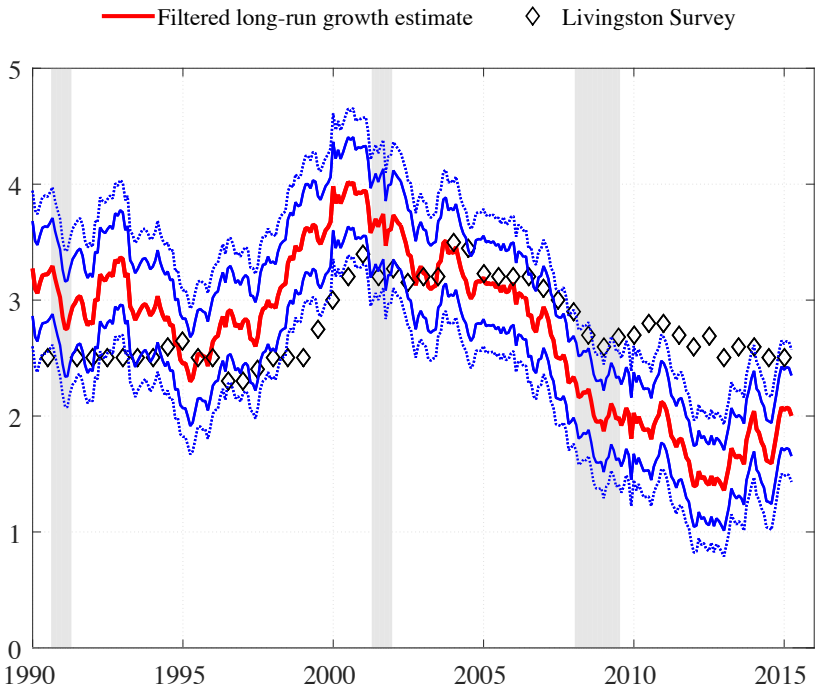
Online Appendix to "Tracking the Slowdown in Long-Run GDP Growth"

by Juan Antolin-Diaz, Thomas Drechsel and Ivan Petrella

Contents

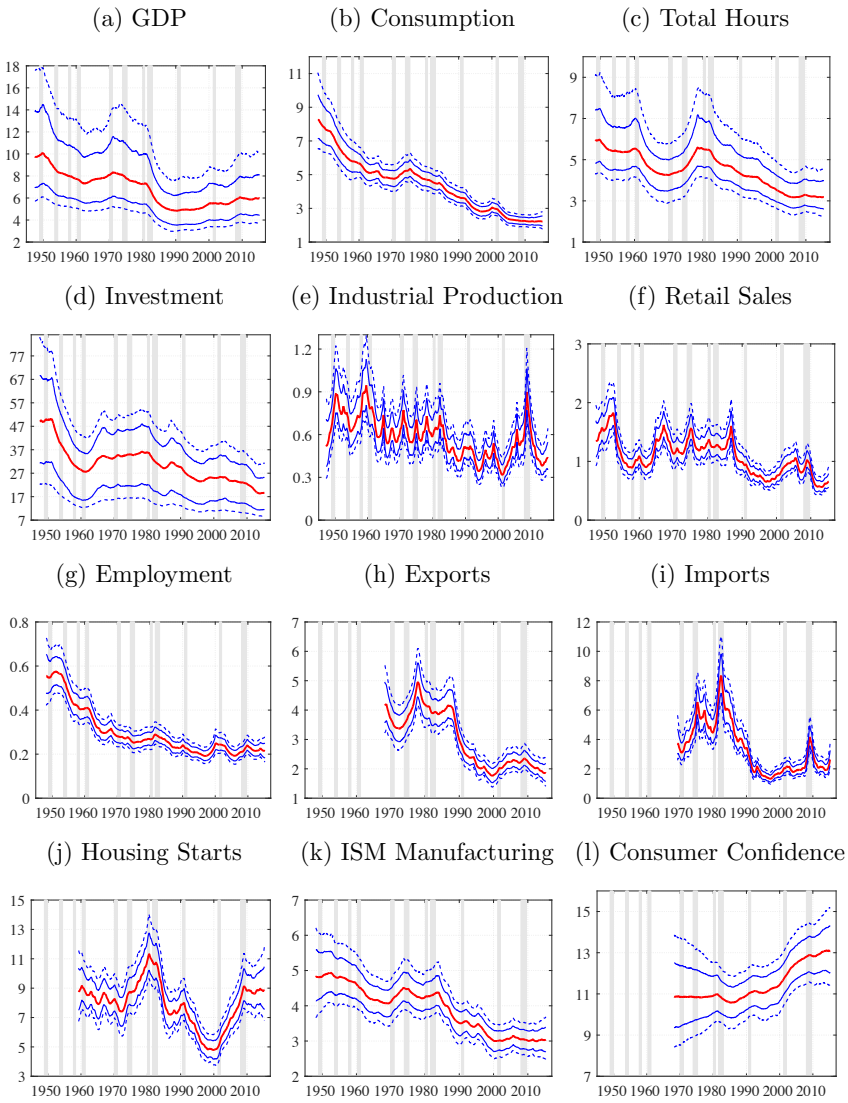
A Additional Figures

Figure A.1: Filtered estimate of long-run growth



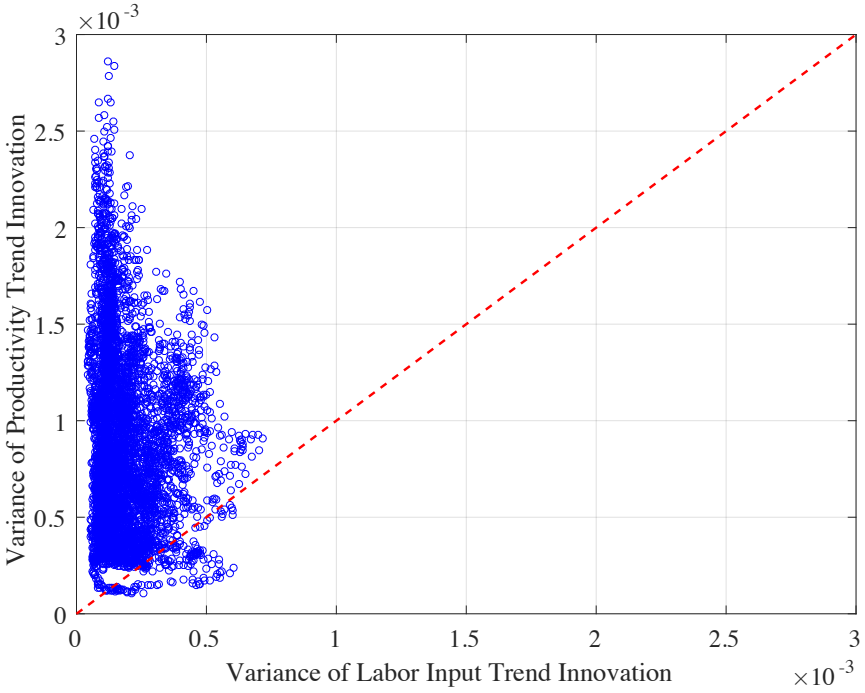
Note: The solid red line is the filtered estimate of the long-run GDP growth rate, $\hat{g}_{t|t}$, using the vintage of National Accounts available as of March 2015. The solid and dotted blue lines capture the corresponding 68% and 90% posterior bands. The black diamonds represent the real-time mean forecast from the Livingston Survey of Professional Forecasters of the average GDP growth rate for the subsequent 10 years.

Figure A.2: Stochastic Volatility of Selected Idiosyncratic Components



Note: Each panel presents the median (solid red), the 68% and the 90% (solid and dashed blue) posterior credible intervals of the volatility of the idiosyncratic component of selected variables. Shaded areas represent NBER recessions. Similar charts for other variables are available upon request.

Figure A.3: Joint Posterior Distribution of Growth Component Innovation Variances



Note: The figure plots 5,000 draws of the joint posterior distribution of the variances of innovations to the labor productivity and hours component. The dashed red line is the 45°line. Under the equal-variance prior the draws would be equally distributed above and below this line. The fact that the bulk of draws lie above indicates that changes in long-run labor productivity drive the variation in long-run output.

B Full Results of Structural Break Tests

B.1 Nyblom Test

Table B.1 reports the result for the Nyblom (1989) test applied to US real GDP growth, as described in Hansen (1992). The sample starts is 1947:Q2. The specification is $y_t = \mu + \rho_1 y_{t-1} + \rho_2 y_{t-1} + \sigma\epsilon_t$, where y_t is real GDP growth. For each parameter of the specification, the null hypothesis is that the respective parameter is constant.

Table B.1:
Test Results of Nyblom Test

	L_c	
	AR(1)	AR(2)
μ	0.518*	0.473*
ρ_1	0.367	0.331
ρ_2		0.094
σ^2	0.843***	0.838***
Joint L_c	2.145***	2.294***

Notes: Results are obtained using Nyblom's L test as described in Hansen (1992). *, ** and *** indicate significance at the 10%, 5% and 1% level.

B.2 Bai and Perron Test

Table B.2 reports the result for the Bai and Perron (1998) test applied to US real GDP growth for the sample starting in 1947:Q2. We apply the $SupF_T(k)$ test for the null hypothesis of no break against the alternatives of $k = 1, 2$, or 3 breaks. Secondly, the test $SupF_T(k+1|k)$ tests the null of k breaks against the alternative of $k+1$ breaks. Finally, the U_dmax statistic tests the null of absence of break against the alternative of an unknown number of breaks. The null hypothesis of no breaks is rejected against the alternative of one break at the 10% level. The null is not rejected against the alternative of two or three breaks. Furthermore, the null hypothesis of one break against two breaks, or the null of only two against three breaks is not rejected. The final test confirms the conclusion that there is some evidence in favor of at least one break, with the null rejected against an unknown number of breaks at the 10% level. The most likely break is identified to have happened in the second quarter of 2000.

Table B.2:
Test Results of Bai-Perron Test

	Sample 1947-2015
	$SupF_T(k)$
$k = 1$	8.379*
	[2000:Q2]
$k = 2$	4.194
	[1968:Q2; 2000:Q2]
$k = 3$	4.337
	[1969:Q1; 1982:Q4; 2000:Q2]
	$SupF_T(k\|k-1)$
$k = 2$	1.109
$k = 3$	2.398
U_dmax	8.379*

Note: Results are obtained using the Bai and Perron (1998) methodology. Dates in square brackets are the most likely break date(s) for each of the specifications. * indicates significance at the 10% level.

C Monte Carlo Evidence

C.1 Setup for Monte Carlo simulations

To assess the performance of our model in the presence of potentially relevant types of misspecification, we carry out a variety of Monte Carlo experiments. In each experiment, we simulate a large number of data sets which are generated from the model under known parameter values, and estimate our model repeatedly over these data sets. This appendix presents the results for two sets of such experiments, which are designed to explore the robustness of crucial assumptions made in the paper.

- In Section C.2 we examine whether the random walk assumption for the time-varying parameters is robust to a different type of structural change. In particular, we verify how the model performs if the underlying long-run growth rate of GDP features one or multiple discrete breaks rather than gradual change. We also estimate our baseline model on data which is generated with a constant instead of a time-varying long-run growth rate of real GDP growth. Furthermore, we repeat this type of experiment for discrete breaks rather than gradual change in the volatilities of both the common factor and the idiosyncratic terms.

- In Section C.3 we explore the robustness of our model to the presence of (unmodeled) change in the long-run growth rate of other series. We entertain the possibility that such unmodeled trends are either independent of the change in the long-run growth real GDP growth or that some series share the trend of GDP. We also verify robustness to both of these types of misspecification simultaneously.

While our simulations feature selected types of misspecification, we aim to ensure a realistic environment for the correctly specified parts of the model. In particular, we set the values of the parameters to their estimated posterior median of the US results. We then take draws for the random disturbances and generate a sample of the vector of 28 observables using equations (1) to (7), and generate 800 periods of data, which corresponds to the monthly sample size in our US application.[1] The four quarterly series are generated by simulating the underlying monthly series and then introducing missing observations by (backwards) applying the polynomial in equation (9). We then estimate the model using the settings described in the paper. The number of simulations (repeatedly drawn samples) per given experiment is set to 100.[2]

[1] While we argue in the paper that the random walk assumption for the estimation of the time-varying parameters is innocuous, it can be problematic to simulate data from parameters that follow random walks. Although we would like the parameters to drift in a non-stationary fashion, i.e. to generate realistic patterns of time-varying volatility, data sets generated from "explosive" processes feature unrealistic properties. To address this issue in the Monte Carlo simulations we discard and regenerate random walks when they drift across a fix threshold. For example, we do not allow the range of (demeaned) time-varying intercept of a given series to exceed the range of its cyclical component.

[2] In certain cases, convergence of the algorithm takes longer in the presence of misspecification, which required us to increase the number of draws of the Gibbs sampler, and thus limited the amount of repetitions that was feasible for a given experiment.

C.2 Results: Sensitivity of random walk specification

The goal of this first set of Monte Carlo experiments is to explore the sensitivity of our modeling choice with respect to the random walk specification of the time-varying parameters. The details about how we justify this modeling assumption can be found in Section 3.1 of the paper. In particular, we aim here to verify whether the model is robust in a context in which there are changes in the long-run growth rate of real GDP growth and in the volatility of business cycles, but these changes occur as discrete breaks rather then as gradual change. Figures C.1 to C.4 present the results of four Monte Carlo experiments.

In the first experiment, the simulated counterpart of real GDP growth features a mean growth rate that is constant but subject to a level shift in the middle of the sample. In Figure C.1, panels (a) and (b), we plot the actual growth rate underlying the data-generating process together with one and two standard deviation percentiles of the 100 simulations of the posterior median, both for the filtered and smoothed estimate. It is reassuring to see that the random walk process "learns" relatively quickly about the underlying change, even in the case of a discrete jump. Panel (c) displays the true, together with the posterior estimate of the common factor for one of the 100 Monte Carlo draws. Panel (d) provides a scatter plot of the true vs. estimated stochastic volatilities. Both pictures show that the models performs well at capturing the simulated objects.

In the second experiment, we repeat the same exercise in the presence of two discrete breaks in the real GDP growth rate. The results are visible in Figure C.2, which tells a very similar story to the first experiment. We omit panels for factor and the stochastic volatility estimates, as they are very similar to the first experiment.

In the third experiment, we verify the consequences of estimating our model in an in environment in which the parameters which we specify as time-varying are in fact constant in the data-generating process. The results, displayed in Figure C.3, confirm that the random walk assumption appears to be entirely innocuous in this setting. Both the long-run GDP growth rate (smoothed and filtered), as well as the volatility of the factor are estimated to be constant, with relatively high precision. In addition, similar to the first experiment, the estimate of the common factor is very precise.

