Prediction with Macroeconomic Models

Gianni **Amisano**[1] and John **Geweke**[2]

Amisano: European Central Bank and Un, of Brescia. **Geweke**: University of Technology Sydney and Un. Colorado

Bank of Canada Forecasting Workshop

Ottawa, 6/7 November 2011

[1]Views expressed are not those of the European Central Bank.

[2]Support from Australian Research Council grant DP110104732.

Main idea

- Econometricians inform policy makers with fiduciary responsibilities.
- Now more than ever it is important that we seriously convey uncertainty conditional on a credible information set.
  - Conditional on a model, do not condition on a pseudo-true parameter vector: use full Bayesian inference.
  - Do not condition on a set of models: recognize that all models are false.
- Whether or not you agree with these ideas, incorporating wider uncertainty produces better predictions.
  - Specifically, it assigns higher probability to events that actually occur.

Outline

1. Overview and main findings
2. Models and data
3. Prediction with individual models
4. Prediction with several models
5. Conclusion

Four levels of uncertainty

1. Intrinsic uncertainty about the future conditional on a model and parameters
2. Extrinsic uncertainty about model parameters conditional on a model
3. Uncertainty about models conditional on a set of models
4. Unconditional uncertainty: All models are false

Here we concentrate on levels 2 and 4.

Objective

- Objective: Incorporate all four levels of uncertainty
- Assess the improvement in the quality of prediction.
- Prediction quality for densities p_{t-1} announced at time $t-1$ for events $\mathbf{y}_t$ at time t, recorded as $\mathbf{y}_t^o$:

$$\left[\prod_{t=1}^{T} p_{t-1} \left(\mathbf{y}_t^o \right) \right]^{1/T}$$

- If $p_{t-1}^{(1)}$ and $p_{t-1}^{(2)}$ are two such predictive densities, then the relative prediction quality is

$$\left[\prod_{t=1}^{T} p_{t-1}^{(1)} \left(\mathbf{y}_t^o \right) \right]^{1/T} \Big/ \left[\prod_{t=1}^{T} p_{t-1}^{(2)} \left(\mathbf{y}_t^o \right) \right]^{1/T} .$$

- These are simple transformations of log scoring functions.

The exercise

- Three models, each representative of their genre:
  - Vector autoregression (VAR)
  - Dynamic stochastic general equilibrium model (DSGE)
  - Dynamic factor model (DFM)
- Vector of macroeconomic time series predicted, $\mathbf{y}_t$:

Real consumption growth rate	Weekly hours worked
Real investment growth rate	Inflation rate
Real output growth rate	Federal funds rate
Real wage growth rate	

- All work is strictly out-of-sample, rolling through one-quarter-ahead predictive densities, 1966:1 – 2010:3.

Main findings, level 2 (parameter uncertainty)

- Comparison
  - Full Bayesian inference
    ... incorporates model-specific extrinsic uncertainty
  - Posterior mode “plug-in”
    ... ignores model-specific extrinsic uncertainty
- Quality of prediction ignoring extrinsic uncertainty (posterior mode) relative to incorporating extrinsic uncertainty (full Bayes) – Geometric mean of probability ratios:
  - DSGE: 0.830
  - VAR: 0.432
  - DFM: 0.721
- There is interesting and rich detail underlying these findings.
  - The advantage of full Bayes over posterior mode prediction is most pronounced in periods of unusual behaviour.
  - This is an implication of the econometrics and what we observe in fact.

Main findings, level 4 (unconditional uncertainty)

- An equally weighted pool of models

$$\frac{1}{3}\left[p_{t-1}^{(DSGE)}\left(\mathbf{y}_t\right)+p_{t-1}^{(VAR)}\left(\mathbf{y}_t\right)+p_{t-1}^{(DFM)}\left(\mathbf{y}_t\right)\right]$$

provides predictions of superior quality.

