Evaluating Forecasts

- Are our forecasts good?
- How do we know?
- How do we assess a historical forecast?
- How do we compare competing forecasts?

Properties of Forecasts

- What are the properties of a good forecast?
- We start by examining optimal forecasts.

Linear Representation

- The Wold representation for y, h steps out, is

$$y_{n+h} = \mu + e_{n+h} + b_1 e_{n+h-1} + b_2 e_{n+h-2} + \cdots$$

- The h-step-ahead optimal forecast is

$$y_{n+h|n} = \mu + b_h e_n + b_{h+1} e_{n-1} + b_{h+2} e_{n-2} + \cdots$$

- The h-step-ahead optimal forecast error is

$$e_{n+h|n} = e_{n+h} + b_1 e_{n+h-1} + b_2 e_{n+h-2} + \cdots + b_{h-1} e_{n+1}$$

Optimal Forecast is Unbiased

- The forecast error is

$$e_{n+h|n} = e_{n+h} + b_1 e_{n+h-1} + b_2 e_{n+h-2} + \cdots + b_{h-1} e_{n+1}$$

- It has expectation

$$E\left(e_{n+h|n}\right) = 0$$

- And thus the optimal forecast is unbiased

One-Step Errors are White Noise

- The one-step forecast error is

$$e_{n+|n} = e_{n+h}$$

- Which is unforecastable white noise
- Thus the optimal one-step-ahead forecast error is white noise and unforecastable

h-step-ahead errors are MA(h-1)

- The h-step forecast error is

$$e_{n+h|n} = e_{n+h} + b_1 e_{n+h-1} + b_2 e_{n+h-2} + \cdots + b_{h-1} e_{n+1}$$

- This is a MA(h-1)
- Thus optimal h-step-ahead forecast errors are correlated, but at most a MA(h-1)

Forecast Variance

- The h-step forecast error is

$$e_{n+h|n} = e_{n+h} + b_1 e_{n+h-1} + b_2 e_{n+h-2} + \cdots + b_{h-1} e_{n+1}$$

- Its variance is the forecast variance, and is

$$\operatorname{var}\left(e_{n+h|n}\right) = \left(1 + b_1^2 + b_2^2 + \cdots + b_{h-1}^2\right)\sigma^2$$

- This is increasing in the forecast horizon *h*
- The variance of optimal forecasts increases with the forecast horizon

Unforecastable Errors

- The forecast errors should be unforecastable from all information available at the time of the forecast
- Not even the optimal forecast
- The coefficients should be zero in the regression

$$e_{n+h|n} = \alpha + \beta y_{n+h|n} + \varepsilon_{n+h}$$

$$\alpha = 0, \beta = 0$$

Formal Comparison

- Since

$$e_{n+h|n} = y_{n+h} - y_{n+h|n}$$

this implies

$$y_{n+h} = \alpha + \beta y_{n+h|n} + e_{n+h|h}$$
$$\alpha = 0, \beta = 1$$

- The regression of the actual value on the ex-ante forecast should have a zero intercept and a coefficient of 1

Mincer-Zarnowitz Regression

- This is called a “Mincer-Zarnowitz” regression, proposed in a paper
  - “The evaluation of economic forecasts”
- Jacob Mincer (1922-2006)
  - Father of modern labor economics
- Victor Zarnowitz (1919-2009)
  - Leading figure in business cycle dating

Mincer-Zarnowitz Test

- Estimate the simple regression

$$y_{n+h} = \alpha + \beta y_{n+h|n} + e_{n+h|h}$$

- Test the joint hypothesis

$$\alpha = 0, \beta = 1$$

- If the coefficients are different, it indicates systematic bias in the historical forecasts

Summary: Properties of Optimal Forecasts

- Unbiased
- 1-step-ahead errors are white noise
- h-step-ahead errors are at most MA(h-1)
- Variance of h-step-ahead error is increasing in h
- Forecast errors should be unforecastable

Forecasting Average Growth

- When we are forecasting future growth, we may be interested in total future growth out to h periods
- For example, the growth rate of GDP during 2010
- This is the average of the growth rates during the four quarters 2010Q1, …, 2010Q4

Average Growth

- If y_t is the growth rate in period t, then the average future h-step growth is

$$y_{n+1:n+h} = \frac{y_{n+1} + \cdots + y_{n+h}}{h}$$

- The forecast of the average growth is

$$y_{n+1:n+h|n} = \frac{y_{n+1|n} + \cdots + y_{n+h|n}}{h}$$

- What are its properties?

