# Empirical Macroeconomics and DSGE Modeling in Statistical Perspective

Daniel J. McDonald[*]
Department of Statistics
University of British Columbia
Vancouver, BC Canada
daniel@stat.ubc.ca

Cosma Rohilla Shalizi
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA USA
cshalizi@cmu.edu

DRAFT: comments welcome; please do not quote or cite

**Abstract**

Write abstract here.

**Keywords:**

## 1 Introduction

Since around 1980, academic macroeconomics has come to be dominated by dynamic stochastic general equilibrium (DSGE) models, and economists have devoted a great deal of attention to their specification, elaboration, mathematical manipulation, estimation, and theoretical refinement. This rise to prominence has certainly not been unopposed, and there has never been a shortage of critics of the whole approach on theoretical grounds, charging that, in one way or another, DSGE models are (or embody, or presume) bad economic theories. TODO: Obviously need citations here, or really a whole lit-review section Rather than pursue the question of whether DSGEs are good *economics*, we look at whether they are good *models*, i.e., whether they meet common-place standards of statistical modeling.

In this paper, we are not primarily concerned with whether DSGE models can predict macroeconomic variables outside of the time period used to estimate their parameters. We avoid this issue not because out-of-sample forecasting isn't a quite elementary test of any statistical model's actual fit to the data (it is), nor because DSGEs do well at it (they are miserable), nor because it's unfair to expect any model to do well (even vector autoregressions do much better). Rather Too many "rathers", we take it for granted that DSGEs are *currently* not useful predictive devices and ask whether they will *ever* be capable of accurate forecasts. By the simple, standard device of simulating a DSGE and then fitting the same model to the simulation output, we show that even when correctly specified in their entirety and provided with centuries of simulated data, these models remain incapable of forecasting. Worse, their parameters remain very badly estimated. To

[*]Both authors were supported by a grant (INO1400020) from the Institute for New Economic Thinking. DJM was partially supported by the National Science Foundation Grants DMS-1407439 and DMS-1753171. CRS wishes to acknowledge support from the National Science Foundation (grants DMS1207759 and DMS1418124), and valuable conversations over many years with Zmarak M. Shalizi.

reliably estimate models of such coelenterate flexibility would require thousands of years (at least) of data from a stationary economy.

We are aware that many economists downplay using DSGEs for prediction on the grounds that the models are instead supposed to inform us about the structure of the economy and about the consequences of prospective policy interventions. Even if this is true, however, one would need some reason to think that *this* DSGE, rather than another, was getting things right. While we agree that models which are capable of accurate statistical prediction may be horrible at causal, counter-factual prediction, the reverse is not true. Cites to counter-factual $\Rightarrow$ unconditional prediction A model which gets the causal structure right, and can make accurate counter-factual predictions, should *a fortiori* be capable of accurate statistical prediction as well. Hence economists' confidence in their favorite DSGEs suitability for policy evaluation cannot be rooted in their statistical predictive ability, since the later does not exist.

We also undermine the notion that DSGE specifications capture the structure of the economy by the simple expedient of swapping the different time series on which they are fit — giving the model as "investment" the series that is really "hours worked," and so forth. Not only does such series swapping do little to degrade the DSGE's performance, in sample or out of sample, in a large fraction of permutations it actually *enhances* predictability. It is, of course, open to an economist to maintain their faith in a favorite DSGE's capturing the structure of the macroeconomy even if it cannot predict and cannot tell the difference between the real data and one with all the series swapped, but then faith truly is maintained on the evidence of things not seen.

The principles to which we appeal — using simulation to assess estimation methods; using randomization to gauge how well a model can appear to fit nonsense data — are neither recondite points of mathematical theory nor controversial questions in the foundations of statistics. Rather, they are readily explained notions, accepted on pretty much all sides within modern statistics, whose force is easily grasped once they are presented. We return to their implications for macroeconomic modeling in the conclusion.

# 2 Specification and Baseline Estimation of the DSGE

I hate having no text in a section prior to the subsection.

## 2.1 Solving DSGEs

A general DSGE model is given by a solution to a constrained stochastic inter-temporal optimization problem

$$z_t^* = \operatorname{argmax} \sum_{t=0}^{\infty} \mathbb{E} g(z_t) \qquad \text{s.t.} \qquad z_t = h(z_{t-1}), \tag{1}$$

for some nonlinear functions $g$ and $h$ parametrized by a $k$-dimensional vector of "deep" parameters $\theta$. The economic agents posited by the model are assumed to solve this problem (optimally) at each time $t$ conditional on all current and previously available information, and we observe part of the solution: that is the observable data are $x_t$ which is a subset of the indices of $z_t^*$.

To estimate a DSGE, the first step is to solve the optimization problem by deriving the first-order conditions for an optimum. The resulting nonlinear system can be written as

$$\Phi(z_t,\ z_{t+1}) = 0, \tag{2}$$

**Algorithm 1** Pseudoalgorithm for estimating linear rational expectations models

```
Write down DSGE as a constrained optimization problem.
Determine the first order conditions for an optima.
Determine steady state path of observables.
Linearize the DSGE around the steady-state yielding a general form for (3).
procedure ESTIMATE THE MODEL(θ)
    Fix a plausible parameter vector θ.
    For this θ, from (3) into state-space form using the method of (Sims, 2002).
    Using equations (4) and (5), evaluate the likelihood of θ using the Kalman filter.
    Maximize the likelihood or explore the posterior, repeating as needed at new values of θ.
end procedure
```

for some $n$-dimensional function $\Phi$. Such systems can rarely be solved analytically for the optimal path; instead, the first-order conditions are used to express the model in terms of stationary variables, and the system is linearized around this steady-state (that is, let $z$ be the vector satisfying $z = z_t = z_{t-1}$). Upon writing $z_t = z_t^* - z$, the (linearized) DSGE can be written as

$$\Gamma_0 z_t = \Gamma_1 z_{t-1} + A + B\epsilon_t + C\eta_t \tag{3}$$

in the notation of (Sims, 2002). Here $\epsilon_t$ are exogenous, possibly serially correlated, random disturbances and $\eta_t = z_t - \mathbb{E}_t z_{t+1}$ are expectational errors determined as part of the model solution. The matrices $\Gamma_0$, $\Gamma_1$, $A$, $B$, and $C$ are functions of $\theta$.

There are a number of ways to estimate $\theta$ using data; a complete treatment is beyond our scope here, but see DeJong and Dave (2011) for details. We will focus on likelihood-based approaches, and so we must solve the model in equation (3). There are many approaches to solving linear rational expectations models (e.g. Blanchard and Kahn, 1980; Klein, 2000), but we will use that in Sims (2002) due to its ubiquity. This method essentially uses a QZ factorization to remove $\Gamma_0$ from the left side of the equation while correctly handling explosive components of the model (multiplying through by the generalized inverse of $\Gamma_0$, $\Gamma_0^\dagger$, can lead to portions of the product $\Gamma_0^\dagger\Gamma_1$ implying nonstationarity). Following this procedure, we retrieve a system of the form

$$z_t = d + Tz_{t-1} + H\varsigma_t \tag{4}$$

as long as there is a unique mapping from equation (3) (there may be multiple solutions or none depending on $\theta$). Since some of the $z_t$ are unobserved, we augment the transition equation in (4) with an observation equation

$$x_t = Zz_t, \tag{5}$$

where the matrix $Z$ subselects the appropriate elements of $z_t$. Collecting equations (4) and (5) gives the form of a linear state-space model. Assuming that the errors $\varsigma_t$ are serially independent multivariate Gaussian allows us to evaluate the likelihood of some parameter vector $\theta$ given observed data $x_1, \ldots, x_T$. Evaluating the likelihood can be done using the Kalman filter (Kalman, 1960) which is readily available in most software packages. The procedure outlined above is summarized in Algorithm 1.

## 2.2 Estimating the Smets and Wouters (2007) model

In this section, we provide a description of the procedure we use for estimating the Smets and Wouters (2007) model. The model is now standard in the macroeconomic forecasting literature, and, though code is readily available (for example in `Dynare` or `Matlab`[1]), our version is implemented fully in `R`. A complete description of the economic implications of the model and its log-linearized form can be found in (Smets and Wouters, 2007) as well as Iskrev (2009) and other sources.

For this model, we use seven observable data series: output growth, consumption growth, investment growth, real wage growth, inflation, hours worked, and the nominal interest rate, which we collect into the vector $x_t$

$$x_t = \begin{bmatrix} y_t - y_{t-1} & c_t - c_{t-1} & i_t - i_{t-1} & w_t - w_{t-1} & \pi_t & l_t & r_t \end{bmatrix}^\top . \tag{6}$$

We describe the specific data and preprocessing routines in Appendix A, but we note that all the data are publicly available from the Federal Reserve Economic Database FRED. We use data from the first quarter of 1956 until the fourth quarter of 2018. Following preprocessing, we are left with 251 available time points.

The model has 52 "deep" parameters as well as 7 parameters representing the standard deviations of the stochastic shocks. Of these 59 total parameters, we estimate 36: 18 are derived from steady-state values as functions of other parameters and 5 are fixed a priori (as in Smets and Wouters 2007). The prior distributions for the 36 estimated parameters are given in Table 1. To estimate the model we minimize the negative log likelihood, penalized by the prior. This is the same as finding the maximum *a posteriori* estimate in a Bayesian setting. Because the likelihood is ill-behaved, having many flat sections as well as local minima, we used `R`'s `optimr` package. We estimated the parameter using both the simulated annealing method, which stochastically explores the likelihood surface in a principled manner, and the conjugate gradient technique. Each procedure was started at 5 random initializations (drawn from the prior distribution) and run for 50,000 iterations (likelihood evaluations) for each starting point. We train the model using only the first 200 time points, saving the remainder to evaluate the model's (pseudo-out-of-sample) predictive performance.

Table 1 presents the posterior mode based on our procedure. Note first that some of the parameter estimates are similar to those presented in (Smets and Wouters, 2007) (shown in the last column), while others differ dramatically. However, comparing the likelihood of of our estimated parameters to those in (Smets and Wouters, 2007), our fit is significantly better. For our dataset, the penalized negative log likelihood of the parameters is 1145 compared to 1232, an improvement of more than 7.1%. The result is similar for the unpenalized negative log likelihood. To check for robustness, we also ran the optimization procedure for 1 million parameter draws and the likelihood only decreased by about 1%. We are therefore confident that we have a parameter combination which nearly achieves the global optimum.[2]

# 3 Permuting the data

One way to assess the predictive ability (and economic content) of the Smets and Wouters (2007) DSGE model is to perform a simple permutation test. That is, rather than giving the model data in the order it expects, we permute the data series and see if the model predicts future data any better. We estimated the model on the properly ordered data (presented in the previous section) as well as all 5039 other permutations of the 7 data series. For each estimation, we used the same estimation procedure as before to minimize the (penalized) negative log likelihood. The model is

[1]See also `https://www.aeaweb.org/articles.php?doi=10.1257/aer.97.3.586&fnd=s`

[2]We have been unable to explain why Smets and Wouters (2007) gives such a different estimate of the posterior mode. It is worth pointing out however, that their MCMC is not necessarily attempting to maximize the posterior, although it may locate one.

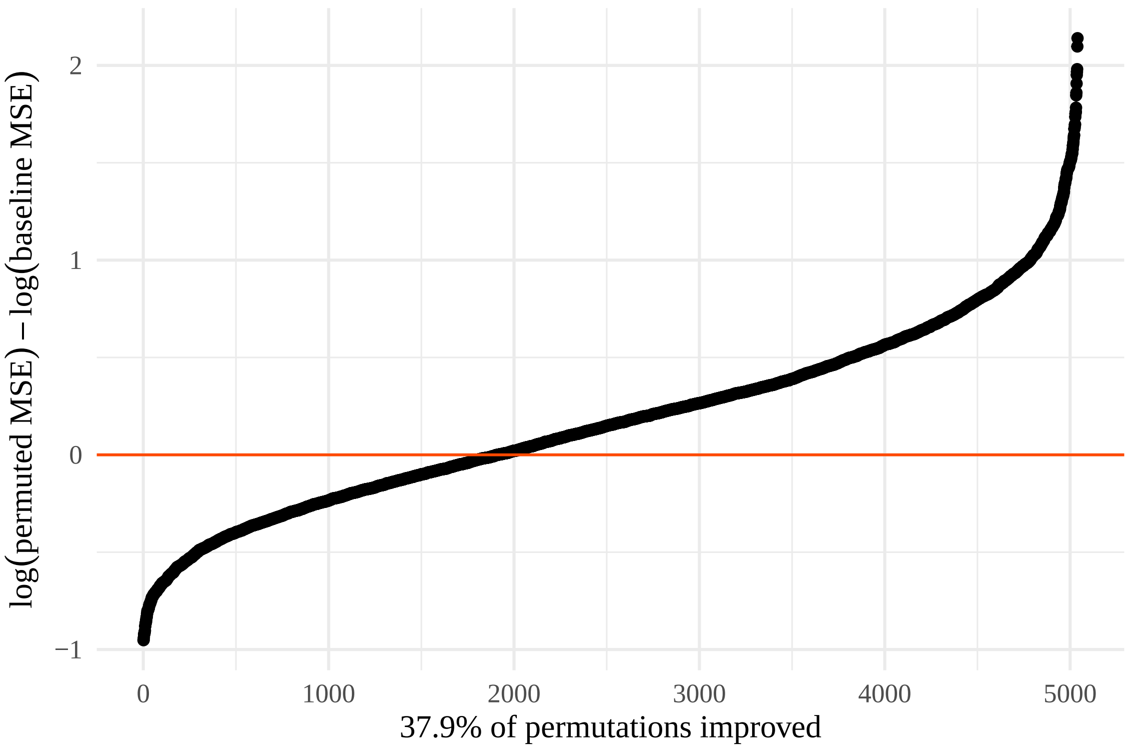


Figure 1: Average percentage improvement in out-of-sample forecast MSE. The horizontal red line represents baseline performance. Lower values are better.

trained using the first 200 time points and its predictive performance is tested on the remaining 51.

