Hi, I’m a Macro professor
and I’m here to help!

Jon Faust

http://e105.org/e607
Very happy to be back at KC Fed
My time as RA here (1981-1983) formative in many ways
e.g., Bryants, no finer sauce
also in Macro

- July 1982: strip mall bank the 10th district made some stupid loans and shook the financial system
  Penn Square caused collapse of Continental Illinois

- August 1982: learned many bigger banks had lent multiples of capital to LDCs who wouldn’t be able to repay
Macro modelling also in crisis

- Sims, 1980
  ‘...claims for identification in these models cannot be taken seriously.’

- Lucas, 1981
  ‘As an advice giving profession we are in way over our heads’
Things were so different . . .

- Fortunately, all that’s behind us
Doh!

They’re doing show trials for DSGE models. . .
Solow:

Especially when it comes to matters as important as macroeconomics, a mainstream economist like me insists that every proposition must pass the smell test: does this really make sense? I do not think that the currently popular DSGE models pass the smell test.
Chari,

A useful aphorism in macroeconomics is: ‘If you have an interesting and coherent story to tell, you can tell it in a DSGE model. If you cannot, your story is incoherent.’
Almost as good as coming home to K.C.

- Two lions of their camps eloquently re-stating positions they’ve held for my whole career
- Both miss the point
Suppose we add a constraint that the model be solvable.

If you have an interesting and coherent story to tell, you can tell it in a DSGE model that we can currently specify, solve, and manipulate...
A Solowesque reply to Solow

- Defended young Hicks against older Hicks’s savage treatment of the IS/LM model
A Solowesque reply to Solow

- Older Hicks was right: IS/LM model doesn’t pass the smell test
- Younger Solow was right: IS/LM was a formalization of some key things and helped advance understanding
- There should be no question that the same is true of DSGE models
Uninteresting question

- Is the DSGE glass nearly empty or virtually full?
More interesting to me

- There is at least enough liquid to drink sparingly...

- ...but what *is* that liquid anyway?
  let’s pay close attention to what it is we are drinking.
More interesting to me

- How best can we use macro models (such as they are) to improve the reliability of the monetary policymaking process?
Main point

- It takes highly disciplined analysis to avoid pitfalls.
- I’ll give 7 suggestions
Enter meeting at $t$ with last period’s optimal path

$$i_t^{*|t-1}, i_{t+1}^{*|t-1}, \ldots$$

Revise this path in light of news arriving b/t $t - 1$ and $t$. 
Policy analysis:

- Update perceived optimal policy path in light of structural interpretation of news.
In a linear Gaussian (DSGE-model) world news is one-step forecast errors in observables, $Z_t$:

$$\nu_t = Z_t - Z_{t|t-1}$$
Revision to policy path a function of the structural interp. of the news.

In the VAR case,

\[ \nu_t = C \varepsilon_t \]

where \( \text{vcov}(\varepsilon) = I \).

Or given any inferred \( \hat{\nu}_t \):

\[ \hat{\varepsilon}_t = C^{-1} \hat{\nu}_t \]
VARMA case a bit different
but not enough to matter for this talk
Model-based policy analysis

- Does the model get the news right?
  Purely a question of reduced form forecasting

- Does the model get the structural interp. right?
  Purely a question of $C$, the impact matrix for structural shocks
Suggestions

- I’ll first focus on the news (pure forecasting)
- Then structural interp. (purely about $C'$)
3 papers

- with Jonathan Wright (FW).
  Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset, REStat 2009.

- Rochelle Edge and Refet Gurkaynak
  How useful are DSGE model forecasts for Central Bankers? forthcoming BPEA

- with Abhishek Gupta,
  Posterior predictive analysis for DSGE modeling, (up shortly on my website)
1. Forecast evaluation: You should do it

- Easy to get excited about wonders of a new model
- But all too often we find shiny new models are worse than useless in forecasting

Meese-Rogoff re: exchange rate models the classic example
2. Real-time/vintage data issues matter

- Serious evaluation for practical real-time forecasting requires real-time data.

- We now know that both model rankings and absolute quality measures may be different in real-time data.

  real-time data often is not a huge issue, but matters in enough cases to be worth the bother.
Old hat

- I’m sure these first two are old hat to this group
3. Nowcast (& backcast) are different

- All ‘forecasts’ start with nowcast and backcast
- For now and backcast, have the option of ‘bean counting’
  replicate the data agency; data construction, not economic modelling
3. Nowcast (& backcast) are different

- Fed does this
  Fed’s nowcast is really good
Greenbook v. univariate AR(4), GDP Growth

RMSE

<table>
<thead>
<tr>
<th>hor:</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB</td>
<td>2.17</td>
<td>2.75</td>
</tr>
<tr>
<td>AR</td>
<td>2.77</td>
<td>2.76</td>
</tr>
</tbody>
</table>

From FW.
The picture is from Edge-Gurkaynak
They label the nowcast as horizon 1.
Note 2: Warning label

- In this talk, I am using selective, provocative reporting not being thorough

- Attempting to motivate you to consider these points
Backcast and nowcast

- Every good forecast should start with a sophisticated nowcast which may involve different machinery than the forecast

