

Example of forecasts: The Economist Poll of Forecasters

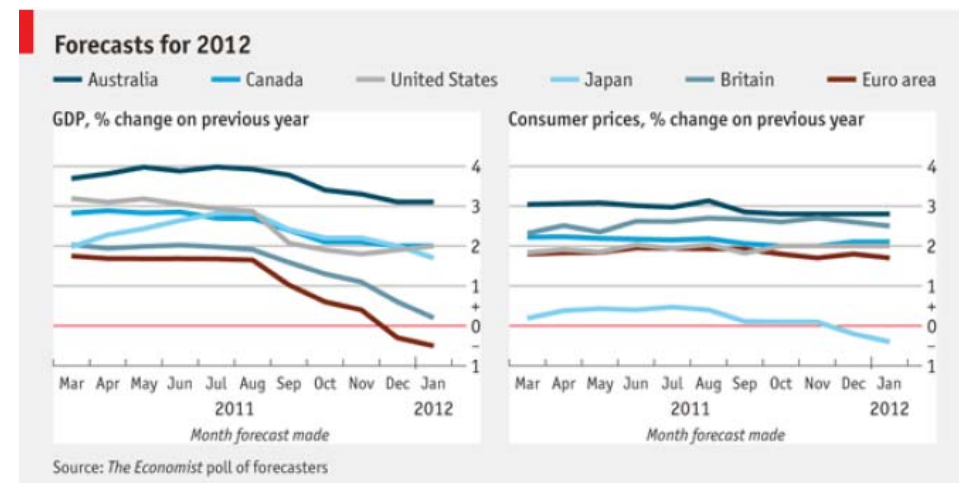
Forecasting with reduced-form models

Empirical Macroeconomics - Lect 3

Dr. Ana Beatriz Galvao

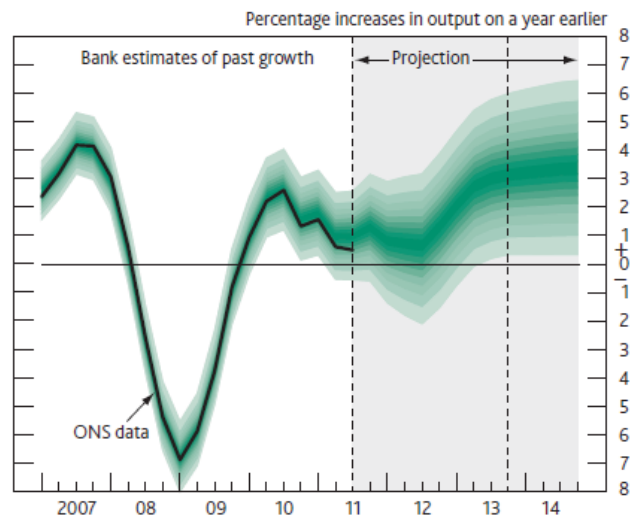
Queen Mary University of London

January 2012



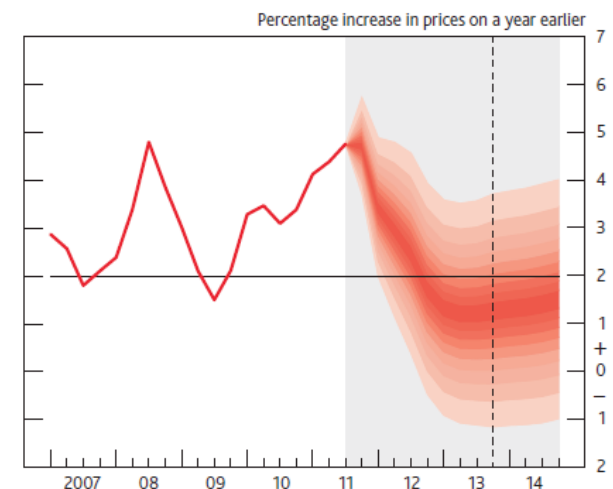
Example of Forecasts: Bank of England

Chart 5.1 GDP projection based on market interest rate expectations and £275 billion asset purchases



Example of Forecasts: Bank of England

Chart 5.6 CPI inflation projection based on market interest rate expectations and £275 billion asset purchases



Which Model for Forecasting?

- Trade-off between theoretical coherence and empirical coherence.
- Reduced-form VARs or DSGE models?
- Additional reduced-form models: Factor models, Autoregressive Distributed Lag Models, Nowcasting state-space models.
- Many papers in the forecasting literature support the claim that models that have good theoretical coherence and are adequate for policy analysis are not very good for forecasting, in particular short-run forecasting.
- Non-models forecasting: survey forecasting and judgemental forecasting.