The next few figures summarize the results. We report three criteria for measuring relative performance. The first is the percent improvement in (out-of-sample) mean-squared test error (MSE) calculated as the natural logarithm of the test error for a particular permutation divided by that of the baseline model and then averaged across all 7 series:

$$\text{Average Percent MSE Improvement (p)} = \frac{1}{7}\sum_{i=1}^{7}\log\left(\frac{\sum_{t=201}^{251}(\widehat{x}_{it}^{(p)} - x_{it})^2}{\sum_{t=201}^{251}(\widehat{x}_{it}^{b} - x_{it})^2}\right), \tag{7}$$

where a superscript $b$ represents the baseline model and $(p)$ the $p^{th}$ permutation. The result is shown in Figure 1. Figure 2 shows boxplots for the percentage improvement separately for each time series. The best model had an average percentage improvement of 95% relative to the baseline model. The horizontal line at zero represents performance equivalent to baseline. About 38% of permutations had better predictive performance than the baseline model.

The second measure of performance is simply the out-of-sample MSE scaled by the observed variance and then averaged across the 7 series:

$$\text{Average Scaled MSE (p)} = \frac{1}{7}\sum_{i=1}^{7}\frac{\sum_{t=201}^{251}(\widehat{x}_{it}^{(p)} - x_{it})^2}{\text{Var}(x_i)}. \tag{8}$$

Figure 3 displays the average of this measure across all 251 series. Again, about 51% of permutations achieved better average scaled MSE than the baseline model. The best permutation achieved an

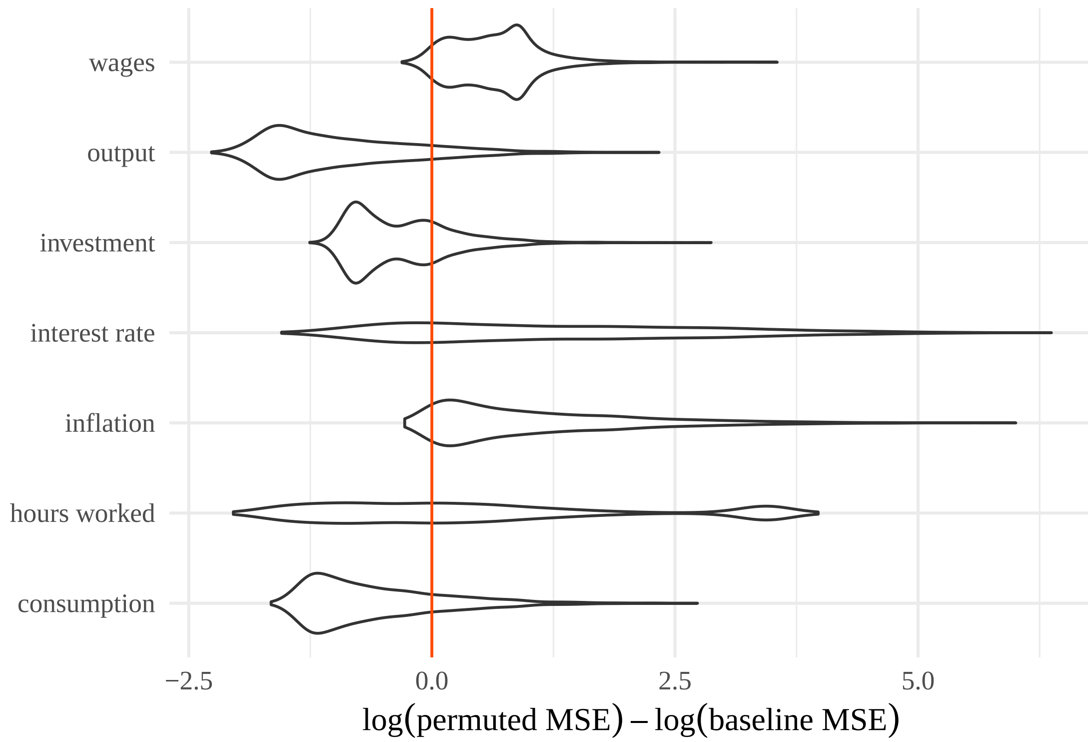


Figure 2: Percentage improvement in out-of-sample forecast MSE for each data series individually. The horizontal red line represents baseline performance. Values to the left are better.

average scaled MSE of 0.53 relative to 1.39 for the baseline model. Figure 4 shows boxplots for the scaled MSE separately for each series. Red dots indicate the performance of the baseline model.

About 0% of permutations resulted in a lower negative penalized log likelihood. A Bayesian interpretation of these results would be that 0% of the permuted "models" are preferable to the true, unpermuted, economic model. It should be noted that this is an in-sample measure of performance and that the Bayesian interpretation is conditional on the model being true and the priors accurately reflecting expert information. An out-of-sample evaluation (without penalty) shows that 8% of the premuted models actually have better prediction performance when evaluated through the likelihood. Figure 5 gives a visual depiction.

Table 2 shows the error measures separately for each series of best permutations (as measured by out-of-sample performance) relative to that of the model fit to the true data. Note that the best permutation is different for the two measures. Tables 3–5 show the 20 best permutations for each evaluation method. There are few commonalities across the best models. One aspect to note is that output shows up in many places *except* where it is supposed to go. Many models perform better with wages in that position. It is unclear whether this is because the model is exceptionally bad at predicting output, whether that slot likes to predict series with small variance, or whether that slot is simply over regularized. Note that hours worked, and to some extent, the interest rate, are the only series that appear consistently in the correct places.

Figure 6 shows boxplots for the parameter estimates across permutations scaled to $[0, 1]$ by the prior range. Red dots indicate the SW model estimates. Clearly, some parameters change very little from permutation to permutation while others change massively. Figure 7 displays the same parameter estimates as percent deviations from the SW model estimates.

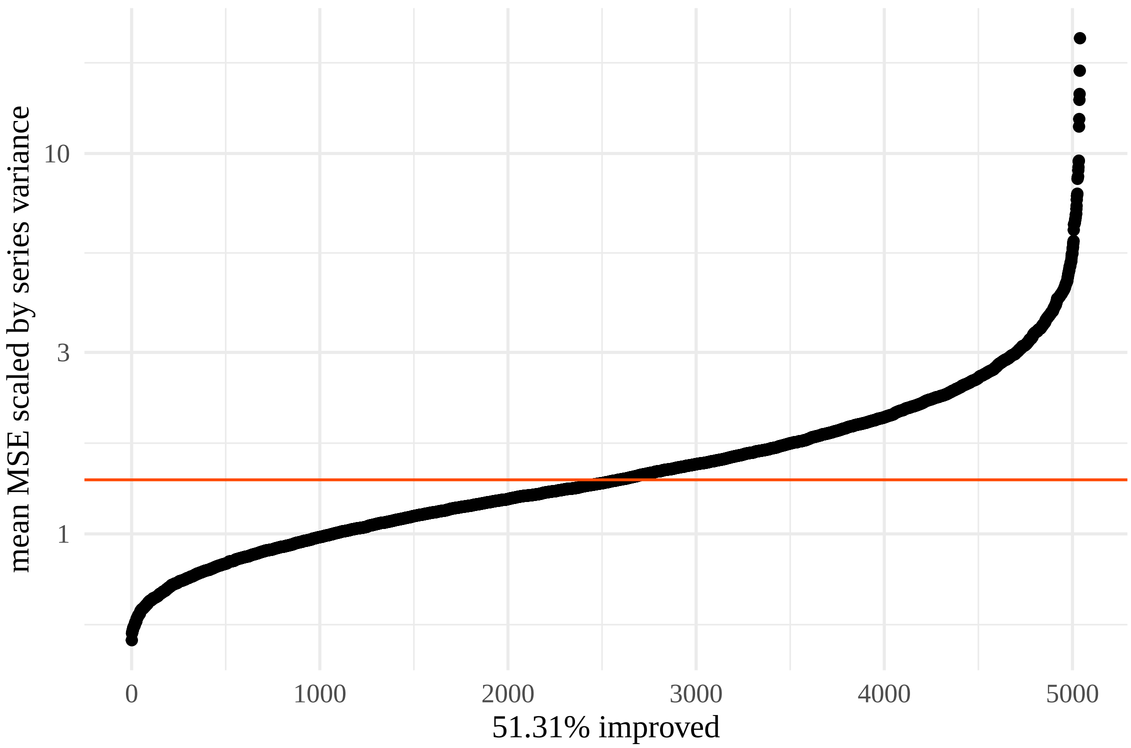


Figure 3: Average improvement in out-of-sample forecast MSE scaled by the variance of the data. The horizontal red line represents baseline performance.

### 3.1 Out-of-sample forecasts

We now examine how well the SW model predicts future data relative to the best models we could have used, had we seen the data. Figure 8 and Figure 9 show the out-of-sample predictions for top 20 flips based on "average percent improvement." This analysis is a post-hoc measure as the best models were selected to make these predictions well, though the parameters were estimated without access to this data. We also show the observed data and the predictions from the SW model. The SW model is quite bad at predicting consumption, investment, output, and wages. It predicts a consistent 0.5% increase in the real wage level that never appears. In fact, wages are far more volatile than any of the permutations can account for. For the case of investment, output, and consumption, the SW model drastically lags the 2009 recession and underestimates its severity. Furthermore, it continues to predict a much stronger recovery, even 10 years later, than has ever materialized. The SW model is quite accurate for the interest rate, though this should perhaps be expected given that the Taylor rule may well drive Federal Reserve decisions rather than describe them.

## 4 Simulate and estimate

Another simple method for evaluating the forecasting ability of a stochastic model is to simulate a long time series from the model and assess how well we can estimate the generating process given more and more data. In a sense, this is a minimal requirement for any model: does it produce consistent estimates of the true parameters? And furthermore, how long will we need? Statistical theory for independent and identically distributed data says that maximum likelihood estimators

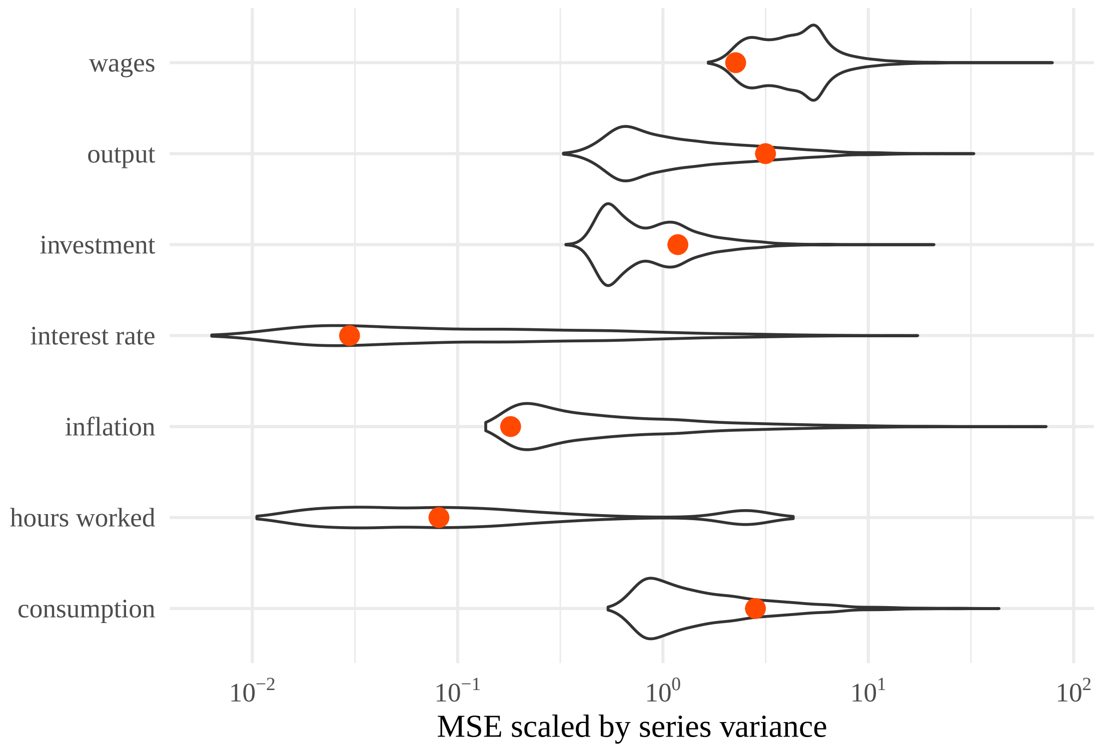


Figure 4: Improvement in out-of-sample forecast MSE scaled by the variance of the data for each data series individually. The red dots represents baseline performance.

for a fixed number of parameters are $n$-consistent in MSE. This is essentially as fast as we could hope. In this exercise, we abstract from the non-linear DSGE: suppose that the data were actually generated by the linearized DSGE represented by equations (4) and (5), can we recover the "deep" parameters which generated the data?

To evaluate this (relatively simple) question, we generate data using using the parameters estimated in the previous section and presented in Table 1. To avoid any dependence on initial conditions, we simulate 3100 data points and discard the first 1000. We then train the model using the first 100 data points (25 years worth of quarterly data), and try to predict the next 1000 to measure out-of-sample performance. We then increment forward 20 time-steps (5 years), train again and predict the next 1000 observations. We continue this process until we are using 1100 data points (275 years of data) to estimate the model. To average over simulation error, we repeat the entire exercise 100 times. For each estimation, we initialize the optimization procedure at the true parameter value and run it for 30000 iterations. This ensures that the true parameters are considered, maximizing the chance that the estimation will return the data generating values. Our goal here is to learn whether we could hope to learn the right parameters in 275 years-time if the linearized DSGE is actually true, and furthermore, how well can we predict future economic movements under this idealized scenario. All figures in this section show the mean in red along with 30%, 60%, 80%, and 90% confidence bands calculated from the replications.