- Corollary: practical forecast comparison should give all models a good nowcast which may involve different machinery than the forecast

- FW does this, and it matters
4. $H^3$: Heavy-handedness helps

- Very strong ad hoc restrictions often help (overfitting hard to avoid)
- Almost impossible to forecast well without draconian restrictions
  This is very consistent across many studies
4. $H^3$: Heavy-handedness helps

- In FW: often hard to benefit from more than 1 data series
  GDP: univariate AR does about best

- Among multi-variate methods: simple average of simple models always among the best
A Faustian Question

- Would you trade all ability to change your forecast for one bit of future knowledge: the ex post mean?
Figure 3. RMSEs of Alternative Forecasts

- x: DSGE, o: mean diamond: GB, box: BVAR

ex post mean; x: DSGE; ◊: GB; □: BVAR
In other work

- I have found similar results for inflation report forecasts.
- Of course, the mean isn't known ex ante, but we can approximate this in realtime.
Dr. Wright’s democratic prior

- Shrink very heavily toward a real-time guess at the *long-run* mean
e.g., a survey long-term expectation

- Works very well
  Jonathan Wright, Evaluating Real-Time VAR Forecasts with an Informative democratic Prior
H\textsuperscript{3} is well-established

Without H\textsuperscript{3} you are toast
5. No evidence formal econ. helps

- Deliberately contentious
  perhaps overstated
- I mean: fairly strict adherence to behavioral restrictions from formal model has never been shown to help
  (no implied indictment of judgemental use of economic wisdom)
5. No evidence formal econ. helps

- Some formal economic models do okay in quasi-realtime work
  Edge-Gurkaynak verify this

- When formal economic models have done well, no evidence that the economics is anything but $H^3$ in disguise
5. No evidence formal econ. helps

- No example exists of an economic model selected *ex ante* doing well for headline numbers
6. Quasi means not

- Generally true
- but especially in ‘quasi-realtime’
- We have one dataset
  * one collection of vintages
- If we search, we can find a model that does well by any criterion
6. Quasi means not

- Remember: we give rewards to optimizing agents who ‘find a model that fits’
  e.g. Smets-Wouters
- Few rewards for those who count the failed attempts...
- What would an economist predict?
Since Meese-Rogoff, economists have been trying to show some formal model has value forecasting exchange rates.

Re-do Mark using 30 vintages surrounding the one Mark used
As if Nelson did his exact work every quarter or so using the latest vintage
Key result

- Plot Mark’s $p$-value for rejecting ‘no predictability’ of $$/DM$ exchange rate
$p$-value against vintage date

Mark missed the minimum $p$-value only slightly
Quasi means not

- Quasi realtime (in an optimizing profession) means not realtime
- Very difficult to adjust for
Quasi means not

- So what do we do?
- Use economics!
7. Stress test the structural interp.

- We need structure for policy analysis
- Pushing on the plausibility of the structural implications can also help distinguish economic wisdom from $H^3$. 
One form of stress test

■ Plug for Faust-Gupta
  Abhishek Gupta, just started at Gettysburg College

■ Papers up shortly
Simple idea

- Key structural relation for policy analytics,

\[ \hat{\varepsilon}_t = C^{-1} \hat{\nu}_t \]

\( \hat{\nu}_t \): estimated news
\( \hat{\varepsilon}_t \): implied structural shocks
Simple idea

- As Bayesians, we can ask:
  How likely would the model have been to produce shocks like the $\hat{\varepsilon}$s implied on the sample

- formally: posterior predictive analysis

- Akin to frequentist residual diagnostics
A bit complex to compute for ‘structural’ elements like structural shocks
Adapt ideas of Gelman et. al on posterior predictive anal.

Contentious to ‘orthodox’ Bayesians
But we argue far less problematic in this context
Under the posterior, the estimated correlation of the structural shocks on the sample was quite high, almost all mass far from zero.

Probability the model would generate a sample where the estimated structural shocks would be as correlated as those estimated on the SW sample: 0.00.
More troubling

- Partition obs. into those in spans of at least 2 neg. quarters of growth and others
- Variance and correlation of shocks is different during the periods of ‘recessions’
- Shocks are bigger and have diff. correlation structure
More troubling

- Probability the model would generate a sample like this is essentially zero.
Interpretation

- We often say we want a structural model b/c it tells a story

- The story of the SW model is that post-War business cycles were a collective freak draw, never to be repeated

  A highly unlikely confluence of abnormally large and abnormally correlated shocks
Aside

- We show how to use this info. to refine the structure...
Macro and policy modeling were in a precarious position as when I was an RA at FRBKC.

And things are not so different now.
But I’m optimistic

- I don’t think we need to start over as we did in 1980
- We have much better tools, data, and models
  And the benefit of hindsight on mistakes of the 1970s
The glass... 

- The glass is far from empty... 
- ...we should take care as we drink
Main point

- Takes very disciplined analysis to avoid silliness
  (and perhaps policy tragedy)

- I’ve tried to highlight some elements of discipline I find useful
Central Banks

- Fed (and other CBs) have been taking the lead in disciplined analysis
- And this conference is another outstanding example of pushing disciplined, policy-relevant work
Central Banks

- I can’t wait to see what the remainder of the conference has in store