Forecasting Models I

Models for forecasting US economic activity:

- 1 Autoregressive Distributed Lag Models using leading indicators as predictors: Stock and Watson (2003).
- 2 Large Bayesian Vector Autoregressive Models: Banbura, Giannone and Reichlin (2010).
- 3 Medium-sized DSGE Models: Smets and Wouters (2007). (only in Lecture 10!).

Both (1) and (2) are reduced-form models, while (3) is a structural model. The VAR model has the Wold theorem as theoretical support, and the success of shrinkage in forecasting as an empirical support. Method (1) requires a set of leading indicators (financial and economic variables) to have predictive content about future economic activity.



ADL models I

- Use leading indicators as predictors of output growth.
- Example of indicators: term spread (long - short-rate), fed funds rate, junk bond spread, return on SP500, real price of oil, new claims for unemployment insurance, housing starts, consumer expectations.
- y_{t+h}^h is the output growth over the next h quarters, expressed at annual rate:

$$y_{t+h}^h = (400/h) \ln(Q_{t+h}/Q_t),$$

where Q_t is real output.

- x_{t-i} for $i = 0, \dots, p - 1$ are past observations on an leading indicator.



ADL models II

- The forecasting model (ADL(q,p)) is:

$$y_{t+h}^h = \alpha + \sum_{i=0}^{p-1} \beta_i x_{t-i} + \sum_{i=0}^{q-1} \gamma_i y_{t-i} + u_{t+h}^h$$

for $t = \max(p, q) + h_{\max} + 1, \dots, T$. It uses information of one leading indicator at a time.

- The forecasting model is estimated for each forecast horizon (assume $p = q = 1$):

$$\begin{aligned} y_{t+1}^1 &= \alpha^1 + \beta_0^1 x_t + \gamma_0^1 y_t + u_{t+1}^1 \\ y_{t+2}^2 &= \alpha^2 + \beta_0^2 x_t + \gamma_0^2 y_t + u_{t+2}^2 \end{aligned}$$



ADL models III

As a consequence, if the above models are estimated until $t = N$ ($N + h_{\max} \leq T$), one- and two-step-ahead forecasts computed with information up to N are:

$$\begin{aligned}\hat{y}_{N+1}^1 &= \alpha^1 + \beta_0^1 x_N + \gamma_0^1 y_N \\ \hat{y}_{N+2}^2 &= \alpha^2 + \beta_0^2 x_N + \gamma_0^2 y_N\end{aligned}$$

- To simulate real-time forecasting, an out-of-sample period of size P is defined such that $t = \dots, N + 1, \dots, N + P, \dots, N + P + h_{\max} = T$. At each forecast origin, from $t = N$ up to $t = N + P - 1$, an ADL model is estimated with one leading indicator and for each forecast horizon.



ADL models IV

- This means that one-step-ahead forecast error is:

$$y_{t+1}^1 - \hat{y}_{t+1|t}^1 \text{ for } t = N, \dots, N + P - 1$$

In general:

$$y_{t+h}^h - \hat{y}_{t+h|t}^h \text{ for } t = N, \dots, N + P - 1.$$

- The *forecast accuracy* of the i^{th} leading indicator is measured using the mean squared forecast error (MSFE). This implies that the forecaster has a *quadratic loss function*.

$$MSFE_i = \frac{\sum_{t=N}^{N+P-1} \left(y_{t+h}^h - \hat{y}_{i,t+h|t}^h \right)^2}{N}$$



ADL models V

- The forecasting accuracy of the model with leading indicator i when forecasting h -step-ahead output growth is compared with the accuracy of a benchmark model using:

$$\frac{MSFE_i}{MSFE_{\text{benchmark}}}$$

if the above ratio < 1 , then the model with the indicator i is more accurate than the benchmark model.

- Stock and Watson (2003) use a model with no leading indicator - an autoregressive model - as the benchmark.



ADL models VI

- "While the decline of the stock market gave some advance warning of the 2001 recession, it was not otherwise a reliable indicator during the 1980s and 1990s. Building permits and consumer confidence, which declined sharply preceding and during the 1990 recession, maintained strength well into the 2001 recession. While the term spread indicated an economic slowdown in 2001, it did not give an early signal in the 1990 recession. The varying performance of these indicators reflects the differences in the shocks and economic conditions prior to the 1990 and 2001 recessions." SW, 2003.



