Forecasting using Economic Systems in Uncertain Environments

Jennifer L. Castle and David F. Hendry

Department of Economics, Oxford University

ESRC Research Methods Festival
3 July 2008

The Salmon of Doubt

“Trying to predict the future is a mug’s game. But ... we need to have some sort of idea of what the future’s actually going to be like because we are going to have to live there,

The Salmon of Doubt

“Trying to predict the future is a mug’s game. But ... we need to have some sort of idea of what the future’s actually going to be like because we are going to have to live there,

probably next week.”

Douglas Adams
MacMillan, 2002.

Introduction

Economic forecasting confronts a **non-stationary, evolving world**, where model and mechanism differ.

Poor historical track record of econometric systems: forecast failures, and out-performed by ‘naive devices’. Problems date from the early history of econometrics.

Such an adverse outcome is surprising: **econometrics uses inter-temporal causal information**.

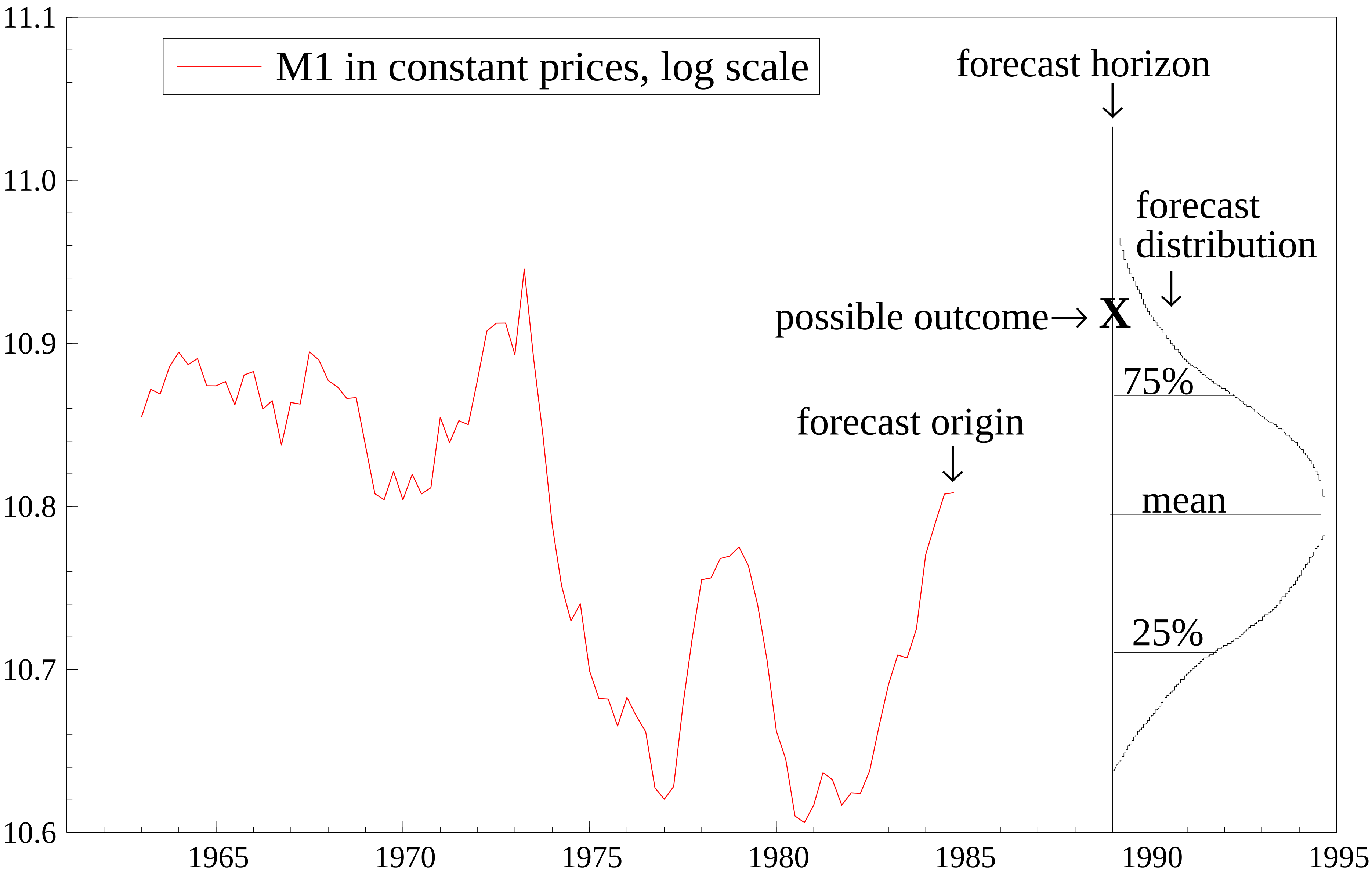
Our aim:

Explain main causes of forecast failure;

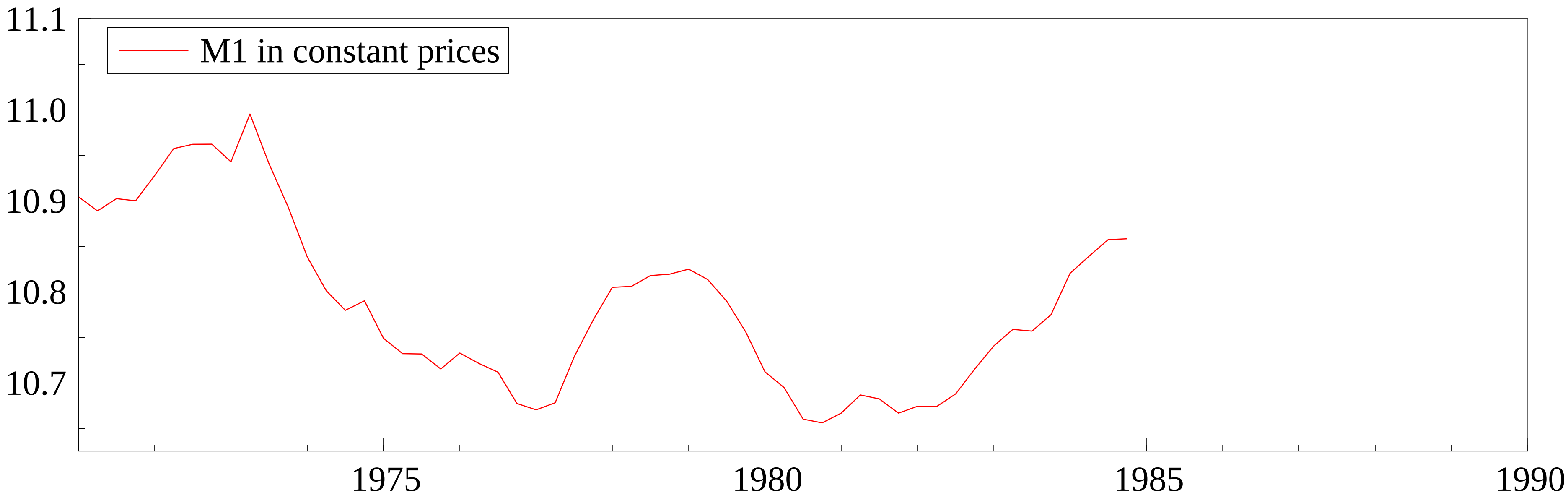
Methods to insure against systematic forecast failure;