Figure 10 shows the average training error (averaged across the 7 series). As we would expect, the variability declines as the size of the training set increases, though not the average. Figure 11 shows the average prediction error over the 7 series. It improves markedly as the training set increases to about 400 observations but then plateaus. This is troubling: as we get more and more data, we can not predict new data any better. This indicates one of three possibilities: (1) that

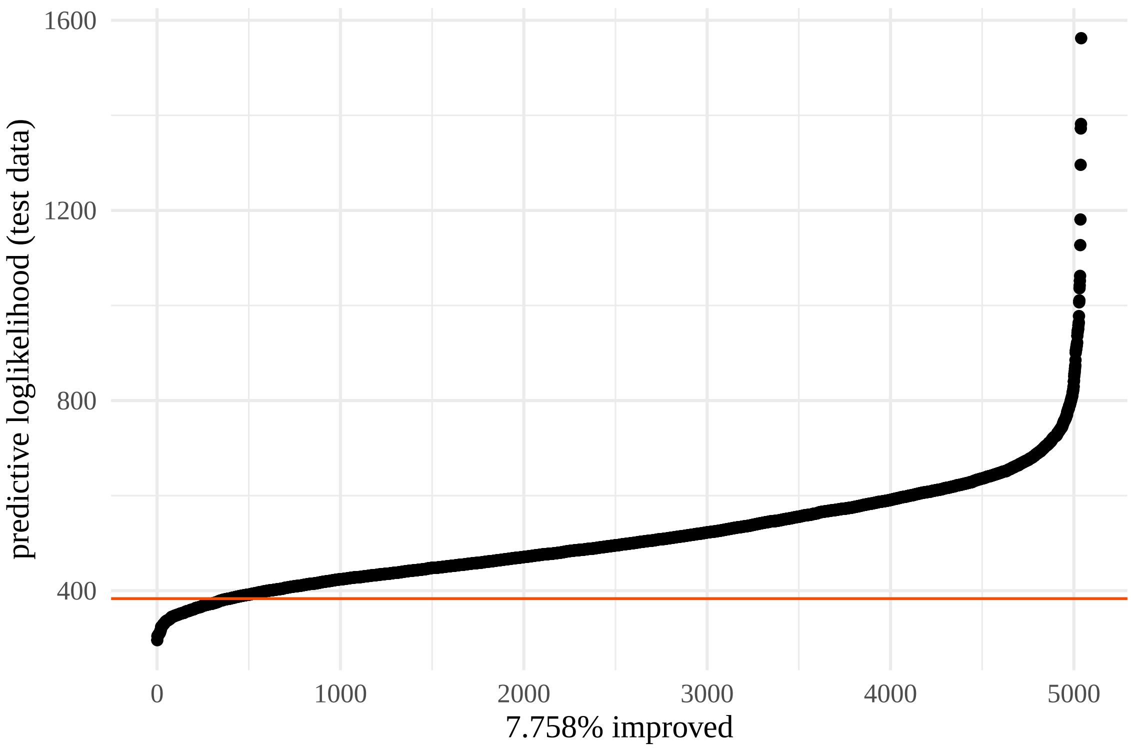


Figure 5: Negative log predictive likelihood for each model. The horizontal red line represents baseline performance. Note that this is an out-of-sample performance measure.

with about 400 observations, we can estimate the parameters nearly perfectly, (2) that the model is poorly identified—some parameters will simply never be well estimated, but we can predict well anyway, or (3) the data are so highly correlated that the range of training observations we consider is far too small—we actually need millions of observations in order to see a meaningful decline in out-of-sample predictive performance. We can better determine which of these three is occurring by comparing with the predictive performance of the true parameters and examining the error in parameter estimates. The red lines in Figure 11 are the out-of-sample mean (solid) and median (dashed) prediction error for the true parameters. The test error is not getting any closer to this ideal scenario, plateauing slightly above the baseline by about 400 training points. This seems to suggest that explanation (2) is the correct one: even with more data, we will never be able to recover the true parameters, though we can predict adequately relatively quickly.

Examining the parameter error serves to confirm this hypothesis. Figure 12 shows the average squared error of the parameter estimates. Improvement in this metric also stagnates despite the size of the training set increasing. Figure 13 and Figure 14 show the change in information per observation (in sample) and per prediction (out of sample). For maximum likelihood inference, we would expect both of these to decline as the amount of training data increases. These figures therefore confirm that the estimation procedure is responding to the increased sample size even though prediction and estimation errors both fail to improve meaningfully. Figure 15 shows the average prediction error for each series individually. From this decomposition, prediction of all time series improves, but only labor ever reaches the minimum (relative to the truth). The other series are all predicted poorly.

Figures 16–20 show the estimates of each parameter individually. The horizontal red lines indicate the truth. Here we can see that while many parameters are well-estimated with decreasing variability as the number of training points increases, others converge to the wrong values, or even fail to converge at all. This phenomenon holds especially for both the parameters of the shock processes, but also for some of the economically-relevant "deep" parameters. For instance, $\sigma_l$, the elasticity of labor supply with respect to the real wage , is consistently underestimated by about -93%. The data provides essentially no information about $\iota_w$, which measures the dependence of real wages on lagged inflation. Other parameters which are poorly estimated include $\varphi$, the steady-state In all these cases, estimation is biased, so using the estimated values from the real data to draw conclusions about the real economy is unwise. It is possible that this bias is due to the linearization procedure, in which case, not only should linearized models be avoided for prediction, but also for drawing any economic inferences. Overall, most parameters are poorly identified as evidenced by their stable (rather than decreasing) variability with more data.

# 5 Things Waldman asked which haven't appeared yet

## 5.1 Does the best (permuted) model swap hours worked?

Examining Tables 3–5, the answer is "not really." Hours worked seems to be in the correct slot frequently. When swapped, it tends to be confused with the interest rate (a flow variable).

## 5.2 Do the deep parameters covary with policy parameters?

The idea here is to assess whether the Lucas critique is really avoided for this model. In other words, if the Fed changes how they manage interest rates, do the other parameters move? Here, we're looking only at the simulate/estimate exercise, examining the distribution of Taylor parameters conditional on the truth generating the model and estimating said parameters. We're not actually considering a distribution of Taylor rule parameters that the Fed might consider. These correlations are shown in Table 6. The Taylor rule parameters are, $\rho$, the autocorrelation in the interest rate, $r_\pi$, the response to inflation, $r_y$, the response to deviations of output from potential, and $r_{\Delta y}$ the response to changes in the deviation. A handful of deep parameters have correlations larger in magnitude than 0.3 (shown in bold).

## 5.3 Changing the truth

What if the permuted model with the highest penalized likelihood were The Truth? How badly would the SW model fit and forecast? Here I actually guess it would do OK. Procedure: (1) generate data out of the best permutation, (2) get forecasts using it and the SW model (say 100 each). Which is better? The average negative log-likelihood of the true model is 875 compared to 953 for the SW model. The MSEs are 0.257 compared to 0.579 respectively.

We actually have that the SW model (unpermuted) has the lowest in-sample negitive log posterior. These were generated using the best unpenalized predicting permutation by log likelihood.

## 5.4 Best model by penalized log-likelihood

The true model is

hours worked, interest rate, inflation, output, consumption, investment, wages

while the best one is

hours worked, interest rate, output, wages, inflation, investment, consumption.

This model is pretty clearly scary to a macroeconomist. I think it's also important to note here, that the penalized negative log-likelihood is very flat relative to permutations. In terms of that metric, 0.22% of permutations are within 10% of the true permutation. In the Appendix, we rewrite the beginning of Smets and Wouters (2007) permuting all the series to reflect the best one by negative log likelihood.

## 6 Conclusions

As we said in the introduction, there are very few who will defend the forecasting record of DSGEs. Rather, their virtues are supposed to lie in their capturing the structure of the economy, and so providing theoretical insight, with meaningful parameters, and an ability to evaluate policy and counterfactuals. We have examined these claims on behalf of DSGEs through series permutation, and through the simple check of seeing how well a DSGE can be estimated from its own simulation output. In both cases, the results are rather negative.

If we randomly re-label the macroeconomic time series and feed them into the DSGE, most of the time we get a model which predicts the (permuted) data *better* than the model predicts the unpermuted data. Even if one disdains forecasting as end in itself, it is hard to see how this is at all compatible with a model capturing something — anything — essential about the structure of the economy[3]. Perhaps even more disturbing, many of the parameters of the model are essentially unchanged under permutation, including "deep" parameters supposedly representing tastes, technologies and institutions.

If we take our estimated model and simulate several centuries of data from it, all in the stationary regime, and then re-estimate the model from the simulation, the results are no more comforting. Forecasting error remains dismal and shrinks very slowly with the size of the data. Much the same is true of parameter estimates, with the important exception that many of the parameter estimates seem to be stuck around values which differ from the ones used to generate the data. These ill-behaved parameters include not just shock variances and autocorrelations, but also the "deep" ones whose presence is supposed to distinguish a micro-founded DSGE from mere time-series analysis. We emphasize that all this is happening in simulations where the model specification is correct, where the parameters are constant, and where the estimation can make use of centuries of stationary data, far more than is available about the actual macroeconomy.

The results of the two tests are grim. Series swapping gives us strong reasons to doubt that the elaborate DSGE machinery manages to capture anything important about the structure of the economy. Even if one dismisses that, and believes (perhaps because "theory is evidence too") that the DSGE must be right, the simulation exercise shows that, even under the most favorable possible circumstances, it is simply infeasible to hope to that the DSGE will give reasonably accurate predictions, or even that it can be reliably estimated.

We do not, of course, have a proof that flaws like this are inherent in the DSGE form. But we emphasize that we have not cherry-picked an obsolete or marginal model[4]. The SW model

[3]It's conceivable that the issue is one of measurement. The variables in the model are defined by their roles in the inter-temporal optimization problem. (Thus labor makes a negative contribution to present utility but a positive contribution to present output, cannot be stored from period to period, etc.) The series gathered by the official statistical agencies use quite distinct definitions, and nothing guarantees that the series with analogous names are good measurements of the theoretical variables. Thus, perhaps, the model is right, but GDPDEF is really a better proxy for *labor* than PRS85006023 is. This would, needless to say, raise its own set of huge problems for the interpretation and use of these models.

[4]We also obtained very similar results for the real-business-cycle model of Kydland and Prescott (1982), but omit them here because that model is obsolete.

is widely regarded as the baseline DSGE for the economy of the United States, which is by far the most important national economy in the world. The SW paper has been cited, as over early October 2019, over 4500 times[5]. What concerns us is that in all that literature, we appear to be the first to have subjected it to such direct, even elementary, tests. We do not assert that all DSGE models must be pathological in the ways we have shown the SW model is. Indeed, readers are free to hope that their favorite DSGE does capture economic structure and can be meaningfully estimated with reasonable amounts of data. But we hope we have persuaded readers that they can, and should, do more than hope: they can *check*.

[5] wrong link Google Scholar (`https://scholar.google.com/scholar?cites=8854430771281116653`), accessed 7 October 2019.

Table 1: Prior distributions, posterior modes, and posterior mode as estimated by (Smets and Wouters, 2007) for the 36 estimated parameters. All values are rounded to two decimal places.

| | prior distribution | prior mean | prior stdev | lower bound | upper bound | posterior mode | SW posterior mode |
|---|---|---|---|---|---|---|---|
| $\sigma_a$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.44 | 0.45 |
| $\sigma_b$ | igamma | 0.10 | 2.00 | 0.02 | 5.00 | 0.30 | 0.24 |
| $\sigma_g$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.57 | 0.52 |
| $\sigma_I$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.48 | 0.45 |
| $\sigma_r$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.22 | 0.24 |
| $\sigma_p$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.15 | 0.14 |
| $\sigma_w$ | igamma | 0.10 | 2.00 | 0.01 | 3.00 | 0.27 | 0.24 |
| $\rho_a$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.97 | 0.95 |
| $\rho_b$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.14 | 0.18 |
| $\rho_g$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.95 | 0.97 |
| $\rho_I$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.66 | 0.71 |
| $\rho_r$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.13 | 0.12 |
| $\rho_p$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.99 | 0.90 |
| $\rho_w$ | beta | 0.50 | 0.20 | 0.00 | 1.00 | 0.95 | 0.97 |
| $\mu_p$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.88 | 0.74 |
| $\mu_w$ | beta | 0.50 | 0.20 | 0.01 | 1.00 | 0.90 | 0.88 |
| $\phi_1$ | gaussian | 4.00 | 1.50 | 2.00 | 15.00 | 5.73 | 5.48 |
| $\sigma_c$ | gaussian | 1.50 | 0.38 | 0.25 | 3.00 | 1.55 | 1.39 |
| $h$ | beta | 0.70 | 0.10 | 0.00 | 0.99 | 0.72 | 0.71 |
| $\xi_w$ | beta | 0.50 | 0.10 | 0.30 | 0.95 | 0.78 | 0.73 |
| $\sigma_l$ | gaussian | 2.00 | 0.75 | 0.25 | 10.00 | 2.01 | 1.92 |
| $\xi_p$ | beta | 0.50 | 0.10 | 0.50 | 0.95 | 0.60 | 0.65 |
| $\iota_w$ | beta | 0.50 | 0.15 | 0.01 | 0.99 | 0.34 | 0.59 |
| $\iota_p$ | beta | 0.50 | 0.15 | 0.01 | 0.99 | 0.26 | 0.22 |
| $\Psi$ | beta | 0.50 | 0.15 | 0.01 | 1.00 | 0.57 | 0.54 |
| $\Phi$ | gaussian | 1.25 | 0.12 | 1.00 | 3.00 | 1.13 | 1.61 |
| $r_\pi$ | gaussian | 1.50 | 0.25 | 1.00 | 3.00 | 2.06 | 2.03 |
| $\rho$ | beta | 0.75 | 0.10 | 0.50 | 0.98 | 0.83 | 0.81 |
| $r_y$ | gaussian | 0.12 | 0.05 | 0.00 | 0.50 | 0.10 | 0.08 |
| $r_{\Delta y}$ | gaussian | 0.12 | 0.05 | 0.00 | 0.50 | 0.21 | 0.22 |
| $\overline{\pi}$ | gamma | 0.62 | 0.10 | 0.10 | 2.00 | 0.63 | 0.81 |
| $100(\beta^{-1}-1)$ | gamma | 0.25 | 0.10 | 0.01 | 2.00 | 0.13 | 0.16 |
| $\bar{l}$ | gaussian | 0.00 | 2.00 | -10.00 | 10.00 | 3.34 | -0.10 |
| $\overline{\gamma}$ | gaussian | 0.40 | 0.10 | 0.10 | 0.80 | 0.47 | 0.43 |
| $\rho_{ga}$ | gaussian | 0.50 | 0.25 | 0.01 | 2.00 | 0.63 | 0.52 |
| $\alpha$ | gaussian | 0.30 | 0.05 | 0.01 | 1.00 | 0.28 | 0.19 |

