_i}, V_\phi)$ where

$$\underline{\hat{\phi}_i} \quad = \quad V_\phi \left(\frac{\mu_\phi}{S_\phi} + \frac{1}{\sigma_{c,i}^2} X_\phi \widehat{c_i} \right) \quad V_\phi = \left[\frac{1}{S_\phi} + \frac{1}{\sigma_{c,i}^2} X_\phi' X_\phi \right]^{-1}$$

where $\underset{(T\times 2)}{X_\phi} = [\underline{c_{i,t-1}} \ , \ \underline{c_{i,t-2}}]$.

(iv) Draw from $p(\sigma^2_{c,i} \mid \Gamma_{-\sigma^2_{c,i}}, \underline{Y_i})$, for each economy i, separately. $p(\frac{1}{\sigma^2_{c,i}} \mid \Gamma_{-\sigma^2_{c,i}}, \underline{Y_i})$ has the Gamma density (α, β) form where $\hat{\alpha} = \nu_{c,i} + \frac{T}{2}$ and $\hat{\beta} = \frac{1}{[S_{c,i} + \frac{[(c_{i,t} - X_\phi \phi)'((c_{i,t} - X_\phi \phi)]}{2}]}$

(v) Draw from $p(\omega_i^2 \mid \Gamma_{-\omega_i^2}, \underline{Y_i})$, for each economy i, separately. $p(\frac{1}{\omega_i^2} \mid \Gamma_{-\omega_i^2}, \underline{Y_i})$ has the Gamma density (α, β) form with parameters $\hat{\alpha} = \nu_{\omega,i} + \frac{T}{2}$ and $\hat{\beta} = \frac{1}{S_{\omega,i} + \frac{[(\underline{\alpha_i} - \alpha_0 1_T)' H' H (\underline{\alpha_i} - \alpha_0 1_T)]}{2}]}$

(vi) For each economy i, draw from $p(\alpha_{0,i} \mid \Gamma_{-\alpha_{0,i}}, Y_i)$. $p(\alpha_{0,i} \mid \Gamma_{-\alpha_{0,i}}, Y_i) \sim \mathcal{N}(\hat{\alpha_{0,i}}, V_{\alpha i 0})$ where

$$\hat{\alpha_{0,i}} \quad = \quad V_{\alpha i 0} \left(\frac{a_0}{b_0} + \frac{\underline{\underline{\alpha_i(1)}}}{\omega_i^2} \right) \qquad V_{\alpha i 0} = \left[\frac{1}{b_0} + \frac{1}{\omega_i^2} \right]^{-1}$$

(vii) Given the parameters $\underline{\alpha}$ from all economies and defining $\underline{Y_i^\star} = \underline{Y_i} - \underline{\alpha_i} - \underline{c_i}$ and stacking this variable across economies to create $Y^\star = [\underline{Y_1^{\star\prime}}, \underline{Y_2^{\star\prime}}, \ldots, \underline{Y_2^{\star\prime}}]'$ then the system becomes a linear regression model $Y^\star = X\underline{\beta} + u$ where X is the stacked system of lagged values of Y. i.e. $X = [Y_{-1}, Y_{-2}, \ldots, Y_{-p}]$, where Y_{-j} is the stacked system of $Y_{i,t}$'s lagged by j time periods. This can be estimated in the standard way so that $p(\underline{\beta} \mid \Gamma_{-\underline{\beta}}, Y) \sim \mathcal{N}(\hat{\beta}, V_\beta)$ where

$$\hat{\beta} \quad = \quad V_\beta \left(S_\beta^{-1} \beta_0 + \frac{1}{\sigma^2} X' X \beta^{OLS} \right) \qquad V_\beta = \left[S_\beta^{-1} + \frac{1}{\sigma^2} X' X \right]^{-1}$$

(viii) Draw from $p(\frac{1}{\sigma^2} \mid \Gamma_{-\sigma^2}, Y)$ using the stacked system. $\frac{1}{\sigma^2}$ will have a Gamma distribution with parameters $\frac{\nu_\sigma + T \times N}{2}$ and $\frac{2}{\frac{\nu_\sigma}{S_\sigma} + [Y^\star - X\underline{\beta})'(Y^\star - X\underline{\beta})]}$.

