SVAR's identified via heteroskedasticity as forecasting models

Chris Sims

June 16, 2021

Specification searches

- Ed Leamer, in his 1978 book *Specification Searches*, argued that the way applied econometricians proceed in practice is disconnected from the way theoretical economists talk and teach.

- He argued that econometric practice is mostly a process of thinking of models, changing the models in response to initial analyses, simplifying or elaborating the models in the light of prior beliefs or new information.

- This process is mostly not discussed explicitly in textbooks. They present methods of inference as if a model were given before the data were even assembled.

Machine learning

- While some of Leamer's discussion is now dated, this basic insight still applies.
- In fact, one way to interpret the rise, outside econometrics, of "machine learning" approaches to data analysis is that those approaches address specification search directly, and partially automate it.
- Econometrics can benefit by more explicit treatment of specification search.

How to do it

- Leamer took a largely Bayesian approach, and saw that approach as necessary to make sense of inference with specification search.
- One development since he wrote is Bayesian non-parametrics, and in particular the idea of a Bayesian sieve.
- A Bayesian sieve places a prior on a countable collection of finite-parameter models.
- In practice, it usually leads to focusing attention on a small subset of the finite-parameter models, but leaves open the possibility of moving to more complex models as new data arrives.

- This is how applied inference usually proceeds; the contribution of thinking of it as Bayesian sieve inference is to allow interpretation of tests, standard errors, and credibility sets in the presence of specification search.

Why this abstract introduction?

- I'm going to argue that SVAR's "identified" through heteroskedasticity, are a natural component of a specification search approach to modeling economic time series.
- Since "structural" VAR's are usually used where the aim is behavioral or causal interpretation of the model, it might otherwise seem odd to invoke SVAR's identified this way as forecasting models.

Substance of the talk

I'll go through the early stages of specification search for a current project, trying to capture possible links from fiscal policy changes to inflation. There are two main areas we'll examine:

- Why the IDH (identified through heteroskedasticity) model is likely to be useful as an extension of the standard VAR model.
- Possible non-stationarity, cointegration, and modeling initial conditions. My views on how to handle these issues have shifted somewhat over time.

