What Do We Really Know about Economic Forecasting?

David F. Hendry

Institute for New Economic Thinking at the
Oxford Martin School, University of Oxford

CIRET: Vienna, September 2012

> "The British, he thought, must be gluttons for satire: even the weather forecast seemed to be some kind of spoof, predicting every possible combination of weather for the next twenty-four hours without actually committing itself to anything specific."

David Lodge: *Changing Places*, 1975

When weather forecasters go awry,
they get a **new super-computer**;
when economists mis-forecast,
we get our budgets cut.

DFH, originally quoted in the UK Press following the 1987 storm.

(A) **Introduction**

(B) What causes forecasting problems?

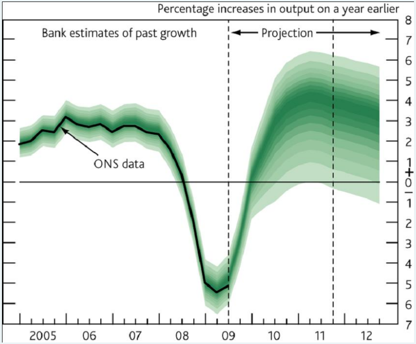
(C) Explaining forecasting

(D) Empirical application of forecasting theory

(E) Forecasting breaks

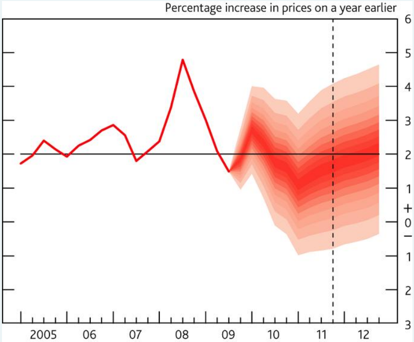
Conclusions

GDP growth projection



Source: Bank of England

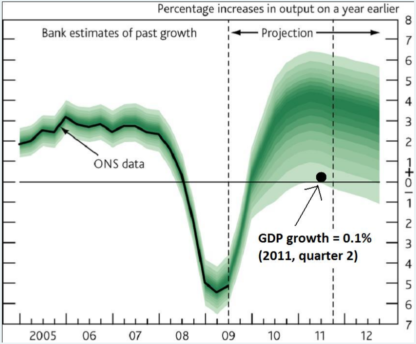
CPI inflation projection



Source: Bank of England

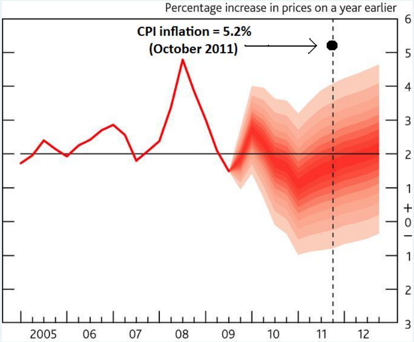
Prerequisite for a forward-looking macroeconomic policy.

GDP growth projection



Source: Bank of England

CPI inflation projection



Source: Bank of England

Prerequisite for a forward-looking macroeconomic policy.

A view of the future is fundamental to economic planning:

unavoidable – but hazardous!

Ancient Egyptians foretold harvests from the level reached by the Nile in the flood season.

The Oracles of Delphi and Nostradamus are early examples of often ambiguous forecasters.

C17th: Sir William Petty discerned a seven-year business cycle, suggesting a basis for systematic economic forecasts.

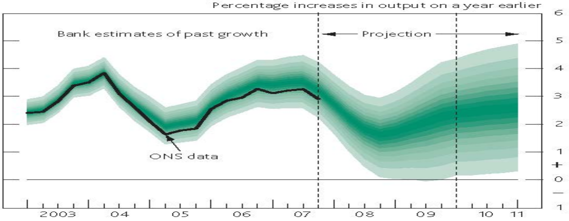
A forecasting industry developed in the USA around 1910–1930, but much of it was wiped out by the Great Depression—
which it failed to foresee!

Did we do any better in 2008?

Bank of England forecasts

Chart shows Bank of England quarterly forecasts for annual changes in UK GDP at February 2008 through end 2010:

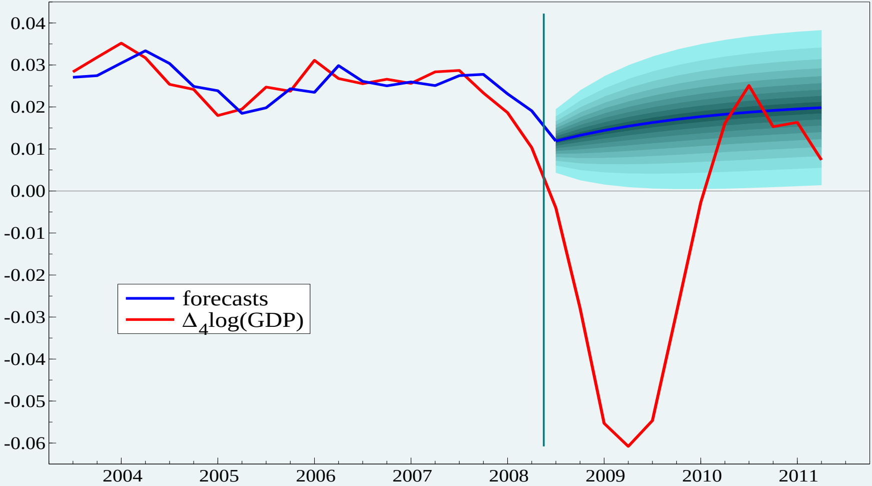
Chart 1 GDP projection based on market interest rate expectations



The fan chart depicts the probability of various outcomes for GDP growth. To the left of the first vertical dashed line, the distribution reflects the likelihood of revisions to the data over the past; to the right, it reflects uncertainty over the evolution of GDP growth in the future. If economic circumstances identical to today's were to prevail on 100 occasions, the MPC's best collective judgement is that the mature estimate of GDP would lie within the darkest central band on only 10 of those occasions. The fan chart is constructed so that outturns are also expected to lie within each pair of the lighter green areas on ten occasions. Consequently, GDP growth is expected to lie somewhere within the entire fan on 90 out of 100 occasions. The bands widen as the time horizon is extended, indicating the increasing uncertainty about outcomes. See the box on page 39 of the November 2007 *Inflation Report* for a fuller description of the fan chart and what it represents. The second dashed line is drawn at the two-year point of the projection.

Distinct slowdown envisaged–

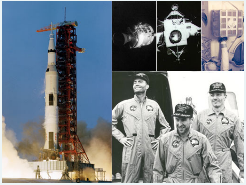
but nothing like the unanticipated 6% fall that materialized.



Massive forecast failure: forecasts up, while data down, then excellent (!), as predicted by our theory of forecasting.

Almost no economic theories allow for unanticipated location shifts: yet empirically occur intermittently.

Analogy: rocket to moon due to land on 4th July, but hit by meteor and knocked off course.



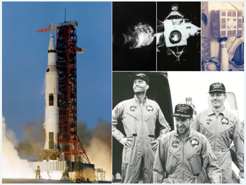
Forecast is badly wrong.

Outcome not due to poor forecasting models; and does not refute Newtonian gravitation theory.

Example of **location shift**: change in previous mean.

Almost no economic theories allow for unanticipated location shifts: yet empirically occur intermittently.

Analogy: rocket to moon due to land on 4th July, but hit by meteor and knocked off course.



Forecast is badly wrong.

Outcome not due to poor forecasting models; and does not refute Newtonian gravitation theory.

Example of **location shift**: change in previous mean.