For the initial values of $\alpha_{0,i}$ we choose the intercept term from an economy i level OLS regression of $Y_{i,t} = \alpha_i + \sum_l^L \beta_l Y_{t-l}$ and for the initial values of $\underline{\beta}$ we use the β_l values from an OLS regression using the stacked system. We set the initial value of c to zero and set $c_{-1} = c_{-2}$ to zero following Chan et al (2019).

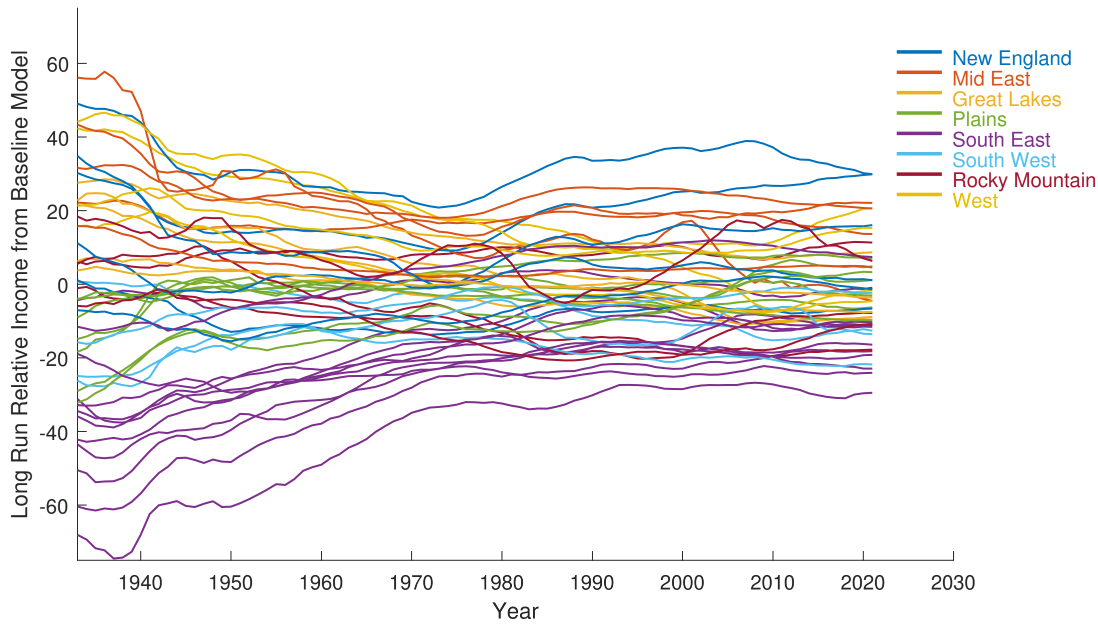
Model 2: The Hierarchical Model

The hierarchical model extends the baseline model by allowing the $\underline{\beta_i}$ coefficients to vary across economies by being randomly drawn from a common higher level distribution. This can be estimated in the same way as in the baseline model except for having a separate draw for each individual economy's $\underline{\beta_i}$ before there is a draw for $\overline{\beta}$ and $\Sigma_{\overline{\beta}}$ as in Chan et al (2019).