Average Forecast Error

- The error of the average forecast is

$$
\begin{aligned}
e_{n+1:n+h|n} &= y_{n+1:n+h|n} - y_{n+1:n+h} \\
&= \frac{y_{n+1|n} + \cdots + y_{n+h|n}}{h} - \frac{y_{n+1} + \cdots + y_{n+h}}{h} \\
&= \frac{\left(y_{n+1|n} - y_{n+1}\right) + \cdots + \left(y_{n+h|n} - y_{n+h}\right)}{h} \\
&= \frac{e_{n+1|n} + \cdots + e_{n+h|n}}{h}
\end{aligned}
$$

- Which is the average of the 1-step through h-step errors

Average Forecast Error Variance

- Since the average forecast error is the **average** of forecast errors, it has a smaller variance than the h-step variance

$$\text{var}\left(e_{n+1:n+h|n}\right) = \text{var}\left(\frac{e_{n+1|n} + \cdots + e_{n+h|n}}{h}\right) \leq \text{var}\left(e_{n+h|n}\right)$$

- So multi-period growth rate forecasts will have smaller variance than h-step ahead growth forecasts
  - The forecasted average growth rate for 2010 has a smaller variance than the forecasted growth rate for 2010Q4

Evaluating Forecasts

- Suppose we have a sequence of real forecasts
- Perhaps they are our own forecasts
- How can we evaluate the forecasts?

Measures of Forecast Performance

- Form the historical sequence of forecasts and actual values.
- Construct the forecast error as the difference

Example

- CBO's Economic Forecasting Record: 2009 Update
  - Link on reading list
- Economic forecasts made by
  - Congressional budget office (CBO)
  - U.S. Administration
  - Private forecasters
    - Blue Chip average
  - CBO regularly assesses their forecasts

CBO Comparison

- Real Output
- Nominal Output
- Inflation
- 3-month T-Bill rate
- 10-year Treasury note rate
- Difference between CPI and GDP inflation
- Both 2-year and 5-year forecasts

Table 4.

Comparison of CBO's, *Blue Chip's*, and the Administration's Forecasts of Two-Year Average Growth Rates for Nominal Output

(Percent, by calendar year)

		CBO		*Blue Chip*[b]		Administration	
	Actual	Forecast	Error[a]	Forecast	Error[a]	Forecast	Error[a]
GNP							
1976–1977	11.5	13.1	1.7	*	*	12.3	0.8
1977–1978	12.1	10.8	-1.3	*	*	11.2	-1.0
1978–1979	12.5	10.9	-1.6	*	*	11.2	-1.3
1979–1980	10.4	11.0	0.5	*	*	10.4	-0.1
1980–1981	10.4	9.7	-0.7	*	*	9.5	-0.8
1981–1982	8.0	12.1	4.1	*	*	11.9	4.0
1982–1983	6.3	9.7	3.4	9.5	3.2	9.8	3.5
1983–1984	9.8	8.2	-1.6	9.0	-0.9	8.0	-1.8
1984–1985	9.0	9.9	0.9	9.6	0.6	9.6	0.6
1985–1986	6.2	7.6	1.3	7.4	1.2	8.2	1.9
1986–1987	5.8	7.1	1.3	6.7	0.9	7.7	1.8
1987–1988	7.0	6.5	-0.5	6.4	-0.5	6.9	-0.1
1988–1989	7.6	6.3	-1.3	6.1	-1.5	6.8	-0.9
1989–1990	6.7	6.8	0.1	6.6	-0.1	7.1	0.4
1990–1991	4.6	6.1	1.5	6.0	1.4	7.1	2.5
1991–1992	4.4	5.7	1.3	5.2	0.8	5.6	1.2