Economic forecasting confronts non-stationary world

Explain main causes of forecast failure;
methods to insure against systematic forecast failure;
progress towards forecasting during breaks.

"Oddly, the industry that is the primary engine of this incredible pace of change – the computer industry – turns out to be rather bad at predicting the future itself. There are two things in particular that it failed to see:

"Oddly, the industry that is the primary engine of this incredible pace of change – the computer industry – turns out to be rather bad at predicting the future itself. There are two things in particular that it failed to see:

***one was the coming of the Internet**, ...;*

"Oddly, the industry that is the primary engine of this incredible pace of change – the computer industry – turns out to be rather bad at predicting the future itself. There are two things in particular that it failed to see:

***one was the coming of the Internet**, ...;*

the other was the end of the century."

Douglas Adams,** The Salmon of Doubt, **2002.

(A) **Introduction**

(B) **What causes forecasting problems?**

(C) Explaining forecasting

(D) Empirical application of forecasting theory

(E) Forecasting breaks

Conclusions

Well developed if
econometric model coincides with stationary data-generating process (DGP).

Consider $n \times 1$ vector $\mathbf{x}_t \sim \mathsf{D}_{\mathbf{x}_t}(\mathbf{x}_t|\mathbf{X}_{t-1}, \boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^k$,
where $\mathbf{X}_{t-1} = (\ldots \mathbf{x}_1 \ldots \mathbf{x}_{t-1})$.

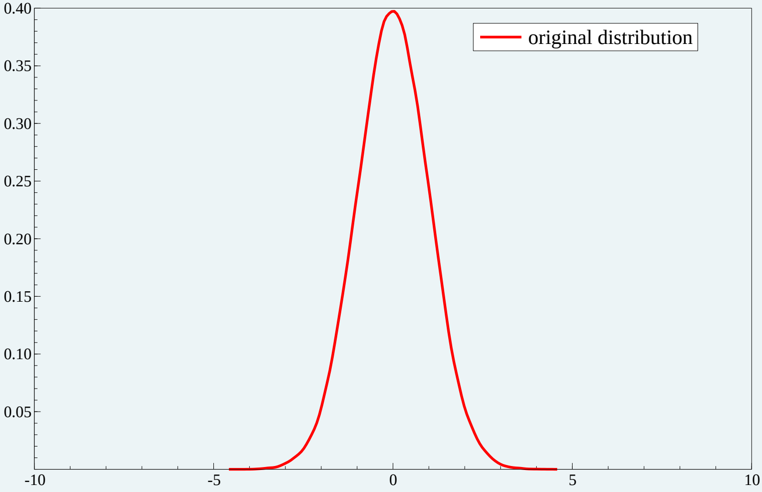
Statistical forecast $\widetilde{\mathbf{x}}_{T+h|T} = \mathbf{f}_h(\mathbf{X}_T)$ for $T+h$ at T.
How to select $\mathbf{f}_h$?

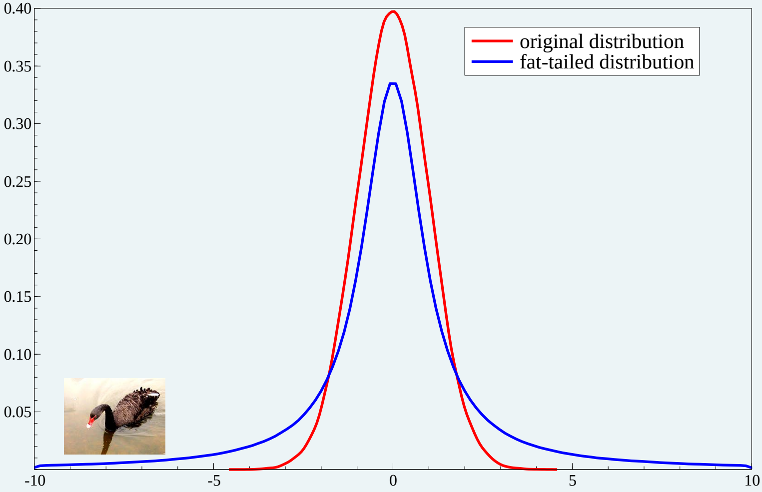
Traditional answer–conditional expectation:
$\widehat{\mathbf{x}}_{T+h|T} = \mathsf{E}[\mathbf{x}_{T+h} \mid \mathbf{X}_T]$ unbiased,
$\mathsf{E}[(\mathbf{x}_{T+h} - \widehat{\mathbf{x}}_{T+h|T}) \mid \mathbf{X}_T] = 0$.

$\widehat{\mathbf{x}}_{T+h|T}$ has smallest mean-square forecast-error matrix.

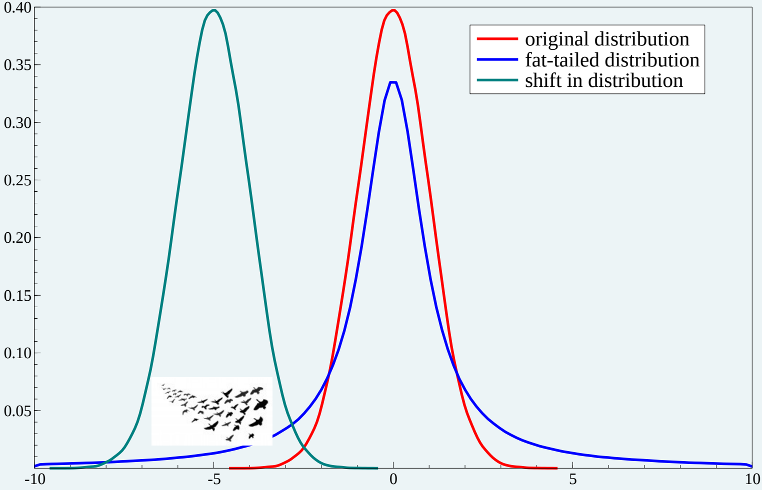
Caveat emptor as we will see!

Reality includes unanticipated shifts




0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
-10
-5
0
5
10
original distribution
fat-tailed distribution

Reality includes unanticipated shifts



First: problems learning $\mathsf{D}_{\mathbf{X}_T^1}(\cdot)$ and $\boldsymbol{\theta}$:

(i) *specification* of the set of relevant variables $\{\mathbf{x}_t\}$,

(ii) *measurement* of the $\mathbf{x}$s,

(iii) *formulation* of $\mathsf{D}_{\mathbf{X}_T^1}(\cdot)$,

(iv) *modeling* of the relationships,

(v) *estimation* of $\boldsymbol{\theta}$, and

(vi) *properties* of $\mathsf{D}_{\mathbf{X}_T^1}(\cdot)$ determine 'intrinsic' uncertainty,

all of which introduce in-sample uncertainties.

Next, over the forecast horizon:

(vii) *properties* of $\mathsf{D}_{\mathbf{X}_{T+H}^{T+1}}(\cdot)$ determine forecast uncertainty,

(viii) *which grows* as H increases,

(ix) especially for *integrated* data,

(x) increased by *changes* in $\mathsf{D}_{\mathbf{X}_{T+H}^{T+1}}(\cdot)$ or $\boldsymbol{\theta}$.

These 10 issues structure analysis of forecasting.