Caveats

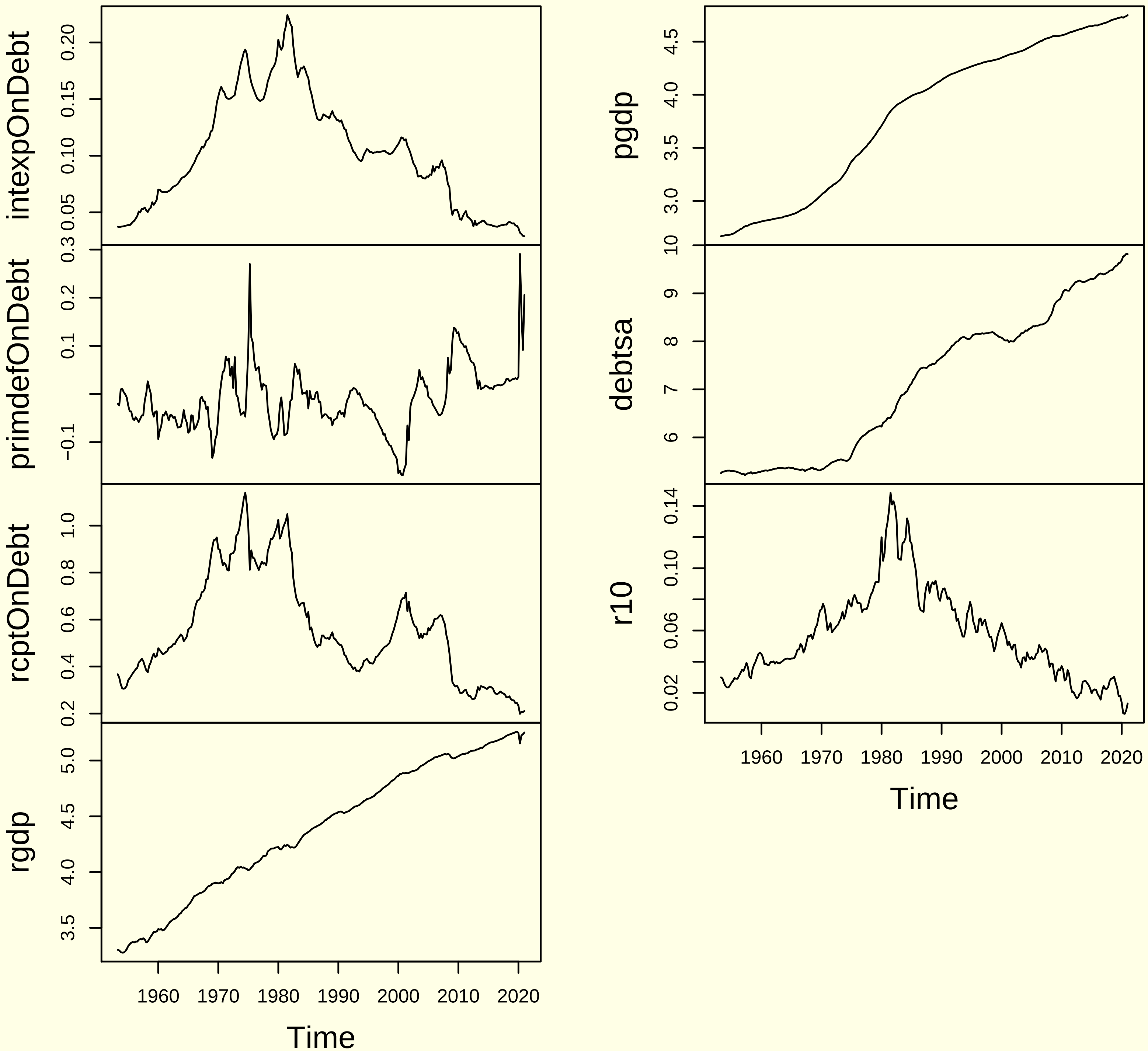
This is all very preliminary. It seems certain I'll need to allow for fat-tailed shock distributions, and I haven't done that yet. Even for the preliminary results shown here, I have not had time to add error bands. For a more finished product illustrating where this work may be headed, see Brunnermeier, Palia, Sastry, and Sims (2018), which is forthcoming in the June AER. It does have t-distributed errors, but did not require as much attention to initial conditions as the current project.