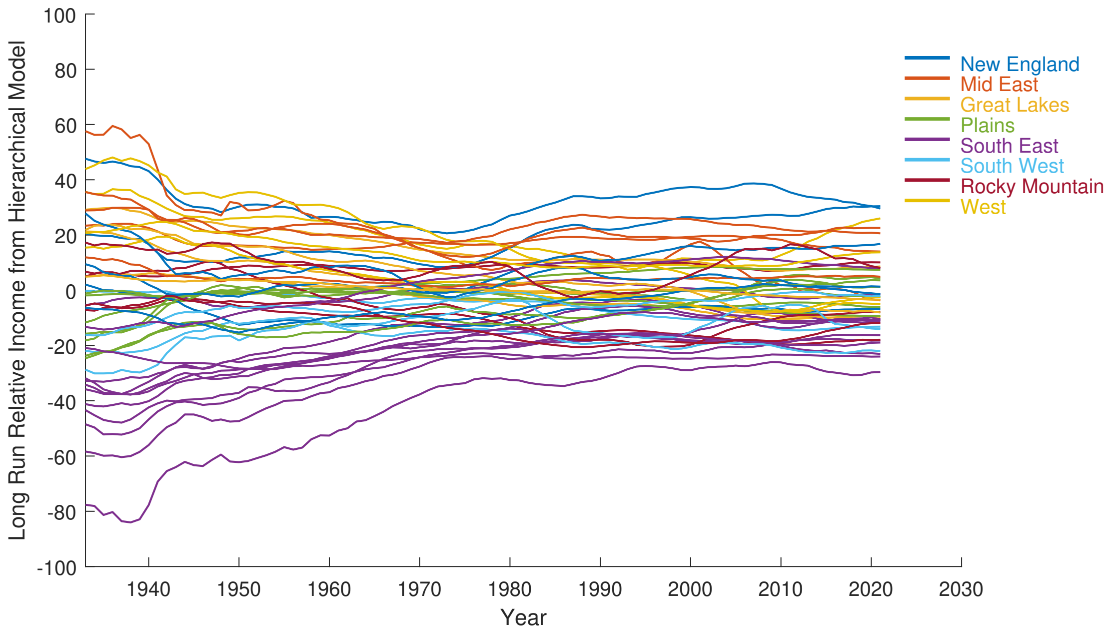
Appendix B

Colour Versions of US states' Growth Paths

Figure B1: Evolution of balanced growth path, US states 1933-1993



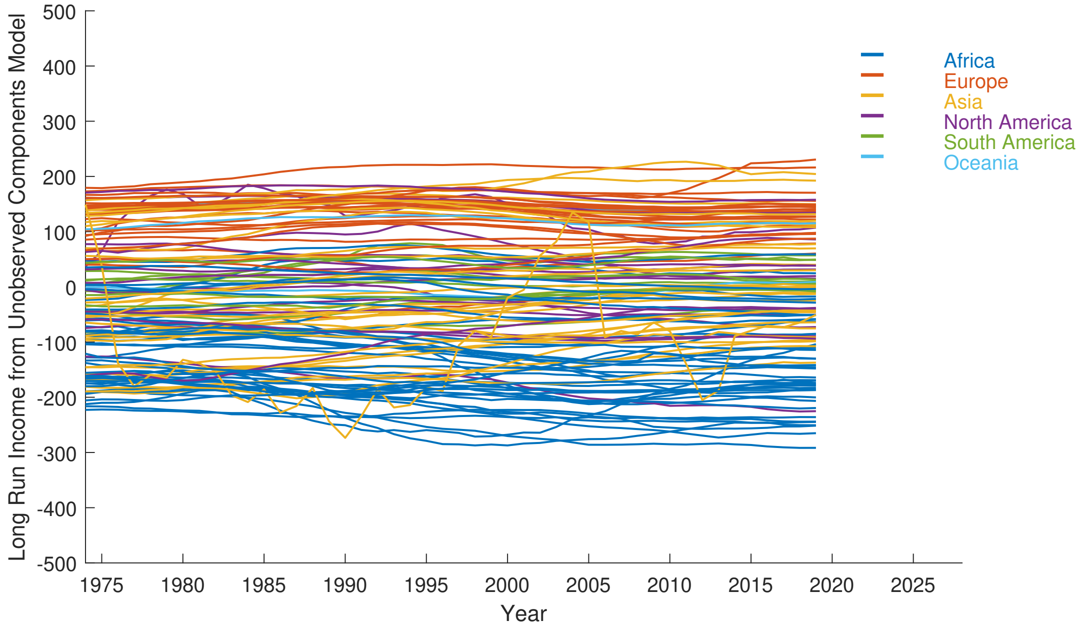
(a) Baseline model



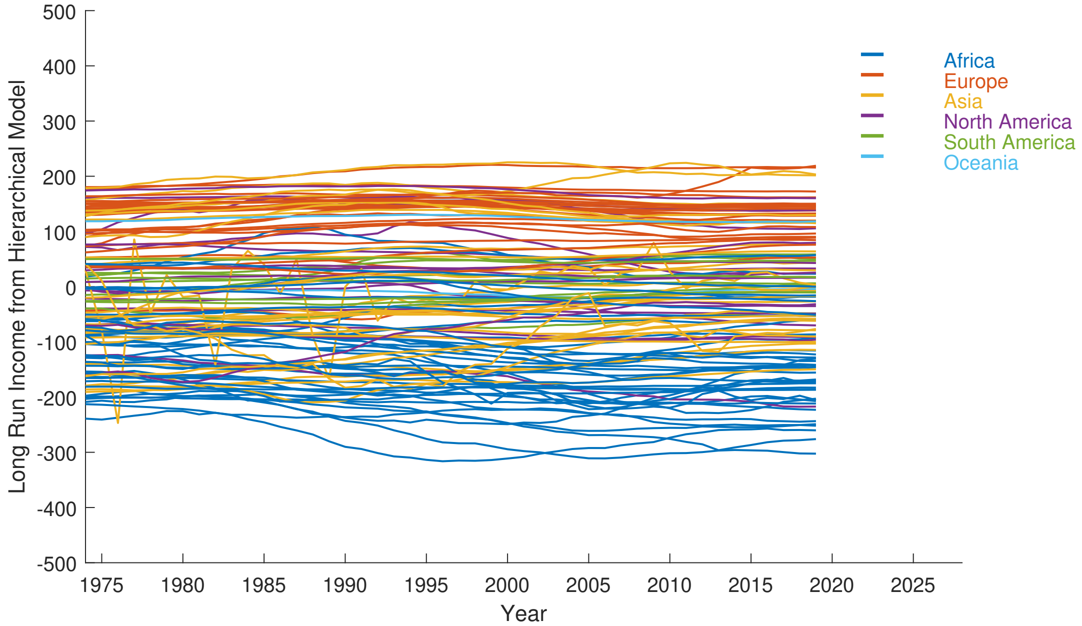
(b) Hierarchical model

Figure B1 plots the evolution of the balanced growth path for the 48 contiguous US states from 1933-2021. Panel a) is a colour version of Figure 2a and Panel b) is a colour version of Figure 3a. The colours highlight different US regions.

Figure B2: Evolution of the balanced growth path in the world economy



(a) Long run growth in the world economy 1974-2019 from the baseline model



(b) Long run growth in the world economy 1974-2019 from the hierarchical model

Figure B2 plots the evolution of the balanced growth path for the 123 countries in the world economy from 1974-2019. Panel a) is a colour version of Figure 5a and Panel b) is a colour version of Figure 6a. The colours highlight different US regions.

Shorter Forecast Horizon Results Tables

Table A1: Five years ahead forecast performances (RMSFE) For US states Dataset

S_c	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^2	10^3	10^5
S_ω									