Comparison

- By showing the actual, forecasts, and forecast errors side-by-side, we can informally see which forecast performs better

Formal Comparison

- The forecasts can be compared by estimating the **bias** and **risk** (expected loss) of the forecasts
- They are estimated from R forecast errors:
  - Bias, Mean Absolute Error, Root Mean Squared Error

$$Bias = \frac{1}{R}\sum_{n=1}^{R} e_{n+h|n}$$

$$MAE = \frac{1}{R}\sum_{n=1}^{R} \left| e_{n+h|n} \right|$$

$$RMSE = \left(\frac{1}{R}\sum_{n=1}^{R} e_{n+h|n}^{2} \right)^{1/2}$$

CBO Comparison

Table 1.

Summary Measures of Performance for Two-Year Average Forecasts

(Percentage points)

	CBO	*Blue Chip*[a]	Administratio
Growth Rate for Real Output (1982-2007)			
Mean error	-0.3	-0.3	-0.1
Mean absolute error	1.0	1.0	1.0
Root-mean-square error	1.2	1.2	1.3
Growth Rate for Nominal Output (1982-2007)			
Mean error	0.1	0.1	0.2
Mean absolute error	1.0	0.9	1.1
Root-mean-square error	1.2	1.1	1.4
Inflation in the Consumer Price Index (1982-2007)			
Mean error	0.2	0.2	0.1
Mean absolute error	0.7	0.7	0.8
Root-mean-square error	0.9	0.9	0.9

Table 2.

Summary Measures of Performance for Five-Year Average Projections

(Percentage points)

	CBO	*Blue Chip*[a]	Administratic
Growth Rate for Real Output (1979-2004)			
Mean error	-0.2	-0.2	0.1
Mean absolute error	0.6	0.6	0.8
Root-mean-square error	0.9	0.8	0.9
Growth Rate for Nominal Output (1982-2004)			
Mean error	0.3	0.5	0.4
Mean absolute error	0.7	0.7	0.8
Root-mean-square error	0.9	0.9	1.0
Difference Between Inflation in the CPI and the GDP Price Index (1983-2004)			
Mean error	-0.2	-0.3	-0.3
Mean absolute error	0.4	0.4	0.5
Root-mean-square error	0.4	0.5	0.5

Data Revision

- A major difficulty with forecast evaluation is that for many series, there are serious data revisions
- The data used for forecasting, and the series published today, are different
- The series forecasted, and the series reported today, are different
- Price series, and real series based on price levels, are rebased every few years
- These rebasing are not scale transformations, because the construction of real output is done at a disaggregate level, and then aggregated.