Survey Forecasts during the 2001 recession

Table 1 Median Forecasts of the Percentage Growth in Quarterly GDP from the Survey of Professional Forecasters

Target Date		Forecasts Made In							
		2000				2001			
Quarter	Actual Growth	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
2000Q4	1.1	2.9	3.1	3.2	3.2				
2001Q1	-0.6	2.8	2.6	3.0	3.3	0.8			
2001Q2	-1.6		2.9	2.7	3.2	2.2	1.2		
2001Q3	-0.3			3.2	3.3	3.3	2.0	1.2	
2001Q4	2.7				3.2	3.7	2.6	2.8	-1.9
2002Q1	5.0					3.7	3.1	2.7	0.1
2002Q2	1.3						3.6	3.0	2.4
2002Q3	4.0							3.9	3.6

◀ ▶ ⏪ ⏩ ⏴ ⏵ ⏶ ⏷ ⏸ ⏹ ⏺ ⏻ ⏼ ⏽ ⏾ ⏿ 🔍 ↻

Relative MSFE: 1999Q1-2002:Q3 (N=11)

Predictor	Transformation	GDP		IP	
		<i>h</i> = 2	<i>h</i> = 4	<i>h</i> = 2	<i>h</i> = 4
		<i>Root Mean Squared Forecast Error</i>			
Univariate autoregression		2.06	2.03	4.34	4.92
		<i>MSFE Rel. to Univariate AR Model</i>			
Random walk	level	1.26	1.11	1.56	1.17
Interest Rates					
Federal funds	Δ	1.01	0.71	0.97	0.78
90-day T-bill	Δ	1.01	0.76	1.02	0.88
1-year T-bond	Δ	1.17	0.96	1.22	1.06
5-year T-bond	Δ	1.37	1.24	1.38	1.23
10-year T-bond	Δ	1.36	1.26	1.21	1.23
Spreads					
Term spread (10 year–federal funds)*	level	0.86	0.65	0.77	0.72
Term spread (10 year–90-day T-bill)	level	0.87	0.62	0.70	0.62
Paper-bill spread (commercial paper–T-bill)	level	1.31	1.17	1.96	1.43
Junk bond spread (high yield–AAA corporate)	level	0.76	0.65	0.67	0.58

◀ ▶ ⏪ ⏩ ⏴ ⏵ ⏶ ⏷ ⏸ ⏹ ⏺ ⏻ ⏼ ⏽ ⏾ ⏿ 🔍 ↻

Relative MSFE: 1999Q1-2002:Q3 (N=11)

Predictor	Transformation	GDP		IP	
		<i>h</i> = 2	<i>h</i> = 4	<i>h</i> = 2	<i>h</i> = 4
		<i>Root Mean Squared Forecast Error</i>			
Univariate autoregression		2.06	2.03	4.34	4.92
Other Financial Variables					
Exchange rate	Δ ln	0.85	0.87	0.85	0.80
Stock prices*	Δ ln	0.83	0.93	0.64	0.71
Output					
Real GDP	Δ ln			0.92	0.96
IP–total	Δ ln	0.98	1.01		
IP–products	Δ ln	1.03	0.99	1.03	0.96
IP–business equipment	Δ ln	1.00	1.01	1.05	1.06
IP–intermediate products	Δ ln	0.89	0.90	0.89	0.88
IP–materials	Δ ln	0.97	1.01	1.04	0.98
Capacity utilization rate	level	0.91	1.01	0.85	1.03
Labor Market					
Employment	Δ ln	0.96	1.00	0.96	0.99
Unemployment rate	Δ	1.24	1.08	1.31	1.09
Average weekly hours in manufacturing*	level	0.87	0.75	0.72	0.87
New claims for unemployment insurance*	Δ ln	0.75	0.84	0.74	0.81

◀ ▶ ⏪ ⏩ ⏴ ⏵ ⏶ ⏷ ⏸ ⏹ ⏺ ⏻ ⏼ ⏽ ⏾ ⏿ 🔍 ↻

Relative MSFE: 1999Q1-2002:Q3 (N=11)

Predictor	Transformation	GDP		IP	
		<i>h</i> = 2	<i>h</i> = 4	<i>h</i> = 2	<i>h</i> = 4
		<i>Root Mean Squared Forecast Error</i>			
Univariate autoregression		2.06	2.03	4.34	4.92
Other Leading Indicators					
Housing starts (building permits)*	Δ ln	1.30	1.07	1.52	1.14
Vendor performance*	level	1.02	0.97	1.19	0.97
Orders–consumer goods and materials*	Δ ln	0.77	0.83	0.81	0.83
Orders–nondefense capital goods*	Δ ln	1.02	1.03	0.92	1.09
Consumer expectations (Michigan)*	level	1.96	2.14	1.33	1.49
Prices and Wages					
GDP deflator	Δ ² ln	1.00	0.94	0.94	0.84
PCE deflator	Δ ² ln	1.01	1.05	0.99	0.99
PPI	Δ ² ln	1.01	1.02	0.96	0.99
Earnings	Δ ² ln	1.00	1.01	0.89	0.98
Real oil price	Δ ² ln	1.13	1.18	1.07	1.11
Real commodity price	Δ ² ln	1.04	1.00	1.12	1.09