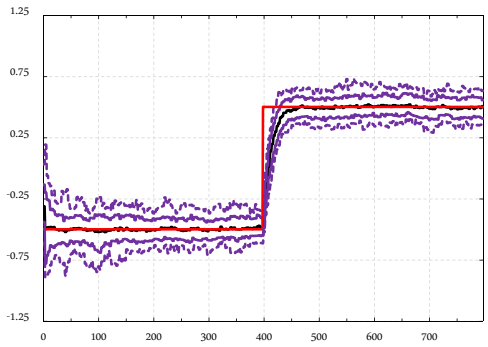
Finally, in the fourth experiment, we again keep the long-run growth rate of real GDP constant but this time introduce a discrete shift in the volatilities of both the common factor and the idiosyncratic terms of all series in the middle of generated data sample. Reassuringly, the shift in the volatilities is well captures in the estimation and does not spill over to the estimate of the long-run growth rate of real GDP.

In conclusion from these experiments, the random walk assumption appears to be flexible enough to accommodate structural change that occurs in discrete steps rather than gradually. This underpins our conclusions about the apparent gradual changes in the long-run growth of the US economy described in the paper.

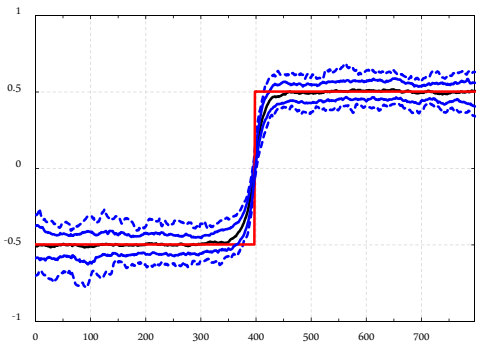
Figure C.1: Simulation Results I

Data-generating process (DGP) with one discrete break in long-run real GDP growth

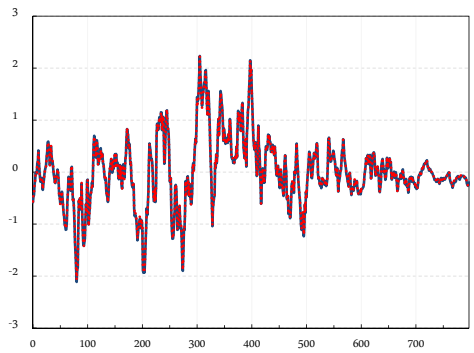
(a) True vs. Estimated Trend (Filtered)



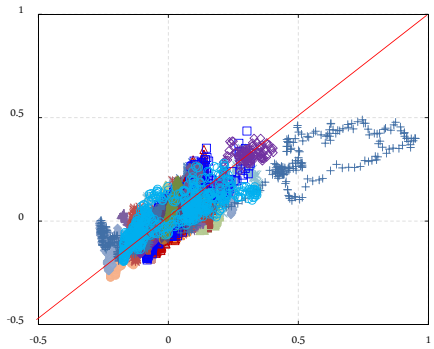
(b) True vs. Estimated Trend (Smoothed)



(c) True vs. Estimated Factor



(d) True vs. Estimated Volatilities of $\mathbf{u}_t$

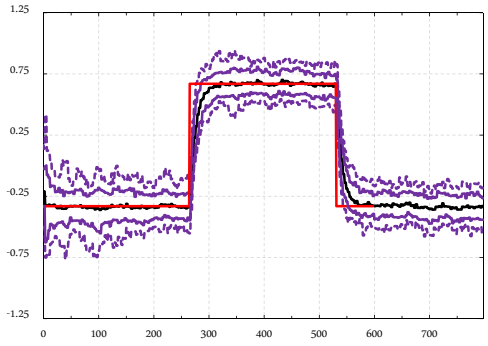


Note: The DGP features a discrete break in the trend of GDP growth occurring in the middle of the sample, as well as stochastic volatility. The sample size is $n = 28$ and $T = 800$, which mimics our US data set. The estimation procedure is the fully specified model as defined by equations (1)-(7) in the text. We carry out 100 simulations drawing from the DGP. Panel (a) presents the long-run growth component as estimated by the Kalman filter, plotted against the actual long-run growth rate generated from the DGP. The corresponding figure for the smoothed estimate is given in panel (b). In both panels, the median (black) as well the 68th (solid) and 90th (dashed) percentile of the 100 simulated outcomes are shown in blue/purple. Panel (c) displays the factor generated by the the DGP (red) and its smoothed estimate (blue) for one draw. Panel (d) provides evidence on the accuracy of the estimation of the SV of the idiosyncratic terms, by plotting the volatilities from the DGP against the estimates for the 24 monthly indicators. Both are normalized by subtracting the average volatility.

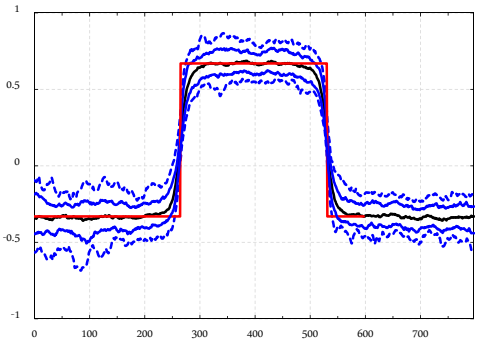
Figure C.2: Simulation Results II

Data-generating process (DGP) with two discrete breaks in in long-run real GDP growth

(a) True vs. Estimated Trend (Filtered)



(b) True vs. Estimated Trend (Smoothed)

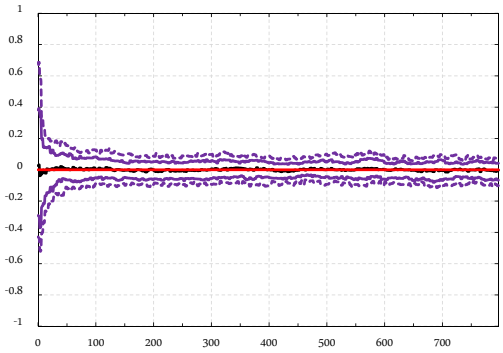


Note: The simulation setup is equivalent to the one in Figure C.1 but features *two* discrete breaks in the trend at 1/3 and 2/3 of the sample. Again, we show the filtered as well as the smoothed trend median estimates and the corresponding 68th and 90th percentiles of the 100 simulated estimates of these objects. Panels (c) and (d) are omitted as they are very similar to Figure C.1.

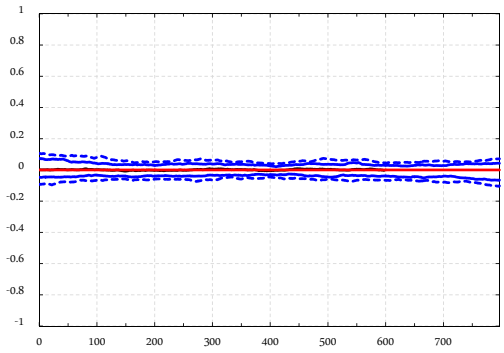
Figure C.3: Simulation Results III

Data-generating process (DGP) without in changes long-run real GDP growth and without SV

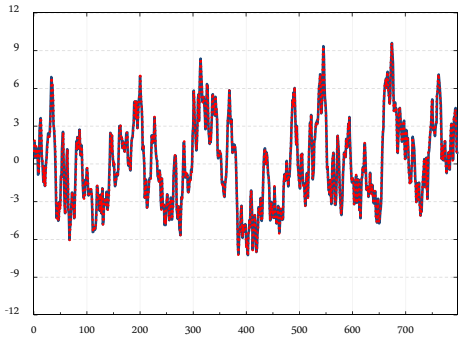
(a) True vs. Estimated Trend (Filtered)



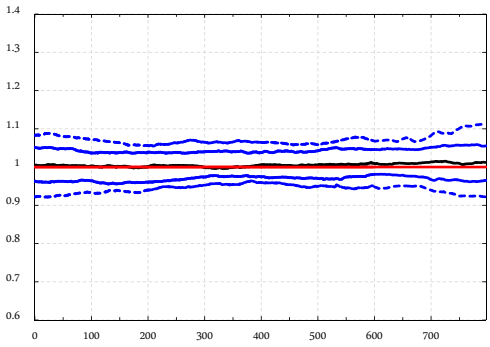
(b) True vs. Estimated Trend (Smoothed)



(c) True vs. Estimated Factor



(d) True vs. Estimated Volatility of Factor

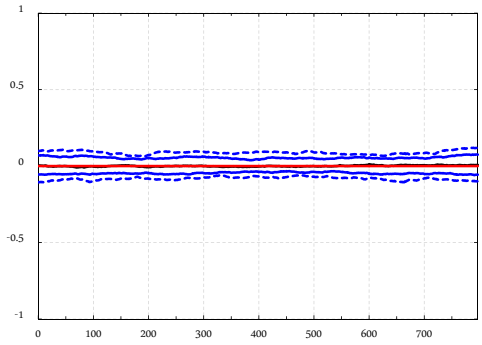


Note: The DGP is the baseline model without trend in GDP growth and without stochastic volatility. The estimation procedure is the fully specified model as explained in the description of Figure C.1. Again, we plot the filtered and smoothed median estimates of the long-run growth rate with 68th and 90th percentiles of the 100 simulated estimates in panels (a) and (b). Panel (c) presents a comparison of the estimated factor and its DGP counterpart for one Monte Carlo draw. Panel (d) in similar to (b), but for the volatility of the common factor.

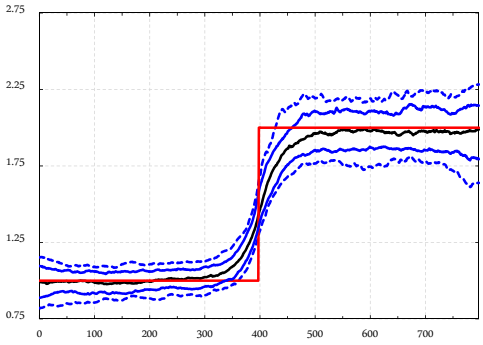
Figure C.4: Simulation Results IV

Data-generating process (DGP) with discrete break in volatility

(a) True vs. Estimated Trend (Smoothed)



(b) True vs. Estimated Volatility of Factor



Note: The DGP does not feature any changes in the trend of GDP growth, but one discrete break in the volatility of the common factor. As in Figures C.1-C.3, the estimation procedure is based on the fully specified mode. Panel (a) displays the smoothed posterior median estimate of the trend component of GDP growth, with 68th and 90th percentiles of the 100 simulations shown as solid and dashed blue lines, respectively. Panel (b) displays the posterior median estimate of the volatility of the common factor (black), with the corresponding percentiles.

C.3 Results: Sensitivity to confounding time-variation

Our model can flexibly accommodate time-varying intercepts in all or a subset of the series contained in our data panel. Given our interest in tracking real GDP growth, we restrict our baseline model to feature a trend in GDP only (shared by consumption) and argue that such unmodeled time-variation is picked up by the idiosyncratic components, which we allow to be persistent. Details about this discussion are contained in Section 3.2 of the paper. The goal of this second set of Monte Carlo experiments is to verify how robust our model is in a setting where time-varying intercepts are indeed present in the data-generating process but not modeled explicitly in the estimation. Figures C.5 to C.7 present the results of three Monte Carlo experiments in which such "confounding trends" are added when generating the data.[3]

In the first experiment, the misspecification arises from the fact that our model explicitly specifies a time-varying mean in the GDP equation only, while the data is generated such that the first 18 series of the panel all feature independent non-stationary means.[4] Figure C.5 presents the estimation results in this setup. Panel (a) shows the percentiles of the deviations of the estimated from the actual real GDP growth rates over the 100 simulations (repeated draws from the DGP). The percentiles are centered relatively tightly around zero, meaning that the trend estimates with 68 and 90% smallest deviations are relatively similar to the original trend process. To illustrate this further, panels (b), (c) and (d) display more detailed results for one of the 100 Monte Carlo simulations, labeled "Median Simulation". This is selected by ordering the outcomes of all repeated samples by the distance of squared deviation of the estimated from the simulated GDP trend and then selecting the median. This essentially means that 49% of the simulations had larger, and 50% smaller deviations than the simulation displayed. The panels plot actual against estimated (black/red) long-run real GDP growth rate, common factor and factor volatility, respectively. In the case of the long-run growth rate the posterior credible intervals are added in blue. These results reveal that in a typical (median) outcome for this type of specification, the model performs well at capturing these objects. Most importantly, the "true" long-run growth rate is contained within the posterior bands throughout the entire estimation sample.

In the second experiment, the data-generating process features a single time-varying mean which is present in the first 6 series, whereas we still only specify it in the first series for the estimation.[5] The results for this experiment are shown in Figure C.6. The panels here are similar to Figure C.5. While the deviations in panel (a)

[3]For simplicity we assume that the estimated model in this section is the one with a trend in GDP only, i.e. $\mathbf{B} = 1$.

[4]Formally, in the DGP $dim(\mathbf{a}_t) = 18$ and $\mathbf{B} = \mathbf{I}_{18}$, while the model for estimation is specified by $\mathbf{a}_t = g_t$ and $\mathbf{B} = 1$. We assume that the remaining 10 of the 28 series are stationary, which mimics the presence of the surveys in our data set.

[5]In our notation this means that in the DGP we have $\mathbf{a}_t = g_t$ and $\mathbf{B} = \mathbf{1}_{6\times 1}$, while the model for estimation is specified by $\mathbf{a}_t = g_t$ and $\mathbf{B} = 1$. We choose 6 series so that both quarterly and monthly variables are affected by the misspecification.

are slightly larger than for the previous figure, indicating that common unmodeled trends are somewhat more challenging to pick up than independent ones, the overall message remains the qualitatively similar. In particular, the results for the "Median Experiment", displayed in panels (b) to (d), are reassuring in that the estimate tracks their data counterpart closely.

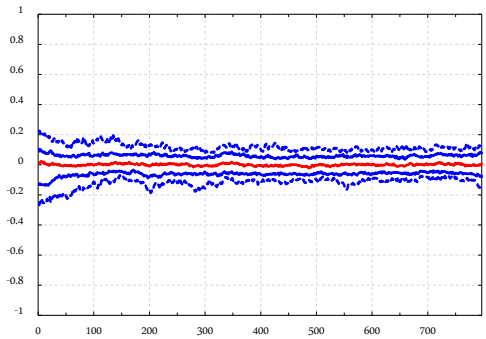
The third experiment introduces both types of misspecification simultaneously, i.e. independent time-varying means in series 1-18 and an additional shared time-varying component in series 1-6. The results are presented in Figure C.7. The take-aways are similar to the previous figures, even in the presence of this heavy type of misspecification.

Overall, these simulation experiments confirm our intuition that the estimate of the time-varying mean of interest is not affected by low frequency movements present in other series that are not explicitly modeled. Despite the extremely unfavorable assumption of a large amount of additional time-variation, the long-run growth rate of real GDP is tracked very well in all settings considered.