Stationary scalar AR(1) DGP with known exogenous $\{z_t\} \sim \mathsf{IN}[0,1]$:

$$x_t = \rho x_{t-1} + \gamma z_t + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma^2_\epsilon\right] \text{ and } |\rho| < 1. \quad (1)$$

When ρ, γ known & constant, forecast from x_T is:

$$\widehat{x}_{T+1|T} = \rho x_T + \gamma z_{T+1} \quad (2)$$

$D_{\mathbf{X}^1_T}(\cdot)$ implies $D_{\mathbf{X}^{T+1}_{T+1}}(\cdot)$, producing unbiased forecast:

$$\mathsf{E}\left[\left(x_{T+1} - \widehat{x}_{T+1|T}\right) | x_T, z_{T+1}\right] = 0$$

with smallest possible variance determined by $D_{\mathbf{X}^1_T}(\cdot)$:

$$\mathsf{V}\left[\left(x_{T+1} - \widehat{x}_{T+1|T}\right)\right] = \sigma^2_\epsilon.$$

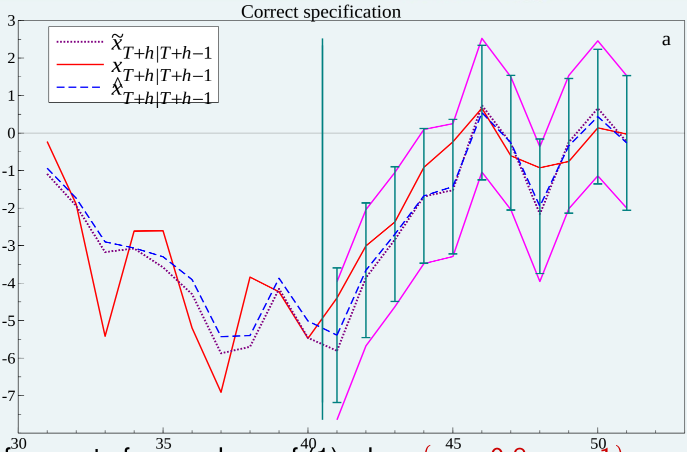
Thus: $D_{\mathbf{X}^{T+1}_{T+1}}(\cdot) = \mathsf{IN}\left[\rho x_T + \gamma z_{T+1}, \sigma^2_\epsilon\right]$.

Issues (i)–(x) ‘assumed away’.

(i) Specification incomplete if (e.g.) vector $\mathbf{x}_t$ not scalar.
(ii) Measurement incorrect if (e.g.) observe $\overline{\mathbf{x}}_t$ not $\mathbf{x}_t$.
(iii) Formulation inadequate if (e.g.) intercept needed.
(iv) Modeling wrong if (e.g.) selected ρx_{t-2}.
(v) Estimating ρ adds bias, $(\rho - \mathsf{E}[\widehat{\rho}])x_T$, and variance $\mathsf{V}[\widehat{\rho}]x_T^2$.
(vi) Properties of $D(\epsilon_t) = \mathsf{IN}\left[0, \sigma_\epsilon^2\right]$ determine $\mathsf{V}[x_t]$.
(vii) Assumed $\epsilon_{T+1} \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right]$, **but** $\mathsf{V}[\epsilon_{T+1}]$ **could differ**.
(viii) Multi-step forecast error $\sum_{h=1}^{H} \rho^{h-1}\epsilon_{T+h}$: $\mathsf{V} = \frac{1-\rho^{2H}}{1-\rho^2}\sigma_\epsilon^2$.
(ix) If $\rho = 1$ have trending forecast variance $H\sigma_\epsilon^2$.
(x) If ρ changes could experience forecast failure.

Must be prepared for risks from (i)–(x).

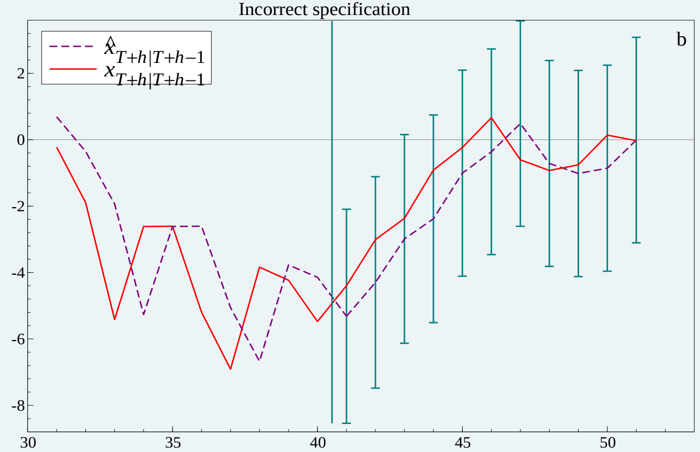
First ‘undo’ **(v)**, estimating (ρ, γ) from sample $t = 1, \ldots, T$
Then **(iv)** by omitting z_t, then **(x)** by changing ρ plus **(iii)**.



Panel a: forecasts from a draw of (1) when $(\rho = 0.8, \gamma = 1)$ are known and constant; ($\hat{x}_{T+h|T+h-1}$ from (2) with error bars of $\pm 2\hat{\sigma}$) and when estimating (ρ, γ) ($\tilde{x}_{T+h|T+h-1}$ with bands).

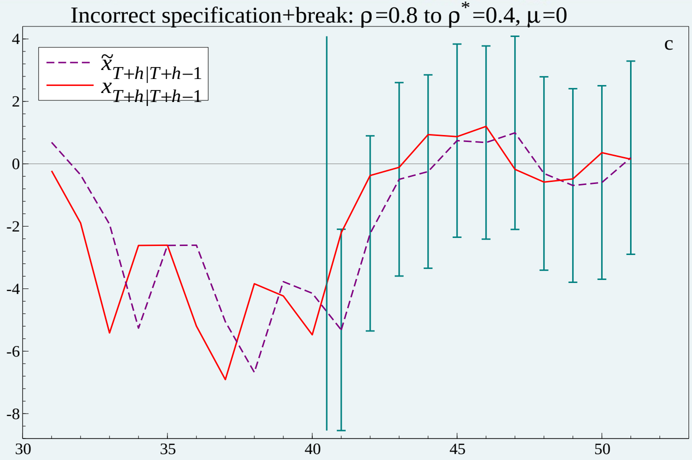
Forecasts almost identical,with small increase in uncertainty.

So not problem (v), estimation uncertainty



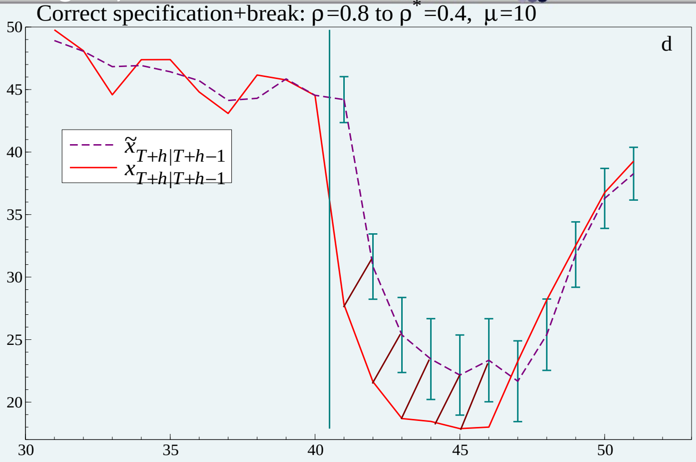
Panel b: forecasts when z_t omitted both in estimation and forecasting: forecasts poorer, but well within *ex ante* forecast intervals.

So not problem (iv), incorrect specification, even with (v)



Panel c: shift in ρ at $T = 41$ to $\rho = 0.4$, then back to $\rho = 0.8$ at $T = 46$, and omitted z_t **so all of (iv), (v) and (x) violated,** yet little noticeable impact from *halving* then *doubling* ρ.