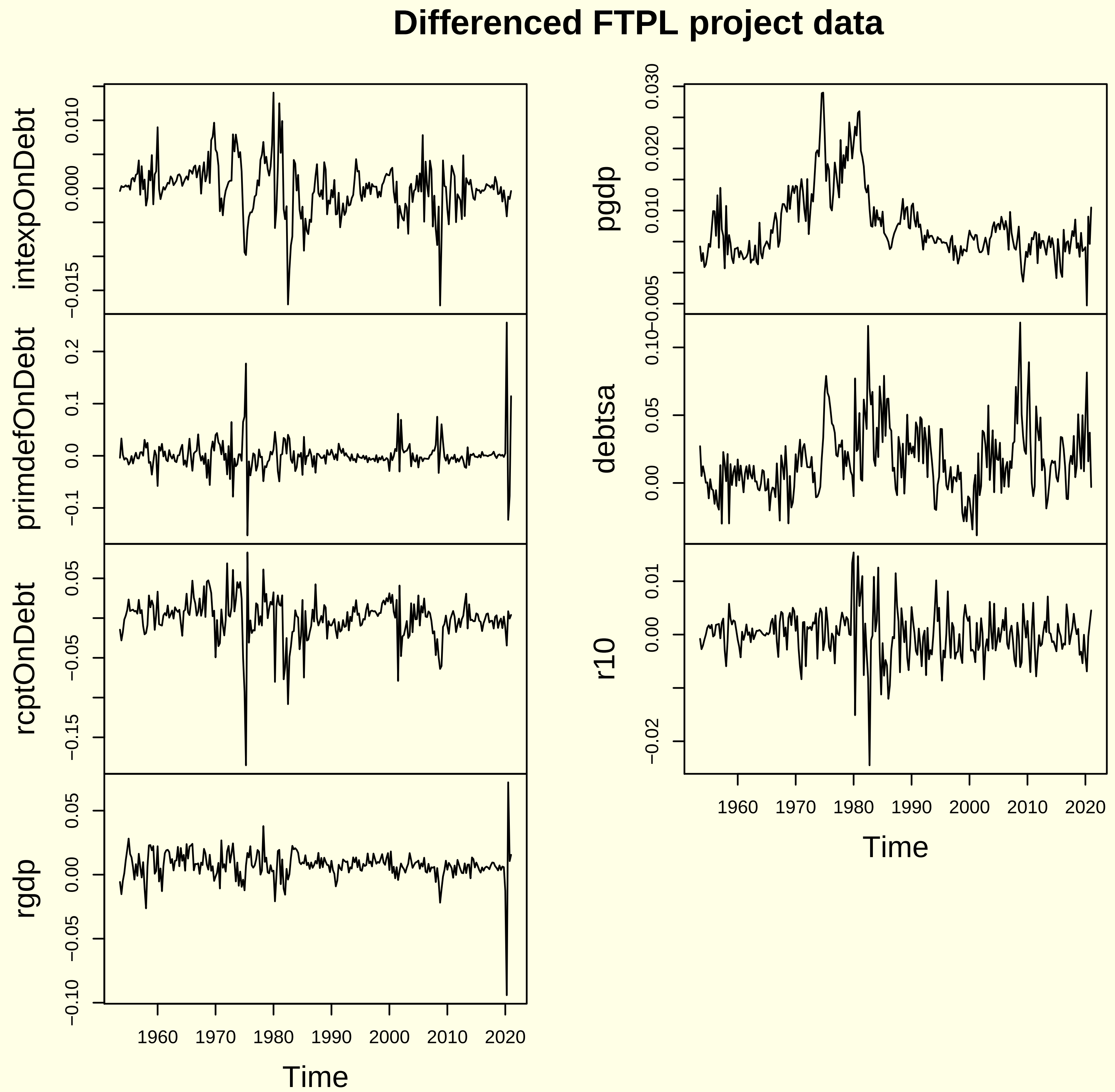
The data, and why IDH is plausible

The data are US, 1953:II-2021:I and 4-6 are logged:

1. the ratio of federal government interest expense to federal debt
2. the ratio of the federal primary deficit to debt
3. the ratio of federal total receipts to debt
4. real GDP
5. GDP deflator
6. federal debt (nominal)
7. ten year government bond interest rate

FTPL project data




Differenced FTPL project data
intexpOnDebt
primdefOnDebt
rcptOnDebt
rgdp
pgdp
debtsa
r10
Time
Time
1960
1970
1980
1990
2000
2010
2020

Takeaways from the data plots

- The primary deficit relative to debt, which over time must average to minus the required real rate of return on debt, spiked up in 1975, before the late-70's burst in inflation. The only comparable spike is the current one.

- The ratios of variances of the differenced data across variables clearly varies over time, giving hope that IDH will produce usably sharp results.

- Whether the results allow economic interpretation or not, though, this seems a natural way to control dimensionality in allowiing for time varying heteroskedasticity.

- And of course for the reasons already discussed in other papers in this conference, allowing time-varying heteroskedasticity might avoid the model's being thrown off track by a period, like the present, where there are extremely large forecast errors.

Fixed relative variance regimes

- The results I'll show are based on assuming constancy of relative structural variances within "regimes" ending at the beginnings of 1960, 1970, 1984, 2000, 2010, and 2021.
- Note that consistency of estimates of the model's constant linear structure does not require that variances of structural shocks must be constant within regimes — only that on average they are different. Sims (2020)

Prior

- Dummy observation (i.e. conjugate) Minnesota prior, including unit root (with low weight) and "co-persistence" (with higher weight) components.

- Prior on A_0, the lead coefficient in the $A(L)y - c + \varepsilon$ specification, has 100 times the identity as mean, standard deviation 200 on all the coefficients.