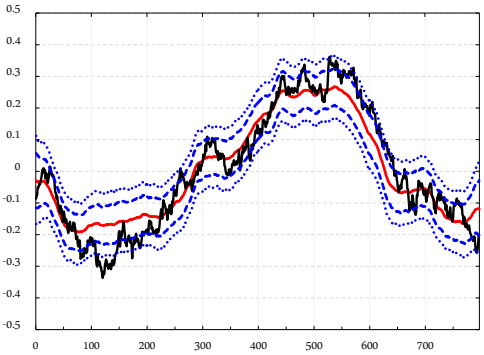
Figure C.5: Simulation Results V

Data-generating process (DGP) with independent unmodeled trends in other series

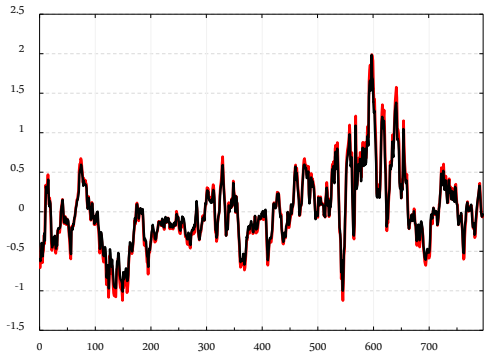
(a) True vs. Est. Trend - Deviation Percentiles



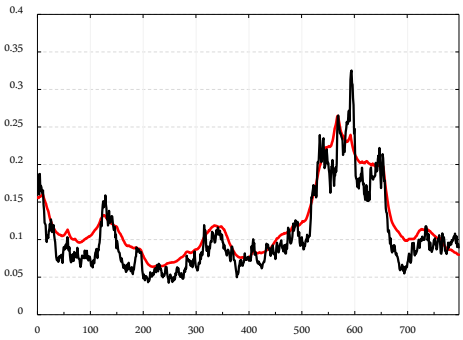
(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)



(d) True vs. Est. Vol (Median Simulation)

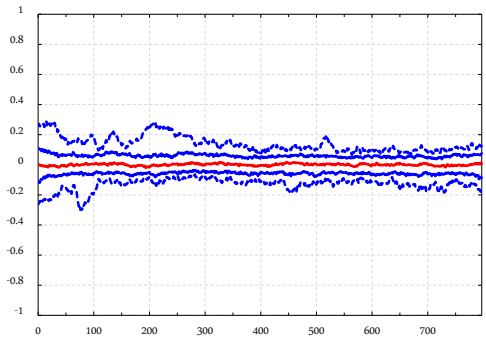


Note: The DGP features independent time-varying means in series 1-18. The sample size is $n = 28$ and $T = 800$, which mimics our US data set. The estimation procedure is the fully specified model as defined by equations (1)-(7) in the text, with a time-varying mean only specified for the real GDP growth equation. We carry out a Monte Carlo simulation with 100 samples repeatedly drawn from the DGP. Panel (a) presents the median (red), as well as the 68 and 90% bands (blue) of the deviation of the estimated long-run growth rate from its actual data counterpart over 100 simulated outcomes. Panel (b) shows the true (black) together with the posterior median estimate (red) of the long-run growth rate of real GDP. The 68% (solid blue) and 90% (dashed blue) posterior credible intervals are also plotted. Panels (c) and (d) plot the median estimate (red) against true (black) common factor and its stochastic volatility.

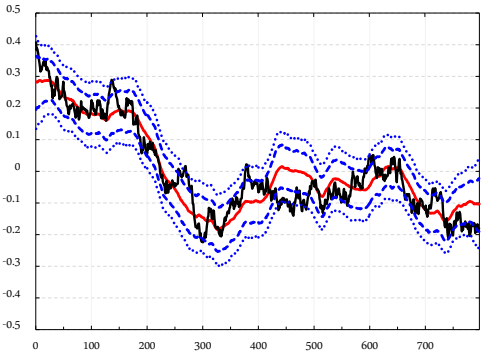
Figure C.6: Simulation Results VI

Data-generating process (DGP) with shared unmodeled trends in other series

(a) True vs. Est. Trend - Deviation Percentiles



(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)

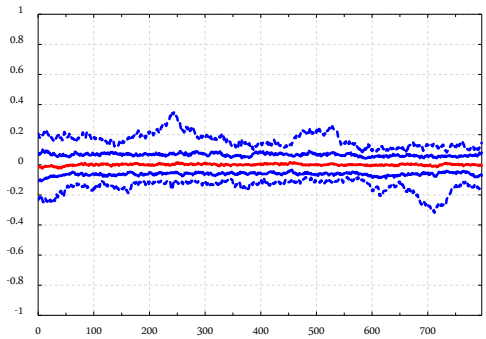
(d) True vs. Est. Vol (Median Simulation)

Note: The DGP features a common time-varying in series 1-6, while the estimation specifies this stochastic trend only in the equation for real GDP growth. The rest of the setup of the simulations, as well as the structure of the panels are similar to Figure C.5.

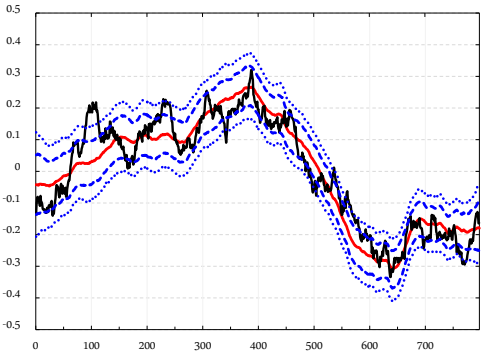
Figure C.7: Simulation Results VII

Data-generating process (DGP) with both independent and shared unmodeled trends in other series

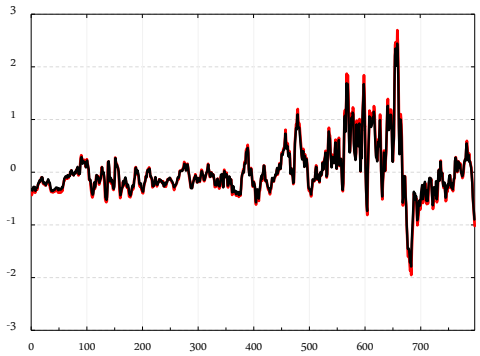
(a) True vs. Est. Trend - Deviation Percentiles



(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)



(d) True vs. Est. Vol (Median Simulation)

The DGP features both independent time-varying components in series 1-18 as well as a common time-varying in series 1-6, while the estimation specifies a stochastic trend only in the equation for real GDP growth. The rest of the setup of the simulations, as well as the structure of the panels are similar to Figure C.5

D Details on Estimation Procedure

D.1 Construction of the State Space System

For expositional clarity, we focus on the baseline case with $\mathbf{B} = 1$ and $\mathbf{a}_t = a_t$ here, so that $m = r = 1$. Recall that in our main specification we choose the order of the polynomials in equations (3) and (4) to be $p = 2$ and $q = 2$, respectively. Let the $n \times 1$ vector $\tilde{\mathbf{y}}_t$, which contains n_q de-meaned quarterly and n_m de-meaned monthly variables (i.e. $n = n_q + n_m$), be defined as

$$\tilde{\mathbf{y}}_t = \begin{bmatrix} y^q_{1,t} \\ \vdots \\ y^q_{n_q,t} \\ y^m_{1,t} - \rho^m_{1,1} y^m_{1,t-1} - \rho^m_{1,2} y^m_{1,t-2} \\ \vdots \\ y^m_{n_m,t} - \rho^m_{n_m,1} y^m_{n_m,t-1} - \rho^m_{n_m,2} y^m_{n_m,t-2} \end{bmatrix},$$

so that the system is written out in terms of the *quasi-differences* of the monthly indicators. Given this re-defined vector of observables, we cast our model into the following state space form:

$$\begin{aligned} \tilde{\mathbf{y}}_\mathbf{t} &= \mathbf{H}\mathbf{X}_t + \tilde{\boldsymbol{\eta}}_\mathbf{t}, & \tilde{\boldsymbol{\eta}}_\mathbf{t} &\sim N(0, \tilde{\mathbf{R}}_\mathbf{t}) \\ \mathbf{X}_t &= \mathbf{F}\mathbf{X}_{t-1} + \mathbf{e}_t, & \mathbf{e}_t &\sim N(0, \mathbf{Q}_t) \end{aligned}$$

where the state vector is defined as $\mathbf{X}'_t = [a_t, \ldots, a_{t-4}, f_t, \ldots, f_{t-4}, \mathbf{u}^{q\prime}_t, \ldots, \mathbf{u}^{q}_{t-4}{}']$. Setting $\lambda_1 = 1$ for identification, the matrices of parameters $\mathbf{H}$ and $\mathbf{F}$, are then constructed as shown below:

$$\mathbf{H} = \left[\begin{array}{c|c|c} \mathbf{H}_a & \begin{matrix} \mathbf{H}_{\lambda_q} \\ \mathbf{H}_{\lambda_m} \end{matrix} & \mathbf{H}_u \end{array} \right],$$

where the respective blocks of $\mathbf{H}$ are defined as

$$\mathbf{H}_a = \left[\begin{array}{ccccc} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ \hline \multicolumn{5}{c}{\mathbf{0}_{(n-1)\times 5}} \end{array} \right], \quad \mathbf{H}_{\lambda_q} = \begin{bmatrix} 1 & \lambda_2 & ... & \lambda_{n_q} \end{bmatrix}' \times \begin{bmatrix} 1/3 & 2/3 & 1 & 2/3 & 1/3 \end{bmatrix},$$

$$\mathbf{H}_{\lambda_m} = \left[\begin{array}{c|c} \lambda_{n_q+1} - \lambda_{n_q+1}\rho^m_{1,1} - \lambda_{n_q+1}\rho^m_{1,2} & \mathbf{0}_{1\times 4} \\ \vdots & \vdots \\ \lambda_n - \lambda_n \rho^m_{n_m,1} - \lambda_n \rho^m_{n_m,2} & \mathbf{0}_{1\times 4} \end{array} \right]$$

$$\mathbf{H}_u = \begin{bmatrix} \bar{\mathbf{H}}_u \\ \mathbf{0}_{n_m \times 5} \end{bmatrix}, \qquad \bar{\mathbf{H}}_u = \mathbf{1}_{n_q \times 1} \times \begin{bmatrix} 1/3 & 2/3 & 1 & 2/3 & 1/3 \end{bmatrix},$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0} & & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_2 & & & \\ \vdots & & \mathbf{F}_{2+1} & & \vdots \\ \vdots & & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{F}_{2+n_q} \end{bmatrix},$$

where the respective blocks of $\mathbf{F}$ are defined as

$$\mathbf{F}_1 = \left[\begin{array}{cc} 1 & \mathbf{0}_{1\times 4} \\ \hline \mathbf{I}_4 & \mathbf{0}_{4\times 1} \end{array}\right] \qquad \mathbf{F}_2 = \left[\begin{array}{ccc} \phi_1 & \phi_2 & \mathbf{0}_{1\times 3} \\ \hline \mathbf{I}_4 & & \mathbf{0}_{4\times 1} \end{array}\right] \qquad \mathbf{F}_{2+j} = \left[\begin{array}{ccc} \rho^q_{j,1} & \rho^q_{j,2} & \mathbf{0}_{1\times 3} \\ \hline \mathbf{I}_4 & & \mathbf{0}_{4\times 1} \end{array}\right]$$

for $j = 1, ..., n_q$.

The error terms are denoted as

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_t &= [\mathbf{0}_{1\times n_q}, \tilde{\boldsymbol{\eta}}^{\boldsymbol{m\prime}}_{\boldsymbol{t}}]' \\ \mathbf{e}_t &= \begin{bmatrix} v_{a_t} & \mathbf{0}_{4\times 1} & \epsilon_t & \mathbf{0}_{4\times 1} & \eta_{1,t} & \mathbf{0}_{4\times 1} & \dots & \eta_{n_q,t} & \mathbf{0}_{4\times 1} \end{bmatrix}', \end{aligned}$$

with covariance matrices

$$\tilde{\mathbf{R}}_{\mathbf{t}} = \begin{bmatrix} \mathbf{0}_{n_q \times n_q} & \mathbf{0}_{n_q \times n_m} \\ \mathbf{0}_{n_m \times n_q} & \mathbf{R}_t \end{bmatrix},$$

where $\mathbf{R}_t = diag(\sigma^2_{\eta^m_{1,t}}, \dots, \sigma^2_{\eta^m_{nm,t}})$ and

$$\mathbf{Q}_t = diag(\omega^2_a, \mathbf{0}_{1\times 4}, \sigma^2_{\epsilon,t}, \mathbf{0}_{1\times 4}, \sigma^2_{\eta^q_{1,t}}, \mathbf{0}_{1\times 4}, \dots, \sigma^2_{\eta^q_{n_q,t}}, \mathbf{0}_{1\times 4}).$$

D.2 Details of the Gibbs Sampler

For ease of notation, we again restrict this description to the case of one time-varying mean specified as $m = r = 1$, $\mathbf{B} = 1$ and $\mathbf{a}_t = a_t$. Let $\boldsymbol{\theta} \equiv \{\boldsymbol{\lambda}, \boldsymbol{\Phi}, \boldsymbol{\rho}, \omega_a, \omega_\varepsilon, \omega_{\eta_1}, \ldots, \omega_{\eta_n}\}$ be a vector that collects the underlying parameters, where $\boldsymbol{\Phi}$ and $\boldsymbol{\rho}$ contain the parameters for factor and idiosyncratic components respectively. The model is estimated using a Markov Chain Monte Carlo (MCMC) Gibbs sampling algorithm in which conditional draws of the latent variables, $\{a_t, f_t\}_{t=1}^T$, the parameters, $\boldsymbol{\theta}$, and the stochastic volatilities, $\{\sigma_{\varepsilon,t}, \sigma_{\eta_i,t}\}_{t=1}^T$ are obtained sequentially. The algorithm has a block structure composed of the following steps.

0. Initialization

The model parameters are initialized at arbitrary starting values $\boldsymbol{\theta}^0$, and so are the sequences for the stochastic volatilities, $\{\sigma^0_{\varepsilon,t}, \sigma^0_{\eta_i,t}\}_{t=1}^T$. Set $j = 1$.

1. Draw latent variables conditional on model parameters and SVs

Obtain a draw $\{a_t^j, f_t^j, \mathbf{u}_t^q\}_{t=1}^T$ from $p(\{a_t, f_t\}_{t=1}^T | \boldsymbol{\theta}^{j-1}, \{\sigma^{j-1}_{\varepsilon,t}, \sigma^{j-1}_{\eta_i,t}\}_{t=1}^T, \mathbf{y})$.

This step of the algorithm uses the state space representation described above (Appendix D.1), and produces a draw from the entire state vector $\mathbf{X}_t$ (which includes the long-run growth components, a_t, the common factor, f_t, and the idiosyncratic components of the quarterly variables, $\mathbf{u}_t^q$) by means of a forward-filtering backward-smoothing algorithm, see Carter and Kohn (1994) or Kim and Nelson (1999b). In particular, we adapt the algorithm proposed by Bai and Wang (2015), which is robust to numerical inaccuracies, and extend it to the case with mixed frequencies and missing data following Mariano and Murasawa (2003), as explained in section 3.3. Like Bai and Wang (2015), we initialise the Kalman Filter step from a normal distribution whose moments are independent of the model parameters, in particular $\mathbf{X}_0 \sim N(0, 10^4\mathbf{I})$.