Table 2: Series MSEs for SW and top models.

| | hours worked | interest rate | inflation | output | consumption | investment | wages |
|---|---|---|---|---|---|---|---|
| SW model | 1.12 | 0.02 | 0.06 | 2.32 | 1.36 | 5.40 | 1.06 |
| Model w/ highest % improvement | 0.24 | 0.01 | 0.06 | 0.39 | 0.34 | 2.64 | 0.84 |
| Model w/ lowest scaled MSE | 0.25 | 0.02 | 0.09 | 0.33 | 0.41 | 1.56 | 0.97 |
| Model w/ lowest out-of-sample likelihood | 0.32 | 0.01 | 0.09 | 0.42 | 0.35 | 2.07 | 1.33 |

Table 3: Permutations with highest % improvement

| hours worked | interest rate | inflation | output | consumption | investment | wages | # different |
|---|---|---|---|---|---|---|---|
| investment | hours worked | interest rate | wages | output | inflation | consumption | 7 |
| hours worked | interest rate | consumption | wages | inflation | output | investment | 5 |
| hours worked | interest rate | output | wages | inflation | consumption | investment | 5 |
| hours worked | investment | inflation | wages | interest rate | consumption | output | 5 |
| hours worked | interest rate | wages | output | inflation | consumption | investment | 4 |
| consumption | hours worked | interest rate | wages | output | investment | inflation | 6 |
| hours worked | investment | interest rate | wages | inflation | output | consumption | 6 |
| inflation | hours worked | interest rate | output | consumption | investment | wages | 3 |
| consumption | hours worked | inflation | wages | interest rate | output | investment | 6 |
| inflation | hours worked | interest rate | consumption | output | investment | wages | 5 |
| hours worked | interest rate | wages | consumption | inflation | output | investment | 5 |
| investment | hours worked | inflation | output | interest rate | consumption | wages | 4 |
| hours worked | inflation | consumption | output | wages | investment | interest rate | 4 |
| consumption | hours worked | interest rate | wages | inflation | output | investment | 7 |
| hours worked | inflation | interest rate | investment | output | consumption | wages | 5 |
| hours worked | inflation | output | consumption | wages | investment | interest rate | 5 |
| inflation | hours worked | interest rate | output | investment | wages | consumption | 6 |
| hours worked | inflation | consumption | wages | interest rate | output | investment | 6 |
| investment | hours worked | interest rate | wages | consumption | inflation | output | 6 |
| inflation | hours worked | interest rate | investment | output | wages | consumption | 7 |

Table 4: Permutations with lowest average scaled MSE

| hours worked | interest rate | inflation | output | consumption | investment | wages | # different |
|---|---|---|---|---|---|---|---|
| hours worked | inflation | consumption | wages | interest rate | output | investment | 6 |
| hours worked | interest rate | output | wages | inflation | consumption | investment | 5 |
| hours worked | interest rate | consumption | wages | inflation | output | investment | 5 |
| hours worked | investment | consumption | wages | interest rate | output | inflation | 6 |
| hours worked | interest rate | wages | output | inflation | consumption | investment | 4 |
| consumption | hours worked | inflation | wages | interest rate | output | investment | 6 |
| hours worked | interest rate | wages | consumption | inflation | output | investment | 5 |
| hours worked | inflation | consumption | output | interest rate | investment | wages | 3 |
| hours worked | output | inflation | wages | interest rate | consumption | investment | 5 |
| inflation | hours worked | consumption | wages | output | interest rate | investment | 7 |
| hours worked | investment | consumption | wages | output | interest rate | inflation | 6 |
| hours worked | investment | output | wages | consumption | interest rate | inflation | 5 |
| investment | hours worked | inflation | output | interest rate | consumption | wages | 4 |
| hours worked | inflation | interest rate | wages | consumption | output | investment | 5 |
| consumption | hours worked | interest rate | wages | inflation | output | investment | 7 |
| hours worked | consumption | interest rate | wages | inflation | investment | output | 5 |
| consumption | hours worked | inflation | wages | output | interest rate | investment | 6 |
| hours worked | consumption | inflation | wages | interest rate | output | investment | 5 |
| hours worked | output | interest rate | consumption | inflation | investment | wages | 4 |
| hours worked | investment | output | wages | inflation | interest rate | consumption | 6 |

Table 5: Permutations with lowest negative log likelihood

| hours worked | interest rate | inflation | output | consumption | investment | wages | # different |
|---|---|---|---|---|---|---|---|
| hours worked | interest rate | output | wages | inflation | investment | consumption | 4 |
| hours worked | interest rate | consumption | wages | inflation | investment | output | 4 |
| hours worked | interest rate | consumption | wages | inflation | output | investment | 5 |
| investment | hours worked | interest rate | wages | output | inflation | consumption | 7 |
| interest rate | hours worked | inflation | wages | consumption | investment | output | 4 |
| inflation | hours worked | interest rate | consumption | output | investment | wages | 5 |
| inflation | hours worked | interest rate | wages | output | investment | consumption | 6 |
| inflation | hours worked | interest rate | wages | consumption | investment | output | 5 |
| inflation | hours worked | interest rate | output | consumption | investment | wages | 3 |
| hours worked | inflation | interest rate | investment | output | consumption | wages | 5 |
| hours worked | consumption | interest rate | wages | inflation | investment | output | 5 |
| hours worked | interest rate | output | wages | inflation | consumption | investment | 5 |
| hours worked | inflation | interest rate | wages | consumption | investment | output | 4 |
| consumption | hours worked | interest rate | wages | output | investment | inflation | 6 |
| hours worked | output | interest rate | wages | inflation | investment | consumption | 5 |
| interest rate | hours worked | inflation | wages | output | investment | consumption | 5 |
| hours worked | consumption | output | wages | inflation | investment | interest rate | 5 |
| interest rate | hours worked | inflation | wages | output | consumption | investment | 6 |
| inflation | hours worked | interest rate | wages | investment | consumption | output | 7 |
| hours worked | inflation | consumption | wages | interest rate | output | investment | 6 |

Table 6: Correlations between 'deep' parameters and Taylor rule parameters. From the 'Simulate and estimate' exercise.

| | $r_\pi$ | $\rho$ | $r_y$ | $r_{\Delta y}$ |
|---|---|---|---|---|
| $\phi_1$ | 0.2 | 0.1 | -0.03 | 0.11 |
| $\sigma_c$ | 0.06 | 0.26 | 0.17 | 0.02 |
| $h$ | **0.33** | 0.08 | 0.02 | **-0.34** |
| $\xi_w$ | -0.06 | **0.43** | 0.02 | 0.08 |
| $\sigma_l$ | 0.01 | 0.19 | 0.03 | **-0.37** |
| $\xi_p$ | -0.26 | **0.36** | 0.13 | 0.19 |
| $\iota_w$ | 0.1 | 0.04 | 0.05 | 0.01 |
| $\iota_p$ | 0.03 | 0.17 | -0.03 | 0.04 |
| $\Psi$ | **0.34** | 0.01 | 0.1 | **0.33** |
| $\Phi$ | 0.02 | 0.03 | 0.03 | 0.02 |
| $r_\pi$ | **1** | **0.31** | 0.24 | 0 |
| $\rho$ | **0.31** | **1** | **0.32** | 0.12 |
| $r_y$ | 0.24 | **0.32** | **1** | -0.04 |
| $r_{\Delta y}$ | 0 | 0.12 | -0.04 | **1** |
| $\overline{\pi}$ | -0.04 | -0.05 | -0.08 | 0.01 |
| $100(\beta^{-1}-1)$ | -0.04 | -0.18 | -0.14 | -0.04 |
| $\overline{l}$ | 0.07 | 0.1 | 0.08 | -0.05 |
| $\overline{\gamma}$ | 0.02 | 0.03 | 0.05 | -0.01 |
| $\rho_{ga}$ | -0.23 | -0.15 | 0 | 0.01 |
| $\alpha$ | 0.18 | 0.11 | 0.16 | -0.04 |

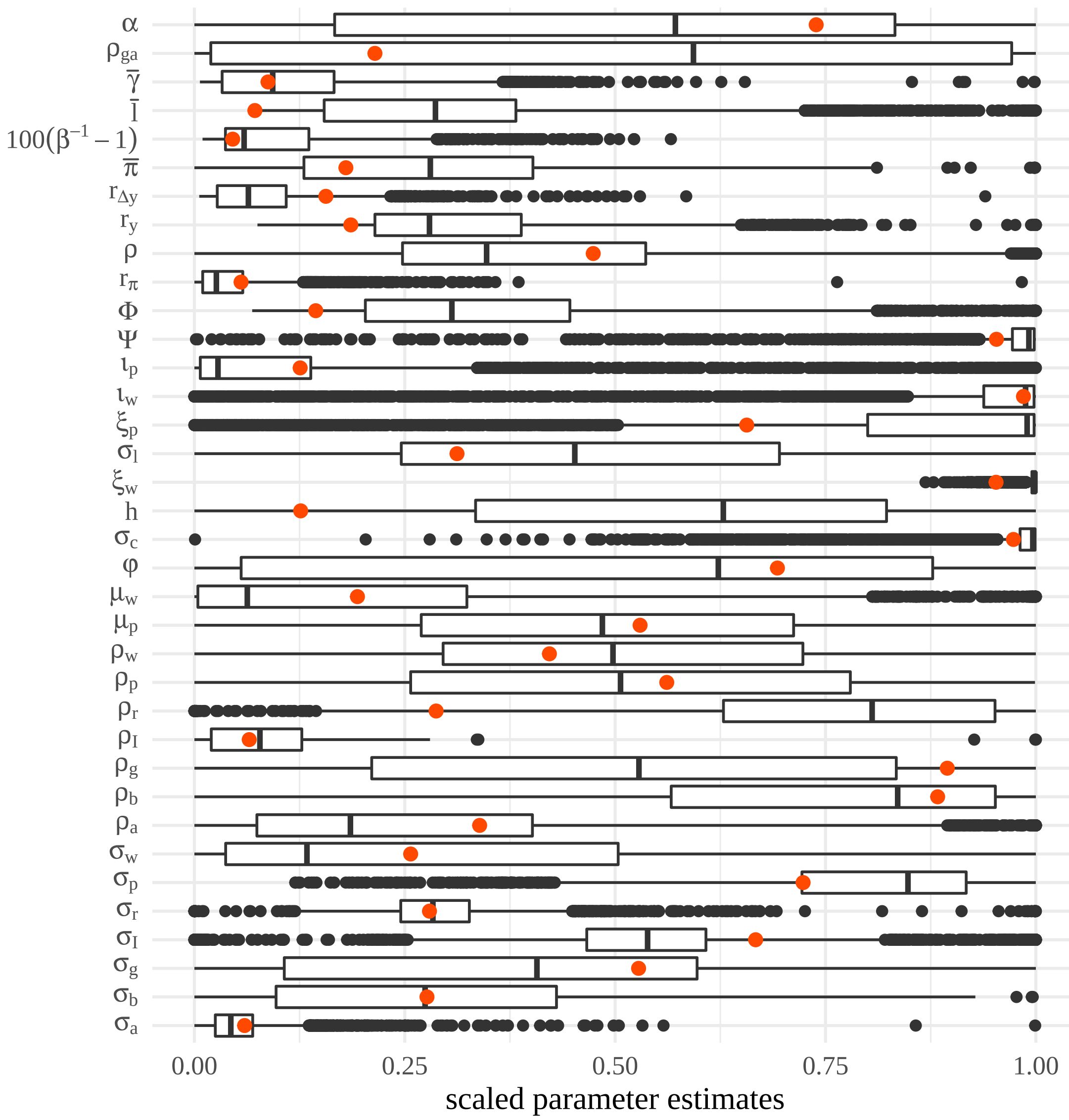


Figure 6: Parameter estimates rescaled to [0,1] based on the prior limits. Red points indicate the SW estimates. Black points are outliers relative to the bulk of the permuted estimates.