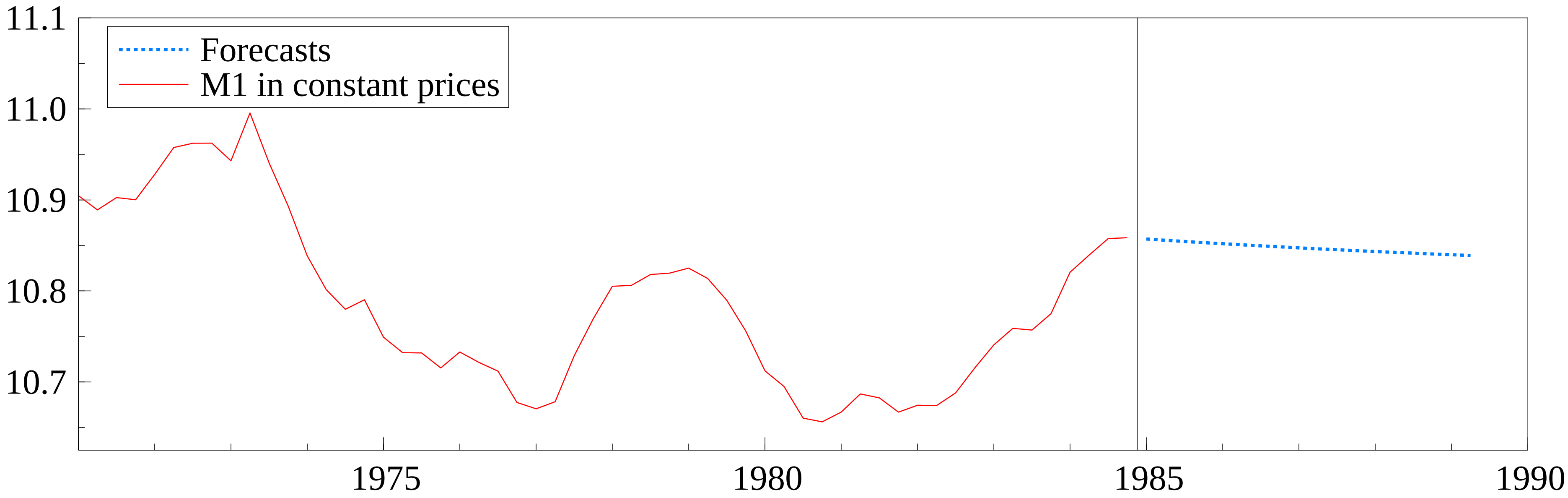
Some progress towards forecasting during breaks.