- Prediction quality relative to this pool:
    - DSGE (full Bayes): 0.681
    - VAR (full Bayes): 0.612
    - DFM (full Bayes): 0.738
    - Bayesian model averaging (real-time weights): 0.734
    - Optimal pool (real-time weights): 0.977

Data: An extension of Smets and Wouters (2007) data set

Quarterly U.S. data, 1951:I - 2010:3

1. Consumption: growth rate in per capita real consumption
2. Investment: growth rate in per capita real investment
3. Output: growth rate in per capita real GDP
4. Hours: log per capita weekly hours
5. Inflation: growth rate in GDP deflator
6. Real wage: growth rate in real wage
7. Interest rate: Federal Funds Rate

Additional series for DFM

1. Stock returns: Growth rate in S&P 500 index
2. Unemployment rate
3. Term premium: 10 year and 3 month bond rates spread
4. Risk premium: BAA and AAA corporate bond spread
5. Money growth: Growth rate in M2

Vector autoregression (VAR) model

- Conventional VAR with Minnesota priors
- Four lags of each variable

Dynamic stochastic general equilibrium (DSGE) model

- Model described in Smets and Wouters, AER 2007
- DSGE model with nominal frictions: price and wage stickiness, monopolistic competition
- Seven structural shocks: total factor productivity, risk premium, investment specific tech shock, wage mark up, price mark up, exogenous government spending, monetary shock
- "The marginal likelihood criterion, which captures the out-of-sample prediction performance, is used to test the [DSGE] model against standard and Bayesian VAR models. We find that the [DSGE] model has a fit comparable to that of Bayesian VAR models." (p. 587)

Dynamic factor model (DFM)

- Model specification following Stock and Watson (2005, NBER working paper).
  - $k = 3$ common factors with VAR dynamics
  - $n = 12$ idiosyncratic terms with AR dynamics
- Structure:
  - $\underset{(12\times 1)}{\mathbf{y}_t} = \Gamma \underset{(3\times 1)}{\mathbf{f}_t} + \mathbf{v}_t$
  - $b_i(L)v_{it} = \varepsilon_{it}, i = 1, 2, ...12;$ lag length 2; $\boldsymbol{\varepsilon}_t \overset{iid}{\sim} N\left(\mathbf{0}, diag(\boldsymbol{\sigma})\right)$
  - $\mathbf{A}(L)\mathbf{f}_t = \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \overset{iid}{\sim} N(\mathbf{0}, \mathbf{I}_3);$ lag length 2
- Marginal predictive distribution for first 7 variables used in this study

Two kinds of prediction

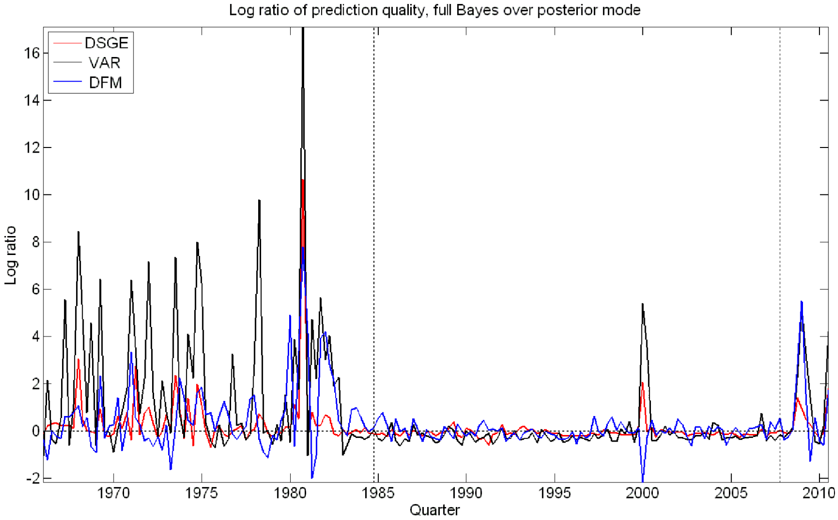
- Abstract representation of the model:

$$\mathbf{y}_t \sim p\left(\mathbf{y}_t \mid \mathbf{Y}^o_{t-1}, \boldsymbol{\theta}_j,\ \mathit{Model}\ j\right)$$