2. Draw the variance of the time-varying GDP growth component

Obtain a draw $\omega_a^{2,j}$ from $p(\omega_a^2 | \{a_t^j\}_{t=1}^T)$.

Taking the sample $\{a_t^j\}_{t=1}^T$ drawn in the previous step as given, and posing an inverse-gamma prior $p(\omega_a^2) \sim IG(S_a, v_a)$ the conditional posterior of ω_a^2 is also drawn inverse-gamma distribution. As discussed in Section 4.2, we choose the scale $S_a = 10^{-3}$ and degrees of freedom $v_a = 1$ for our baseline specification.

3. Draw the autoregressive parameters of the factor VAR

Obtain a draw $\boldsymbol{\Phi}^j$ from $p(\boldsymbol{\Phi} | \{f_t^{j-1}, \sigma^{j-1}_{\varepsilon,t}\}_{t=1}^T)$.

Taking the sequences of the common factor $\{f_t^{j-1}\}_{t=1}^T$ and its stochastic volatility $\{\sigma_{\varepsilon,t}^{j-1}\}_{t=1}^T$ from previous steps as given, and posing a non-informative prior, the corresponding conditional posterior is drawn from the Normal distribution, see, e.g. Kim and Nelson (1999b). In the more general case of more than one factor, this step would be equivalent to drawing from the coefficients of a Bayesian VAR. Like Kim and Nelson (1999b), or Cogley and Sargent (2005), we reject draws which imply autoregressive coefficients in the explosive region.

4. Draw the factor loadings

Obtain a draw of $\boldsymbol{\lambda}^j$ from $p(\boldsymbol{\lambda}|\rho^{j-1}, \{f_t^{j-1}, \sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T, \mathbf{y})$.

Conditional on the draw of the common factor $\{f_t^{j-1}\}_{t=1}^T$, the measurement equations reduce to n independent linear regressions with heteroskedastic and serially correlated residuals. By conditioning on $\boldsymbol{\rho}^{j-1}$ and $\sigma_{\eta_{i,t}}^{j-1}$, the loadings can be estimated using GLS and non-informative priors. When necessary, we apply restrictions on the loadings using the formulas provided by de Wind and Gambetti (2014), see Appendix D.3 for further information.

5. Draw the serial correlation coefficients of the idiosyncratic components

Obtain a draw of $\boldsymbol{\rho}^j$ from $p(\boldsymbol{\rho}|\lambda^{j-1}, \{f_t^{j-1}, \sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T, \mathbf{y})$.

Taking the sequence of the common factor $\{f_t^{j-1}\}_{t=1}^T$ and the loadings drawn in previous steps as given, the idiosyncratic components for the monthly variables can be obtained as $u_{i,t} = y_{i,t} - \lambda^{j-1} f_t^{j-1}$. For the quarterly variables, a draw of the idiosyncratic components has been obtained directly from Step 1. Given a sequence for the stochastic volatility of the i^{th} component, $\{\sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T$, the residual is standardized to obtain an autoregression with homoskedastic residuals whose conditional posterior can be drawn from the Normal distribution as in step 2.3.

6. Draw the stochastic volatilities

Obtain a draw of $\{\sigma_{\varepsilon,t}^j\}_{t=1}^T$ and $\{\sigma_{\eta_{i,t}}^j\}_{t=1}^T$ from $p(\{\sigma_{\varepsilon,t}\}_{t=1}^T|\boldsymbol{\Phi}^{j-1}, \{f_t^{j-1}\}_{t=1}^T)$, and from $p(\{\sigma_{\eta_{i,t}}\}_{t=1}^T|\boldsymbol{\lambda}^{j-1}, \boldsymbol{\rho}^{j-1}, \{f_t^{j-1}\}_{t=1}^T, \mathbf{y})$ respectively.

Finally, we draw the stochastic volatilities of the innovations to the factor and the idiosyncratic components independently, using the algorithm proposed by Kim et al. (1998), which uses a mixture of normal random variables to approximate the elements of the log-variance. This is a more efficient alternative to the exact Metropolis-Hastings algorithm previously proposed by Jacquier et al. (2002). For the general case in which there is more than one factor, the volatilities of the factor VAR can be drawn jointly, see Primiceri (2005).

Increase j by 1, go to Step 2.1 and iterate until convergence is achieved.

D.3 Implementing linear restrictions on the factor loadings

To impose linear restrictions on the factor loadings $\boldsymbol{\lambda}$ in equation (1) of the paper, we follow de Wind and Gambetti (2014). For linear restrictions of the form

$$\mathbf{R}\boldsymbol{\lambda} = \mathbf{r} \tag{11}$$

these authors consider the special case with $\mathbf{r} = \mathbf{0}$ in equation (54) in the appendix to their paper. For $\mathbf{r} \neq \mathbf{0}$, this equation is amended as shown here. Let $\boldsymbol{\lambda}^u$ and $\boldsymbol{\lambda}^r$ denote the unrestricted and restricted loading matrix, respectively. $\boldsymbol{\lambda}^r$ is then drawn from a posterior distribution defined by (12) to (14):

$$\boldsymbol{\lambda}^r \sim N\left(\overline{\boldsymbol{\lambda}}^r, \mathbf{P}_\lambda^r\right), \tag{12}$$

where

$$\begin{aligned} \overline{\boldsymbol{\lambda}}^r &= \boldsymbol{\lambda}^u - \mathbf{P}_\lambda^u \mathbf{R}' \left(\mathbf{R}\mathbf{P}_\lambda^u \mathbf{R}'\right)^{-1} \left(\mathbf{R}\boldsymbol{\lambda}^u - \mathbf{r}\right) && (13) \\ \mathbf{P}_\lambda^r &= \mathbf{P}_\lambda^u - \mathbf{P}_\lambda^u \mathbf{R}' \left(\mathbf{R}\mathbf{P}_\lambda^u \mathbf{R}'\right)^{-1} \mathbf{R}\mathbf{P}_\lambda^u. && (14) \end{aligned}$$

E Details on the Construction of the Data Base

E.1 US (Vintage) Data Base

For our US real-time forecasting evaluation, we consider data vintages since 11 January 2000 capturing the real activity variables listed in the text. For each vintage, the start of the sample is set to January 1960, appending missing observations to any series which starts after that date. All times series are obtained from one of these sources: (1) Archival Federal Reserve Economic Data (ALFRED), (2) Bloomberg, (3) Haver Analytics. Table E.1 provides details on each series, including the variable code corresponding to the different sources.

For several series, in particular Retail Sales, New Orders, Imports and Exports, only vintages in nominal terms are available, but series for appropriate deflators are available from Haver, and these are not subject to revisions. We therefore deflate them using, respectively, CPI, PPI for Capital Equipment, and Imports and Exports price indices. Additionally, in several occasions the series for New Orders, Personal Consumption, Vehicle Sales and Retail Sales are subject to methodological changes and part of their history gets discontinued. In this case, given our interest in using long samples for all series, we use older vintages to splice the growth rates back to the earliest possible date.

For *soft* variables real-time data is not as readily available. The literature on real-time forecasting has generally assumed that these series are unrevised, and therefore used the latest available vintage. However while the underlying survey responses are indeed not revised, the seasonal adjustment procedures applied to them do lead to important differences between the series as was available at the time and the latest vintage. For this reason we use seasonally un-adjusted data and re-apply the Census-X12 procedure in real time to obtain a real-time seasonally adjusted version of the surveys. We follow the same procedure for the initial unemployment claims series. We then use Bloomberg to obtain the exact date in which each monthly data point was first published.

Table E.1:
Detailed Description of Data Series

	Frequ.	Start Date	Vintage Start	Trans-formation	Publ. Lag	Data Code
Real Gross Domestic Product	Q	Q2:1947	Dec 91	%QoQ Ann	26	GDPC1(F)
Real Consumption (ex. durables)	Q	Q2:1947	Dec 91	%QoQ Ann	26	
Hours worked	Q	Q2:1948	Dec 91	%QoQ Ann	28	
Real Investment (incl. durable cons.)	Q	Q2:1947	Dec 91	%QoQ Ann	26	
Real Industrial Production	M	Jan 47	Jan 97	% MoM	15	INDPRO(F)
Real Manufacturers' New Orders Nondefense Capital Goods Excluding Aircraft	M	Mar 68	Mar 97	% MoM	25	NEWORDER(F)[1] PPICPE(F)
Real Light Weight Vehicle Sales	M	Feb 67	Mar 97	% MoM	1	ALTSALES(F)[2] TLVAR(H)
Real Personal Income less Transfer Payments	M	Feb 59	Dec 97	% MoM	27	DSPIC96(F)
Real Retail Sales Food Services	M	Feb 47	Jun 01	% MoM	15	RETAIL(F) CPIAUCSL(F) RRSFS(F)[3]
Real Exports of Goods	M	Feb 68	Jan 97	% MoM	35	BOPGEXP(F)[4] C111CPX(H) TMXA(H)
Real Imports of Goods	M	Feb 69	Jan 97	% MoM	35	BOPGIMP(F)[4] C111CP(H) TMMCA(H)
Building Permits	M	Feb 60	Aug 99	% MoM	19	PERMIT(F)
Housing Starts	M	Feb 59	Jul 70	% MoM	26	HOUST(F)
New Home Sales	M	Feb 63	Jul 99	% MoM	26	HSN1F(F)
Total Nonfarm Payroll Employment (Establishment Survey)	M	Jan 47	May 55	% MoM	5	PAYEMS(F)
Civilian Employment (Household Survey)	M	Feb 48	Feb 61	% MoM	5	CE16OV(F)
Unemployed	M	Feb 48	Feb 61	% MoM	5	UNEMPLOY(F)
Initial Claims for UE	M	Feb 48	Jan 00*	% MoM	4	LICM(H)

(Continues on next page)

Detailed Description of Data Series (Continued)

Markit Manufacturing PMI	M	May 07	Jan 00*	-	-7	S111VPMM(H)[5] H111VPMM(H)
ISM Manufacturing PMI	M	Jan 48	Jan 00*	-	1	NMFBAI(H) NMFNI(H) NMFEI(H) NMFVDI(H)[6]
ISM Non-manufacturing PMI	M	Jul 97	Jan 00*	-	3	NAPMCN(H)
Conference Board: Consumer Confidence	M	Feb 68	Jan 00*	Diff 12 M.	-5	CCIN(H)
University of Michigan: Consumer Sentiment	M	May 60	Jan 00*	Diff 12 M.	-15	CSENT(H)[5] CONSSENT(F) Index(B)
Richmond Fed Manufacturing Survey	M	Nov 93	Jan 00*	-	-5	RIMSXN(H) RIMNXN(H) RIMLXN(H)[6]
Philadelphia Fed Business Outlook	M	May 68	Jan 00*	-	0	BOCNOIN(H) BOCNONN(H) BOCSHNN(H) BOCDTIN(H) BOCNENN(H)[6]
Chicago PMI	M	Feb 67	Jan 00*	-	0	PMCXPD(H) PMCXNO(H) PMCXI(H) PMCXVD(H)[6]
NFIB: Small Business Optimism Index	M	Oct 75	Jan 00*	Diff 12 M.	15	NFIBBN (H)
Empire State Manufacturing Survey	M	Jul 01	Jan 00*	-	-15	EMNHN(H) EMSHN(H) EMDHN(H) EMDSN(H) EMESN(H)[6]

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. In the last column, (B) = Bloomberg; (F) = FRED; (H) = Haver;
1) deflated using PPI for capital equipment; 2) for historical data not available in ALFRED we used data coming from HAVER; 3) using deflated nominal series up to May 2001 and real series afterwards; 4) nominal series from ALFRED and price indices from HAVER. For historical data not available in ALFRED we used data coming from HAVER; 5) preliminary series considered; 6) NSA subcomponents needed to compute the SA headline index. * Denotes seasonally un-adjusted series which have been seasonally adjusted in real time.

E.2 Data Base for Other G7 Economies

Table E.2:
CANADA

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Industrial Production: Manuf., Mining, Util.	M	Jan-1960	% MoM
Manufacturing New Orders	M	Feb-1960	% MoM
Manufacturing Turnover	M	Feb-1960	% MoM
New Passenger Car Sales	M	Jan-1960	% MoM
Real Retail Sales	M	Feb-1970	% MoM
Construction: Dwellings Started	M	Feb-1960	% MoM
Residential Building Permits Auth.	M	Jan-1960	% MoM
Real Exports	M	Jan-1960	% MoM
Real Imports	M	Jan-1960	% MoM
Unemployment Ins.: Initial and Renewal Claims	M	Jan-1960	% MoM
Employment: Industrial Aggr. excl. Unclassified	M	Feb-1991	% MoM
Employment: Both Sexes, 15 Years and Over	M	Feb-1960	% MoM
Unemployment: Both Sexes, 15 Years and Over	M	Feb-1960	% MoM
Consumer Confidence Indicator	M	Jan-1981	Diff 12 M.
Ivey Purchasing Managers Index	M	Jan-2001	Level
ISM Manufacturing PMI	M	Jan-1960	Level
University of Michigan: Consumer Sentiment	M	May-1960	Diff 12 M.