Time series of M1 in constant prices

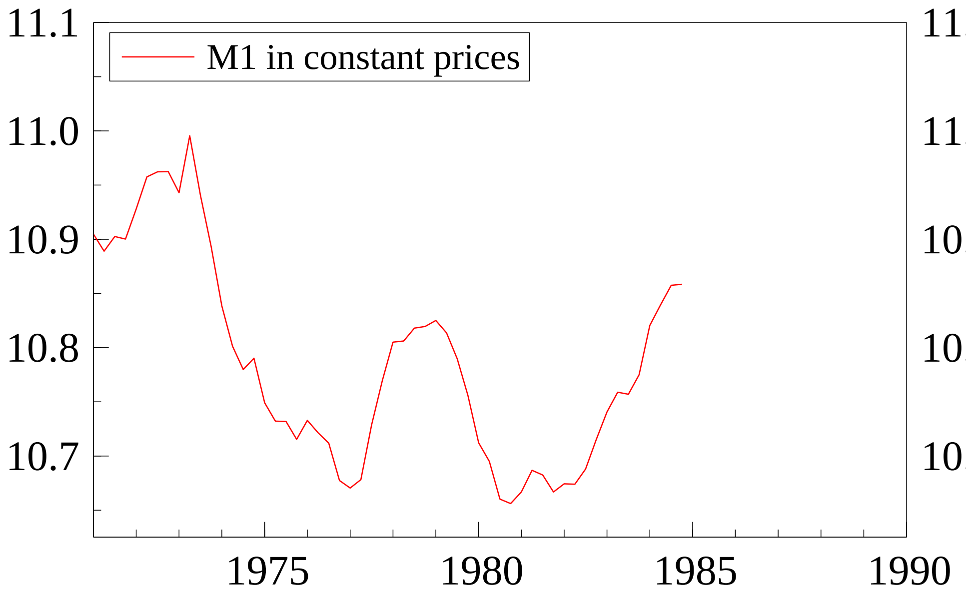


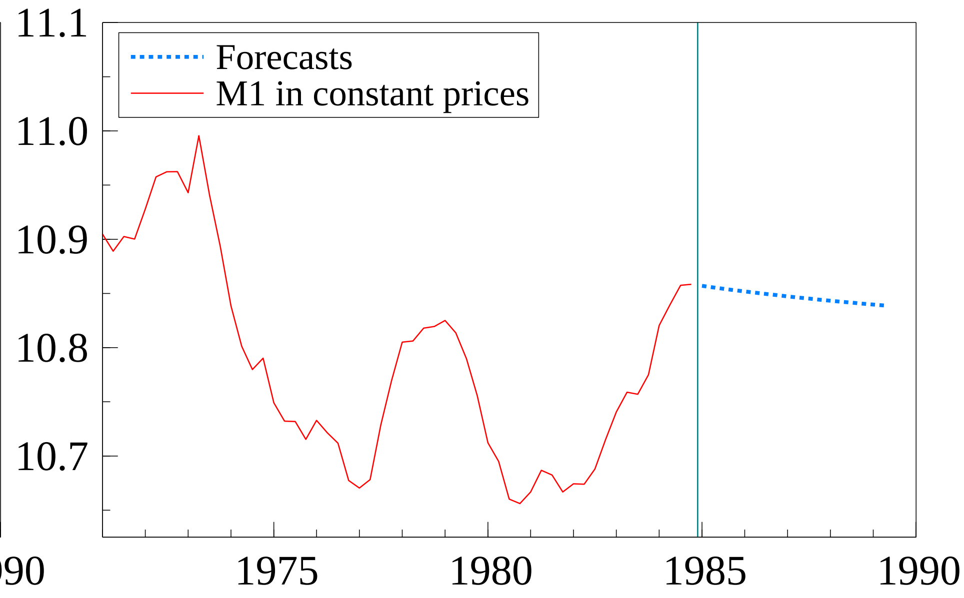
Forecasts of M1 in constant prices

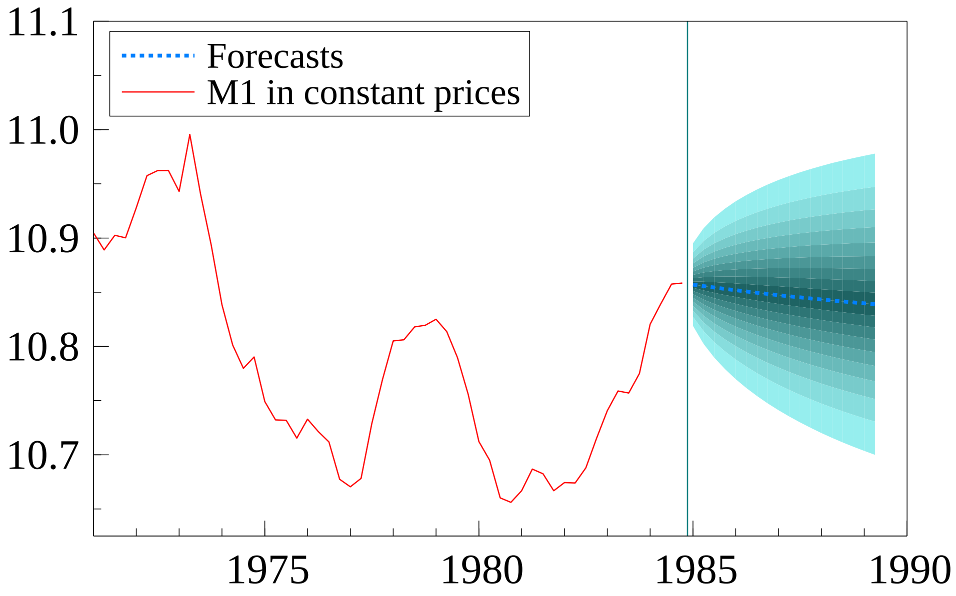




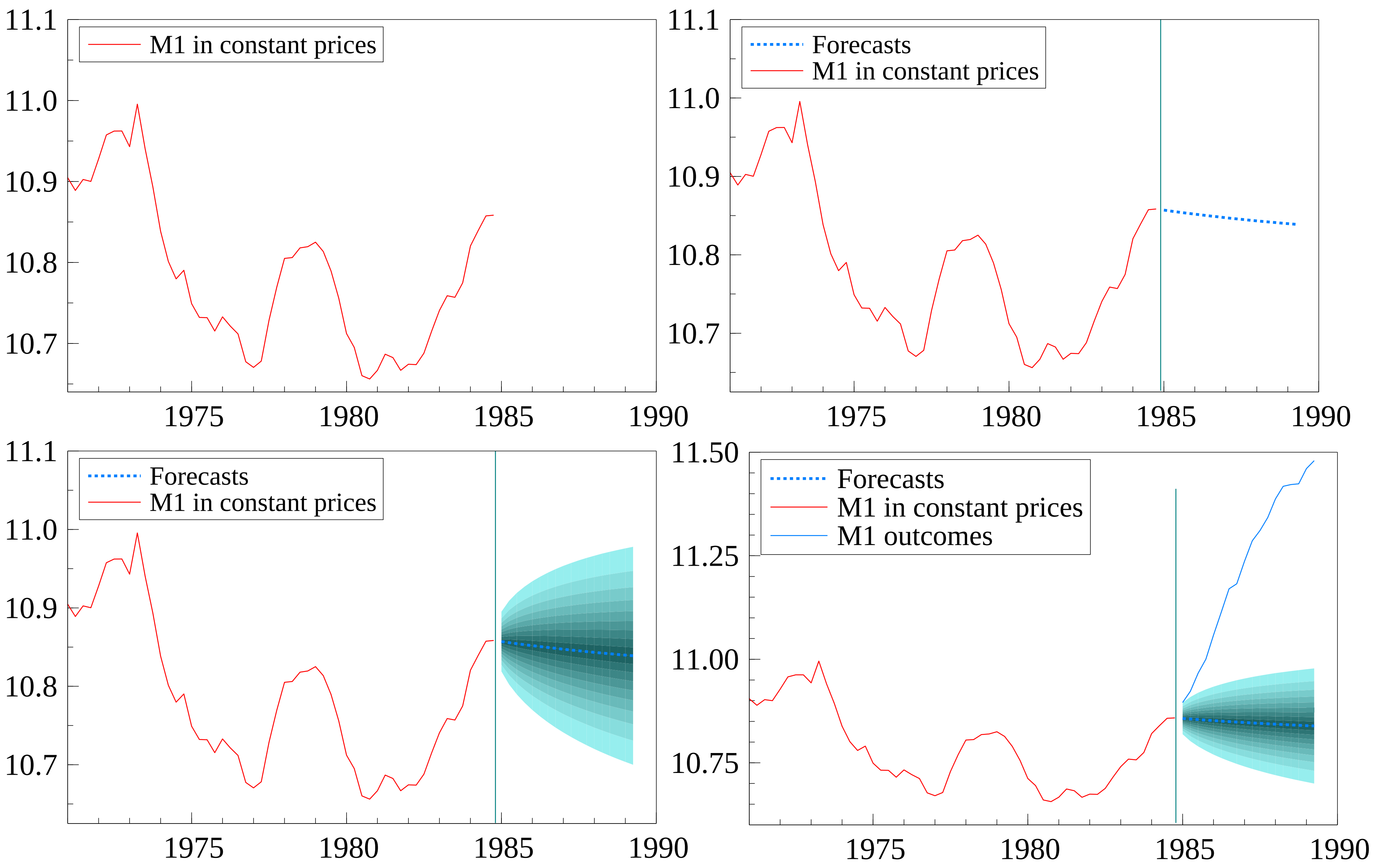
Forecasts of M1 with uncertainty







Outcomes of M1 with uncertainty!