  - $j =$ VAR, DSGE or DFM
  - $\boldsymbol{\theta}_j$ is the parameter vector of model j
  - $\mathbf{Y}^o_{t-1}$ is the data through quarter $t-1$
- Posterior mode prediction: for each $t = 1, \ldots, 179$:
  - $\widehat{\boldsymbol{\theta}}_{j,t-1} = \arg\max_{\boldsymbol{\theta}_j} p\left(\boldsymbol{\theta}_j \mid \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)$
  - Evaluate $p\left(\mathbf{y}^o_t \mid \mathbf{Y}^o_{t-1}, \widehat{\boldsymbol{\theta}}_{j,t-1}, \mathit{Model}\ j\right)$
- Full Bayes prediction: for each $t = 1, \ldots, 179$:
  - $\boldsymbol{\theta}^{(m)}_{j,t-1} \sim p\left(\boldsymbol{\theta}_j \mid \mathbf{Y}^o_{t-1}, j\right) \quad \left(m = 1, \ldots, 10^4\right)$
  - Evaluate $p\left(\mathbf{y}^o_t \mid \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)$

$$\approxeq 10^{-4} \sum_{m=1}^{10^4} p\left(\mathbf{y}_t \mid \mathbf{Y}^o_{t-1}, \boldsymbol{\theta}^{(m)}_{j,t-1}, \mathit{Model}\ j\right)$$

Full Bayes vs posterior mode



Prediction quality in the three models

- Prediction quality of posterior mode relative to full Bayes:

$$\left[\prod_{t\in\tau}\frac{p\left(\mathbf{y}_t^o \mid \mathbf{Y}_{t-1}^o, \widehat{\boldsymbol{\theta}}_j, \textit{Model } j\right)}{p\left(\mathbf{y}_t^o \mid \mathbf{Y}_{t-1}^o, \textit{Model } j\right)}\right]^{1/\#(\tau)}$$

  - This is the geometric mean of the relative predictive densities in a particular set of quarters τ.

Set of quarters τ	DSGE	VAR	DFM
Entire period 1966:1-2010:3	0.830	0.432	0.721
Pre-moderation 1966:1-1984:4	0.641	0.151	0.514
Great moderation 1985:1-2007:4	1.055	1.104	1.015
Post global financial crisis 2008:1-2010:3	0.737	0.293	0.401
Expansions (NBER)	0.860	0.527	0.879
Contractions (NBER)	0.708	0.180	0.300

Prediction quality for individual series

- Prediction quality for posterior mode relative to full Bayes for series i $(i = 1, \ldots, 7)$:

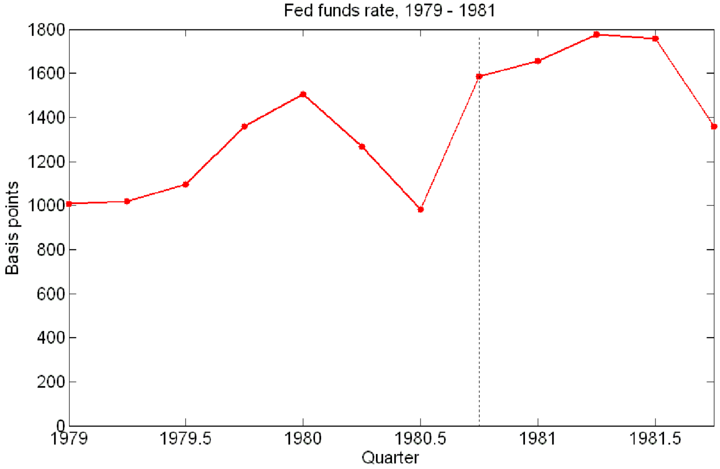
$$\left[\prod_{t=1}^{T} \frac{p\left(y_{it}^{o} \mid \mathbf{Y}_{t-1}^{o}, \widehat{\boldsymbol{\theta}}_{j}, Model\ j\right)}{p\left(y_{it}^{o} \mid \mathbf{Y}_{t-1}^{o}, Model\ j\right)} \right]^{1/\#(\tau)}$$