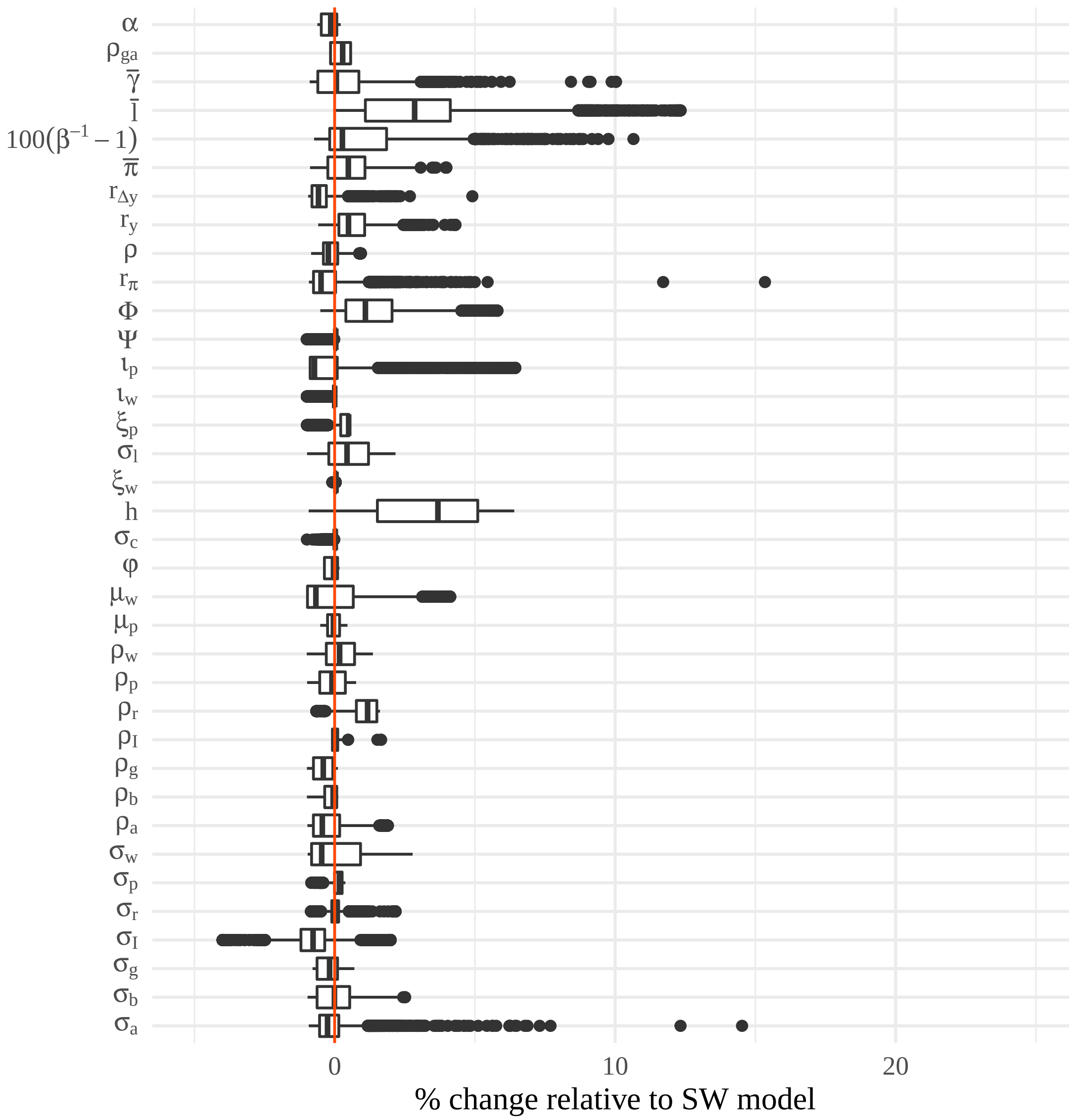


Figure 7: Percentage change in estimated parameter relative to the SW estimates. Some large outliers have been removed to better show the bulk.

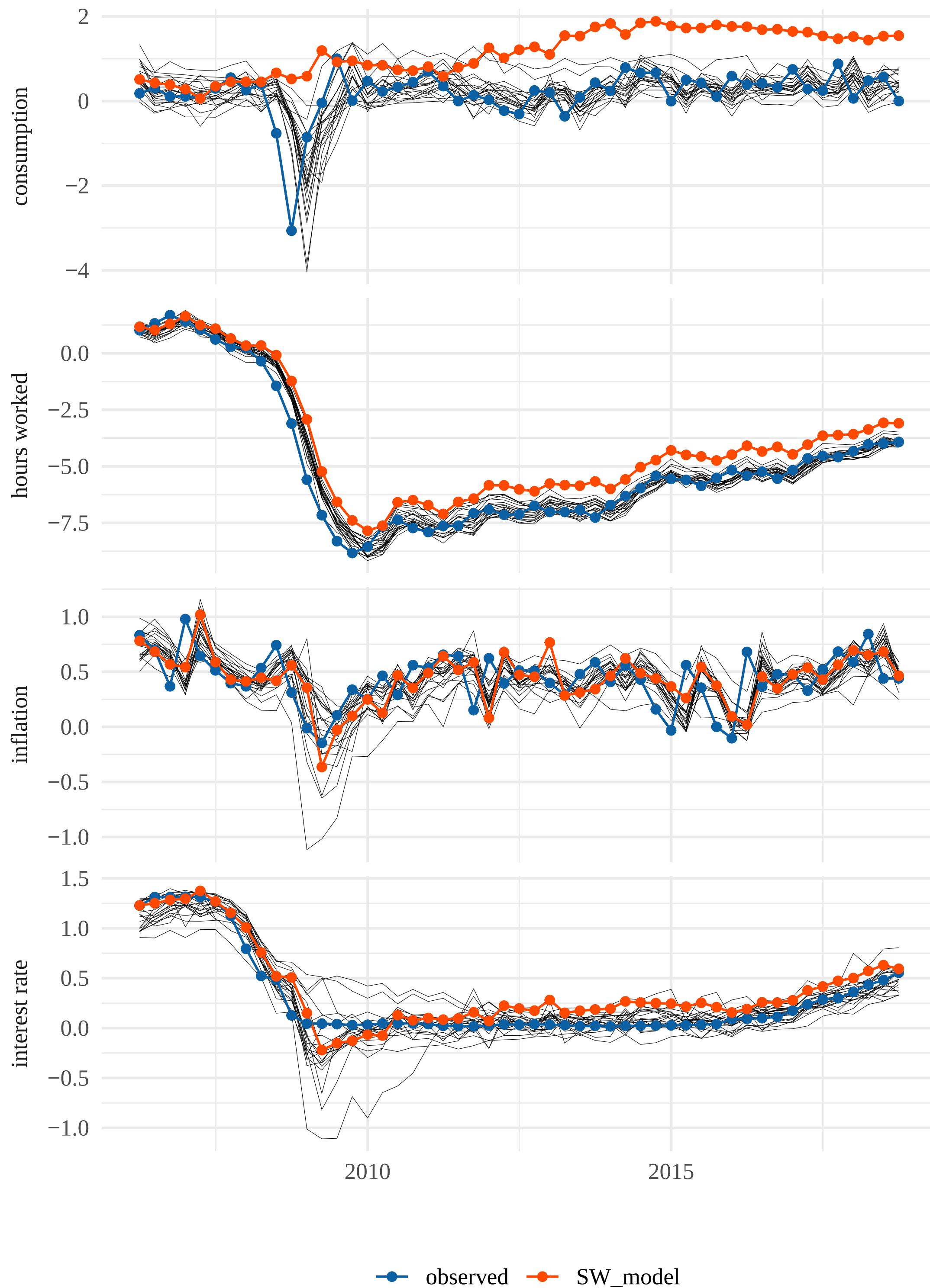


Figure 8: Out-of-sample predictions for the best predicting permutations (black), the SW model (red), and the observed data (blue).

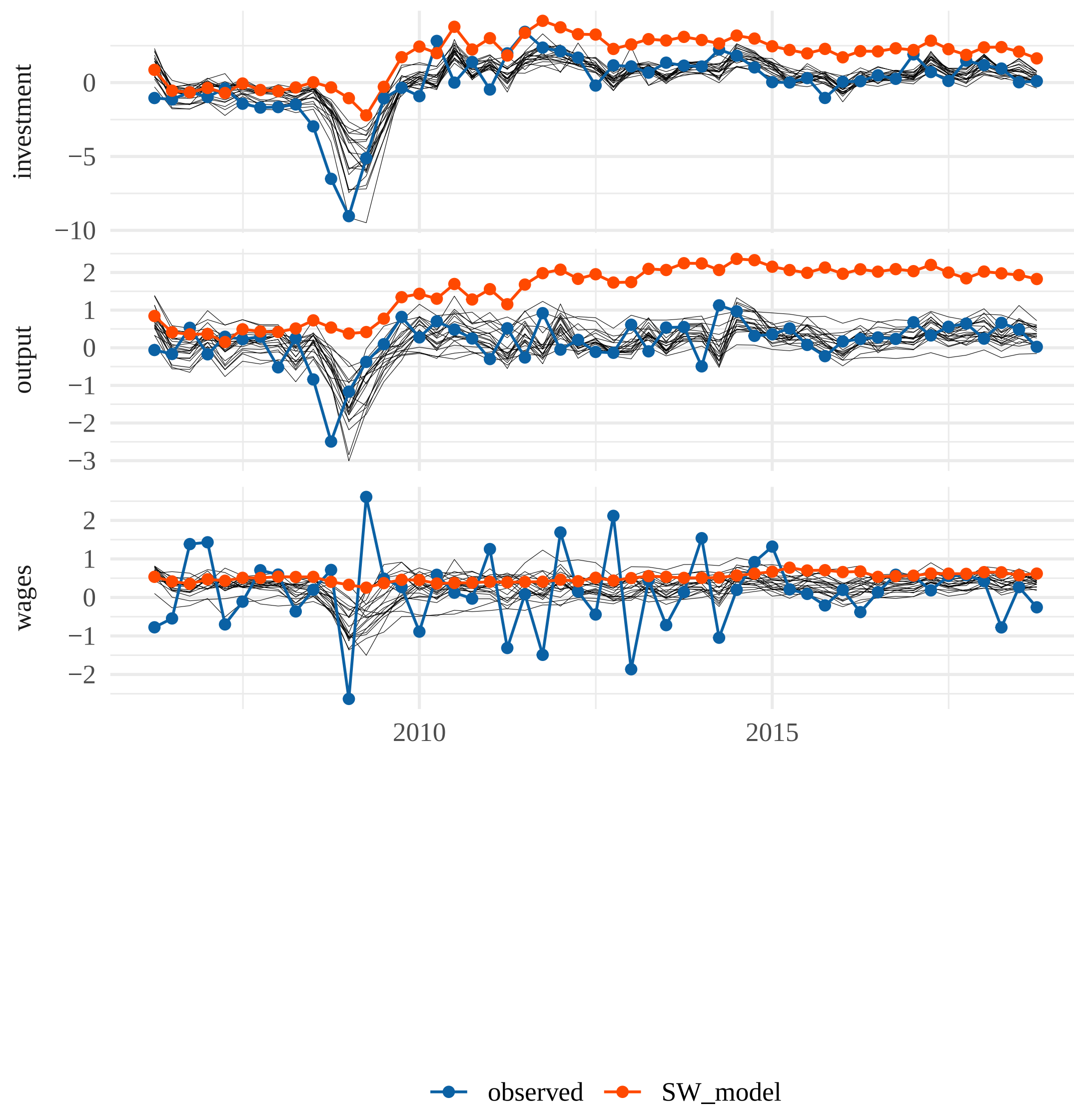


Figure 9: Out-of-sample predictions for the best predicting permutations (black), the SW model (red), and the observed data (blue).

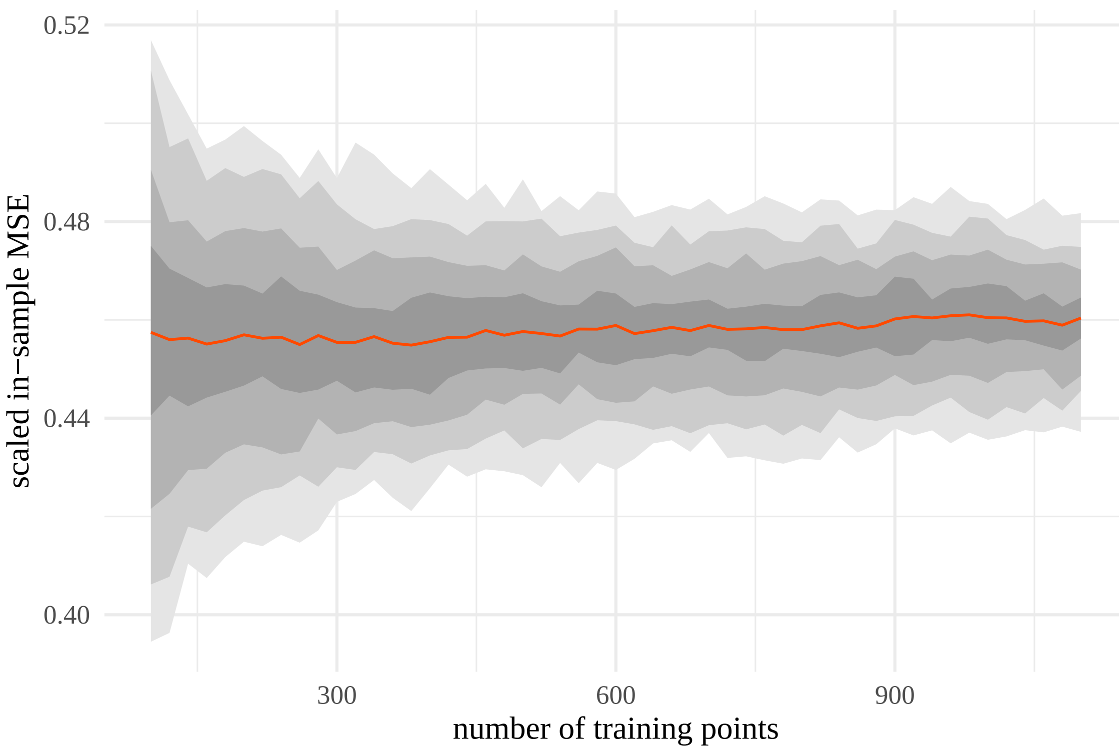


Figure 10: Training (in sample) error averaged across series as the number of training points increases. We would expect the variability to decline while the average remains constant. The mean is shown in red along with 30%, 60%, 80%, and 90% confidence bands calculated from the replications.