Forecast uncertainty

Problem with forecasting is: **future is uncertain**.
Forecast uncertainty is intrinsic; but two sources:
one we know is present and understand the probabilities;
and one due to factors we do not even know exist.

> **“Because of the things we don’t know [that] we don’t know, the future is largely unpredictable.”** Maxine Singer, 1997, *Thoughts of a Nonmillenarian*, p. 39.

In tossing 2 dice, the two sources are:

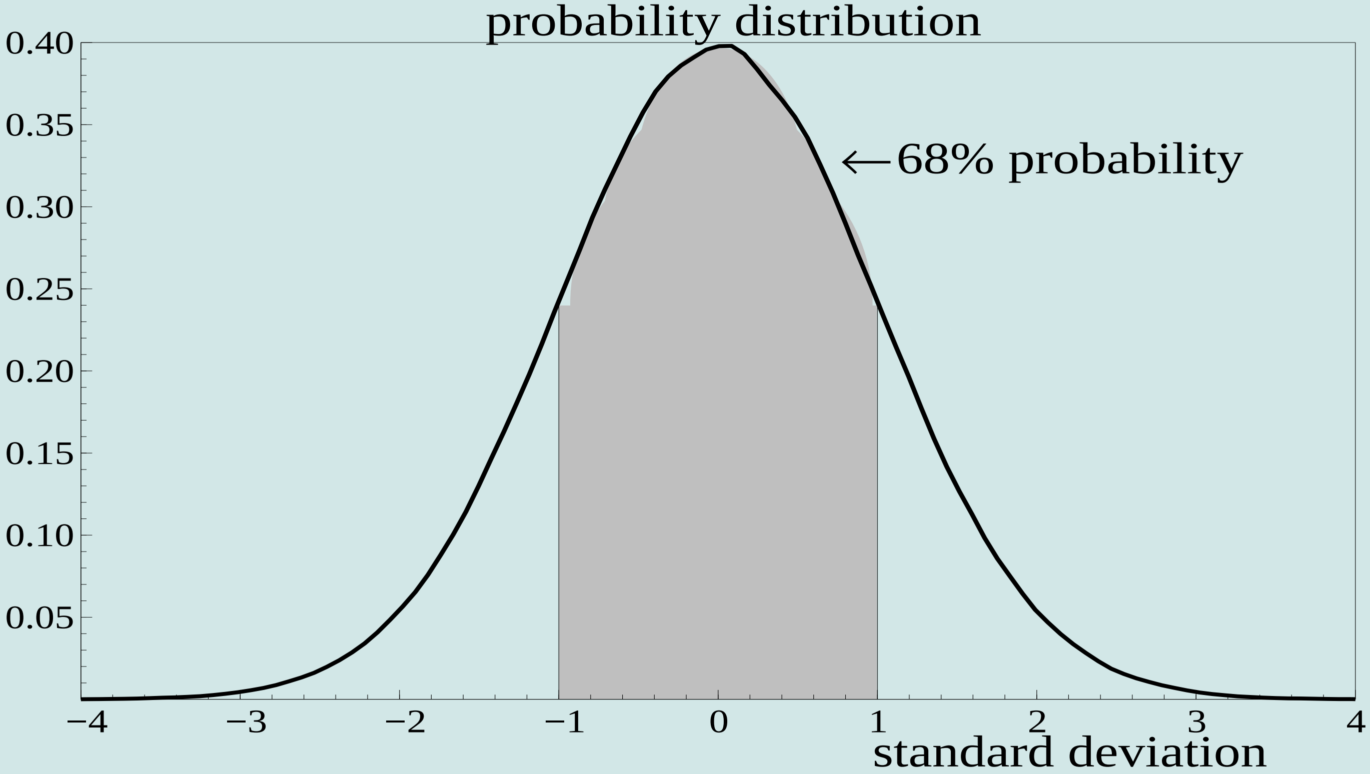
- the probability any pair of numbers will be face up
- **not knowing that the dice are loaded**.

Second is type of problem in above quote by Singer.

Statistical science

Based on:

individually unpredictable events are ‘regular’ on average.



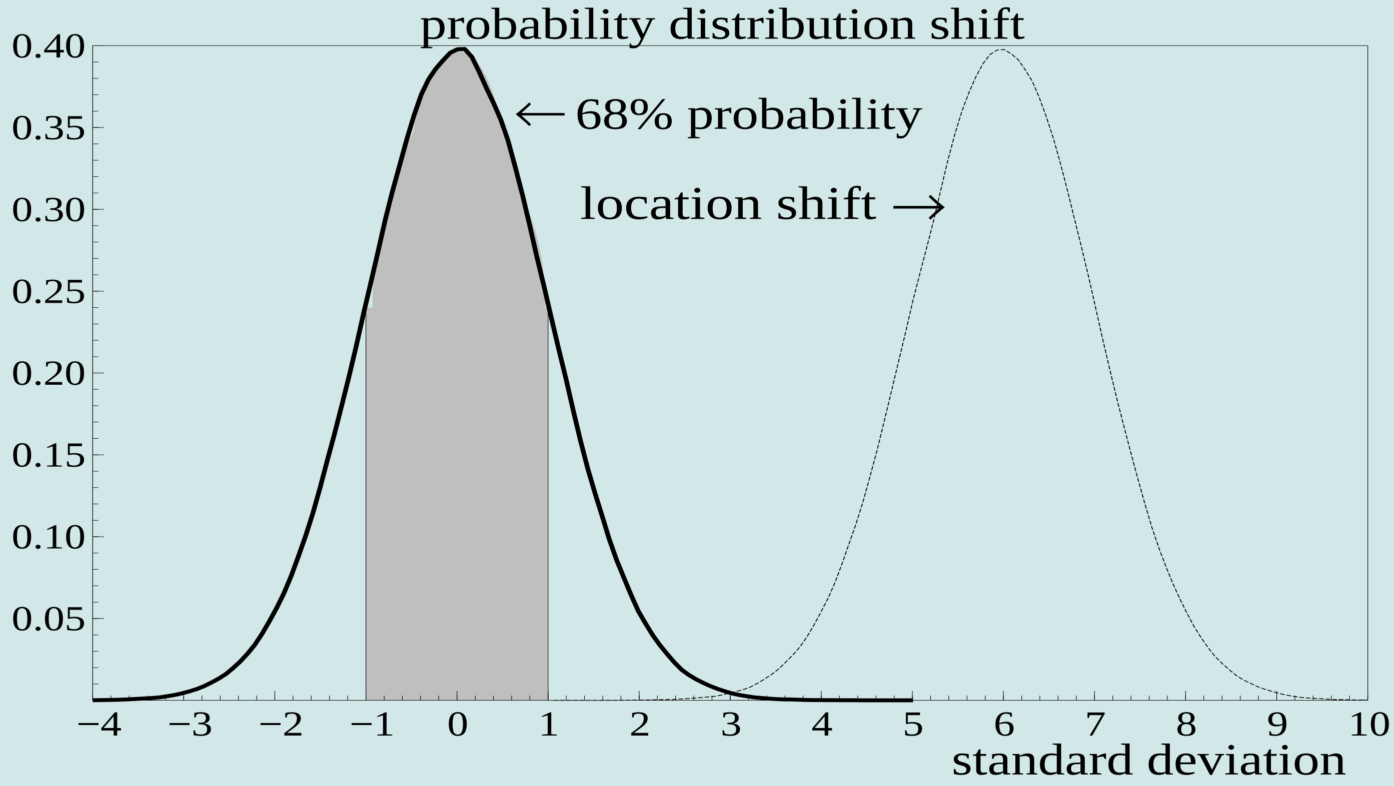
Economic forecasting similar:

models of economy ‘average’ over possible future ‘shocks’.

Works well for ‘measurable uncertainty’: **but—**

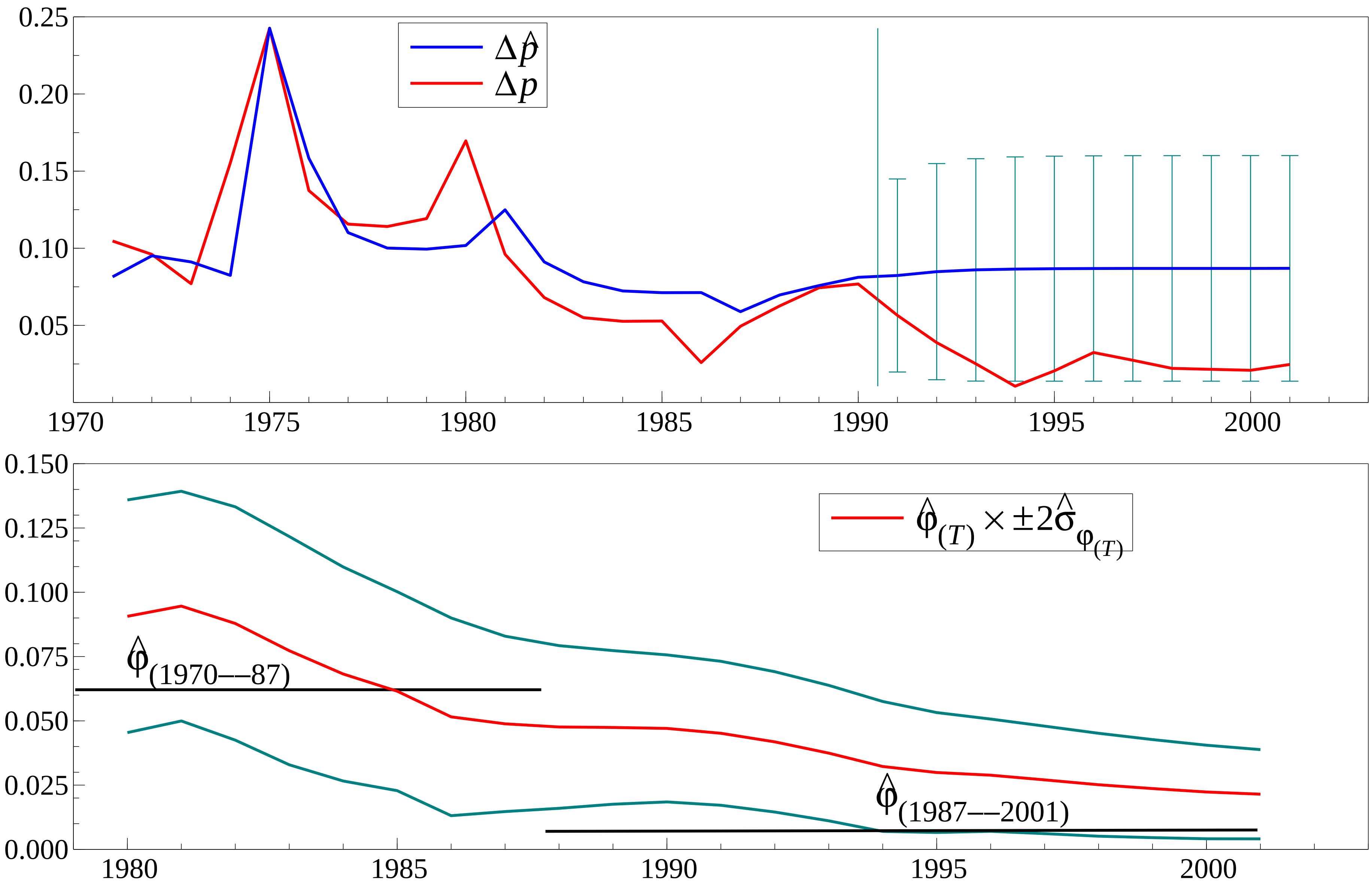
Forecasting difficulties

Economic forecasters confront a difficult environment.
Impossible to conceive of all possibilities:
economic ‘earthquakes’ seem to occur all too often.
Unmeasured uncertainty important for the future.



Good guides sparse when future is not like past.
Distributions of events change over time: **non-stationarity**.

AR(1) inflation forecasts



Example

Stationary scalar first-order autoregressive example:

$$x_t = \rho x_{t-1} + v_t \quad \text{where} \quad v_t \sim \mathsf{IN}\left[0, \sigma_v^2\right] \quad \text{and} \quad |\rho| < 1.$$

With ρ known and constant, forecast from x_T is:

$$\widehat{x}_{T+1|T} = \rho x_T$$

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

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

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

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

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

But inflation example illustrates main problem:
Mean has changed over time.