- Prior on the 7 by 6 matrix of relative variances is scaled Dirichlet for each row, independent across rows, mean 1 for all elements.

Initial conditions, cointegration, non-stationarity

- A Bayesian approach avoids the hall of mirrors that frequentist inference encounters in the presence of possible non-stationarity, because the Gaussian likelihood does not change form when non-stationarity is present.

- However, there is still a messy issue for Bayesian inference with possible non-stationarity.

- Differencing, or pre-processing to remove trend, is tremendously wasteful of information, particular (as in the current project) we are interested in slow-moving aspects of the data.

Distribution of the initial conditions

- In stationary cases, the distribution of initial conditions can be derived from the VAR model coefficients. Using the full sample by using the marginal distribution of the initial conditions makes likelihood computation harder, but it is possible. In long stationary time series it may not make much difference.

- But if we condition on initial conditions, we are in effect saying that we are happy with model estimates that imply the initial conditions are so far from steady state that a similar deviation is not likely over the whole subsequent sample period.

- This creates the possibility of estimates that imply that the main low frequency patterns in the data were predictable from transients generated by the rare initial conditions.

What to do about this

- The Minnesota prior by itself tends to reduce the problem. It pulls complex roots toward 1, and with its "co-persistence" dummy observations penalizes large deviations between mean initial conditions and the implied unconditional mean.

- One can estimate the model using such a prior, but conditioning on initial conditions, then check to see if there is a problem. Most of the time, I have found there isn't a problem.

Definition of "there isn't a problem"

- We know that when its roots are distinct we can transform the system from one in y to one in $z = P^{-1}y$, where P is the matrix of right eigenvectors of the system matrix and each element of the z vector dollows a univariate AR.

- Roots within $1/T$ of 1, where T is sample size, imply z components that generate near-linear trends over the sample period. They do this by implying an unconditional mean for that z_i that is outside range of sample observed values, with the trend generated as a slow convergence to this extremely long run mean.

- Such components make sense, and we would not want to eliminate them by forcing the initial value of z_i to be close to its long run mean.

- z_i values corresponding to roots much farther from 1 than 1/T — e.g. $2/T$ — imply a half life or a period that is well within the sample period. If such a z_i is far from its unconditional mean, it generates complicated behavior in the forecast of the full sample based on initial conditions only.

- But if there is a clear separation between roots closer than $1/T$ to 1 and the other roots correspond to initial z_i's not far, in unconditional standard deviation units, from their unconditional means, then again there isn't a problem.