So not problem (x), changed parameter, even with (iv) and (v)



Panel d: same shift in ρ but:

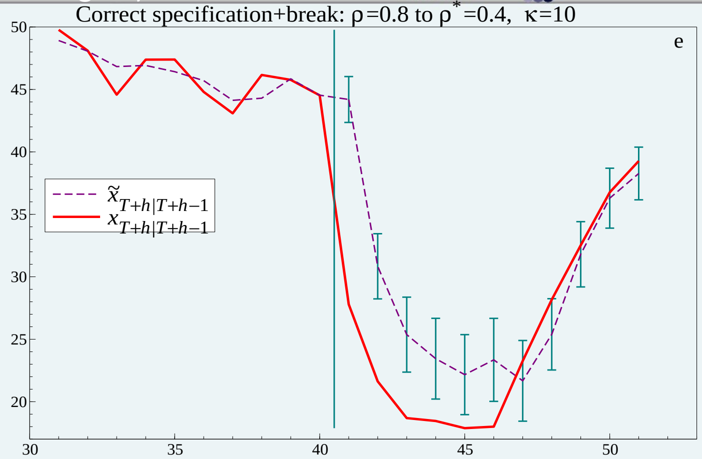
$$x_t = \mu + \rho x_{t-1} + \gamma z_t + \epsilon_t \quad \text{where} \quad \mu = 10 \qquad (3)$$

Catastrophic impact from halving ρ, yet not from *doubling* again.

First 5 forecasts upward relative to previous outcome.

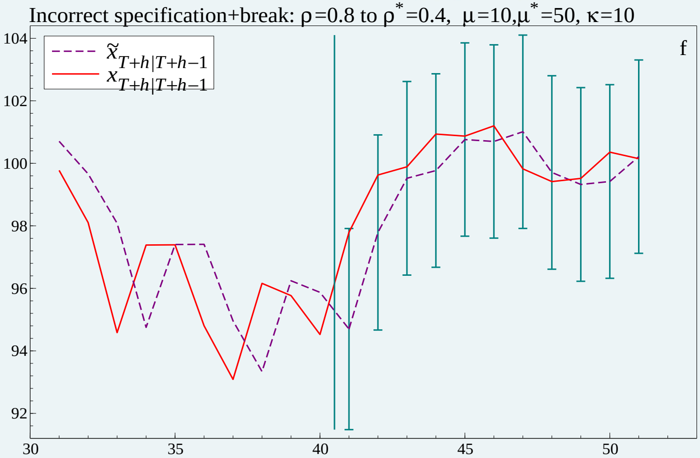
So is problem intercept–or mis-specification?

Correct specification, $\mu = 0$ and changed ρ



Panel e: **model correctly specified in-sample**,
forecasts for same break, $\mu = 0$, but $\mathsf{E}[z_t] = \kappa = 10$.
forecast failure is manifest–and essentially identical to *d*.
In-sample correct specification need not help
even with a zero intercept and known future z_{T+h}.

Incorrect specification, $\mu = 10$ & ρ changed twice



Panel f: **model incorrectly specified**, forecasts after same breaks in ρ to ρ^*, & **both** $\mu = 10,\ \kappa = 10$ with $\mu^* = 50$ at $T = 41$ then back to $\mu = 10$ at $T = 46$ so $x_{T+h} = \mu^* + \rho^* x_{T+h-1} + \gamma z_{T+h} + \epsilon_{T+h}$
yet no forecast failure when $\widehat{x}_{T+h|T+h-1} = \widehat{\mu} + \widehat{\rho} x_{T+h-1}$.

(A) **Introduction**

(B) **What causes forecasting problems?**

(C) **Explaining forecasting**

(D) Empirical application of forecasting theory

(E) Forecasting breaks

Conclusions

DGP is:

$$x_t = \theta + \rho\left(x_{t-1} - \theta\right) + \gamma\left(z_t - \kappa\right) + \epsilon_t \qquad (4)$$

$\epsilon_t \sim \mathsf{IN}[0, \sigma_\epsilon^2]$, $\mathsf{E}[x_t] = \theta$ and $\mathsf{E}[z_t] = \kappa$ with $\gamma \neq 0$.
Mis-specified forecasting model:

$$\widehat{x}_{T+1|T} = \widehat{\theta} + \widehat{\rho}\left(\widehat{x}_T - \widehat{\theta}\right) \qquad (5)$$

estimated over $t = 1, \ldots, T$, with parameter estimates $(\widehat{\theta}, \widehat{\rho})$ where $\mathsf{E}[\widehat{\theta}] = \theta_e$ and $\mathsf{E}[\widehat{\rho}] = \rho_e$, from estimated $\widehat{x}_T$ at forecast origin.

Forecast error is $\widehat{\epsilon}_{T+1|T} = x_{T+1} - \widehat{x}_{T+1|T}$.

Break occurs at T, with post-break DGP, $t = T+1, \ldots$:

$$x_t = \theta^* + \rho^*\left(x_{t-1} - \theta^*\right) + \gamma^*\left(z_t - \kappa^*\right) + \epsilon_t \qquad (6)$$

UNIVERSITY OF OXFORD | OXFORD MARTIN SCHOOL | Institute for New Economic Thinking

Problems due to effects on $\mathsf{E}[x_t] = (\mu + \gamma\kappa)/(1-\rho) = \theta$:

$$\Delta x_t = (\rho - 1)(x_{t-1} - \theta) + \gamma(z_t - \kappa) + \epsilon_t \qquad (7)$$

No forecast failure if $\mathsf{E}[x_t] = \theta$ before and after shift in ρ.

Forecast failure if $\mathsf{E}[x_{T+h}] = \theta^* \neq \theta$ **changes**.

$\mathsf{E}[x_{T+h}]$ shifts from $\theta = 50$ to $\theta^* = 17$ in both cases *d* and *e*
but $\theta = \theta^*$ **in** *f*.

All models in this class are **equilibrium correction**:
fail systematically if $\mathsf{E}[\cdot]$ changes to θ^*, as forecasts converge back to θ, **irrespective of new parameter values**.

Huge class of equilibrium-correction models (EqCMS):
regressions; dynamic systems; VARs; DSGEs;
ARCH; GARCH; some other volatility models.

Pervasive and pernicious problem affecting all members.

All main sources of forecast errors occur using (5) when (6) is DGP:

$$\widehat{\epsilon}_{T+1|T} = \theta^* - \widehat{\theta} + \rho^* \left(x_T - \theta^*\right) - \widehat{\rho}\left(\widehat{x}_T - \widehat{\theta}\right) + \gamma^* \left(z_{T+1} - \kappa^*\right) + \epsilon_{T+1} \quad (8)$$

(*ia*) **deterministic shifts**: $(\theta,\ \kappa)$ to $(\theta^*,\ \kappa^*)$;

(*ib*) **stochastic breaks**: $(\rho,\ \gamma)$ to $(\rho^*,\ \gamma^*)$;

(*iia,b*) **inconsistent parameter estimates**: $\theta_e \neq \theta,\quad \rho_e \neq \rho$;

(*iii*) **forecast origin uncertainty**: $\widehat{x}_T$;

(*iva,b*) **estimation uncertainty**: $\mathsf{V}[\widehat{\rho},\ \widehat{\theta}]$;

(*v*) **omitted variables**: z_{T+1};

(*vi*) **innovation errors**: ϵ_{T+1}.