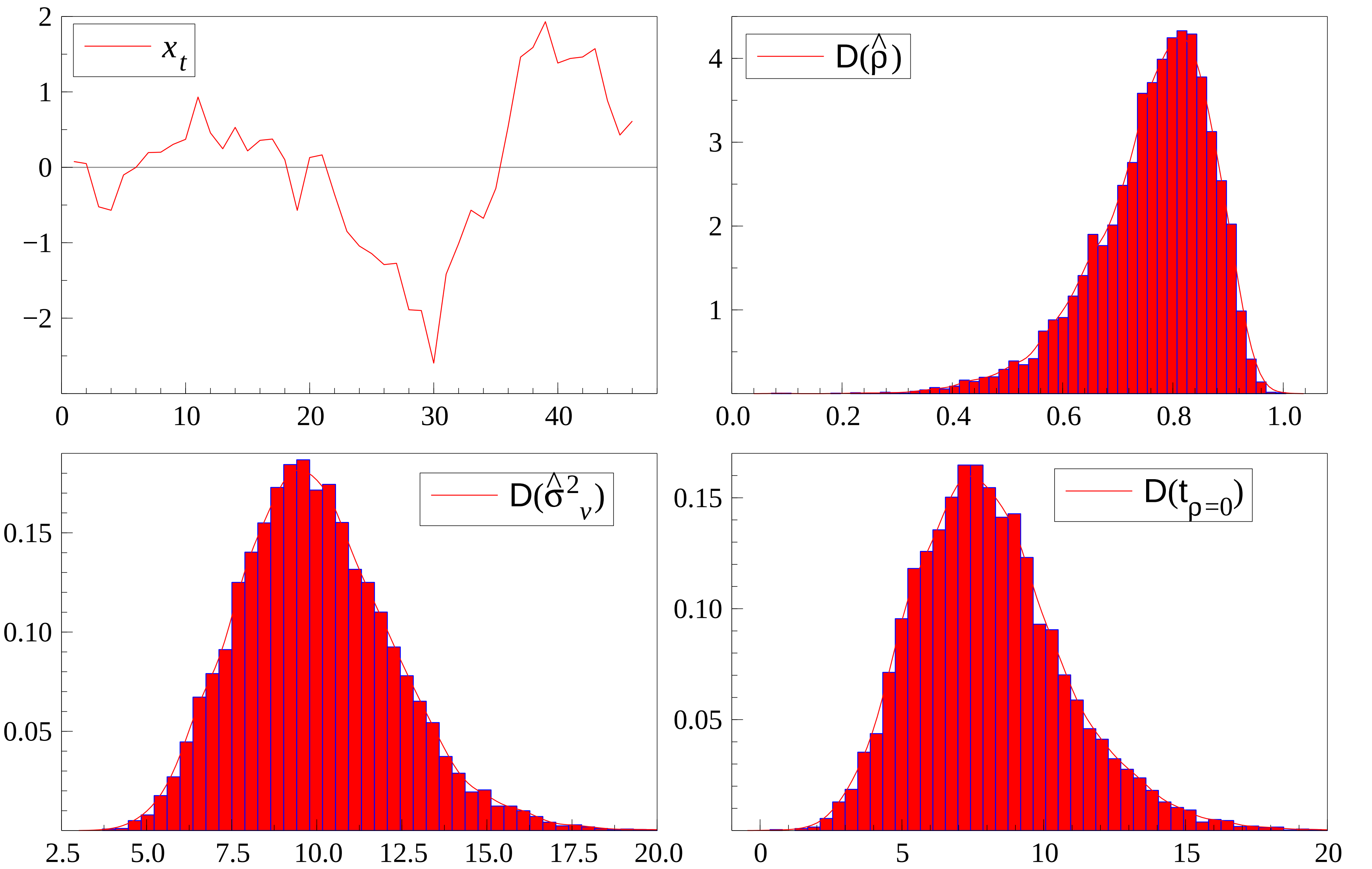
Potential problems

(1) Specification incomplete if (e.g.) vector $\mathbf{x}_t$ not scalar.
(2) Measurement incorrect if (e.g.) observe $\widetilde{\mathbf{x}}_t$ not $\mathbf{x}_t$.
(3) Formulation inadequate if (e.g.) intercept needed.
(4) Modelling wrong if (e.g.) selected ρx_{t-2}.
(5) Estimating ρ adds bias, $(\rho - \mathsf{E}[\widehat{\rho}])x_T$, and variance $\mathsf{V}[\widehat{\rho}]x_T^2$.
(6) Properties of $\mathsf{D}(v_t) = \mathsf{IN}\left[0, \sigma_v^2\right]$ determine $\mathsf{V}[x_t]$.