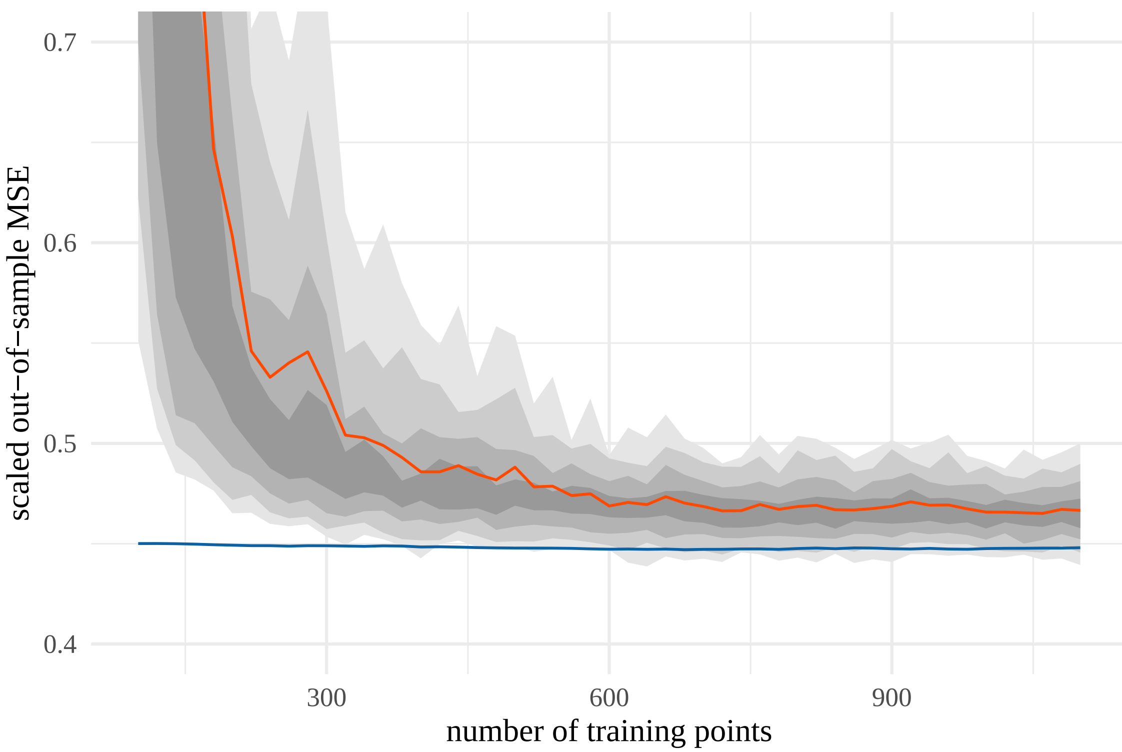


Figure 11: Out-of-sample error averaged across series as the number of training points increases. We would expect both the variability and the average to decline. The blue line is the test error for the true model.

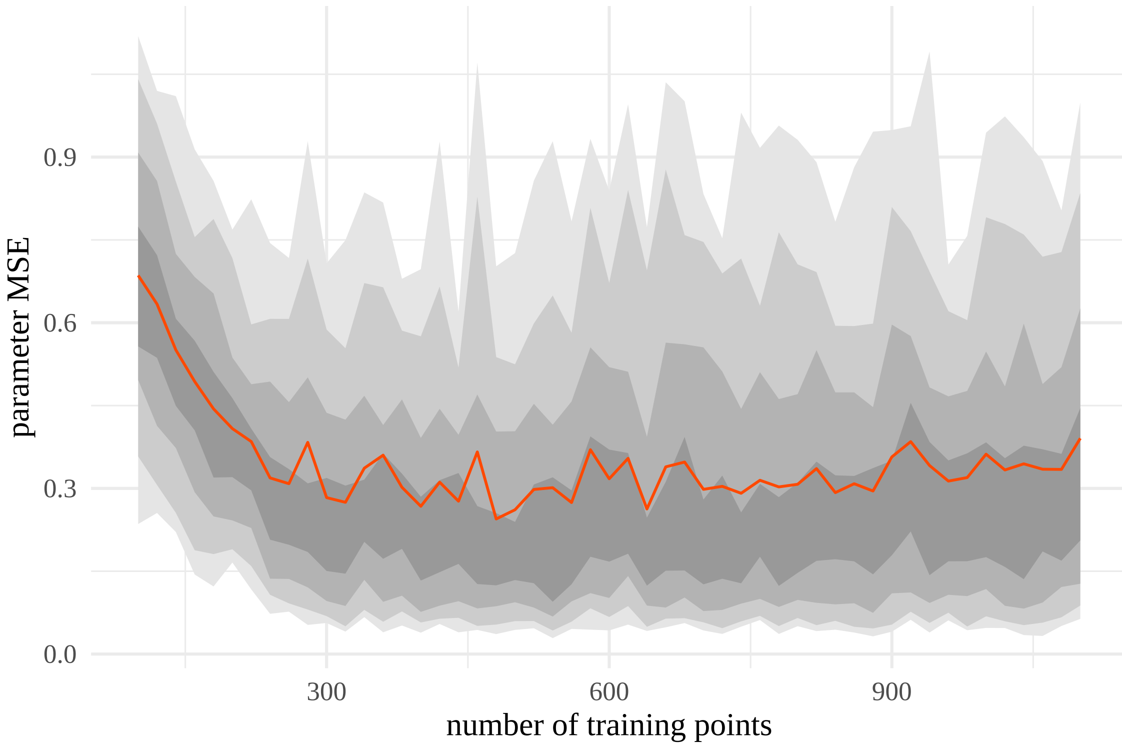


Figure 12: The parameter MSE shows similar behavior to the predictive MSE: steep initial decline toward an asymptote greater than that of the true model.

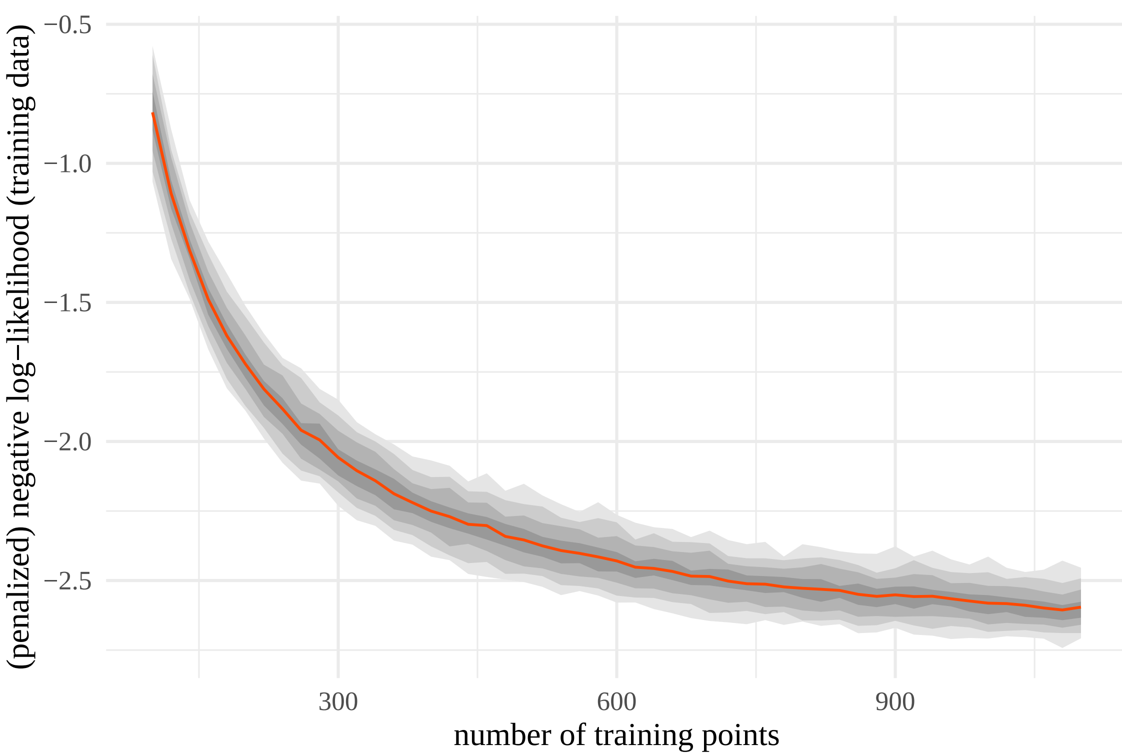


Figure 13: The negative log-likelihood (in-sample) per observation.

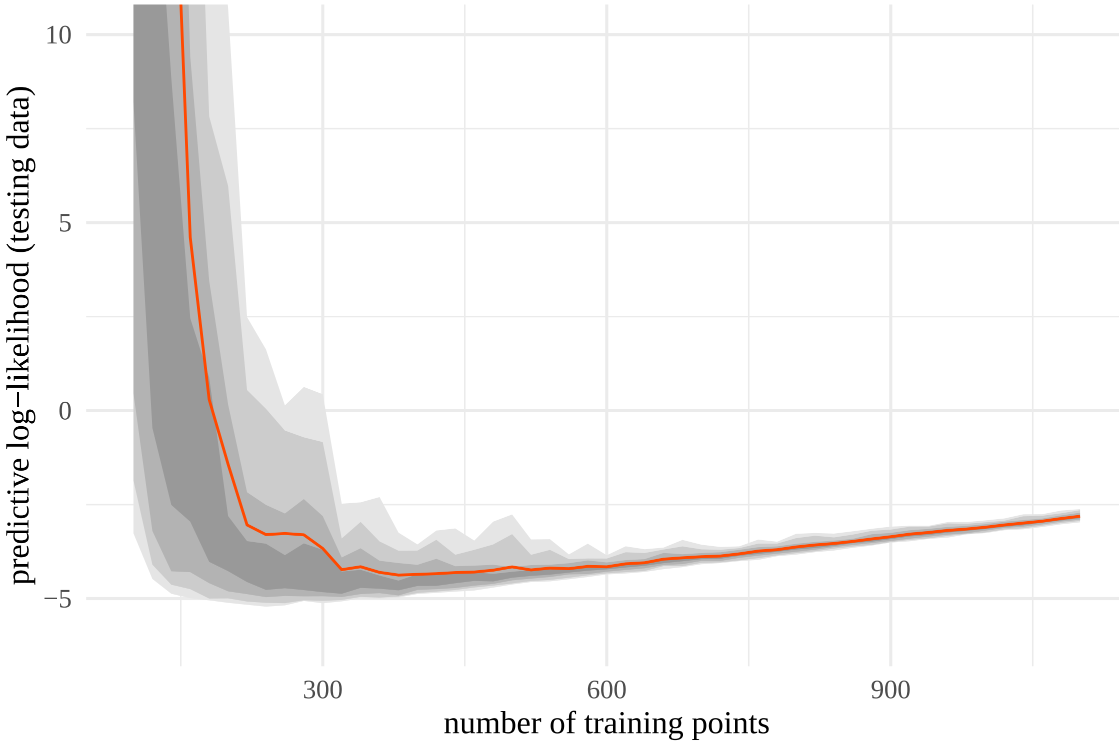


Figure 14: The negative predictive log-likelihood (out-of-sample) per observation.


Figure 15: Mean-squared prediction error for each series.

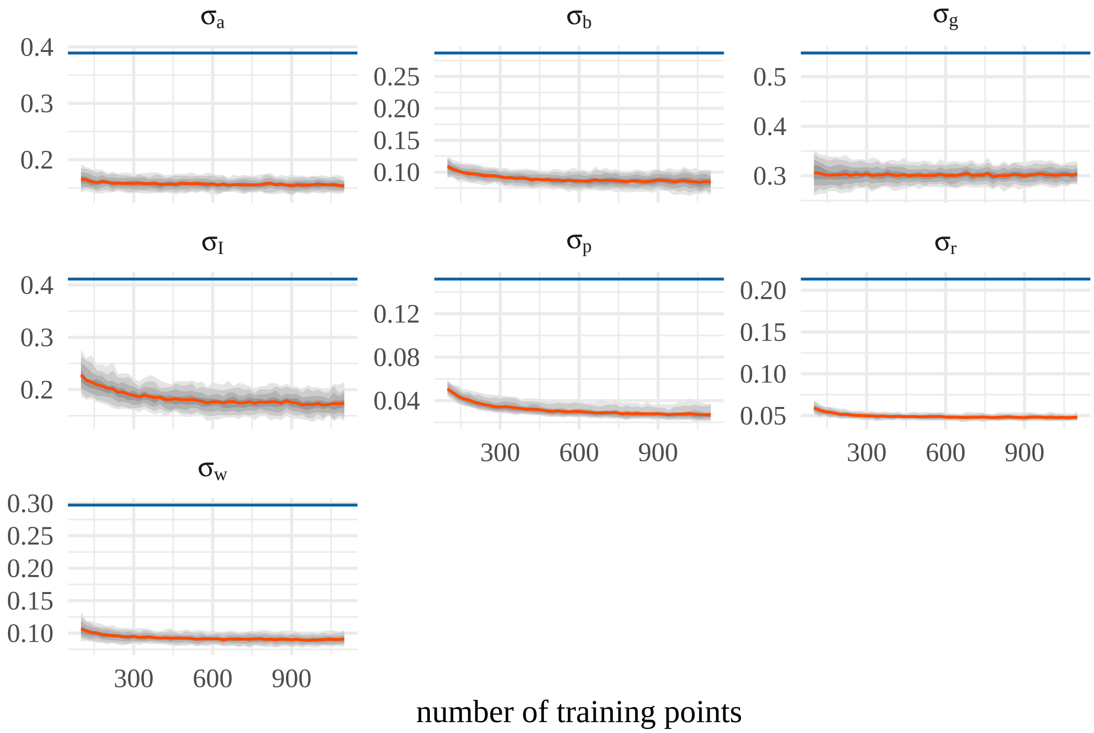


Figure 16: Estimates for the standard deviations of stochastic shocks.