Addressing mistakes in reverse order:

(*vi*): **innovation error** $\mathsf{E}[\epsilon_{T+1}] = 0$ and $\mathsf{V}[\epsilon_{T+1}] = \sigma^2_\epsilon$ so no bias, and $O_p(1)$ variance (irreducible if $\{\epsilon_t\}$ an innovation).

(*v*): **omitted variable** $\mathsf{E}[z_{T+1} - \kappa^*] = 0$ and $\mathsf{V}[(z_{T+1} - \kappa^*)] = \sigma^2_z$, so no bias despite omission and change in parameter values, and $O_p(1)$ variance, reducible by including $\{z_t\}$, with estimation variance of $O_p(T^{-1})$.

(*ivb*): **slope estimation** $\mathsf{E}[\widehat{\rho} - \rho_e] = 0$, plus estimation variance of $O_p(T^{-1})$.

(*iva*): **equilibrium-mean estimation** $\mathsf{E}[\widehat{\theta} - \theta_e] = 0$ with an estimation variance of $O_p(T^{-1})$.

(*iii*): **forecast-origin uncertainty** $\mathsf{E}[\widehat{x}_T - x_T] = 0$ only if forecast origin unbiasedly estimated, with variance $O_p(1)$.

(*iib*) **slope mis-specification** $\mathsf{E}[(\rho - \rho_e)(x_T - \theta)] = 0$ as $\mathsf{E}[x_T - \theta] = 0$, and an $O_p(1)$ variance unconditionally.

(*iia*) **equilibrium-mean mis-specification**: $\theta \neq \theta_e$ possible if in-sample location shifts not modelled.

(*ib*) **slope change** $\mathsf{E}[(\rho^* - \rho)\,(x_T - \theta)] = 0$ irrespective of $\rho^* \neq \rho$.

(*ia*) **equilibrium-mean change**–fundamental problem: $\theta^* \neq \theta$ induces forecast failure.

Once in-sample breaks removed, from good forecast origin estimates, irrespective of model mis-specification, still have:

$$\mathsf{E}[\widehat{\epsilon}_{T+1|T}] \simeq (1-\rho^{*})(\theta^{*}-\theta) \quad (9)$$

and that bias persists at $\widehat{\epsilon}_{T+2|T+1}$ etc., so long as (5) is used, even though no further breaks ensue.

Power of insight exemplified by:

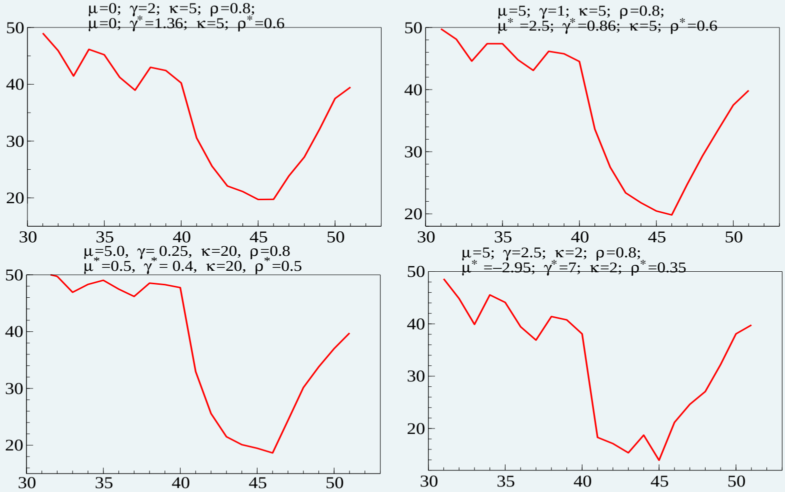
(a) change *both* μ and ρ by large magnitudes with $\theta=\theta^{*}$: outcome is isomorphic to $\mu=\mu^{*}=0$, so no break is detected;

(b) even if $\mu=\mu^{*}=0$ & z_t correctly included but needs forecast, get **8 more problems**: $\rho\neq\rho^{*}$ still induces forecast failure by shifting θ;

(c) benefit from including z_t primarily if κ alone shifts to κ^{*} which induces a shift in θ to θ^{*} that is captured; and

(d) if z_{T+1} has to be forecast, must be closer to κ^{*} than κ for a smaller forecast-error bias than omission.

Result applies to all EqCMs: they fail systematically when $\mathsf{E}[x]$ changes as models' forecasts converge to θ irrespective of value of θ^{*}.



Can essentially replicate break by changing μ, γ and ρ in many combinations: **economic agents could not tell what had shifted till long afterwards.**

“Here is Edward Bear, coming downstairs now,
bump, bump, bump, on the back of his head,
behind Christopher Robin. It is, as far as he knows, the only way of coming downstairs, but sometime he feels there really is another way, if only he could stop bumping for a moment and think of it.”

A.A. Milne, Winnie-the-Pooh, 1926.

*"Here is Edward Bear, coming downstairs now,
bump, bump, bump, on the back of his head,
behind Christopher Robin. It is, as far as he knows, the
only way of coming downstairs, but sometime he feels
there really is another way, if only he could stop
bumping for a moment and think of it."*
A.A. Milne, Winnie-the-Pooh, 1926.

That is how forecasters must often feel!

(*iia*) showed need to remove in-sample location shifts
or have systematic mis-forecasting.
Numbers, timings and magnitudes of breaks in models usually unknown: 'portmanteau' approach required to detect location shifts anywhere in sample, **while also selecting over other variables**.

To check the null of no outliers or location shifts in a model, **impulse-indicator saturation** (IIS) creates complete set of indicator variables: $\{1_{\{j=t\}}\} = 1$ when $j = t$ and 0 otherwise for $j = 1, \ldots, T$ observations, then adds T impulse indicators during model selection.

Many well-known procedures are variants of IIS.
Chow (1960) test is sub-sample IIS over $T - k + 1$ to T.

Salkever (1976) tests parameter constancy by indicators.

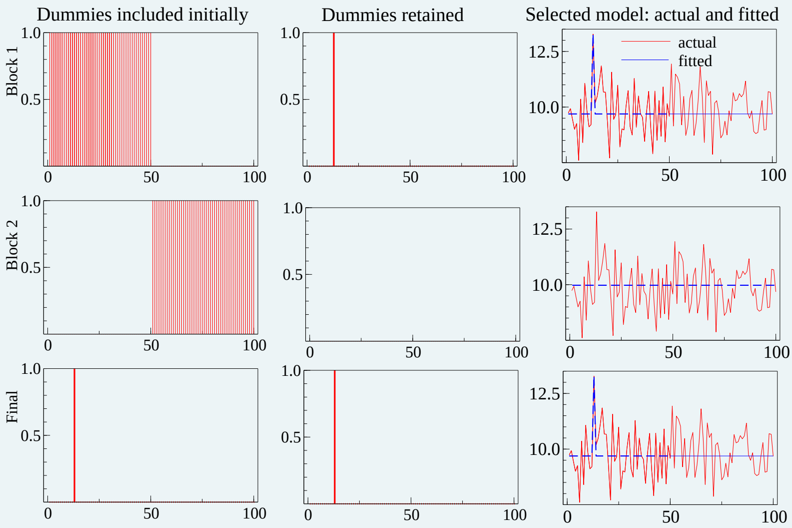
Recursive estimation equivalent to IIS over future sample, reducing indicators one at a time.

Feasible 'split-sample' IIS algorithm in Hendry, Johansen, and Santos (2008). First, include half of indicators, record significant: **just 'dummying out'** $T/2$ **observations**.
Then omit, include other half, record again.
Combine sub-sample indicators, & select significant.