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table E.3:
GERMANY

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Mfg Survey: Production: Future Tendency	M	Jan-1960	Level
Ifo Demand vs. Prev. Month: Manufact.	M	Jan-1961	Level
Ifo Business Expectations: All Sectors	M	Jan-1991	Level
Markit Manufacturing PMI	M	Apr-1996	Level
Markit Services PMI	M	Jun-1997	Level
Industrial Production	M	Jan-1960	% MoM
Manufacturing Turnover	M	Feb-1960	% MoM
Manufacturing Orders	M	Jan-1960	% MoM
New Truck Registrations	M	Feb-1963	% MoM
Total Unemployed	M	Feb-1962	% MoM
Total Domestic Employment	M	Feb-1981	% MoM
Job Vacancies	M	Feb-1960	% MoM
Retail Sales Volume excluding Motor Vehicles	M	Jan-1960	% MoM
Wholesale Vol. excl. Motor Veh. and Motorcycles	M	Feb-1994	% MoM
Real Exports of Goods	M	Feb-1970	% MoM
Real Imports of Goods	M	Feb-1970	% MoM

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table E.4:
JAPAN

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
TANKAN All Industries: Actual Business Cond.	Q	Sep-1974	Diff 1 M.
Markit Manufacturing PMI	M	Oct-2001	Level
Small Business Sales Forecast	M	Dec-1974	Level
Small/Medium Business Survey	M	Apr-1976	Level
Consumer Confidence Index	M	Mar-1973	Level
Inventory to Sales Ratio	M	Jan-1978	Level
Industrial Production: Mining and Manufact.	M	Jan-1960	% MoM
Electric Power Consumed by Large Users	M	Feb-1960	% MoM
New Motor Vehicle Registration: Trucks, Total	M	Feb-1965	Diff 1 M.
New Motor Vehicle Reg: Passenger Cars	M	May-1968	% MoM
Real Retail Sales	M	Feb-1960	% MoM
Real Department Store Sales	M	Feb-1970	% MoM
Real Wholesale Sales: Total	M	Aug-1978	% MoM
Tertiary Industry Activity Index	M	Feb-1988	% MoM
Labor Force Survey: Total Unemployed	M	Jan-1960	% MoM
Overtime Hours / Total Hours (manufact.)	M	Feb-1990	% MoM
New Job Offers excl. New Graduates	M	Feb-1963	% MoM
Ratio of New Job Openings to Applications	M	Feb-1963	% MoM
Ratio of Active Job Openings and Active Job Appl.	M	Feb-1963	% MoM
Building Starts, Floor Area: Total	M	Feb-1965	% MoM
Housing Starts: New Construction	M	Feb-1960	% MoM
Real Exports	M	Feb-1960	% MoM
Real Imports	M	Feb-1960	% MoM

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table E.5:
United Kingdom

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Mar-1960	% QoQ Ann.
Dist. Trades: Total Vol. of Sales	M	Jul-1983	Level
Dist. Trades: Retail Vol. of Sales	M	Jul-1983	Leve
CBI Industrial Trends: Vol. of Output Next 3 M.	M	Feb-1975	Level
BoE Agents' Survey: Cons. Services Turnover	M	Jul-1997	Level
Markit Manufacturing PMI	M	Jan-1992	Level
Markit Services PMI	M	Jul-1996	Level
Markit Construction PMI	M	Apr-1997	Level
GfK Consumer Confidence Barometer	M	Jan-1975	Diff 12 M.
Industrial Production: Manufacturing	M	Jan-1960	% MoM
Passenger Car Registrations	M	Jan-1960	% MoM
Retail Sales Volume: All Retail incl. Autom. Fuel	M	Jan-1960	% MoM
Index of Services: Total Service Industries	M	Feb-1997	% MoM
Registered Unemployment	M	Feb-1960	% MoM
Job Vacancies	M	Feb-1960	% MoM
LFS: Unemployed: Aged 16 and Over	M	Mar-1971	% MoM
LFS: Employment: Aged 16 and Over	M	Mar-1971	% MoM
Mortgage Loans Approved: All Lenders	M	May-1993	% MoM
Real Exports	M	Feb-1961	% MoM
Real Imports	M	Feb-1961	% MoM

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table E.6:
FRANCE

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Industrial Production	M	Feb-1960	% MoM
Total Commercial Vehicle Registrations	M	Feb-1975	% MoM
Household Consumption Exp.: Durable Goods	M	Feb-1980	% MoM
Real Retail Sales	M	Feb-1975	% MoM
Passenger Cars	M	Feb-1960	% MoM
Job Vacancies	M	Feb-1989	% MoM
Registered Unemployment	M	Feb-1960	% MoM
Housing Permits	M	Feb-1960	% MoM
Housing Starts	M	Feb-1974	% MoM
Volume of Imports	M	Jan-1960	% MoM
Volume of Exports	M	Jan-1960	% MoM
Business Survey: Personal Prod. Expect.	M	Jun-1962	Level
Business Survey: Recent Output Changes	M	Jan-1966	Level
Household Survey: Household Conf. Indicator	M	Oct-1973	Diff 12 M.
BdF Bus. Survey: Production vs. Last M., Ind.	M	Jan-1976	Level
BdF Bus. Survey: Production Forecast, Ind.	M	Jan-1976	Level
BdF Bus. Survey: Total Orders vs. Last M., Ind.	M	Jan-1981	Level
BdF Bus. Survey: Activity vs. Last M., Services	M	Oct-2002	Level
BdF Bus. Survey: Activity Forecast, Services	M	Oct-2002	Level
Markit Manufacturing PMI	M	Apr-1998	Level
Markit Services PMI	M	May-1998	Level

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table E.7:
ITALY

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Markit Manufacturing PMI	M	Jun-1997	Level
Markit Services PMI: Business Activity	M	Jan-1998	Level
Production Future Tendency	M	Jan-1962	Level
ISTAT Services Survey: Orders, Next 3 M-	M	Jan-2003	Level
ISTAT Retail Trade Confidence Indicator	M	Jan-1990	Level
Industrial Production	M	Jan-1960	% MoM
Real Exports	M	Jan-1960	% MoM
Real Imports	M	Jan-1960	% MoM
Real Retail Sales	M	Feb-1990	% MoM
Passenger Car Registrations	M	Jan-1960	% MoM
Employed	M	Feb-2004	% MoM
Unemployed	M	Feb-1983	% MoM

Notes: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

F Choice of Priors

As explained in the paper, we use non-informative priors for loadings and serial correlation coefficients of factor and idiosyncratic components in order to aide comparability with the previous literature, which has generally used classical estimation methods. With respect to the choice of priors on the new parameters of our specification, namely ω_a^2, ω_ε^2 and $\omega_{\eta,i}^2$ in equations (5)-(7), we closely follow the related literature, in particular Cogley and Sargent (2005) and Primiceri (2005), by setting relatively conservative priors, which shrink the model towards the benchmark with no time-variation, but are still loose enough for the data to be able to speak. In particular, in all the inverse-gamma (IG) distributions we set the number of degrees of freedom to 1, the minimum required to make the prior distributions proper while keeping the weight of the prior low. As to the choice of the scale parameter of the IG distributions, it is worth pointing out that this does not parametrize time variation itself, but rather incorporates a prior belief about the amount of time variation. To gain an intuition about the prior on ω_a^2, in Section 4.2 we note that the chosen value of 0.001 implies that over a period of ten years the random walk process of the long-run growth rate is expected to vary with a standard deviation of around 0.4 percentage points in annualized terms, which we believe is a fairly conservative prior in terms of economic magnitudes. The choice of 10^{-4} for the prior on ω_ε^2 and $\omega_{\eta,i}^2$ is similar to the approach of Primiceri (2005).

To shed some light on the robustness of our results to the choice of priors, in what follows we explore the sensitivity of our main results to varying the tightness of the respective priors. To summarize the most notable finding, we find that the data strongly drives the result of time variation both in the long-run growth rate and the volatilies: a dogmatically large amount of shrinkage is needed in order to make either of them disappear.

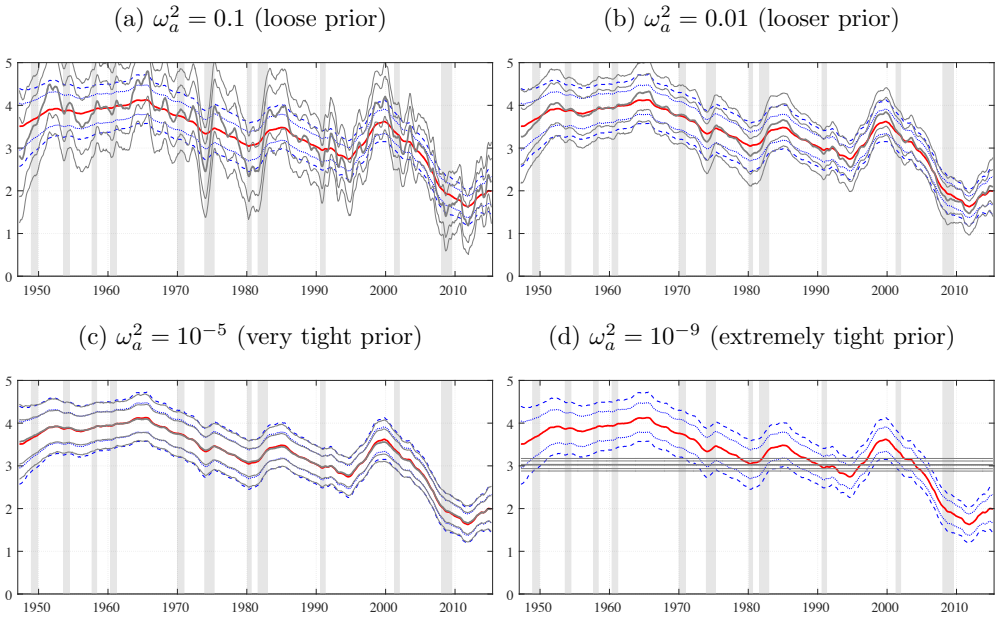
F.1 Robustness checks on prior choice

Prior on innovation variance to the time-varying long-run growth rate

In Figure F.1 we explore the sensitivity of our key results to the choice of the scale parameter of the prior on the innovation variance to the time-varying long-run growth rate of real GDP, ω_a^2. Each panel plots our baseline estimate of g_t, which has been obtained with a prior scale of 10^{-3} (red/blue). We then successively compare this baseline estimate with alternative estimates obtained when imposing both looser and tighter prior scales, respectively (gray). Panel (a) of the figure reveals that with a prior implying a very large variance the estimated trend is pinned down with relatively more uncertainty and evolves in bumpy fashion, yet the qualitative pattern around the evolution of long-run growth, in particular the recent slowdown, remains clearly visible. Panels (b) and (c) show that using a ten times looser prior (0.01) and a hundred times tighter prior (10^{-5}) than the one in our baseline setting gives very similar results

to ours. In the later case, the estimate is almost identical. Finally, a dogmatically tight prior (10^{-9}) is required to make variation in the long-run growth rate disappear entirely, which is visible in Panel (d).

Figure F.1: Comparison Across Different Prior Scales of ω_α^2



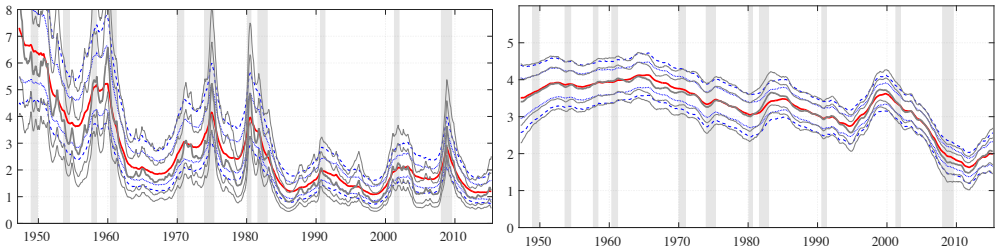
Note: In each panel our baseline the median estimate of real GDP growth based on a scale of 10^{-3} is presented (red), with corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals. The corresponding estimates based an different prior scales are superimposed in gray in each panel.

Prior on innovation variance to the SV

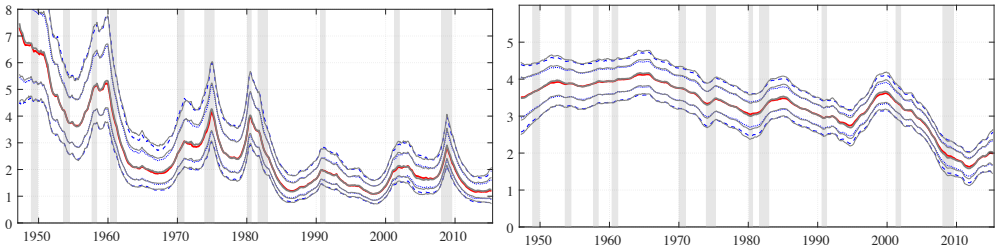
Figure F.2 presents the sensitivity of the results to the choice of the scale parameter of the prior on the innovation variance to the SV in both the common factor and the idiosyncratic components. Similar to Figure F.1 we compare our baseline estimates (red/blue), where we set $\omega_\varepsilon^2 = \omega_{\eta,i}^2 = 10^{-4}$, with estimates obtained under a range of varied prior scales (gray). In each case, the figure shows both the estimated SV of the factor as well as the estimate of the long-run growth rate of real GDP growth. Panel (a) displays the results for a very loose prior (1), while Panel (b) for a prior which is ten times looser than the baseline (10^{-3}). Finally, the estimates shown in Panel (c) are obtained under a tighter prior (10^{-5}). Again, the results reported in the paper do not seem to be affected. Both the estimates of the SV and the long-run growth rate of real GDP are almost identical to our main results.

Figure F.2: Comparison Across Different Prior Scales of ω_{ε}^2 and $\omega_{\eta,i}^2$

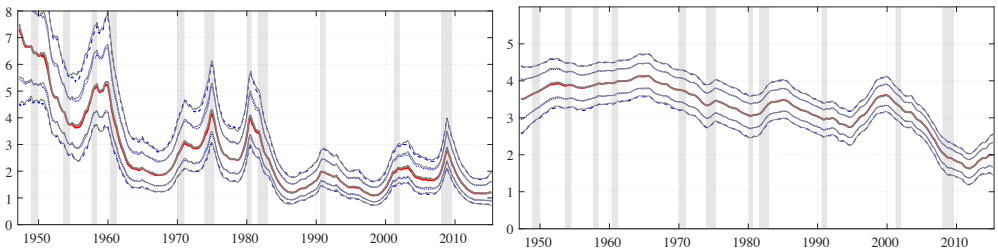
(a) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 1$ (very loose prior)



(b) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 10^{-3}$ (looser prior)



(c) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 10^{-5}$ (tighter prior)



Note: In each panel our baseline estimate of the SV of the common factor based on a scale of 10^{-4} is presented (red) in the left chart. The right chart plots the estimate of the long-run growth rate of real GDP based on the same scale. Corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals are also plotted. The analogue estimates based on the alternative prior scales are superimposed in gray in each panel.

Prior on serial correlation in factor and idiosyncratic components

As a final robustness check, we consider "Minnesota"-style priors on the autoregressive coefficients of the factor as well as shrinking the coefficients of the serial correlation towards zero. To be precise, we center the prior on the first lag of the factor around 0.9 and all other lags at zero. The motivation for these priors is to express a preference for a more parsimonious model where the factors capture the bulk of the persistence of the series and the idiosyncratic components are close to iid, that is closer to true measurement error. These alternative priors do not meaningfully affect the posterior estimates of our main objects of interest, so we omit additional figures. Note that we have found some evidence that the use of such priors might at times improve the convergence of the algorithm. Specifically, when we apply the model to the other G7 economies (see Section 5), we find that for some countries where few monthly indicators are available, shrinking the serial correlations of the idiosyncratic components towards zero helps obtaining a common factor that is persistent.