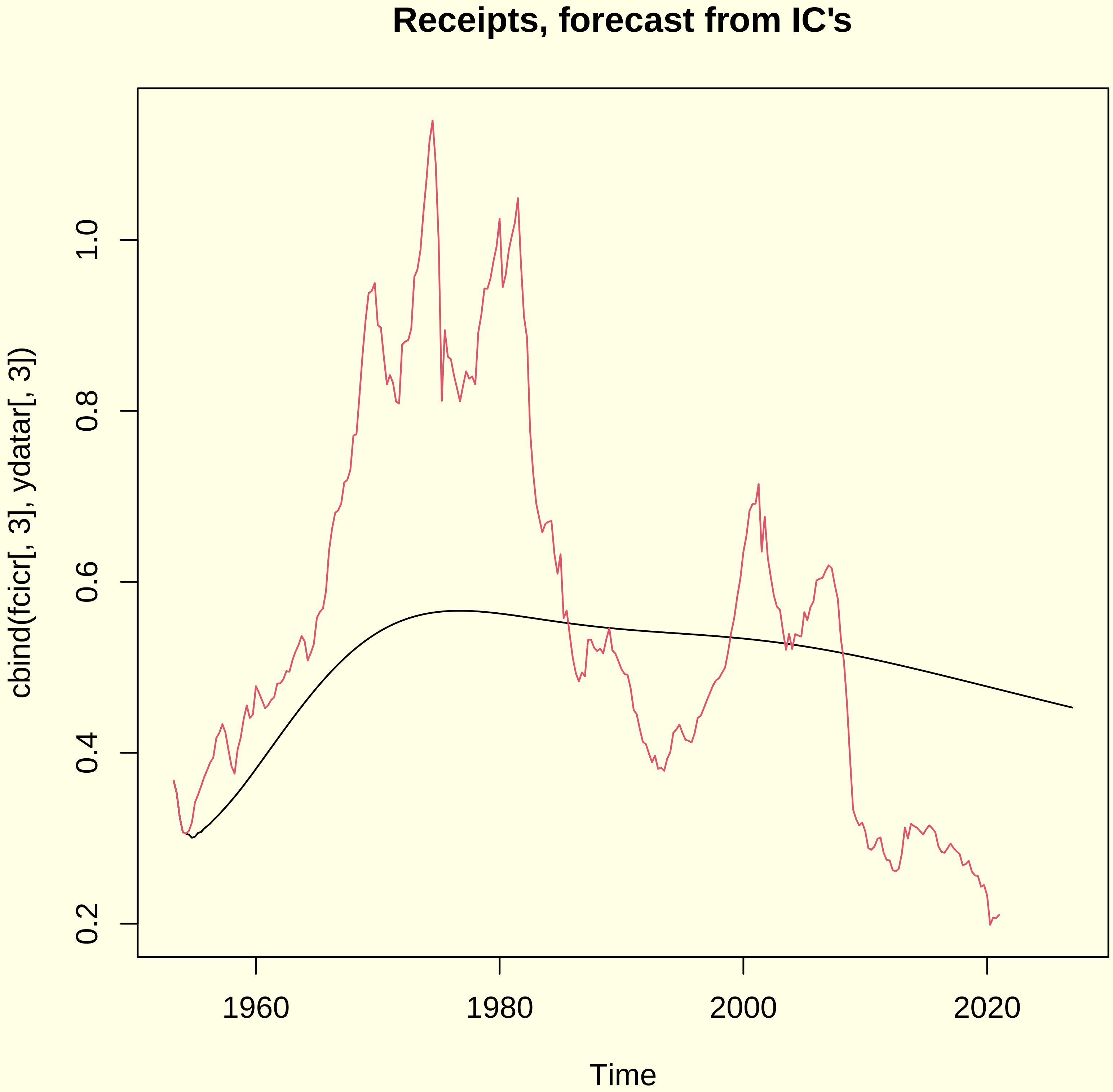
Alas, in this data set, there is a problem