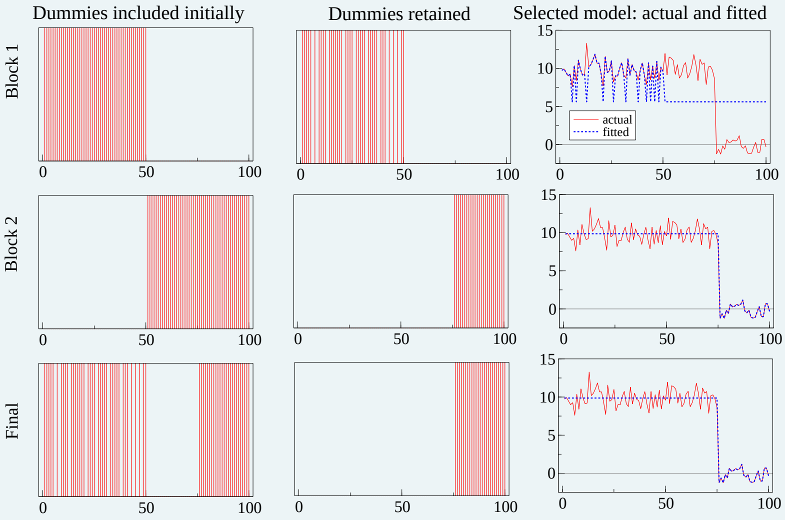
αT indicators selected on average under null at significance level α, so (e.g.) 99% 'efficient' at, say, $T = 100$ when $\alpha = 1/T = 0.01$.

Next Figure shows IIS in action for null:
following Figure records IIS for 10σ shift occurring at $0.75T = 75$.
Rows show first half; second half; final combination:
columns show the dummies entered; retained; and the data outcome.

'Split-sample' search under null



'Split-sample' search under alternative



Initially, many indicators now retained (top row),
considerable discrepancy between the first-half and second-half means.

When second set entered, all indicators for location shift period are retained.

Once combined set entered, despite large number of dummies, selection reverts to just those for break period.

Under null, indicators significant in sub-sample would remain so overall.
For non-null, sub-sample significance can be transient, due to unmodeled features that occur elsewhere in data.

Castle, Doornik, and Hendry (2012) show IIS can detect multiple location shifts and outliers,
including breaks close to start and end of sample,
as well as correcting for non-normality.

Hendry, Johansen, and Santos (2008) split-sample algorithm

Impulse indicators ordered $I_1, ..., I_T$. For $k = 2$ blocks:

1. Partition as $\mathcal{B}_1 = \{I_1, ..., I_{T/2}\}$ & $\mathcal{B}_2 = \{I_{T/2+1}, ..., I_T\}$
2. Estimate parameters in two sub-samples of $T/2$
3. Run model selection with each block to get $\mathcal{B}_1^*$ & $\mathcal{B}_2^*$
4. Form union $\mathcal{S} = \mathcal{B}_1^* \cup \mathcal{B}_2^*$
5. Run model selection on $\mathcal{S}$ for $\mathcal{S}^*$

Distribution of IIS under null known for several, and unequal, splits.

Johansen and Nielsen (2009) extend to stationary and unit-root autoregressive-distributed lag models.

Autometrics handles model selection with more variables than observations, tackling multiple breaks by IIS: see Doornik (2009).

When ρ was changed back to $\rho = 0.8$, the old equilibruim was restored, and forecasts rapidly converged back to $\mathsf{E}[x]$.
Suggests original model 'recovers' when DGP reverts.

Even so robust forecasts may do better.

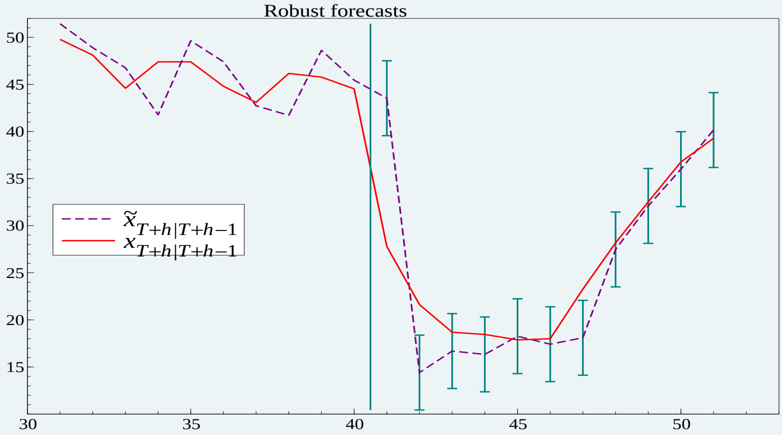
Difference the mis-specified model (5) **after estimation**:

$\Delta\widetilde{x}_{T+h|T+h-1} = \widehat{\rho}\Delta x_{T+h-1}$ so:

$$\widetilde{x}_{T+h|T+h-1} = x_{T+h-1} + \widehat{\rho}\Delta x_{T+h-1} \quad (10)$$

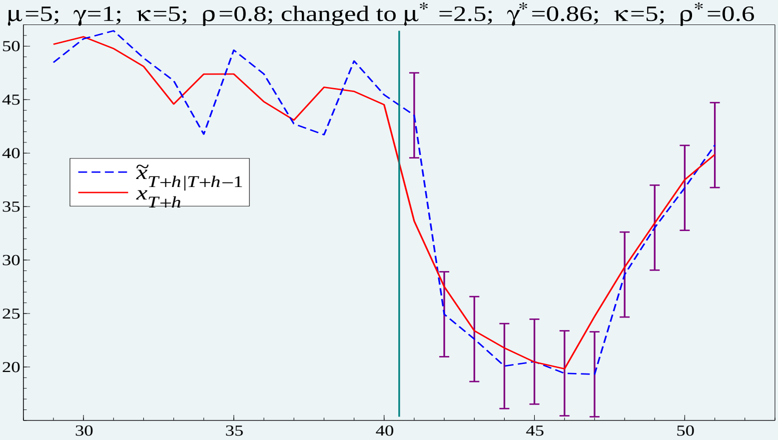
Uses 'wrong' $\widehat{\rho}$ for first 5 forecasts;
incorrectly differenced;
and omits relevant variable.

But avoids systematic forecast failure once $h > 2$.



Robust forecasting device (10) for DGP in Panel c:
avoids most of last 9 forecast errors.
RMSFE of *Panel d* is 6.6 versus 5.5 here; but 3.8 versus 2.0 over last 9 forecasts.

Apply to 'all parameters change' DGP.



Robust device **avoids almost all but first forecast error.**

Stark contrast to in-sample DGP forecasts.....

At $h > 2$-periods after the break, using:

$$\widetilde{x}_{T+h|T+h-1} = x_{T+h-1} + \widehat{\rho}\Delta x_{T+h-1} \quad (11)$$

so:

$$\widetilde{\epsilon}_{T+h|T+h-1} = (1 - \widehat{\rho})\,\Delta x_{T+h-1} \quad (12)$$

where from equation (7):

$$\Delta x_{T+h-1} = (\rho^* - 1)\,(x_{T+h-2} - \theta^*) + \gamma^*(z_{T+h-1} - \kappa^*) + \epsilon_{T+h-1} \quad (13)$$

so (11) contains everything you ever wanted to know when forecasting:
a] corrects to the new equilibrium through $(x_{T+h-2} - \theta^*)$;
b] includes the effect from z_{T+h-1} even though that is omitted from the forecasting device;
c] has the correct adjustment speed $(\rho^* - 1)$;
d] has the correct parameters γ^* and κ^*;
e] uses the in-sample well-determined estimate $\widehat{\rho}$, albeit that ρ has shifted.