Baseline Model									
10^{-5}	5.597	5.917	6.301	5.976	6.286	5.701	7.670	6.989	12.065
10^{-3}	4.931	4.923	4.879	4.947	5.022	4.991	4.958	5.367	16.713
10^{-2}	5.098	5.185	5.053	5.376	4.986	4.705	4.626	4.964	24.218
10^{-1}	**4.952**	**4.986**	**4.667**	**4.598**	**4.647**	**4.578**	**4.485**	**4.621**	**22.395**
1	5.134	4.676	5.060	4.747	4.908	4.753	4.591	4.630	33.179
10	4.856	4.893	4.956	4.805	4.880	4.833	4.716	5.013	31.743
10^2	4.577	4.621	4.570	4.607	4.559	4.584	4.633	5.183	24.357
10^3	4.537	4.530	4.546	4.521	4.524	4.533	4.624	5.571	29.702
10^4	4.531	4.548	4.548	4.531	4.540	4.525	4.632	6.896	67.832
Hierarchical Model									
10^{-5}	5.380	5.160	5.316	5.295	5.353	5.447	6.012	6.488	14.842
10^{-3}	4.747	4.788	4.777	4.734	4.973	4.866	5.066	5.753	15.569
10^{-2}	4.644	4.633	4.643	4.680	4.836	4.829	4.918	5.340	18.039
10^{-1}	4.566	4.562	4.588	4.585	4.575	4.545	4.629	5.158	21.847
1	**4.604**	**4.596**	**4.577**	**4.571**	**4.602**	**4.518**	**4.574**	**5.353**	**22.331**
10	4.590	4.597	4.591	4.586	4.592	4.597	4.572	5.256	21.881
10^2	4.579	4.589	4.569	4.574	4.578	4.574	4.616	5.133	24.495
10^3	4.566	4.557	4.569	4.584	4.581	4.568	4.653	5.259	31.958
10^5	5.750	5.907	6.141	5.521	5.512	5.610	6.189	10.672	68.776