Real Output

	Actual				CBO		*Blue Chip* [e]		Administration	
	1972 Dollars[a]	1982 Dollars[b]	1987 Dollars[c]	Chain-Type Annual-Weighted Index	Forecast	Error[d]	Forecast	Error[d]	Forecast	Error[d]
Real GNP										
1976–1977	6.7	4.8	4.8	5.1	6.2	1.1	*	*	5.9	0.9
1977–1978	5.2	5.0	4.7	5.1	5.5	0.4	*	*	5.1	0.1
1978–1979	3.9	3.9	3.8	4.5	4.7	0.3	*	*	4.7	0.3
1979–1980	1.3	1.1	1.1	1.6	2.7	1.2	*	*	2.9	1.3
1980–1981	1.1	0.9	0.5	1.0	0.5	-0.5	*	*	0.5	-0.5
1981–1982	0.2	-0.3	-0.4	0.2	2.1	1.9	*	*	2.6	2.4
1982–1983	0.7	0.5	0.7	1.2	2.1	0.9	2.0	0.8	2.7	1.4
1983–1984	5.2	5.2	4.9	5.7	3.4	-2.3	3.5	-2.2	2.6	-3.1
1984–1985	*	5.1	4.4	5.4	4.7	-0.7	4.3	-1.1	4.7	-0.7
1985–1986	*	3.0	2.8	3.5	3.3	-0.2	3.2	-0.3	3.9	0.4
1986–1987	*	3.1	2.9	3.3	3.1	-0.1	3.0	-0.3	3.7	0.4
1987–1988	*	3.9	3.5	3.8	2.9	-0.9	2.8	-0.9	3.3	-0.5
1988–1989	*	3.5	3.3	3.9	2.4	-1.4	2.1	-1.7	3.0	-0.9
1989–1990	*	1.7	2.0	2.8	2.5	-0.3	2.2	-0.6	3.2	0.4
1990–1991	*	*	0.3	0.9	2.0	1.2	1.9	1.1	2.8	1.9
1991–1992	*	*	0.7	1.5	1.6	0.1	1.2	-0.3	1.4	-0.1

Meese-Rogoff Puzzle

- The most influential paper using the method of forecast model comparison is
  - "Empirical exchange rate models of the seventies"
  - Richard Meese and Kenneth Rogoff
  - *Journal of International Economics*, 1983

Meese-Rogoff

- Ken Rogoff (currently Harvard)
  - Recent book
  - *This Time is Different: Eight Centuries of Financial Folly*
- Dick Meese (formerly Berkeley, now Barclay Global Investors)
  - 1978 UW Ph.D.
  - Economics Dept Advisory Board

UW News Item

- **Economics grad gives $1 million to endow professorship**
- Richard Meese, an economist and expert on foreign currency exchange rates, has given $1 million to the University of Wisconsin-Madison to establish a professorship in applied econometrics.
- "The university laid the foundation for my work success," says Meese, who manages research and strategy on global currency products for Barclays Global Investors and a former economics professor at the University of California, Berkeley.
- Richard A. Meese Chair in Applied Econometrics
  - Held by Chris Taber

Meese-Rogoff paper

- They compare the RMSE and bias of 1-month, 6-month and 12-month forecasts of a set of exchange rates, using structural models
- They compare the performance of the economic models with the performance of a random walk
- They found the random walk beat the economic models
- Very influential paper

Root mean square forecast errors.[a]

Exchange rate	Model: / Horizon	Random walk	Forward rate	Univariate autoregression	Vector autoregression	Frenkel-Bilson[b]	Dornbusch-Frankel[b]	Hooper-Morton[b]
$/mark	1 month	3.72	3.20	3.51	5.40	3.17	3.65	3.50
	6 months	8.71	9.03	12.40	11.83	9.64	12.03	9.95
	12 months	12.98	12.60	22.53	15.06	16.12	18.87	15.69
$/yen	1 month	3.68	3.72	4.46	7.76	4.11	4.40	4.20
	6 months	11.58	11.93	22.04	18.90	13.38	13.94	11.94
	12 months	18.31	18.95	52.18	22.98	18.55	20.41	19.20
$/pound	1 month	2.56	2.67	2.79	5.56	2.82	2.90	3.03
	6 months	6.45	7.23	7.27	12.97	8.90	8.88	9.08
	12 months	9.96	11.62	13.35	21.28	14.62	13.66	14.57
Trade-weighted dollar	1 month	1.99	N.A.	2.72	4.10	2.40	2.50	2.74
	6 months	6.09	N.A.	6.82	8.91	7.07	6.49	7.11
	12 months	8.65	14.24	11.14	10.96	11.40	9.80	10.35

Summary

- Evaluation of forecasts achieve by comparing the bias, MAE and RMSE of forecast errors
- Most influential paper is Meese-Rogoff, because they showed that naïve random walk model has lower forecast risk than structural economic models
- This established a challenge for economic modeling and forecasting.
  - Can we beat simple naïve models?!