Series	DSGE	VAR	DFM
Consumption	1.015	0.999	0.982
Investment	1.001	1.008	0.975
Output	1.020	1.005	0.953
Hours	1.008	0.994	0.978
Inflation	1.004	1.016	0.997
Real wage	0.990	0.943	0.947
Fed funds rate	0.991	0.736	0.856

Decomposition by quarter: Most extreme quarters shown

	All 7 series			Fed funds rate		
	DSGE	VAR	DFM	DSGE	VAR	DFM
1968:1	0.0468					
1971:2	0.0637	0.0002				
1973:3	0.0928	0.0006		0.111	0.0045	
1974:4		0.0003		0.219	0.0023	0.159
1978:2		5×10^{-5}				
1980:1			0.0074			
1980:4	2×10^{-5}	4×10^{-8}	0.0004	0.0003	4×10^{-8}	0.0002
1981:1			0.0155			
1982:1			0.0148			
2000:1	0.1260					
2009:1			0.0041			

Role of fed funds rate variations (79-81)



Analysis of predictive variance

- Basic decomposition (Rao-Blackwell Theorem):

$$\begin{aligned} \mathrm{var}\left(\mathbf{y}_t \mid \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right) &= \mathrm{var}_{\boldsymbol{\theta}_j}\left[\mathrm{E}\left(\mathbf{y}_t \mid \boldsymbol{\theta}_j, \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)\right] \\ &\quad + \mathrm{E}_{\boldsymbol{\theta}_j}\left[\mathrm{var}\left(\mathbf{y}_t \mid \boldsymbol{\theta}_j, \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)\right] \end{aligned}$$

- Extrinsic variance:

$$\mathrm{var}_{\boldsymbol{\theta}_j}\left[\mathrm{E}\left(\mathbf{y}_t \mid \boldsymbol{\theta}_j, \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)\right]$$

- Intrinsic variance:

$$\mathrm{E}_{\boldsymbol{\theta}_j}\left[\mathrm{var}\left(\mathbf{y}_t \mid \boldsymbol{\theta}_j, \mathbf{Y}^o_{t-1}, \mathit{Model}\ j\right)\right]$$

Analysis of predictive variance: Simple example

- Linear regression model

$$y_t = \beta' \mathbf{x}_t + \varepsilon_t \quad \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right)$$

plus a proper prior distribution for $\left(\boldsymbol{\beta}, \sigma^2\right)$

- Then

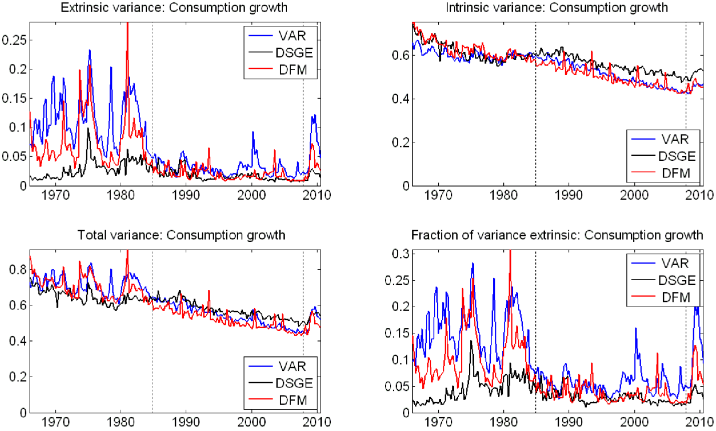
$$\begin{aligned} & \mathrm{var}\left(y \mid \mathbf{x}^*, Model, Data\right) \\ = \quad & \underbrace{\mathbf{x}^{*\prime}\left[\mathrm{var}\left(\boldsymbol{\beta} \mid Model, Data\right)\right]\mathbf{x}^*}_{Extrinsic\ variance} \\ & + \underbrace{\mathrm{E}\left(\sigma^2 \mid Model, Data\right)}_{Intrinsic\ variance} \end{aligned}$$