Notes: Table A1 reports the Root-Mean-Squared Errors for the five period ahead out-of-sample forecast. The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.

Table A2: Five years ahead forecast performances (RMSFE) for the world economy dataset

$\boldsymbol{S_c}$	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^{2}	10^{3}	10^{5}
$\boldsymbol{S_\omega}$									
Baseline Model									
10^{-5}	25.58	23.22	23.08	23.17	23.94	22.89	21.79	20.47	32.16
10^{-3}	23.90	24.04	24.04	24.14	24.25	23.99	22.43	22.21	39.97
10^{-2}	23.51	23.45	23.50	23.47	23.46	23.29	21.81	21.09	47.67
10^{-1}	22.94	22.98	23.05	22.89	23.07	22.58	21.33	19.37	55.71
1	22.40	22.48	22.33	22.22	22.62	22.06	21.18	18.69	57.13
10	21.94	21.96	21.93	21.86	21.84	21.66	21.27	18.83	48.45
10^{2}	18.95	19.12	18.80	18.93	19.25	19.32	19.78	19.22	45.33
10^{3}	17.96	17.92	17.98	17.95	18.01	18.02	18.42	19.10	48.64
10^{5}	**18.08**	**18.16**	**17.90**	**18.00**	**18.06**	**18.01**	**18.99**	**21.97**	**120.57**
Hierarchical Model									
10^{-5}	21.86	21.95	22.01	21.94	21.81	22.03	21.52	20.64	40.48
10^{-3}	21.59	21.77	21.49	21.24	21.56	22.18	21.16	20.02	41.93
10^{-2}	21.40	21.79	21.32	21.45	22.07	21.94	21.91	21.21	46.01
10^{-1}	20.94	20.88	20.85	20.75	21.09	21.59	21.93	21.22	55.72
1	19.58	19.52	19.72	19.68	19.51	20.29	21.15	21.15	57.70
10	**19.47**	**19.20**	**19.54**	**19.75**	**19.23**	**19.44**	**19.65**	**20.64**	**55.35**
10^{2}	19.54	19.72	19.86	19.96	20.05	19.70	19.83	20.40	54.40
10^{3}	21.01	20.43	20.75	20.77	21.00	20.83	20.84	22.02	67.11
10^{5}	24.57	25.42	25.13	25.42	25.34	25.25	25.48	29.66	162.15

Notes: Table A2 reports the Root-Mean-Squared Errors for five period ahead out-of-sample forecast.The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.