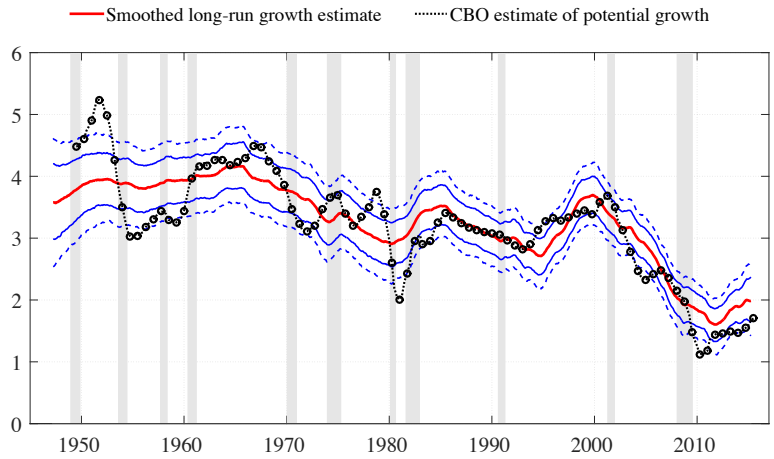
G Comparison with CBO Measure of Potential Output

Our estimate of long-run growth and the CBO's potential growth estimate capture related but not identical concepts. The CBO measures the growth rate of potential output, i.e. the level of output that could be obtained if all resources were used fully, whereas our estimate, similar to Beveridge and Nelson (1981), measures the component of the growth rate that is expected to be permanent. Moreover, the CBO estimate is constructed using the so-called "production function approach", which is radically different from the DFM methodology.[6]

As a simple sanity check, it is interesting to see that despite employing different statistical methods they produce qualitatively similar results, visible in Figure G.1, with the CBO estimate displaying a more marked cyclical pattern but remaining for most of the sample within the 90% credible posterior interval of our estimate. As in our estimate, most of the slowdown occurred prior to the Great Recession. The CBO's estimate was below ours immediately after the recession, reaching an unprecedented low level of about 1.25% in 2010, and remains in the lower bound of our posterior estimate since then. Section 4.6 expands on the reason for this divergence and argues that this is likely to stem from the larger amount of information incorporated in the DFM. In fact, the CBO estimate of potential growth is noticeably more cyclical.

[6]Essentially, the production function approach calculates the trend components of the supply inputs to a neoclassical production function (the capital stock, total factor productivity, and the total amount of hours) using statistical filters and then aggregates them to obtain an estimate of the trend level of output. See CBO (2001).

Figure G.1: Long-run GDP Growth Estimate in Comparison to CBO



Note: The figure displays the posterior median estimate of long-run GDP growth with the corresponding credible intervals, as displayed in Figure 2 Panel (a) in the main body of the paper, in comparison with the CBO's measure of potential output growth, which is shown in black circles.

H Details About the Forecast Evaluation

H.1 Setup

Using our real-time database of US vintages, we re-estimate the following three models each day in which new data is released: a benchmark with constant long-run GDP growth and constant volatility (Model 0, similar to Banbura and Modugno (2014)), a version with constant long-run growth but with stochastic volatility (Model 1, similar to Marcellino et al. (2014)), and the baseline model put forward in the paper with both time-variation in the long-run growth of real GDP and SV (Model 2). Allowing for an intermediate benchmark with only SV allows us to evaluate how much of the improvement in the model can be attributed to the addition of the long-run variation in GDP as opposed to the SV. We evaluate the point and density forecast accuracy relative to the initial ("Advance") release of GDP, which is released between 25 and 30 days after the end of the reference quarter.[7]

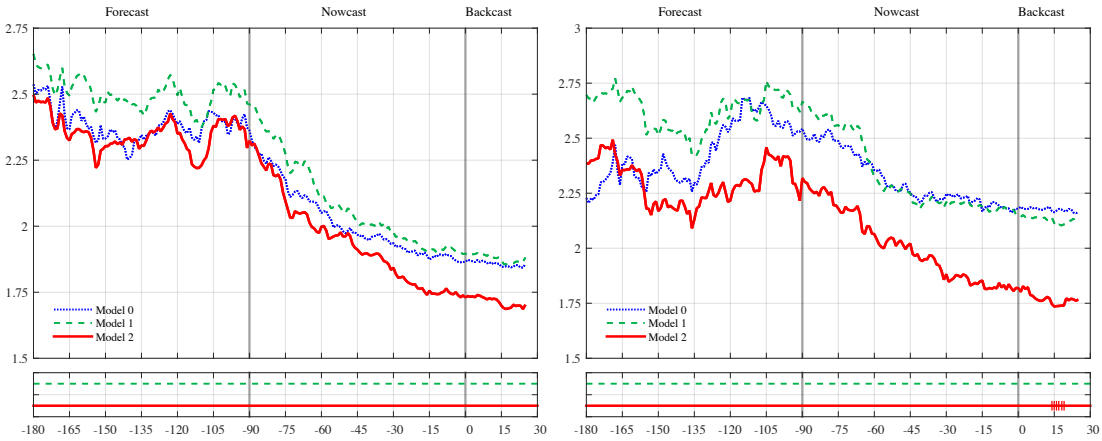
When comparing the three different models, we test the significance of any improvement of Models 1 and 2 relative to Model 0. This raises some important econometric complications given that (i) the three models are nested, (ii) the forecasts are produced using an expanding window, and (iii) the data used is subject to revision. These three issues imply that commonly used test statistics for forecasting accuracy, such as the one proposed by Diebold and Mariano (1995) and Giacomini and White (2006) will have a non-standard limiting distribution. However, rather than not reporting any test, we follow the "pragmatic approach" of Faust and Wright (2013) and Groen et al. (2013), who build on Monte Carlo results in Clark and McCracken (2012). Their results indicate that the Harvey et al. (1997) small sample correction of the Diebold and Mariano (1995) statistic results in a good sized test of the null hypothesis of equal finite sample forecast precision for both nested and non-nested models, including cases with expanded window-based model updating. Overall, the results of the tests should be interpreted more as a rough gauge of the significance of the improvement than a definitive answer to the question. We compute various point and density forecast accuracy measures at different moments in the release calendar, to assess how the arrival of information improves the performance of the model. In particular, the computations are carried out starting 180 days before the end of the reference quarter, and every subsequent day up to 25 days after its end, when the GDP figure for the quarter is usually released. This means that we will evaluate the forecasts of the next quarter, current quarter (nowcast), and the previous quarter (backcast). We consider two different samples for the evaluation: the full sample (2000:Q1-2015:Q1) and the sample covering the recovery since the Great Recession (2009:Q2-2015:Q1).

[7] We have explored the alternative of evaluating the forecasts against subsequent releases, or the latest available vintages. The relative performance of the three models is broadly unchanged, but all models do better at forecasting the initial release. If the objective is to improve the performance of the model relative to the first official release, then ideally an explicit model of the revision process would be desirable. The results are available upon request.

Figure H.1: Point Forecast Accuracy Evaluation

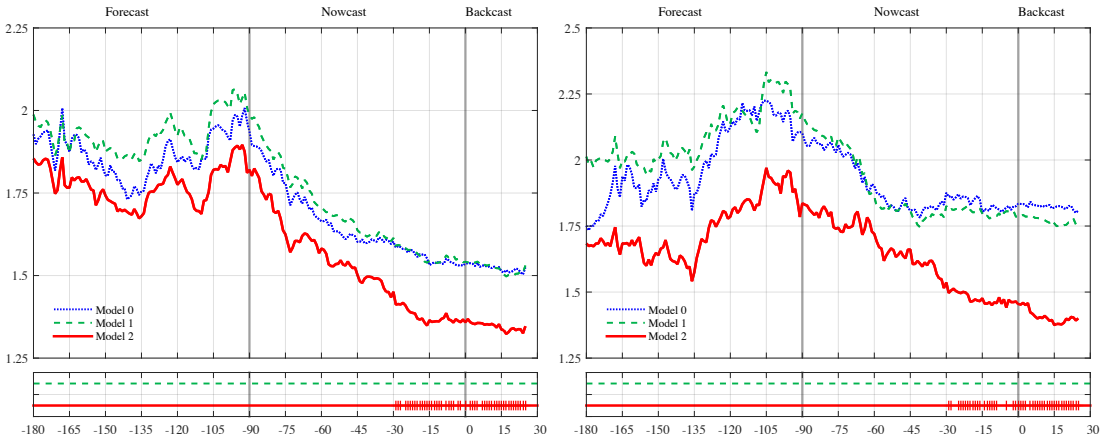
(a) Root Mean Squared Error

Full Sample: 2000:Q1-2015:Q1 Recovery Sample: 2009:Q2-2015:Q1



(b) Mean Absolute Error

Full Sample: 2000:Q1-2015:Q1 Recovery Sample: 2009:Q2-2015:Q1



Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

H.2 Point Forecast Evaluation

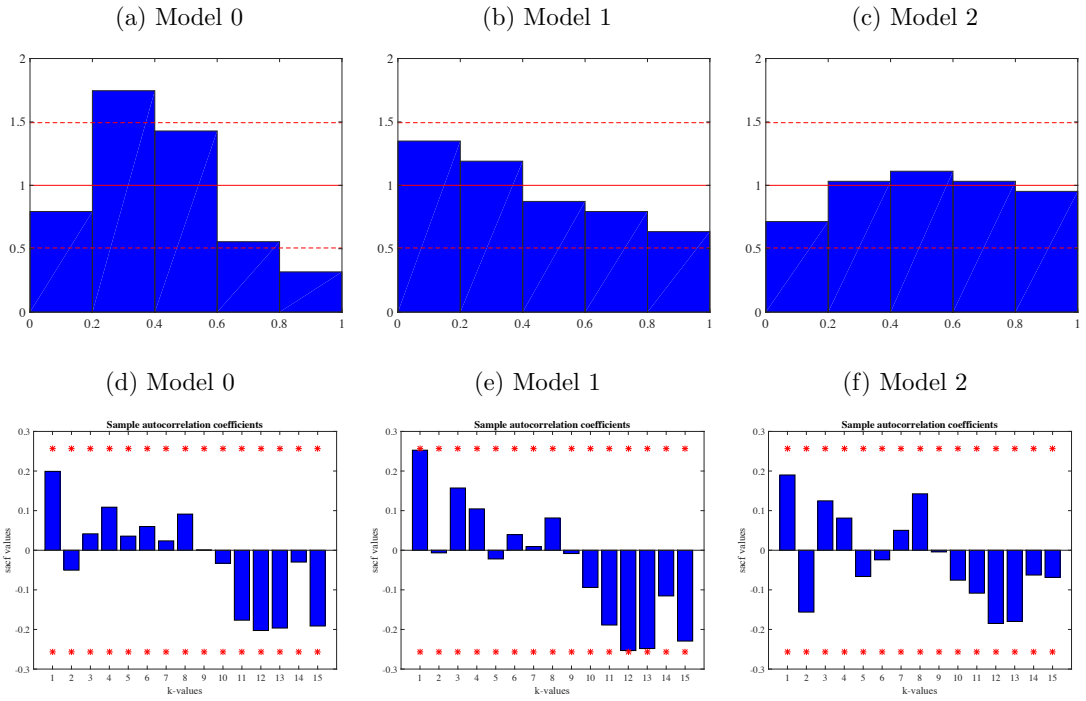
Figure H.1 shows the results of evaluating the posterior mean as point forecast. We use two criteria, the root mean squared error (RMSE) and the mean absolute error (MAE). As expected, both of these decline as the quarters advance and more information on monthly indicators becomes available, see e.g. Banbura et al. (2012). Both the RMSE and the MAE of Model 2 are lower than that of Model 0, particularly so from the start of the nowcasting period, while Model 1 is somewhat worse overall. Our gauge of significance indicates that these differences in nowcasting performance are significant at the 10% level for the overall sample in the case of the MAE, but not the RMSE. The improvement in performance is much clearer in the recovery sample. In fact, the inclusion of the time varying long run component of GDP helps anchor GDP predictions at a level consistent with the weak recovery experienced in the past few years and produces nowcasts that are 'significantly' superior to those of the reference model from around 30 days before the end of the reference quarter. In essence, ignoring the variation in long-run GDP growth would have resulted in being on average around 1 percentage point too optimistic from 2009 to 2015.

H.3 Density Forecast Evaluation

Density forecasts can be used to assess the ability of a model to predict unusual developments, such as the likelihood of a recession or a strong recovery given current information. The adoption of a Bayesian framework allows us to produce density forecasts from the DFM that consistently incorporate both filtering and estimation uncertainty. Figure H.2 reports the probability integral transform (PITs) and the associated autocorrelation functions (ACFs) for the 3 models calculated with the nowcast of the last day of the quarter. Diebold et al. (1998) highlight that well calibrated densities are associated with uniformly distributed and independent PITs. Figure H.2 suggests that the inclusion of SV is paramount to get well calibrated densities, whereas the inclusion of the long-run growth component helps to get a more appropriate representation of the right side of the distribution, as well as making sure that the first order autocorrelation is not statistically significant.

There are several measures available for density forecast evaluation. The (average) log score, i.e. the logarithm of the predictive density evaluated at the realization, is one of the most popular, rewarding the model that assigns the highest probability to the realized events. Gneiting and Raftery (2007), however, caution against using the log score, emphasizing that it does not appropriately reward values from the predictive density that are close but not equal to the realization, and that it is very sensitive to outliers. They therefore propose the use of the (average) continuous rank probability score (CRPS) in order to address these drawbacks of the log-score. Figure H.3 shows that by both measures our model outperforms its counterparts. Interestingly, the comparison of Model 1 and Model 2 suggests that failing to properly account for the long-run growth component might give a misrepresentation of the GDP densities,

Figure H.2: Probability Integral Transform (PITs)



Note: The upper three panels display the cdf of the Probability Integral Transforms (PITs) evaluated on the last day of the reference quarter, while the lower three display the associated autocorrelation functions.

resulting in poorer density forecasts.

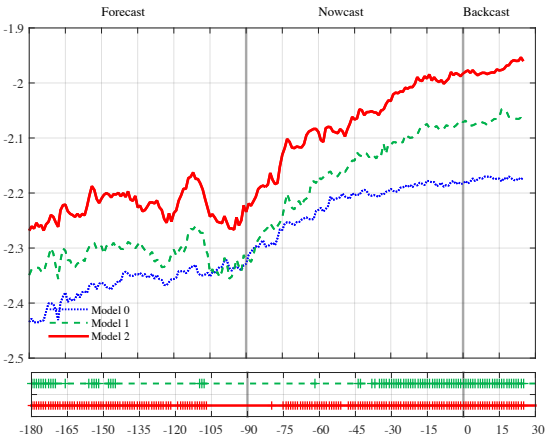
In addition to the above results, we also assess how the three models fare when different areas of their predictive densities are emphasized in the forecast evaluation. To do that we follow Groen et al. (2013) and compute weighted averages of Gneiting and Raftery (2007) quantile scores (QS) that are based on quantile forecasts that correspond to the predictive densities from the different models (Figure H.4).[8] Our results indicate that while there is an improvement in density nowcasting for the entire distribution, the largest improvement comes from the right tail. For the full sample, Model 1 is very close to Model 0, suggesting that being able to identify the location of the distribution is key to the improvement in performance. In order to appreciate the importance of the improvement in the density forecasts, and in particular in the right side of the distribution, we calculated a recursive estimate of the likelihood of a 'strong recovery', where this is defined as the probability of an average growth rate of GDP (over the present and next three quarters) above the historical average. Model 0 and Model 2 produce very similar probabilities up until 2011 when, thanks to the

[8] As Gneiting and Ranjan (2011) show, integrating QS over the quantile spectrum gives the CRPS.