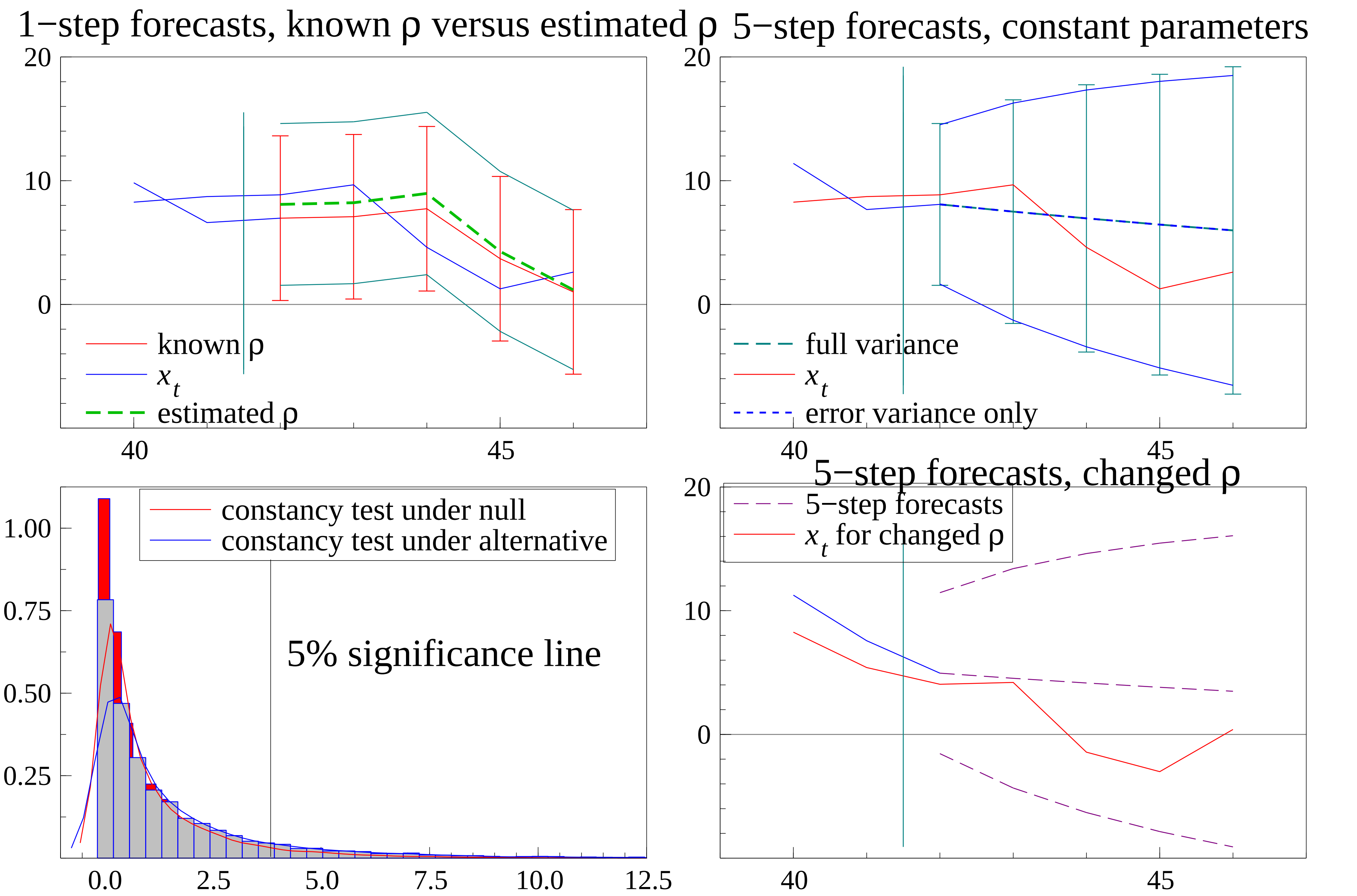
(7) Assumed $v_{T+1} \sim \mathsf{IN}\left[0, \sigma_v^2\right]$ **but** $\mathsf{V}[v_{T+1}]$ **could differ**.
(8) Multi-step forecast error $\sum_{h=1}^{H} \rho^{h-1} v_{T+h}$ has $\mathsf{V} = \frac{1-\rho^{2H}}{1-\rho^2}\sigma_v^2$.
(9) If $\rho = 1$ have trending forecast variance $H\sigma_v^2$.
(10) If ρ changes could experience forecast failure.

Must be prepared for risks from (1)–(10).

Estimated AR1: $\rho = 0.8, T = 40, \sigma_v^2 = 10$



AR1 forecasts: break in $\rho = 0.4$ at $T = 40$



Problems hardly disastrous

Small increase in uncertainty from estimating ρ;
forecast intervals grow quite slowly as H increases.
Little noticeable impact from halving ρ at $T = 40$.
Constancy test hardly rejects false null.

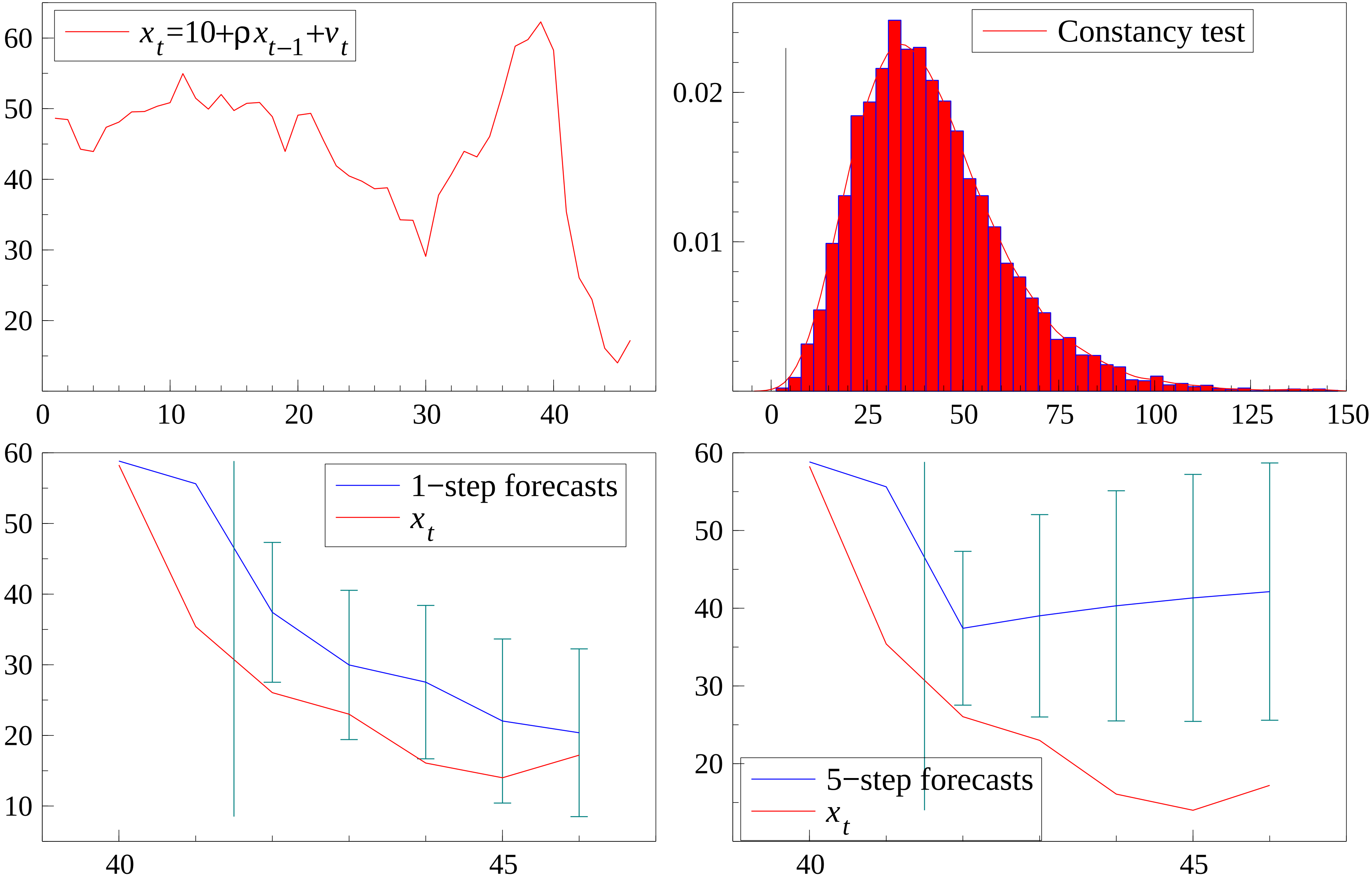
But, slight change to model:

$$x_t = \alpha + \rho x_{t-1} + v_t \text{ where } v_t \sim \mathsf{IN}\left[0, \sigma_v^2\right] \text{ and } |\rho| < 1.$$