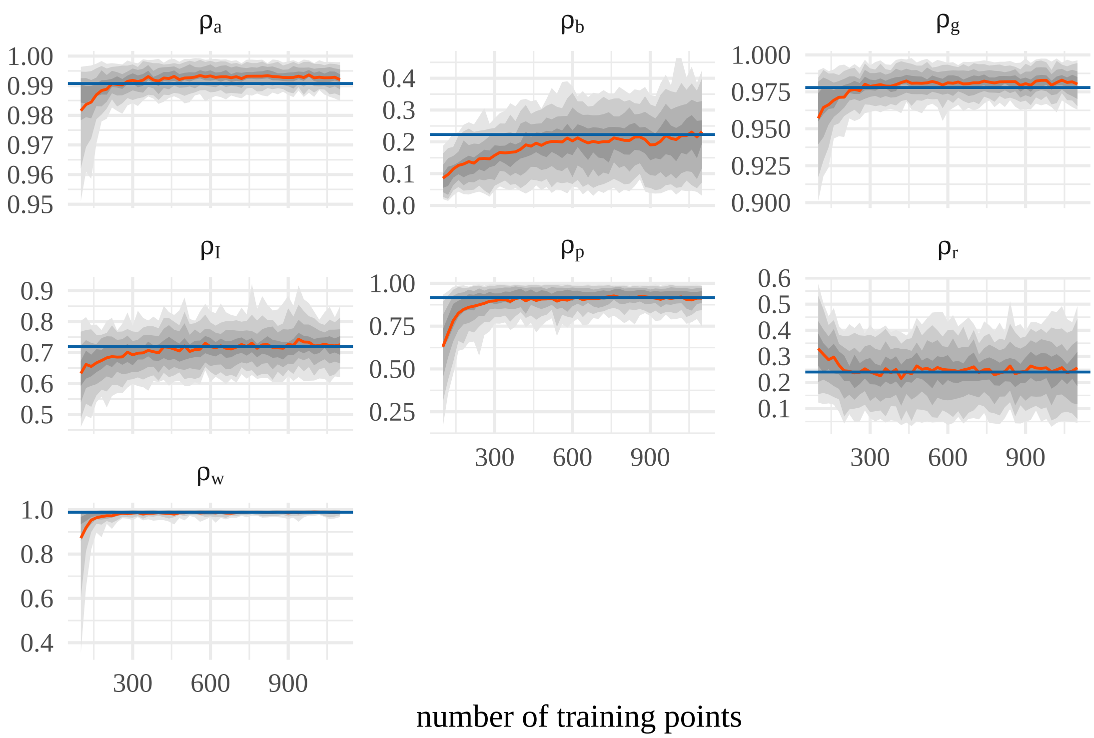


Figure 17: Estimates for the autocorrelation parameters in the shock processes.

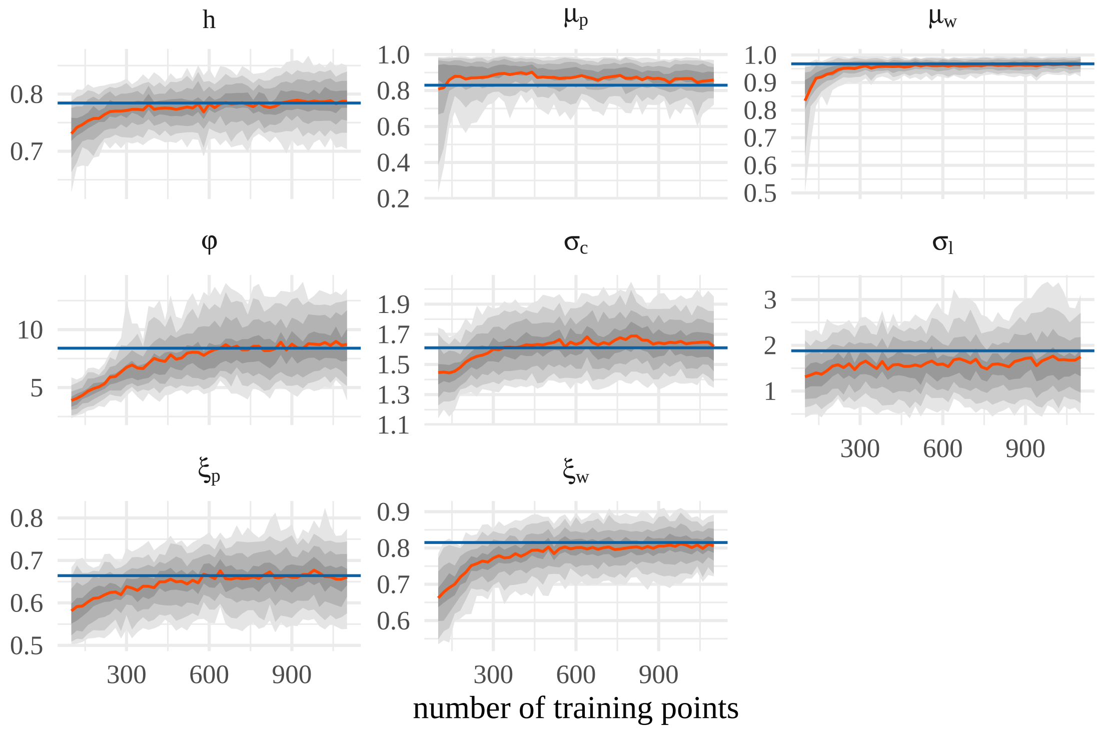


Figure 18: Estimates of "deep" parameters

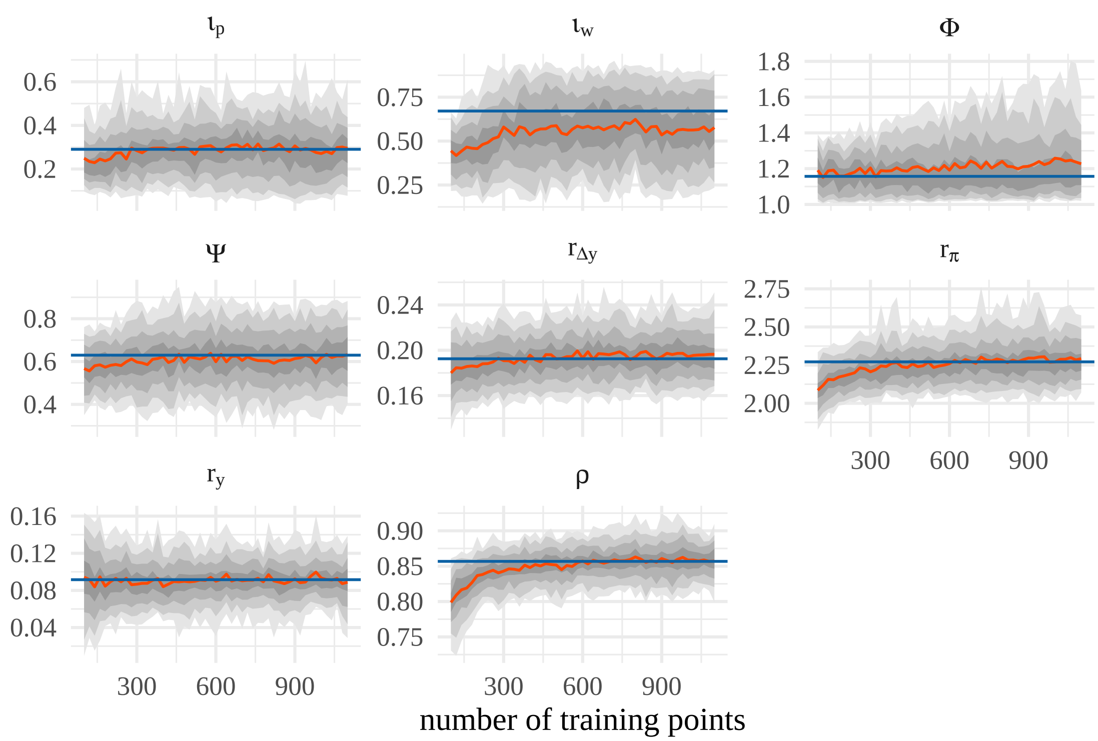


Figure 19: Estimates of "deep" parameters

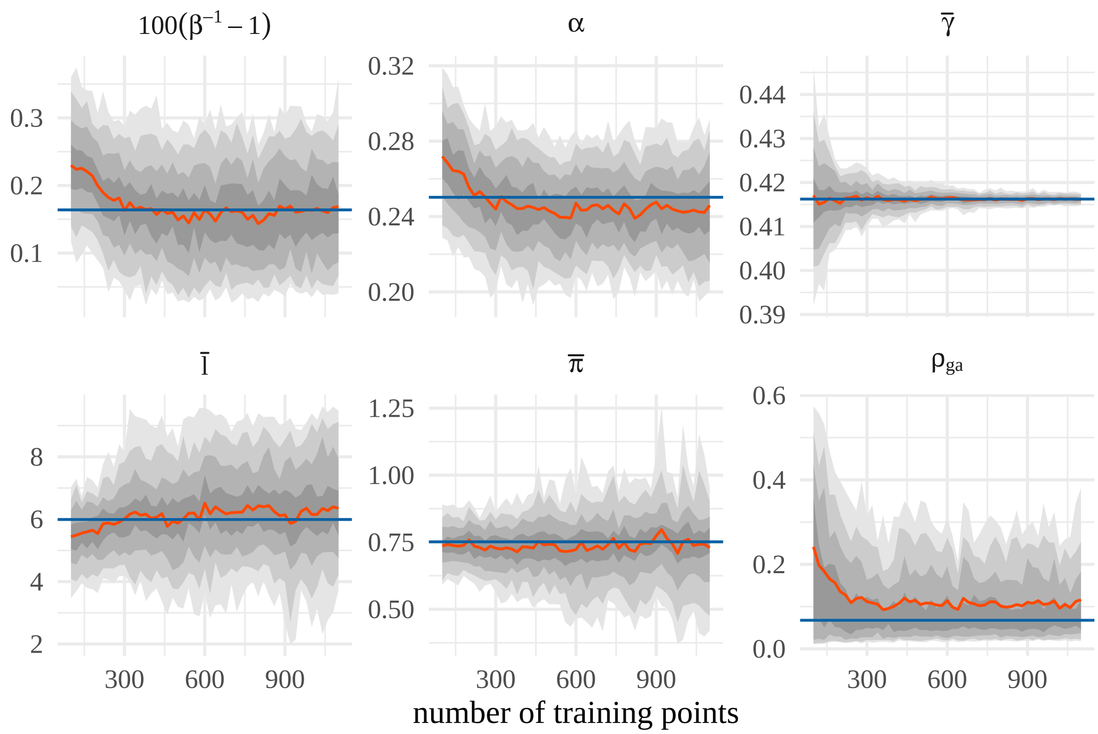


Figure 20: Estimates of "deep" parameters

# A Data preprocessing

The necessary series are shown in Table 7. All of the data are quarterly. The required series are GDPC1, GDPDEF, PCEC, FPI, CE16OV, FEDFUNDS, CNP16OV, PRS85006023, and COMPNFB. These nine series are used to create $x_t$ as follows:

$$\begin{aligned}
pop_t^{ind} &= CNP16OV_t/CNP16OV_{2012Q3},\\
emp_t^{ind} &= CE16OV_t/CE16OV_{2012Q3},\\
y_t &= 100\ln\left(\frac{GDPC1_t}{pop_t^{ind}}\right),\\
c_t &= 100\ln\left(\frac{PCEC_t/GDPDEF_t}{pop_t^{ind}}\right),\\
i_t &= 100\ln\left(\frac{FPI_t/GDPDEF_t}{pop_t^{ind}}\right),\\
l_t &= 100\ln\left(\frac{PRS85006023_t/emp_t^{ind}}{pop_t^{ind}}\right),\\
\pi_t &= 100\ln\left(\frac{GDPDEF_t}{GDPDEF_{t-1}}\right),\\
w_t &= 100\ln\left(\frac{COMPNFB_t}{GDPDEF_t}\right),\\
r_t &= FEDFUNDS/4.
\end{aligned}$$

| Series ID | Description | Unit |
|---|---|---|
| GDPC1 | Real Gross Domestic Product | Billions of Chained 2012 $ |
| GDPDEF | GDP Implicit Price Deflator | Index: 2012=100 |
| PCEC | Personal Consumption Expenditures | Billions of $ |
| FPI | Fixed Private Investment | Billions of $ |
| CE16OV | Civilian Employment | Thousands of persons |
| FEDFUNDS | Effective Federal Funds Rate | Percent |
| CNP16OV | Civilian Noninstitutional Population | Thousands of persons |
| PRS85006023 | Nonfarm business sector: average weekly hours | Index: 2012=100 |
| COMPNFB | Nonfarm business sector: Compensation per hour | Index: 2012=100 |

Table 7: Data series from FRED for estimating the DSGE.

# B Rewriting Smets and Wouters (2007)

To be updated again.

In the rest of this section, we describe the log-linearized version of the DSGE model that we subsequently estimate using US data. All variables are log-linearized around their steady-state balanced growth path. Starred variables denote steady-state values. We first describe the aggregate demand side of the model and then turn to the aggregate supply.

The aggregate resource constraint is given by

$$y_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g. \tag{9}$$

Output ($y_t$) is absorbed by consumption ($c_t$), investment ($i_t$), capital-utilization costs that are a function of the capital utilization rate ($z_t$), and exogenous spending ($\epsilon_t^g$); $c_y$ is the steady-state share of consumption in output and equals $1 - g_y - i_y$, where $g_y$ and $i_y$ are respectively the steady-state exogenous spending-output ratio and investment-output ratio. The steady-state investment-output ratio in turn equals $(\gamma - 1 + \delta)k_y$, where $\gamma$ is the steady-state growth rate, $\delta$ stands for the depreciation rate of capital, and $k_y$ is the steady-state capital-output ratio. Finally,

$z_y = R_*^k ky$, where $R_*^k$ is the steady-state rental rate of capital. We assume that exogenous spending follows a first-order autoregressive process with an IID-Normal error term and is also affected by the productivity shock as fol- lows: $\epsilon_t^g = \rho_g \epsilon_{t-1}^g + \eta_t^g + \rho_{ga}\eta_t^a$. The latter is empirically motivated by the fact that, in estimation, exogenous spending also includes net exports, which may be affected by domestic productivity developments.