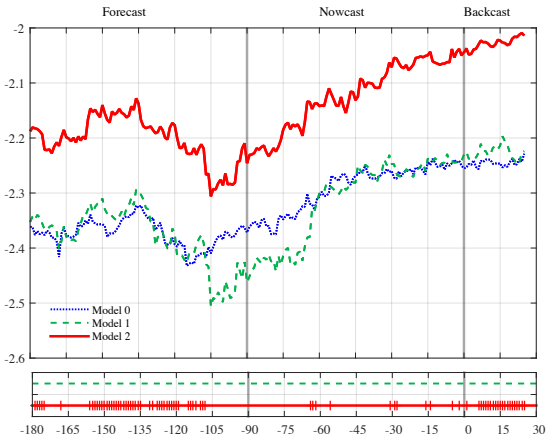
Figure H.3: Density Forecast Accuracy Evaluation

(a) Log Probability Score

Full Sample: 2000:Q1-2015:Q1

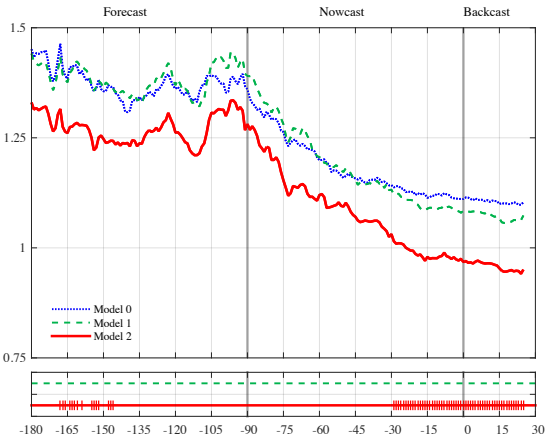


Recovery Sample: 2009:Q2-2015:Q1

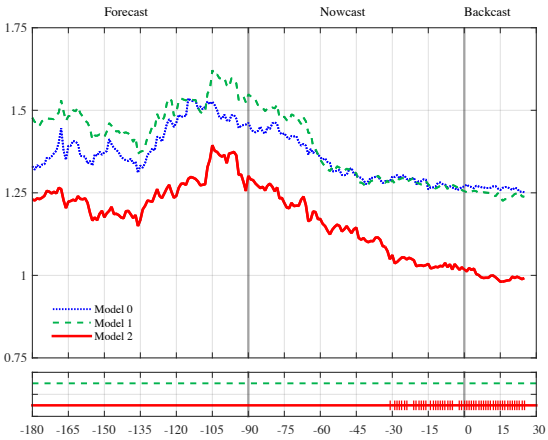


(b) Continuous Rank Probability Score

Full Sample: 2000:Q1-2015:Q1



Recovery Sample: 2009:Q2-2015:Q1



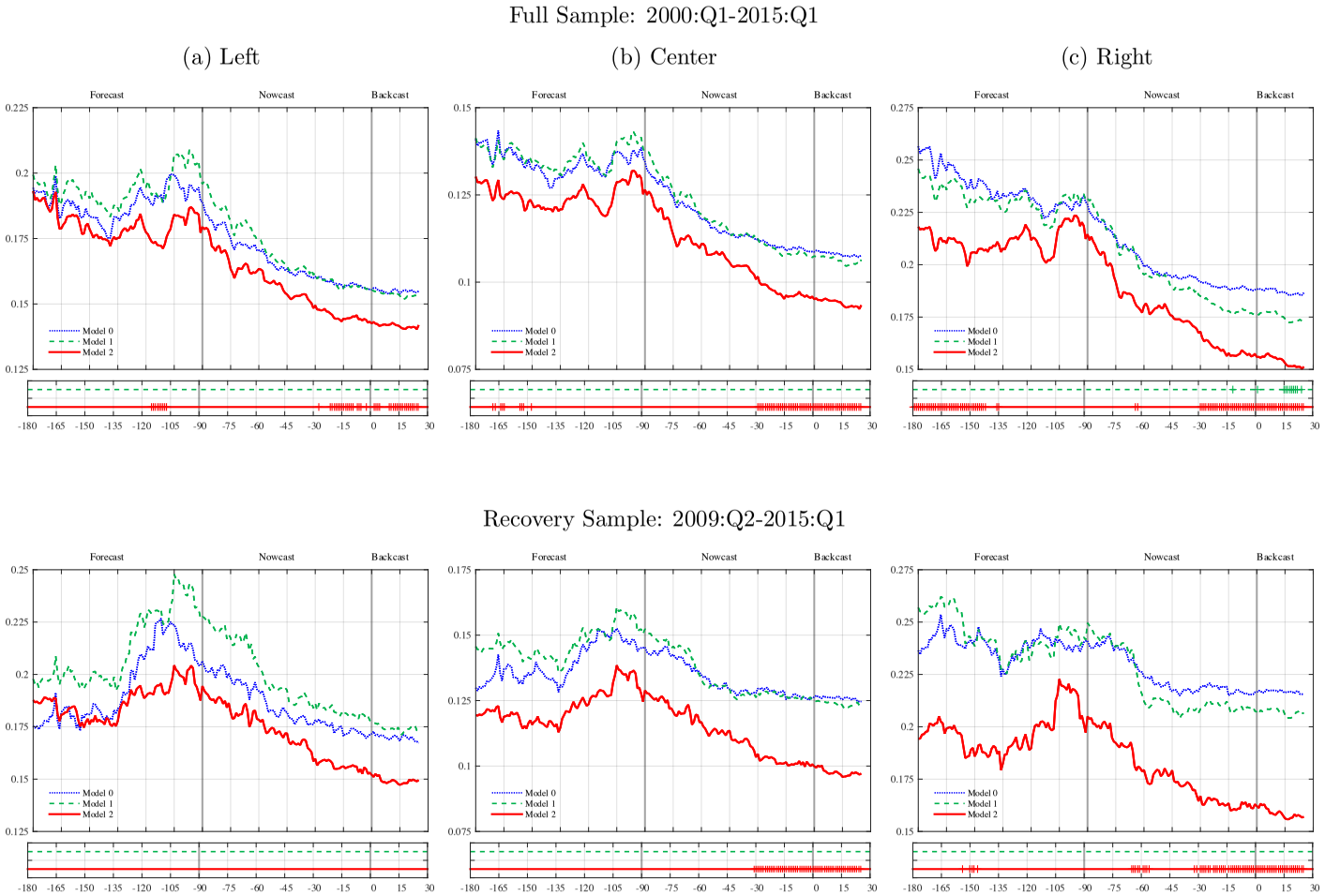
Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

downward revision of long-run GDP growth, Model 2 starts to deliver lower probability estimates consistent with the observed weak recovery. The Brier score for Model 2 is 0.186 whereas the score for Model 0 is 0.2236 with the difference significantly different at 1% (Model 1 is essentially identical to Model 0).[9]

In sum, the results of the out-of-sample forecasting evaluation indicate that a model that allows for time-varying long-run GDP growth and SV produces short-run forecasts that are on average (over the full evaluation sample) either similar to or improve upon the benchmark model. The performance tends to improve substantially in the sub-sample including the recovery from the Great Recession, coinciding with the significant downward revision of the model's assessment of long-run growth. Furthermore, the results indicate that while there is an improvement in density nowcasting for the entire distribution, the largest improvement comes from the right tail.

[9]The results are available upon request.

Figure H.4: Density Forecast Accuracy Evaluation: Quantile Score Statistics



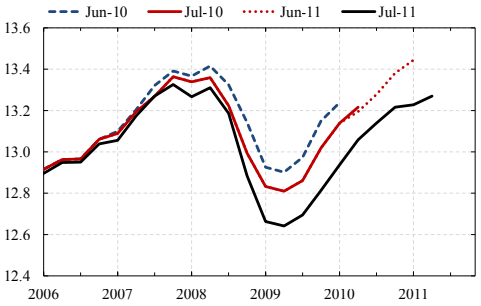
Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

I Case Study - The Decline of The Long-Run Growth Estimate in Mid-2010 and Mid-2011

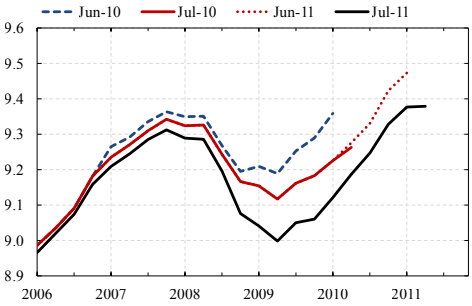
Figure I.1 looks in more detail at the specific information that, in real time, led the model to reassess its estimate of long-run growth. There are large reassessments of long-run growth around July 2010 and July 2011, coinciding with the publication by the Bureau of Economic Analysis of the annual revisions to the National Accounts, which each year incorporate previously unavailable information for the previous three years. In both cases, the revisions implied substantial downgrades both to GDP (Panel a) and in particular to the growth of consumption (Panel b) in the first years of the recovery, from 2.5% to 1.6%, and instead allocated much of the growth in GDP during the recovery to inventory accumulation. The estimate of long-run growth produced by our model is downgraded sharply as information about these revision is coming in, reflecting the role of consumption as the most persistent and forward looking component of GDP. This is clearly visible in Panels (c) and (d) of Figure I.1. In particular, Panel (c) presents the evolution of the GDP nowcast for 2010 produced by Model 2 (black line), in comparison with the counterfactual nowcast that would result if there had been no revisions to long-run growth (dashed line). It is evident that the bulk of the revisions to GDP growth that year are the consequence of a large downward revision to long-run growth. Panel (d) plots the annual nowcast of GDP produced by Model 0 (dashed line), which does not allow for changes in long-run GDP growth, and Model 2 (solid line), our baseline specification. Up to mid-2010, both models produce similar nowcasts (not shown). After 2010, however, it is clearly visible that Model 0 begins each year expecting robust growth of above 3%, only to be disappointed by incoming data. The nowcasts by Model 2, which has incorporated the decline in long-run growth, do not suffer from the same pattern of systematic downward surprises.

Figure I.1: Case Study: Impact of National Accounts Revision

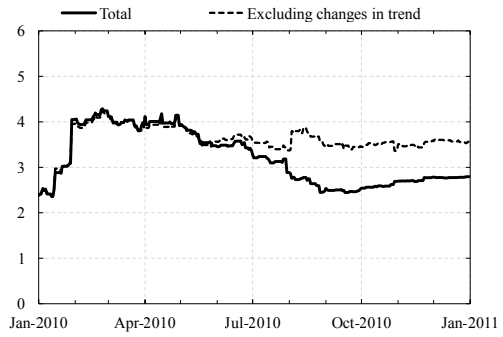
(a) Vintages of Real GDP



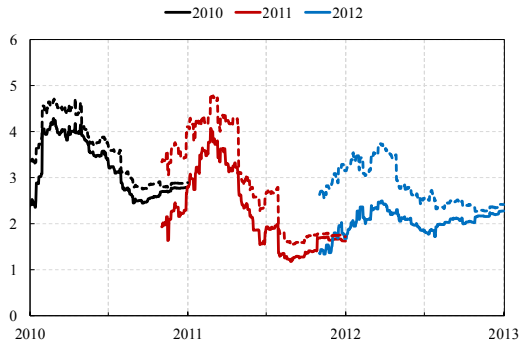
(b) Vintages of Real Personal Consumption Exp.



(c) Breakdown of Model 2's nowcast for 2010



(d) Evolution of calendar-year GDP nowcasts



Note: Panels (a) and (b) compare several vintages of data on real GDP and real personal consumption expenditures around the time of important revisions by the BEA. Panel (c) presents the evolution of the GDP nowcast for 2010 produced by Model 2 (black line), in comparison with the counterfactual nowcast that would result if there had been no revisions to long-run growth (dashed line). The evolution of calendar-year nowcasts of real GDP growth produced by Model 0 (dashed) and Model 2 (solid) are presented in Panel (d).

J Inspecting Data Set Size and Composition - More Details

J.1 Extended Model: Estimation Using a Very Large Panel

With regards to the size of the data set, in Section 4.1 of the main text we argue in favor of excluding disaggregated series within the various categories of real activity. This is because of the fact that the strong correlation across series within the same category might be a source of model misspecification. This is for two reasons: first, strong correlation in the idiosyncratic terms of series between the same category, and second, the fact that finer disaggregation levels are available for certain categories can lead to oversampling, see Boivin and Ng (2006) and Alvarez et al. (2012) for more details.

It is possible, however, to consider a more general specification of our model that can alleviate this problem, once we take into account the fact that persistent idiosyncratic movements common across series of the same category usually reflect differences in phase relative to the common activity factor. For example, all series related to employment respond to innovations to real activity with a lag. An interesting question is how our results are affected if one were to aim to make the dimension of $\mathbf{y}_t$ as large as possible, instead of carefully making variable selection based on the criteria discussed in the paper. In order to illustrate this point, we construct a "universe" of potentially available real activity time series for inclusion, based on a systematic attempt to find as many as possible US real activity time series. This is the "extended model" introduced in Section 4.6 of the paper.

J.1.1 Construction of the Extended Data Set

To construct the data panel for the extended model, we proceed as follows. First, we obtain all of the monthly real activity variables contained in the data set used by Stock and Watson (2012), which results in 75 time series.[10] Second, we exhaustively expand the monthly series contained in our original data set along all levels and dimensions of disaggregation available through Haver Analystics.[11] Out of this collection of expanded series, we then select any series that is not already contained in the 75 Stock and Watson indicators. Overall, this procedure results in a data set of as many as 155 time series capturing US real activity.[12]

[10] Details on this data set can be found in the online supplement to Stock and Watson (2012), available on Mark Watson's website. The only variable we were not able to obtain is Construction Contracts, which is not publicly available.

[11] This includes for example disaggregation along sectoral, regional and demographic characteristics.

[12] A detailed list of variables is available upon request.

J.1.2 Extended Model Specification

Maintaining the specification with a single factor (i.e. $k = 1$) we modify equation (1) of the paper as follows:

$$\mathbf{y}_t = \mathbf{c}_t + \mathbf{\Lambda}(L) f_t + \mathbf{u}_t, \tag{15}$$

such that the loading matrix $\Lambda(L)$ is now a polynomial in the lag operator of order s, i.e. contains the loadings on the contemporaneous common factor and its lags. In the special case where $s = 0$ we obtain our baseline specification. For the extended model, we set the maximum lag length, $s = 5$. The remaining equations of the model remain unchanged.