Taxonomy explains outcomes of all six forecast scenarios:

a model matches DGP: only (iv) and (vi) matter;
b model mis-specified: but (v) only adds to variance;
c model non-constant & mis-specified, but zero long-run mean;
d shift in non-zero long-run mean induces forecast failure;
e even if model matches DGP in-sample;
f model non-constant & mis-specified, but long-run mean constant despite changes in intercepts and means–no failure.

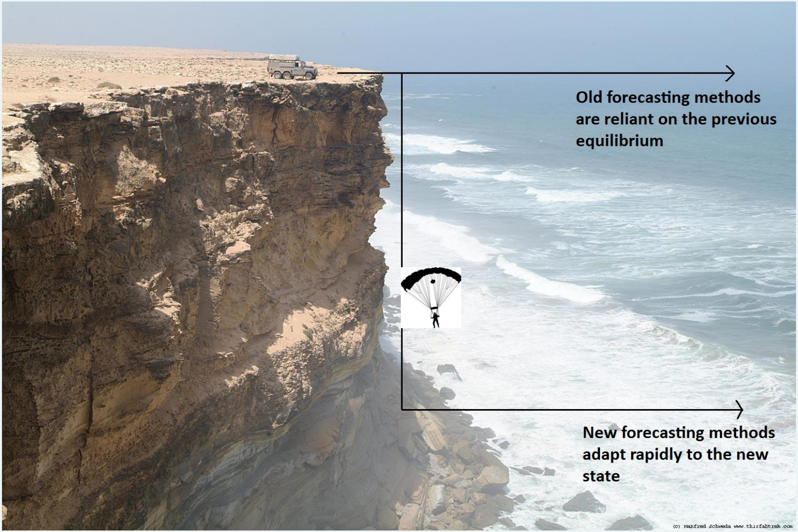
Conclude: location shifts are primary cause of forecast failure. Implies models robust after location shifts should avoid **systematic** mis-forecasting.

a shows that a model which matches the DGP can forecast well;
b shows that a mis-specified model can forecast quite well;
c shows that a non-constant, mis-specified model can still forecast quite well;
d or very badly depending on a non-zero long-run mean;
e as can a model which matches the DGP in-sample;
f whereas a non-constant, mis-specified model with a non-zero long-run mean can also forecast quite well, absent a location shift; and a few periods after a location shift, a differenced model with no economic content, can forecast well.

No connection between in-sample verisimilitude and later presence or absence of forecast failure.

Changing equilibrium

Changing equilibrium


Old forecasting methods are reliant on the previous equilibrium
New forecasting methods adapt rapidly to the new state
(c) Manfred Schweda www.thisfabtrek.com

(A) **Introduction**

(B) **What causes forecasting problems?**

(C) **Explaining forecasting**

(D) **Empirical application of forecasting theory**

(E) Forecasting breaks

Conclusions

Selected autoregressive model by *Automatics* 1989(2)–2007(4):

$$\widehat{y}_t = \underset{(0.04)}{0.93}\, y_{t-1} + \underset{(0.001)}{0.002} - \underset{(0.005)}{0.014}\, 1_{1990(3)}$$

$$\widehat{\sigma} = 0.0047 \quad \chi^2(2) = 2.27 \quad F_{ar}(5,67) = 2.60^{*} \quad R^2 = 0.89$$

$$F_{het}(2,71) = 2.73 \quad F_{arch}(5,67) = 2.33 \quad F_{reset}(2,70) = 1.03$$

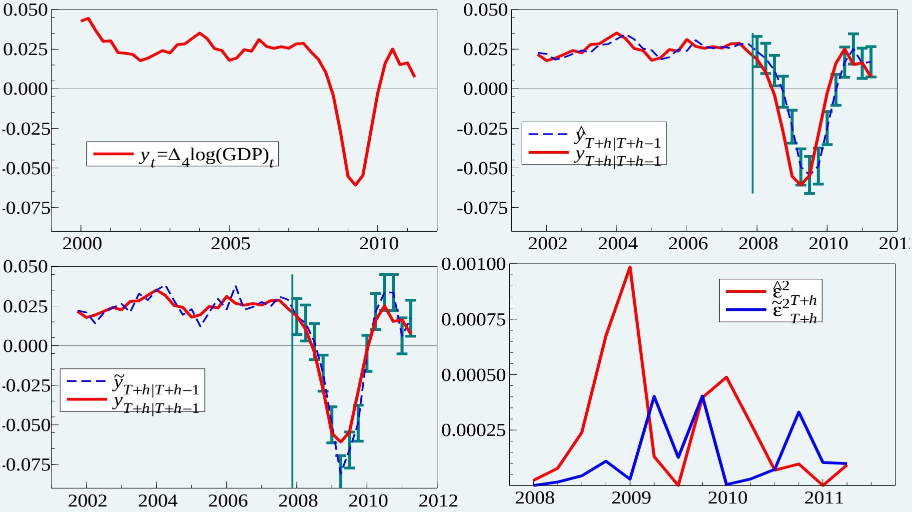
$1_{1990(3)}$ is indicator for $1990(3)$ selected by IIS.

Corresponding robust device, therefore, was:

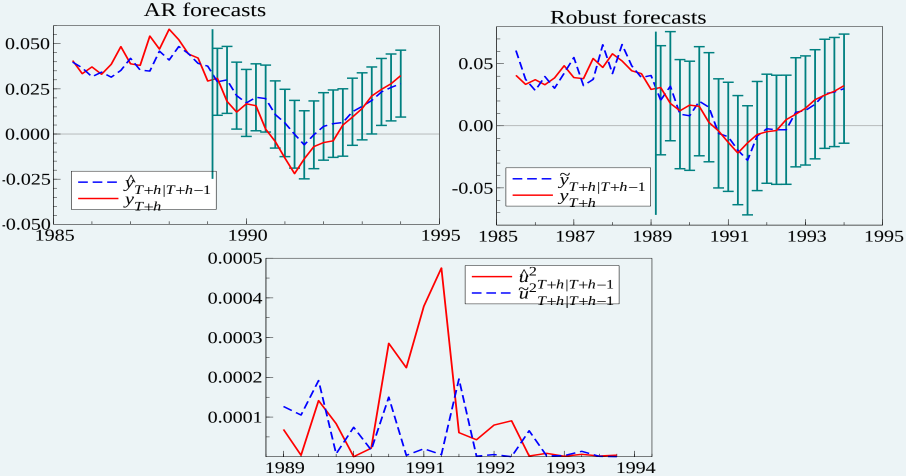
$$\widetilde{y}_t = y_{t-1} + 0.93\Delta y_{t-1} \qquad \widetilde{\sigma} = 0.0057$$

Their respective forecasts follow.

Data, forecasts and squared forecast errors



RMSFEs over first 5 of 0.122 v. 0.062 so nearly halved.



HM Treasury badly mis-forecast earlier recession 1989(1)–1993(4): similar to AR forecasts here: but robust again performs well.

Can wait to use robust after forecast failure has occurred: but should always calculate. Even better–**forecast the break**.