The dynamics of consumption follows from the consumption Euler equation and is given by

$$c_t = c_1 c_{t-1} + (1 - c_1)\mathbb{E}_t[c_{t+1}] + c_2(l_t - \mathbb{E}_t[l_{t+1}]) - c_3(r_t - \mathbb{E}_t[\pi_{t+1}]) + \epsilon_t^b), \tag{10}$$

where $c_1 = (\lambda/\gamma)/(1 + \lambda/\gamma)$, $c_2 = [(\sigma_c - 1)(W_*^h L_*/C_*)]/[\sigma_c(1 + \lambda/\gamma)]$, and $c_3 = (1 - \lambda/\gamma)/[\sigma_c(1 + \lambda/\gamma)]$. Current consumption ($c_t$ depends on a weighted average of past and expected future consumption, and on expected growth in hours worked $l_t - \mathbb{E}_t[l_{t+1}]$), the ex ante real interest rate ($r_t - \mathbb{E}_t[\pi_{t+1}]$, and a disturbance term $\epsilon_t^b$ . Under the assumption of no external habit formation ($\lambda = 0$) and log utility in consumption ($\sigma_c = 1$), $c1 = c2 = 0$ and the traditional purely forward-looking consumption equation is obtained. With steady-state growth, the growth rate $\gamma$ marginally affects the reduced-form parameters in the linearized consumption equation. When the elasticity of intertemporal substitution (for constant labor) is smaller than one ($\sigma_c > 1$), consumption and hours worked are complements in utility and consumption depends positively on current hours worked and negatively on expected growth in hours worked (see Susanto Basu and Kimball 2002). Finally,the disturbance term $\epsilon_t^b$ represents a wedge between the interest rate controlled by the central bank and the return on assets held by the households. A positive shock to this wedge increases the required return on assets and reduces current consumption. At the same time, it also increases the cost of capital and reduces the value of capital and investment, as shown below.[6] This shock has similar effects as so-called net-worth shocks in Ben S. Bernanke, Gertler, and Simon Gilchrist (1999) and Christiano, Roberto Motto, and Massimo Rostagno (2003), which explicitly model the external finance premium. The disturbance is assumed to follow a first-order autoregressive process with an IID-Normal error term: $\epsilon_t^b = \rho_b \epsilon_{t-1}^b + \eta_t^b$.

The dynamics of investment comes from the investment Euler equation and is given by

$$i_t = i_1 i_{t-1} + (1 - i_1)\mathbb{E}[i_{t+1}] + i_2 q_t + \epsilon_t^i, \tag{11}$$

where $i_1 = 1/(1 + \beta\gamma^{1-\sigma_c})$, $i_2 = 1/[(1 + \beta\gamma^{1-\sigma_c})\gamma^2\varphi]$, $\varphi$ is the steady-state elasticity of the capital adjustment cost function, and is the discount factor applied by households. As in CEE (2005), a higher elasticity of the cost of adjusting capital reduces the sensitivity of investment ($i_t$) to the real value of the existing capital stock ($q_t$). Modeling capital adjustment costs as a function of the change in investment rather than its level introduces additional dynamics in the investment equation, which is useful in capturing the hump-shaped response of investment to various shocks. Finally, $\epsilon_t^i$ represents a disturbance to the investment-specific technology process and is assumed to follow a first-order autoregressive process with an IID-Normal error term: $\epsilon_t^i = \rho_i \epsilon_{t-1}^i + \eta_t^i$.

The corresponding arbitrage equation for the value of capital is given by

$$q_t = q_1 \mathbb{E}_t[q_{t+1}] + (1 - q_1)\mathbb{E}_t[r_{t+1}^k] - (r_t - \mathbb{E}[\pi_{t+1}] + \epsilon_t^b), \tag{12}$$

where $q_1 = \beta\gamma^{-\sigma_c}(1 - \delta) = [(1 - \delta)/(R_*^k + (1 - \delta))]$. The current value of the capital stock ($q_t$) depends positively on its expected future value and the expected real rental rate on capital ($\mathbb{E}_t[r_{t+1}^k]$) and negatively on the ex ante real interest rate and the risk premium disturbance.

Turning to the supply side, the aggregate production function is given by

$$y_t = \phi_p(\alpha k_t^2 + (1 - \alpha)l_t + \epsilon_t^a). \tag{13}$$

Output is produced using capital ($k_t^s$) and labor services (hours worked, $l_t$). Total factor productivity ($\epsilon_t^a$) is assumed to follow a first-order autoregressive process: $\epsilon_t^a = \rho_a \epsilon_{t-1}^a + \eta_t^a$. The parameter $\alpha$ captures the share of capital in production, and the parameter $\phi_p$ is one plus the share of fixed costs in production, reflection the presence of fixed costs in production.

As newly installed capital becomes effective only with a one-quarter lag, current capital services used in production ($k_t^s$) are a function of capital installed in the previous period ($k_{t-1}$) and the degree of capital utilization ($z_t$):

$$k_t^s = k_{t-1} + z_t. \tag{14}$$

Cost minimization by the households that provide capital services implies that the degree of capital utilization is a positive function of the rental rate of capital,

$$z_t = z_1 r_t^k, \tag{15}$$

where $z_q = (1 - \psi)/\psi$ and $\psi$ is a positive function of the elasticity of the capital utilization adjustment cost function and normalized to be between zero and one. When $\psi = 1$, it is extremely costly to change the utilization of capital and, as a result, the utilization of capital remains constant. In contrast, when $\psi = 0$, the marginal cost of changing

[6] Include?

the utilization of capital is constant and, as a result, in equilibrium the rental rate on capital is constant, as is clear from equation (15).

The accumulation of installed capital ($k_t$) is a function not only of the flow of investment but also of the relative efficiency of these investment expenditures as captured by the investment-specific technology disturbance

$$k_t = k_1 k_{t-1} + (1-k_1) i_t + k_2 \epsilon_t^i, \tag{16}$$

with $k_1 = (1-\delta)/\gamma$ and $k_2 = (1-(1-\delta)/\gamma)(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi$.

Turning to the monopolistic competitive goods market, cost minimization by firms implies that the price mark-up ($\mu_t^p$), defined as the difference between the average price and the nominal marginal cost or the negative of the real marginal cost, is equal to the difference between the marginal product of labor ($mpl_t$) and the real wage ($w_t$):

$$\mu_t^p = mpl_t - w_t = \alpha(k_t^s - l_t) + \epsilon_t^a - w_t. \tag{17}$$

As implied by the second equality in (17), the marginal product of labor is itself a positive function of the capital-labor ratio and total factor productivity.

Due to price stickiness, as in Calvo (1983), and partial indexation to lagged inflation of those prices that can not be reoptimized, as in Smets and Wouters (2003), prices adjust only sluggishly to their desired mark-up. Profit maximization by price-setting firms gives rise to the following New-Keynesian Phillips curve:

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 \mathbb{E}_t[\pi_{t+1}] - \pi_3 \mu_t^p + \epsilon_t^p, \tag{18}$$

where $\pi_1 = \iota_p/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, $\pi_2 = \beta\gamma^{1-\sigma_c}/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, and $\pi_3 = 1/(1+\beta\gamma^{1-\sigma_c}\iota_p)[(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\xi_p((\phi_p-1)\epsilon_p+1)]$. Inflation ($\pi_t$) depends positively on past and expected future inflation, negatively on the current price mark-up, and positively on a price mark-up disturbance ($\epsilon_t^p$). The price mark-up disturbance is assumed to follow an ARMA(1,1) process: $\epsilon_t^p = \rho_p \epsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p$, where $\eta_t^p$ is an IID-Normal pricemark-up shock. The inclusion of the MA term is designed to capture the high-frequency fluctuations in inflation.

When the degree of indexation to past inflation is zero ($\iota_p = 0$), equation (18) reverts to a standard, purely forward-looking Phillips curve ($\pi_1 = 0$). The assumption that all prices are indexed to either lagged inflation or the steady-state inflation rate ensures that the Phillips curve is vertical in the long run. The speed of adjustment to the desired mark-up depends, among others, on the degree of price-stickiness ($\xi_p$), the curvature of the Kimball goods market aggregator ($\epsilon_p$), and the steady-state mark-up, which in equilibrium is itself related to the share of fixed costs in production ($\phi_p - 1$) through a zero-profit condition. A higher $\epsilon_p$ slows down the speed of adjustment because it increases the strategic complementarity with other price setters. When all prices are flexible ($\xi_p = 0$) and the price-mark-up shock is zero, equation (18) reduces to the familiar condition that the price mark-up is constant, or equivalently that there are no fluctuations in the wedge between the marginal product of labor and the real wage.

Cost minimization by firms will also imply that the rental rate of capital is negatively related to the capital-labor ratio and positively to the real wage (both with unitary elasticity):

$$r_t^k = -(k_t - l_t) + w_t \tag{19}$$

In analogy with the goods market, in the monopolistically competitive labor market, the wage mark-up will be equal to the difference between the real wage and the marginal rate of substitution between working and consuming ($mrs_t$),

$$\mu_t^w = w_t - mrs_t = w_t - (\sigma_l l_t + \frac{1}{1-\lambda/\gamma}\left(c_t - \frac{\lambda c_{t-1}}{\gamma}\right), \tag{20}$$

where $\sigma_l$ is the elasticity of labor supply with respect to the real wage and $\lambda$ is the habit parameter in consumption.

Similarly, due to nominal wage stickiness and partial indexation of wages to inflation, real wages adjust only gradually to the desired wage mark-up:

$$w_t = w_1 w_{t-1} + (1-w_1)(\mathbb{E}_t[w_{t+1}] + \mathbb{E}_t[\pi_{t+1}]) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \epsilon_t^w, \tag{21}$$

with $w_1 = 1/(1+\beta\gamma^{1-\sigma_c})$, $w_2 = (1+\beta\gamma^{1-\sigma_c}\iota_w)/(1+\beta\gamma^{1-\sigma_c})$, $w_3 = \iota_w/(1+\beta\gamma^{1-\sigma_c})$, and $w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/(\xi_w((\phi_w-1)\epsilon_w+1))]$.

The real wage $w_t$ is a function of expected and past real wages, expected, current, and past inflation, the wage mark-up, and a wage-markup disturbance ($\epsilon_t^w$). If wages are perfectly flexible ($\xi_w = 0$), the real wage is a constant mark-up over the marginal rate of substitution between consumption and leisure. In general, the speed of adjustment to the desired wage mark-up depends on the degree of wage stickiness ($\xi_w$) and the demand elasticity for labor, which itself is a function of the steady-state labor market mark-up ($\phi_w - 1$) and the curvature of the Kimball labor market aggregator ($\epsilon_w$). When wage indexation is zero ($\iota_w = 0$), real wages do not depend on lagged inflation ($w_3 = 0$). The wage-markup disturbance ($\epsilon_w^t$) is assumed to follow an ARMA(1, 1) process with an IID-Normal error term:

$\epsilon_t^w = \rho_w \epsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w$. As is the case of the price mark-up shock, the inclusion of an MA term allows us to pick up some of the high-frequency fluctuations in wages.[7]

Finally, the model is closed by adding the following empirical monetary policy reaction function:

$$r_t = \rho r_{t-1} + (1-\rho)\left[r_\pi \pi_t + r_y (y_t - y_t^p)\right] + \\ + r_{\Delta y}\left[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)\right] + \epsilon_t^r. \tag{22}$$

The monetary authorities follow a generalized Taylor rule by gradually adjusting the policy-controlled interest rate ($r_t$) in response to inflation and the output gap, defined as the difference between actual and potential output (John B. Taylor 1993). Consistently with the DSGE model, potential output is defined as the level of output that would prevail under flexible prices and wages in the absence of the two "mark-up' ' shocks.[8]

The parameter $\rho$ captures the degree of interest rate smoothing. In addition, there is a short-run feedback from the change in the output gap. Finally, we assume that the monetary policy shocks ($\epsilon_t^r$) follow a first-order autoregressive process with an IID-Normal error term: $\epsilon_t^r = \rho_r \epsilon_{t-1}^r + \eta_t^r$.

Equations (9) to (22) determine 14 endogenous variables: $y_t$, $c_t$, $i_t$, $q_t$, $k_t^s$, $k_t$, $z_t$, $r_t^k$, $\mu_t^p$, $\pi_t$, $\mu_t^w$, $w_t$, $l_t$, and $r_t$. The stochastic behavior of the system of linear rational expectations equations is driven by seven exogenous disturbances: total factor productivity ($\epsilon_t^a$), investment-specific technology ($\epsilon_t^i$), risk premium ($\epsilon_t^b$), exogenous spending ($\epsilon_t^g$), price mark-up ($\epsilon_t^p$), wage mark-up ($\epsilon_t^w$), and monetary policy ($\epsilon_t^r$) shocks. Next we turn to the estimation of the model.

# References


Olivier Jean Blanchard and Charles M Kahn. The solution of linear difference models under rational expectations. *Econometrica: Journal of the Econometric Society*, pages 1305–1311, 1980.

D.N. DeJong and C. Dave. *Structural macroeconometrics*. Princeton Univ Press, Princeton, 2 edition, 2011.

Nikolay Iskrev. Local identification in DSGE models. Technical report, Banco de Portugal, August 2009.

Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1): 35–45, 1960.

Paul Klein. Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423, 2000.

Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370, November 1982.

Christopher A. Sims. Solving linear rational expectations models. *Computational Economics*, 20(1-2):1–20, October 2002.

Frank Smets and Rafael Wouters. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, 97(3):586–606, June 2007.


[7] Alternatively, we could interpret this disturbance as a labor supply disturbance coming changes in preferences for leisure.

[8] In practical terms, we expand the model consisting of equations (9) to (22) with a flexible-price-and-wage version in order to calculate the model-consistent output gap. Note that the assumption of treating the wage equation disturbance as a wage mark-up disturbance rather than a labor supply disturbance coming from changed preferences has implications for our calculation of potential output.