J.1.3 Priors and Model Settings

The data is standardized prior to estimation. "Minnesota"-style priors are applied to the coefficients in $\mathbf{\Lambda}(L)$, $\phi(L)$ and $\rho_i(L)$. More specifically:

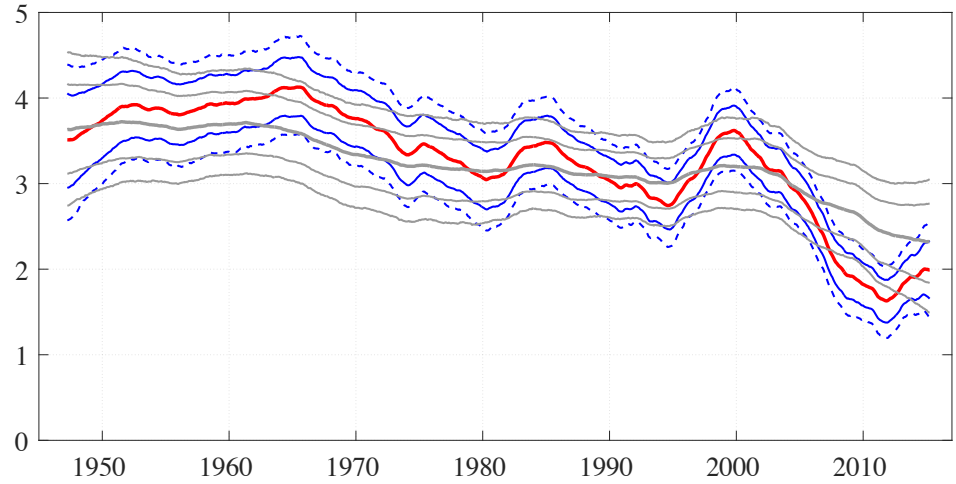
- For the autoregressive coefficients of the factor dynamics, $\phi(L)$, the prior mean is set to 0.9 for the first lag, and to zero in subsequent lags. This reflects a belief that the common factor captures a highly persistent but stationary business cycle process.
- For the factor loadings, $\mathbf{\Lambda}(L)$, the prior mean is set to 1 for the first lag, and to zero in subsequent lags. This shrinks the model towards the factor being the cross sectional average of the variables, see D'Agostino et al. (2015).
- For the autoregressive coefficients of the idiosyncratic, $\rho_i(L)$ the prior is set to zero for all lags, thus shrinking the model towards a model with no serial correlation in $u_{i,t}$.

In all cases, the variance on the priors is set to $\frac{\tau}{h^2}$, where τ is a parameter governing the tightness of the prior, and h is equal to the lag number of each coefficent, ranging $1 : p$, $1 : q$ and $1 : s + 1$. Following D'Agostino et al. (2015), we set $\tau = 0.2$, a value which is standard in the Bayesian VAR literature.

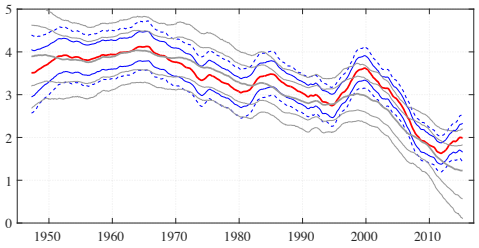
J.2 Results Across Alternative Data Sets

Figure J.1: Comparison Across Alternative Data Sets/Models

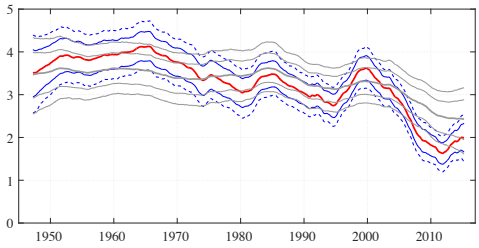
(a) Baseline With And Without Including Consumption



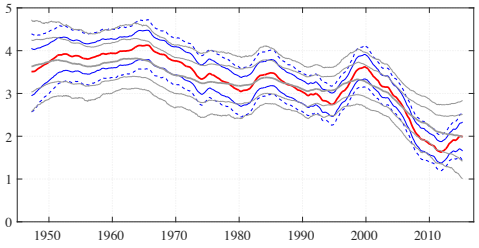
(b) Okun



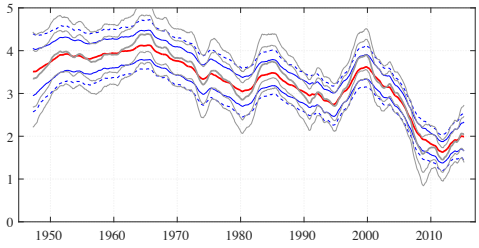
(c) Mariano-Murasawa



(d) Extended model ex. consumption



(e) Extended model



Note: In each panel our baseline the median estimate of real GDP growth is presented (red), with corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals. The corresponding estimates for the respective alternative data sets are superimposed in gray.

K A Growth Accounting Exercise

The decomposition exercise carried out in Section 5 of the paper provides a first step towards giving an economically more meaningful interpretation of the movements in long-run real GDP growth detected by our model. While our equation $g_t = z_t + h_t$ follows from a simple identity, we demonstrate in this appendix how it can be related to the standard growth accounting framework.

To illustrate this point, consider two versions of the standard neoclassical growth model. In the first version, assume a standard Cobb-Douglas production function with Hicks-neutral technological change and constant returns to scale. In growth rates, this can be written as

$$dlogY_t = dlogTFP_t + \alpha dlogK_t + (1-\alpha)dlogH_t, \tag{16}$$

where Y_t, K_t and H_t denote the level of output, the capital stock and labor input (total hours), respectively. α is the capital share and TFP_t is total factor productivity. Rearranging this relation gives

$$dlogY_t = dlogTFP_t + dlogH_t + \alpha(dlogK_t - dlogH_t), \tag{17}$$

so that the growth rate of real GDP is the sum of long-run growth in technology, total hours and a third term which captures differential growth in input factors which implies changes in the capital-labor ratio ("capital deepening"). In the second version of the neoclassical growth model, consider adding growth in labor-augmenting technology in the form of labor quality, denoted Q_t. In this case, the relation between growth rates is rearranged to

$$dlogY_t = dlogTFP_t + dlogH_t + \alpha(dlogK_t - dlogH_t) + (1-\alpha)dlogQ_t. \tag{18}$$

Both relations (17) and (18) can be captured in our econometric framework. We define the first four elements of our vector of observables $\mathbf{y}_t$ in equation (1) to be the growth rate in real GDP, real consumption, TFP and total hours worked. As in the baseline model, *transitory* fluctuations in inputs (due to temporary shocks) would still be captured by the cyclical dactor, f_t, whereas the various sources of *permanent* changes in the growth rate of inputs (say, the long-run growth rate of technology, or the long-run growth rate of the population) would be included in $\mathbf{a}_t$. In order to mimic the relations prescribed by the two versions of the neoclassical growth model, we specify the long-run time variation in our model, $\mathbf{a}_t$ as a composite of three terms. While $\tilde{h}_t$ captures long-run movements in hours, the movements in long-run labor productivity are now further decomposed into a "technology" term $\tilde{z}_t$ and a "non-technology" term $\tilde{x}_t$. Formally, $\mathbf{c}_t$ in equation (2) is constructed as

$$\mathbf{a}_t = \begin{bmatrix} \tilde{z}_t \\ \tilde{h}_t \\ \tilde{x}_t \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{19}$$

What the non-technology term corresponds to depends on the underlying structure that is assumed. For instance, in the first version of the neoclassical growth model above

$$\tilde{x}_t \equiv \alpha(dlogK_t - dlogH_t) \tag{20}$$

and in the second case

$$\tilde{x}_t \equiv \alpha(dlogK_t - dlogH_t) + (1-\alpha)dlogQ_t. \tag{21}$$

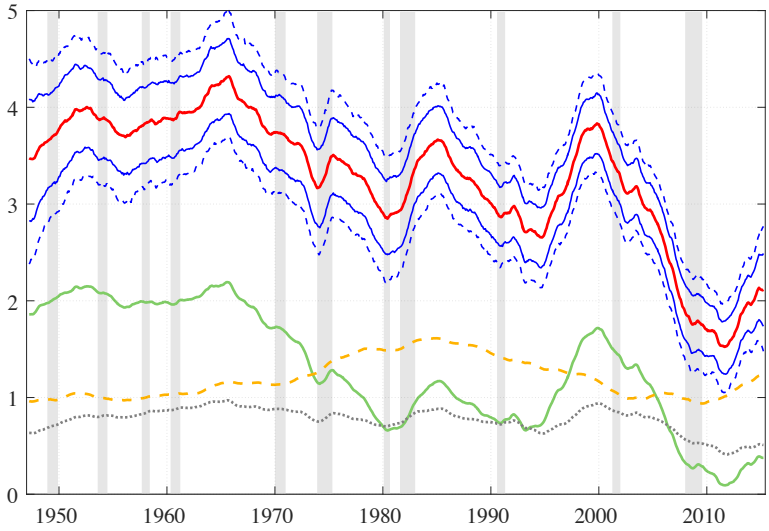
In both cases, $\tilde{x}_t$ subsumes potential long-run factors other than TFP that may explain changes in the long-run labor productivity trend we discuss in the paper.[13] $\tilde{z}_t$ is intended to capture changes in the long-run technology growth rate.

Figure K.1 presents the results for US data when defining $\mathbf{c}_t$ by (19), and the measure of utilization-adjusted TFP from Fernald (2012) is used as an additional observable. Panel (a) shows the posterior estimate of long-run real GDP growth (including bands), together with the decomposition into long-run total hours growth, long-run technology growth and long-run non-technological growth. Reassuringly, the evolution of the total long-run growth component, g_t (red) is virtually identical to the estimate from our baseline model. The estimate of long-run hours growth (orange) is also very similar to its counterpart in Section 5 of the paper. Interestingly, the non-technological term (dashed gray) is positive on average and is relatively stable over the sample. Finally, the key insight from panel (a) is that our estimate of the long-run technology (green) displays strong movements that are very similar to the long-run growth rate of labor productivity which we have extracted in the simpler decomposition in Section 5. Under the assumption of a neoclassical structure, changes in long-run technology growth appear to be the main driver behind the recent slowdown in long-run real GDP growth. Panel (b) plots the growth rate of the utilization-adjusted TFP measure by Fernald (2012) in black, together with its long-run counterpart as estimated by our model (green, with blue bands). It is visible that the DFM approach is capable of extracting a smooth low frequency trend from the volatile series of TFP, which captures well-known episodes such as the 1970's slowdown and the IT boom of the 1990's. Overall, our framework is capable of providing an interesting angle on real-time movements in technology trends.

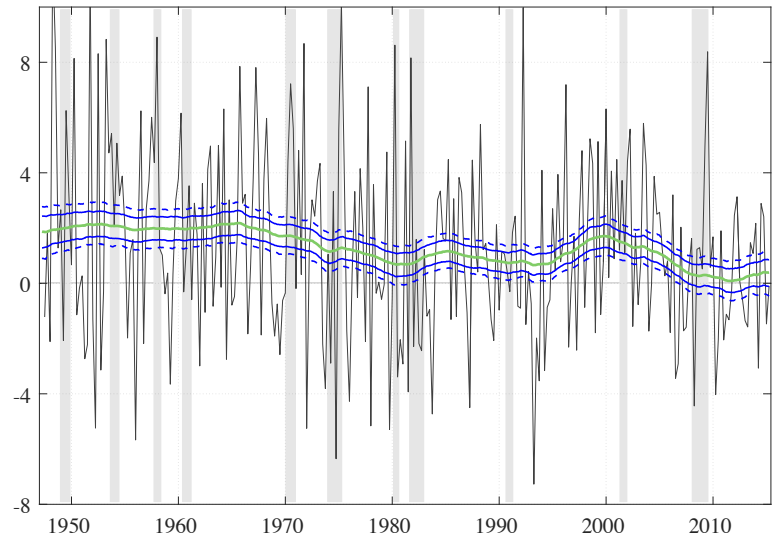
[13]Note here that in the first case we could also directly capture the technological parameter α into the matrix $\mathbf{B}$ by setting its (1,4) and (2,4) elements to α and interpreting $\tilde{z}_t$ directly as capital deepening. The specification above is somewhat more appealing in that it allows for a non-constant capital share. One can easily impose a constant value for α by scaling the posterior estimate of $\tilde{x}_t$ by that value.

Figure K.1: Results of Growth Accounting Exercise

(a) Further Decomposition of US Long-Run US Output Growth



(b) Fernald (2012) TFP Growth: Data and Long-Run Estimate



Note: Panel (a) displays the posterior median estimates of long-run real GDP growth in red, together with the posterior median estimates of its components, long-run hours growth, long-run TFP growth and long-run non-technological growth (orange dashed, green, gray dotted). For long-run real GDP growth the corresponding with corresponding 68% and 90% posterior credible intervals are shown in solid and dashed blue. Panel (b) plots the growth rate of utilization-adjusted TFP by Fernald (2012) in black, together with its long-run counterpart in our econometric framework, i.e. the estimate of $\tilde{z}_t$, with corresponding 68% and 90% posterior credible intervals (green/blue).

References

Alvarez, R., Camacho, M., and Perez-Quiros, G. (2012). Finite sample performance of small versus large scale dynamic factor models. CEPR Discussion Papers 8867, C.E.P.R. Discussion Papers.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–68.

Bai, J. and Wang, P. (2015). Identification and Bayesian Estimation of Dynamic Factor Models. *Journal of Business & Economic Statistics*, 33(2):221–240.

Banbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2012). Now-casting and the real-time data flow. Working Papers ECARES 2012-026, ULB.

Banbura, M. and Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1):133–160.

Beveridge, S. and Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7(2):151–174.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194.

Carter, C. K. and Kohn, R. (1994). On Gibbs Sampling for State Space Models. *Biometrika*, 81(3):541–553.

CBO (2001). CBOs Method for Estimating Potential Output: An Update. Working papers, The Congress of the United States, Congressional Budget Office.

Clark, T. E. and McCracken, M. W. (2012). In-sample tests of predictive ability: A new approach. *Journal of Econometrics*, 170(1):1–14.

Cogley, T. and Sargent, T. J. (2005). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics*, 8(2):262–302.

D'Agostino, A., Giannone, D., Lenza, M., and Modugno, M. (2015). Nowcasting business cycles: A bayesian approach to dynamic heterogeneous factor models. Technical report.

de Wind, J. and Gambetti, L. (2014). Reduced-rank time-varying vector autoregressions. CPB Discussion Paper 270, CPB Netherlands Bureau for Economic Policy Analysis.

Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998). Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review*, 39(4):863–83.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2.

Fernald, J. (2012). A quarterly, utilization-adjusted series on total factor productivity. *Federal reserve bank of San Francisco working paper*, 19:2012.

Giacomini, R. and White, H. (2006). Tests of Conditional Predictive Ability. *Econometrica*, 74(6):1545–1578.

Gneiting, T. and Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102:359–378.

Gneiting, T. and Ranjan, R. (2011). Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules. *Journal of Business & Economic Statistics*, 29(3):411–422.

Groen, J. J. J., Paap, R., and Ravazzolo, F. (2013). Real-Time Inflation Forecasting in a Changing World. *Journal of Business & Economic Statistics*, 31(1):29–44.

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modeling*, 14(4):517–533.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291.

Jacquier, E., Polson, N. G., and Rossi, P. E. (2002). Bayesian Analysis of Stochastic Volatility Models. *Journal of Business & Economic Statistics*, 20(1):69–87.

Kim, C.-J. and Nelson, C. R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. The MIT Press.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies*, 65(3):361–93.

Marcellino, M., Porqueddu, M., and Venditti, F. (2014). Short-term gdp forecasting with a mixed frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics, Forthcoming*.

Mariano, R. S. and Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4):427–443.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, 84(405):223–230.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852.

Stock, J. and Watson, M. (2012). Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, Spring:81–156.