◀ ▶ ⏪ ⏩ ⏴ ⏵ ⏶ ⏷ ⏸ ⏹ ⏺ ⏻ ⏼ ⏽ ⏾ ⏿ 🔍 ↻

Large Bayesian VARs I

- VAR models with a large dimension: up to 131 variables.
- The prior distribution is a modification of the Minnesota prior: the degree of tightness increases with the dimension of the VAR.
- Forecasts of employment with large Bayesian VARs are more accurate than small VARs. Large BVAR forecasts are also more accurate than models that include 3 factors (factor models are another parsimonious way of dealing with a large data set).
- The VAR model is estimated to monthly time series from 1959 up to 2003. Data set includes traditional leading and coincident indicators and includes all series in Stock and Watson (2003).



Large Bayesian VARs II

- Recall the reduced form VAR:

$$y_t = c + A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t$$

where A_i are $m \times m$ coefficients matrices and c is a $m \times 1$ of intercepts.

- ε_t is a $m \times 1$ vector of disturbances that have the following properties with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Sigma_\varepsilon$.
- The prior mean on the autoregressive coefficients is:

$$E[(A_k)_{ij}] = \begin{cases} \delta_i, j = i, k = 1 \\ 0, \text{ otherwise} \end{cases}$$

where $k = 1, \dots, p$. $\delta_i = 1$ if variable is non-stationary and $\delta_i = 0$ if it is stationary (unit root testing?). Time series are considered in logs.



Large Bayesian VARs III

- The prior variance is:

$$V[(A_k)_{ij}] = \begin{cases} \frac{\lambda^2}{k^2}, j = i \\ \nu \frac{\lambda^2 \sigma_i^2}{k^2 \sigma_j^2}, \text{ otherwise} \end{cases}$$

where λ controls the overall tightness of the prior ($\lambda = 0$, posterior = prior; $\lambda = \infty$, OLS estimates); ν is how much lags of other variables are less important ($0 < \nu < 1$; assume $\nu = 1$ because of structural analysis).

- The tightness of the prior has to be chosen such that λ decreases with m (VAR dimension). The authors compute a λ in a pre-evaluation period (60-69) such that the fit of the model of dimension m fits as the VAR with only three variables (benchmark case with employment, cpi and interest rate).



Large Bayesian VARs IV

- Forecasts are computed by iteration as it is usual in the VAR literature (assume $p = 1$ for simplicity):

$$\begin{aligned} \hat{y}_{t+1|t}^{(\lambda, m)} &= \hat{c}^{(\lambda, m)} + \hat{A}_1^{(\lambda, m)} y_t \\ \hat{y}_{t+2|t}^{(\lambda, m)} &= \hat{c}^{(\lambda, m)} + \hat{A}_1^{(\lambda, m)} \hat{y}_{t+1|t}^{(\lambda, m)} \end{aligned}$$

- The out-of-sample period is 1970-2003, and forecasts are computed up to 12 months-ahead: $32 * 12 = 384$ observations.
- BVAR is re-estimated at each forecast origin with a window of 10 years of data (sample size is fixed).
- The MSFE is employed to measure forecast accuracy and the benchmark model is a random walk with drift. Forecasts of employment, CPI and fed fund rate are evaluated.



Large Bayesian VARs V

- "Results show that a standard Bayesian VAR model is an appropriate tool for large panels of data. Not only a Bayesian VAR estimated over 100 variables is feasible, but it produces better forecasting results than the typical seven variables VAR considered in the literature". BGR, 2010.

Large Bayesian VAR Forecasting Performance

Table I. BVAR, Relative MSFE, 1971–2003

		<i>SMALL</i>	<i>CEE</i>	<i>MEDIUM</i>	<i>LARGE</i>
$h = 1$	EMPL	1.14	0.67	0.54	0.46
	CPI	0.89	0.52	0.50	0.50
	FFR	1.86	0.89	0.78	0.75
$h = 3$	EMPL	0.95	0.65	0.51	0.38
	CPI	0.66	0.41	0.41	0.40
	FFR	1.77	1.07	0.95	0.94
$h = 6$	EMPL	1.11	0.78	0.66	0.50
	CPI	0.64	0.41	0.40	0.40
	FFR	2.08	1.30	1.30	1.29
$h = 12$	EMPL	1.02	1.21	0.86	0.78
	CPI	0.83	0.57	0.47	0.44
	FFR	2.59	1.71	1.48	1.93
λ		∞	0.262	0.108	0.035