Table A3: One year ahead forecast performances (RMSFE) for US states dataset

$\boldsymbol{S_c}$	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^2	10^3	10^5
$\boldsymbol{S_\omega}$									
Baseline Model									
10^{-5}	1.984	2.518	2.718	2.458	2.360	2.178	2.175	2.370	5.169
10^{-3}	1.830	1.838	1.844	1.842	1.862	1.883	1.861	1.880	4.265
10^{-2}	1.813	1.834	1.828	1.827	1.817	1.819	1.792	1.832	4.120
10^{-1}	1.787	1.762	1.743	1.716	1.736	1.767	1.743	1.739	4.578
1	1.816	1.754	1.762	1.692	1.790	1.747	1.724	1.713	5.030
10	1.669	1.670	1.676	1.676	1.681	1.688	1.683	1.703	4.318
10^2	**1.621**	**1.625**	**1.619**	**1.621**	**1.610**	**1.592**	**1.633**	**1.697**	**4.191**
10^3	1.625	1.623	1.615	1.617	1.610	1.615	1.650	1.763	4.780
10^5	1.608	1.627	1.631	1.622	1.627	1.623	1.749	2.507	8.728
Hierarchical Model									
10^{-5}	1.885	1.815	1.906	1.867	1.912	2.008	2.076	2.050	3.625
10^{-3}	1.776	1.749	1.776	1.745	1.811	1.868	1.883	1.956	3.691
10^{-2}	1.742	1.722	1.731	1.719	1.800	1.839	1.860	1.897	3.966
10^{-1}	1.711	1.738	1.701	1.733	1.700	1.789	1.799	1.842	4.369
1	1.695	1.713	1.698	1.686	1.692	1.712	1.756	1.828	4.620
10	1.685	1.687	1.686	1.678	1.684	1.692	1.713	1.801	4.510
10^2	1.677	1.671	1.673	1.669	1.672	1.676	1.713	1.795	4.533
10^3	**1.678**	**1.668**	**1.673**	**1.668**	**1.673**	**1.668**	**1.707**	**1.824**	**5.589**
10^4	1.729	1.784	1.795	1.802	1.749	1.784	1.900	2.721	10.931

Notes: Table A3 reports the Root-Mean-Squared Errors for the one period ahead out-of-sample forecast.The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.

Table A4: One year ahead forecast performances (RMSFE) for the world economy dataset

S_c	10^{-5}	10^{-3}	10^{-2}	10^{-1}	1	10	10^2	10^3	10^5
S_ω									
Baseline Model									
10^{-5}	6.75	7.28	7.21	7.08	8.39	6.84	6.66	6.76	9.44
10^{-3}	6.52	6.51	6.54	6.54	6.59	6.63	6.59	6.49	9.11
10^{-2}	6.45	6.46	6.45	6.49	6.54	6.55	6.54	6.29	9.25
10^{-1}	6.40	6.40	6.42	6.43	6.50	6.46	6.57	6.20	9.93
1	6.36	6.37	6.36	6.38	6.43	6.44	6.60	6.18	10.24
10	6.38	6.40	6.38	6.43	6.45	6.49	6.58	6.11	9.04
10^2	6.15	6.16	6.12	6.16	6.19	6.28	6.25	6.17	8.42
10^3	5.91	5.93	5.92	5.93	5.95	5.98	6.05	6.11	9.51
10^5	**5.90**	**5.90**	**5.87**	**5.90**	**5.91**	**5.94**	**6.08**	**6.74**	**16.89**
Hierarchical Model									
10^{-5}	6.39	6.38	6.38	6.38	6.36	6.46	6.45	6.33	8.46
10^{-3}	6.35	6.34	6.34	6.33	6.36	6.53	6.43	6.25	8.65
10^{-2}	6.28	6.30	6.30	6.30	6.36	6.51	6.47	6.36	9.07
10^{-1}	6.24	6.20	6.20	6.19	6.28	6.44	6.45	6.34	9.88
1	6.06	6.07	6.06	6.05	6.11	6.26	6.46	6.43	10.08
10	5.99	5.99	6.00	6.03	6.00	6.06	6.25	6.45	10.20
10^2	**5.96**	**5.99**	**6.02**	**6.02**	**6.03**	**6.00**	**6.13**	**6.23**	**10.13**
10^3	6.11	6.00	6.08	6.08	6.12	6.07	6.12	6.31	11.85
10^5	6.41	6.50	6.37	6.52	6.51	6.51	6.66	7.33	25.16

Notes: Table A4 reports the Root-Mean-Squared Errors for the one period ahead out-of-sample forecast.The estimates are produced recursively using 20,000 draws from the posterior and the mean parameter estimates are used to generate the forecast.