Analysis of variance computation

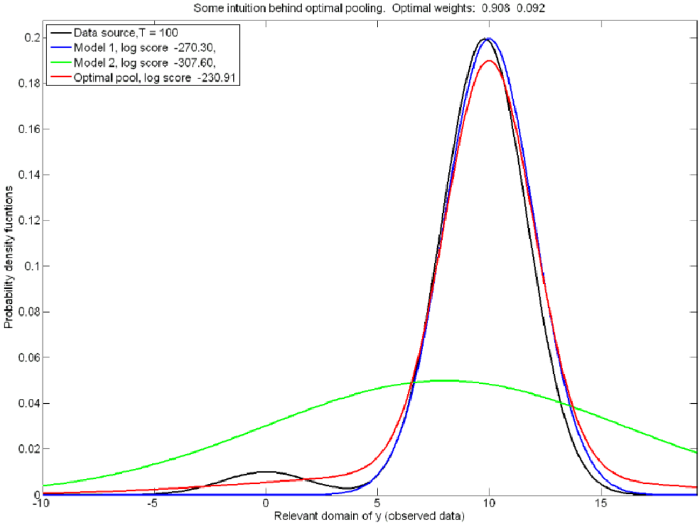
- Thin the original MCMC sample from $p\left(\boldsymbol{\theta} \mid \mathbf{Y}_{t-1}^{0},\ Model\ j\right)$ to 100 draws $\boldsymbol{\theta}^{(m)}$ $(m = 1, \ldots, 100)$
- For each $\theta^{(m)}$, draw $\mathbf{y}_t^{(m,i)} \sim p\left(\mathbf{y}_t \mid \mathbf{Y}_{t-1}^{0}, \boldsymbol{\theta}^{(m)}\ Model\ j\right)$ $(i = 1, \ldots, 100)$.
- Select a particular element y_{ts} (e.g. $s =$ consumption growth, output growth, ...) for study
- Undertake a classical one-way analysis of variance on $\left\{y_{ts}^{(m,i)}\right\}$ treating the parameter vector as the single factor.
- Decompose total variance into extrinsic variance (variance due to the factor) and intrinsic variance (the remainder).

Variance components for

ΔC



Model pooling: the intuition



Model pooling

- A pool of predictive densities from models $\mathcal{M}_i$ $(i = 1, \ldots, n)$ is the sequence

$$p\left(\mathbf{y}_t; \mathbf{Y}^o_{t-1}\right) = \sum_{i=1}^{n} w_i p\left(\mathbf{y}_t; \mathbf{Y}^o_{t-1}, \mathcal{M}_i\right)$$

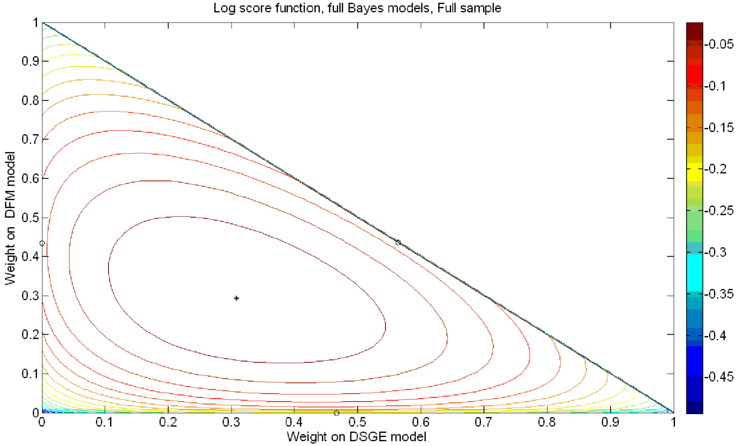
  - $p\left(\mathbf{y}_t; \mathbf{Y}^o_{t-1}, \mathcal{M}_i\right)$ is the predictive density from model $\mathcal{M}_i$;
  - $w_i \geq 0$ $(i = 1, \ldots, n)$, $\sum_{i=1}^{n} w_i = 1$, i.e. $\mathbf{w} = (w_1, \ldots, w_n)'$ is contained in the unit simplex.
  - For more details see Geweke and Amisano (*J Econometrics* 2011).