(A) **Introduction**

(B) **What causes forecasting problems?**

(C) **Explaining forecasting**

(D) **Empirical application of forecasting theory**

(E) Forecasting breaks

Conclusions

- **Essentially require a crystal ball to foresee shifts:** but worth investigating what would be required.
- **First must analyze 'unpredictability': especially breaks.**

Type of unpredictable break matters

- **'Internal' break changes the model in use.**
- **'External' shift alters 'forecast conditions', leaving model unchanged.**

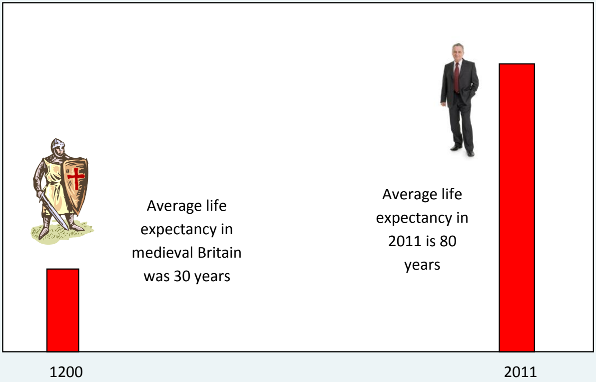
Both can cause forecast failure.

- Potential role for **many different information sources**, including **surveys**, *Google Trends* & volatility measures.
- **If fail to forecast break, could model the change process: predict impact of an 'internal' break during its progress.**
- **Can also mitigate forecast failure by robust devices.**

Types of unpredictability

- **Intrinsic unpredictability** in a known distribution: chance distribution sampling, ‘random errors’, etc.;
- **Instance unpredictability** (known unknowns): outliers at unanticipated times from fat-tailed distributions;
- **Extrinsic unpredictability** (unknown unknowns): unanticipated shifts of distributions.

Change is endemic:
The world is vastly different today compared to yesterday.



Future will see more large, unanticipated shocks;
New thinking to adapt to such challenges

Each type of unpredictability has different effects on:

- **economic analyses**;
- **forecasting**;
- **empirical modeling**.

First two go awry from **extrinsic** unpredictability as:
(a) conditional expectations today are then **biased** for outcomes tomorrow, and are **not** minimum variance predictors,
so mathematical basis of inter-temporal economics is invalid after shifts–major risks from using models based on such derivations;
(b) **forecast failure occurs**–major risks from not using robust devices;
(c) yet, **those outcomes are susceptible to modeling *ex post***–major risks from not doing so.

Analyzed in **Clements and Hendry (2005)**

Develop methods for forecasting breaks, with robust strategies if breaks incorrectly predicted

First requires that:

(1) breaks are predictable

(2) there is information relevant to that predictability

(3) such information is available at forecast origin

(4) we have a forecasting model that embodies it

(5) we have a method for selecting that model

(6) resulting forecasts are usefully accurate

Analyzed in **Castle, Fawcett, and Hendry (2011)**

2004 Indian Ocean Boxing Day tsunami

(1) Failure to predict undersea earthquake off Sumatra: potentially predictable, with wide margin of timing– see Stein, Barka, and Dieterich (1997) for predictions of 1999 earthquake at İzmit;

(2) no relevant information at time to predict earthquake, but once tsunami started, advance warning feasible: Holliday, Rundle, Tiampo, and Turcotte (2006) show stress tension and its release were measured;

(3) but unfortunately, not noted at time;

(4) forecasting models for tsunamis make timings and locations of impacts predictable;

(5) selecting appropriate model based on physical theory, calibrated once tsunami warning system in place;

(6) then uncertainty lies within fairly small intervals.

Several possibilities:

'leading indicators'–but historical record unimpressive;

non-linear functions of variables already in models–same.

Surveys can offer timely, and sometimes advance, information.

Also consider information outside usual subject matter:
Castle, Fawcett, and Hendry (2011) discuss survey data, *Google Trends*: see Choi and Varian (2009) & prediction markets (Iowa).

Software like *Autometrics* can handle hundreds of candidate variables, non-linear functions and multiple location shifts: see Doornik (2009)

Rapid information updates at forecast origin using **high-frequency data should help**.

May detect breaks sooner, so adapt better
But higher-frequency data can be noisier...

(A) **Introduction**

(B) **What causes forecasting problems?**

(C) **Explaining forecasting**

(D) **Empirical application of forecasting theory**

(E) **Forecasting breaks**

Conclusions

Conclusions on model-based forecasting

In non-stationary economies subject to unanticipated structural breaks, where models differ from DGPs in unknown ways, selected from unreliable data,
forecasting implications differ considerably from
model = DGP in a constant mechanism.

Unanticipated location shifts pernicious for forecasting:
systematic mis-forecasting in all forms of equilibrium-correction models.

Yet every DGP parameter shifted without any noticeable effect if no location shift.

Many risks from not forecasting location shifts

- **forecasting–**
  failure primarily due to location shifts;
- **black swans cease to be ‘independent’–**
  so flocks occur;
- **difficult to predict location shifts–**
  worth trying;
- **difficulties of prediction remain during breaks–**
  even if known form;
- **but can mitigate failure after location shifts–**
  by robust devices.
- **economic analyses–**
  mathematical basis is invalid after location shifts;
- **verisimilitude of a model not checked by forecasting success–**or failure.

References

Castle, J. L., J. A. Doornik, and D. F. Hendry (2012). Model selection when there are multiple breaks. *Journal of Econometrics 169*, 239–246.

Castle, J. L., N. W. P. Fawcett, and D. F. Hendry (2010). Forecasting with equilibrium-correction models during structural breaks. *Journal of Econometrics 158*, 25–36.

Castle, J. L., N. W. P. Fawcett, and D. F. Hendry (2011). Forecasting Breaks and During Breaks. In M. P. Clements and D. F. Hendry (Eds.), *Oxford Handbook of Economic Forecasting*, pp. 315–353. Oxford: Oxford University Press.

Castle, J. L. and N. Shephard (Eds.) (2009). *The Methodology and Practice of Econometrics*. Oxford: OUP.

Choi, H. and H. Varian (2009). Predicting the present with Google Trends. Economics Research Group, Google.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica 28*, 591–605.

Clements, M. P. and D. F. Hendry (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Clements, M. P. and D. F. Hendry (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, Mass.: MIT Press.

Clements, M. P. and D. F. Hendry (2005). Guest Editors' introduction: Information in economic forecasting. *Oxford Bulletin of Economics and Statistics 67*, 713–753.

Doornik, J. A. (2009). Autometrics. See Castle and Shephard (2009), pp. 88–121.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics 135*, 399–426.

Hendry, D. F., S. Johansen, and C. Santos (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics 33*, 317–335. Erratum, 337–339.

Hendry, D. F. and G. E. Mizon (2010). On the mathematical basis of inter-temporal optimization. Discussion paper 497, Economics Department, Oxford.

Hendry, D. F. and G. E. Mizon (2012). Open-model forecast-error taxonomies. In X. Chen and N. R. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, DOI: 10.1007/978–1–4614–1653–1_9. New York: Springer.

Holliday, J. R., J. B. Rundle, K. F. Tiampo, and D. L. Turcotte (2006). Using earthquake intensities to forecast earthquake occurrence times. *Nonlinear Processes in Geophysics 13*, 585–593.

Johansen, S. and B. Nielsen (2009). An analysis of the indicator saturation estimator as a robust regression estimator. See Castle and Shephard (2009), pp. 1–36.

Salkever, D. S. (1976). The use of dummy variables to compute predictions, prediction errors and confidence intervals. *Journal of Econometrics 4*, 393–397.

Stein, R., A. Barka, and J. Dieterich (1997). Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International 128*, 594–604.

(A) Introduction

(B) What causes forecasting problems?

(C) Explaining forecasting

(D) Empirical application of forecasting theory

(E) Forecasting breaks

Conclusion