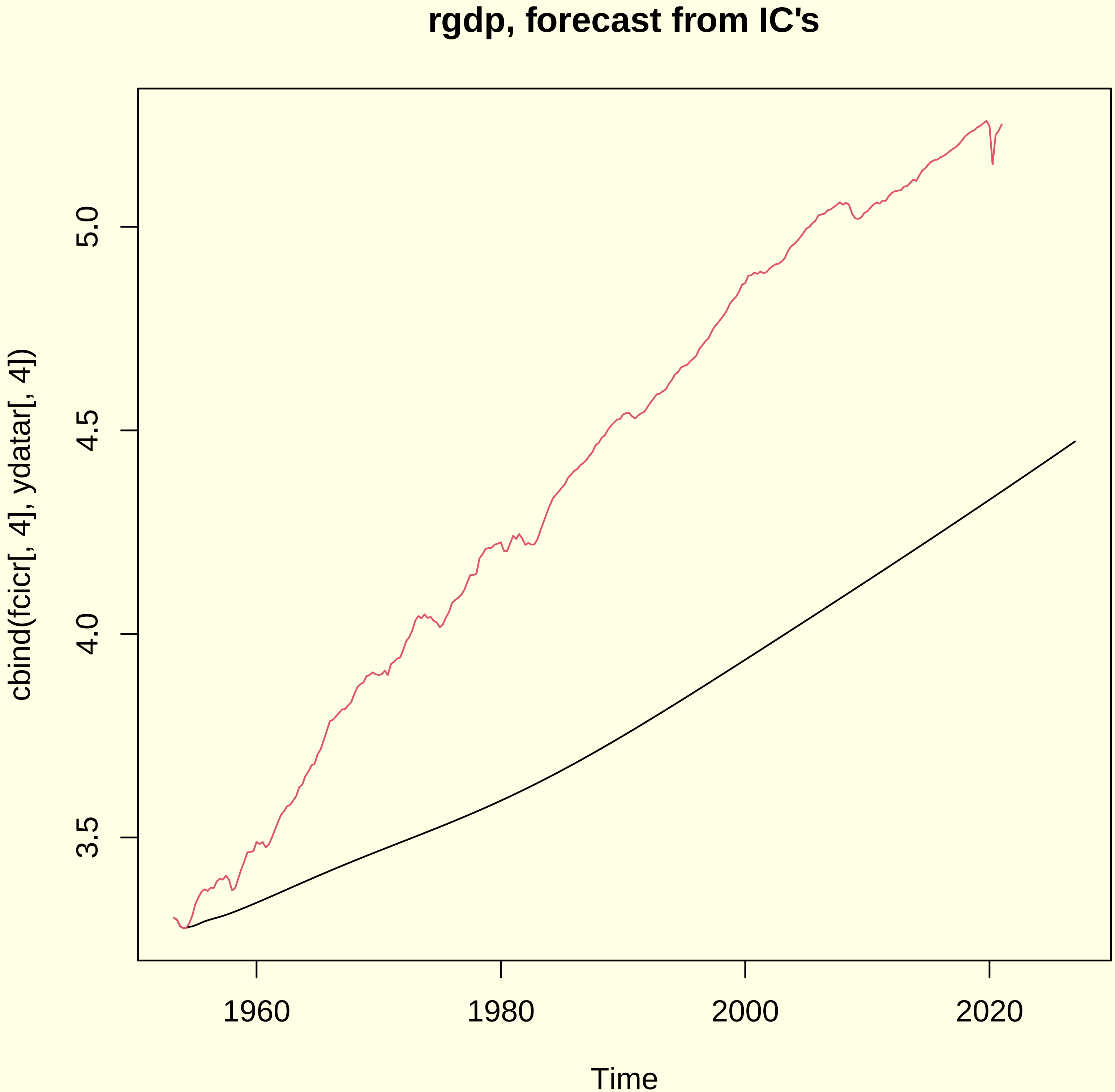
For the 7 largest roots r_i,

$$\frac{1}{|1-r_i|} = 330, , 129, 129, 22, 22, 32, 6$$
$$\frac{z_{i0}-\bar{z}}{\sigma_i} = \text{NA}, 2.5, 2.2, 6.3, 6.2, 7.5, 3.2$$

For the remaining roots, the deviations from steady state are all less than 3.3 standard deviation units.

The problem shows up in full-sample forecasts


Receipts, forecast from IC's
cbind(fcicr[, 3], ydatar[, 3])
0.2
0.4
0.6
0.8
1.0
1960
1980
2000
2020
Time



rgdp, forecast from IC's
cbind(fcicr[, 4], ydatar[, 4])
5.0
4.5
4.0
3.5
1960
1980
2000
2020
Time

What next?

- Try tightening the co-persistence dummy observation
- Try using unconditional distribution of stationary z_i's in the likelihood.
- The first is easy. The second prevents analytic computation of the marginal data density conditional on A_0 and the relative variance parameters, which will greatly slow computation.

Conclusion

- No conclusion, yet. But I hope this has given you an idea of the process by which I might arrive at one.

- In case you're interested, the model at this stage does not see 2021 as like mid-1975: It forecasts inflation over the next 10 years as around 2%, while the debt over that period grows to 222% of GDP.

*

References

Brunnermeier, M., D. Palia, K. A. Sastry, and C. A. Sims (2018): "Feedbacks: Financial Markets and Economic Activity," Discussion paper, Princeton University, http://sims.princeton.edu/yftp/bpss/.

Leamer, E. E. (1978): *Specification Searches: Ad Hoc Inference with Non Experimental Data*. John Wiley and Suns.

Sims, C. A. (2020): "SVAR Identification Through Heteroskedasticity with Misspecified Regimes," Discussion paper, Princeton University, http://sims.princeton.edu/yftp/bpss/IDHmsspcfdRgms.pdf.