- Since we have $n = 3$ models, the average log score of the pool

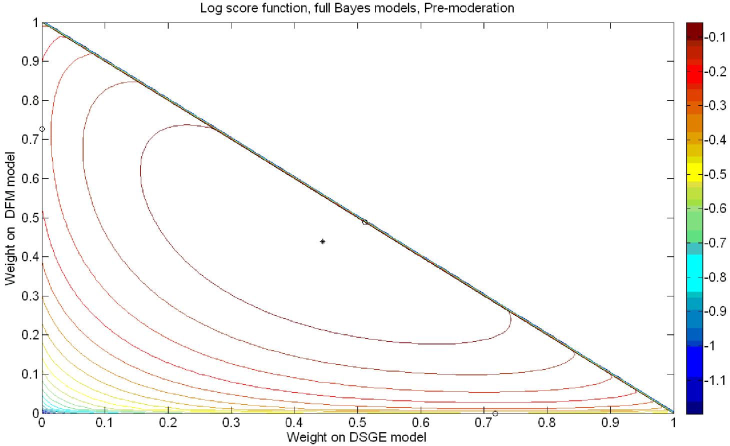
$$f\left(\mathbf{w}\right) = T^{-1} \sum_{t=1}^{T} \log \left[\sum_{i=1}^{n} w_i p\left(\mathbf{y}^o_t; \mathbf{Y}^o_{t-1}, \mathcal{M}_i\right) \right]$$

can be shown graphically

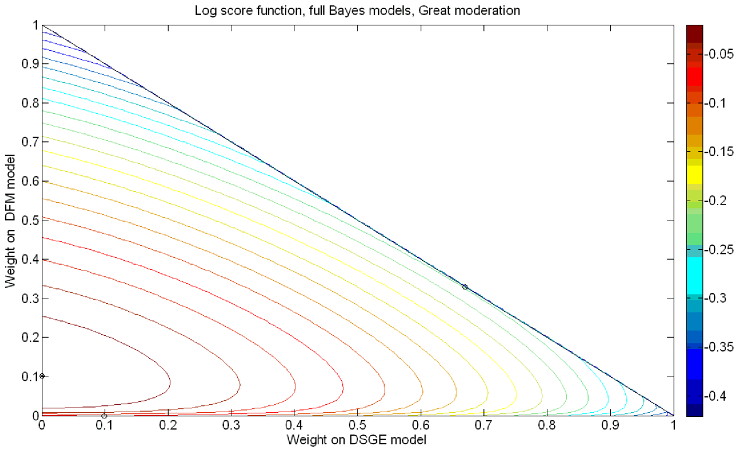
Optimal weights, full sample



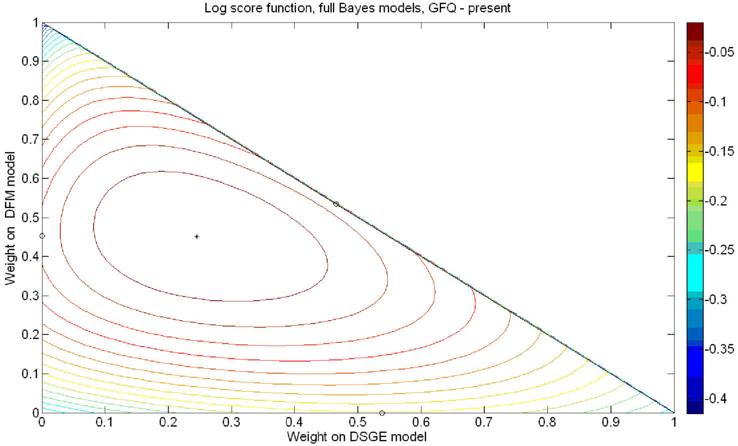
Optimal weights, pre-Great Moderation



Optimal weights, Great Moderation



Optimal weights, Great Financial Crisis



The equally-weighted pool of the three models

- This will serve as a benchmark for the rest of the presentation.
- Easy to understand:

$$p\left(\mathbf{y}_t \mid \mathbf{Y}_{t-1}^o, Pool\right) = \frac{1}{3}\sum_{j=1}^{3} p\left(\mathbf{y}_t \mid \mathbf{Y}_{t-1}^o, Model\ j\right)$$

- Outperforms
  - Each model (by a lot)
  - Bayesian model averaging (by a lot)
  - Real time optimal pooling (by a little)
- We will look at the value of each model as well.

Performance relative to equally weighted pool

(Geometric mean ratio over quarters)

Period	DSGE	VAR	DFM	Opt Pool	BMA
1966:1-2010:3	0.681	0.612	0.738	0.977	0.734
1966:1-1984:4	0.603	0.312	0.768	0.982	0.757
1985:1-2007:4	0.766	1.017	0.729	0.970	0.729
2008:1-2010:3	0.600	0.902	0.619	1.002	0.619
Expansions	0.666	0.676	0.727	0.962	0.712
Contractions	0.751	0.391	0.791	1.045	0.836

Model value

- The value of model j in an equally weighted pool of n models is

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\log\left[\frac{1}{n}\sum_{i=1}^{n}p\left(\mathbf{y}_t \mid Model\ \ i\right)\right] - \log\left[\frac{1}{n-1}\sum_{i\neq j}^{n}p\left(\mathbf{y}_t \mid Model\ \ i\right)\right]\right\} \quad (1)$$

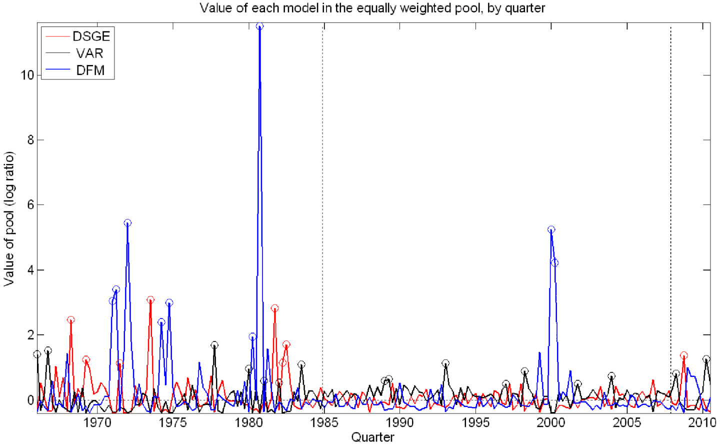
- It's the incremental improvement in the predictive likelihood when the model is added to an equally weighted pool.
- Value can be decomposed observation by observation as indicated in (1).
  - Each term is bounded below by $\log\left[\left(n-1\right)/n\right]$.
  - There is no upper bound.

Model values

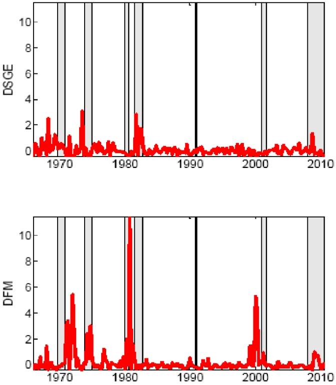
(Log scale, quarterly averages)

Period	DSGE	VAR	DFM
1966:1-2010:3	0.080	0.081	0.215
1966:1-1984:4	0.201	-0.017	0.450
1985:1-2007:4	-0.008	0.150	0.033
2008:1-2010:3	-0.020	0.189	0.116
Expansions	0.059	0.102	0.208
Contractions	0.171	-0.011	0.246

Model values, graphically (I)



Model values, graphically (II)



Conclusion: Be more realistic about uncertainty when predicting

- Models provide predictive distributions for the future conditional on the past.
- When recent behavior has been unusual, relative to the past, extrinsic uncertainty will be high.
- During such periods fully incorporating extrinsic uncertainty in predictive distributions is critical.
  - Follows from a straightforward econometric argument.
  - Consistent with the record of the US economy and three canonical models.
- When recent behavior has been unusual, differences in predictive distributions across models will be most pronounced.
- During such periods the returns to model pooling are highest.