Everything else the same, except $\alpha = 10$.
Little change in estimation distributions or forecasts:
until non-constant ρ, **for same size and time of break**.

Then – **catastrophe!!**

AR1 forecasts: intercept & break in ρ



Problems now disastrous

Change due to effect on $\mathsf{E}[x_t]$.

In first case $\mathsf{E}[x_t] = 0$ before and after shift in ρ.

In second: $\mathsf{E}[x_t] = \alpha/(1-\rho)$.
Shifts markedly from 50 to 17.

All models in this class are **equilibrium correction**:
so fail systematically if $\mathsf{E}[\cdot]$ changes.

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

Pervasive and pernicious problem.

Explanation

Must write conditional expectation as:
$\widehat{\mathbf{x}}_{T+h|T} = \mathsf{E}_{T+h}[\mathbf{x}_{T+h}|\mathbf{X}_T]$.
Fine if stationary: $\mathsf{E}_{T+h} = \mathsf{E}_T$.
But paradox if $\mathsf{D}_{\mathsf{x}_t}(\cdot)$ not constant:
need to know whole future distribution to derive forecast.
Cannot prove $\widetilde{\mathbf{x}}_{T+h|T} = \mathsf{E}_T[\mathbf{x}_{T+h}|\mathbf{X}_T]$ is useful.

Empirically-relevant theory needs to allow for:

model mis-specified for DGP
parameters estimated from inaccurate observations,
on an integrated-cointegrated system,
which intermittently alters unexpectedly
from structural breaks.

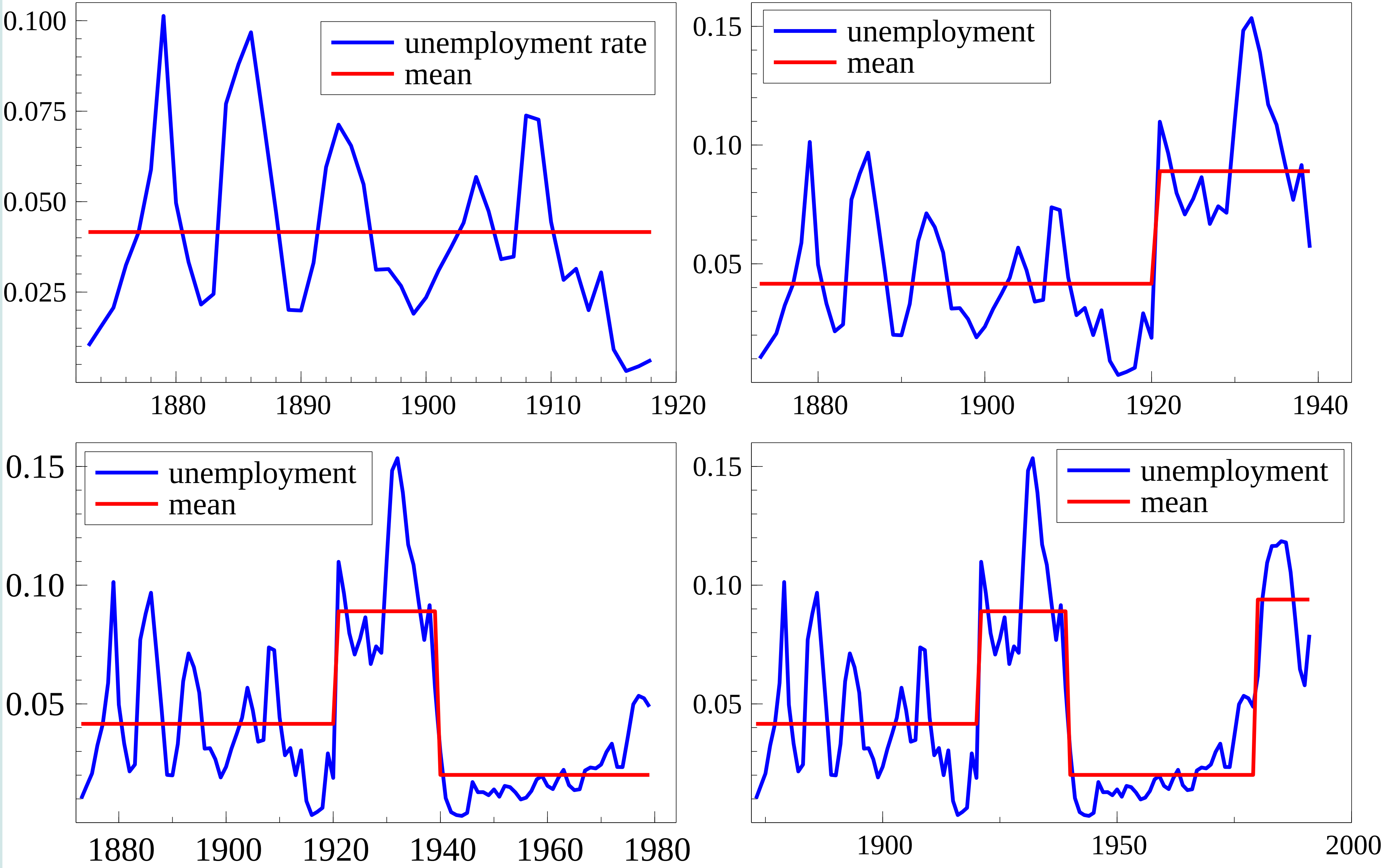
Possible forecasting problems

Mis-specification, mis-estimation, non-constancy,
of **deterministic**, **stochastic**, or **error** components,
all could induce forecast failure.

But location shifts are the key problem,
namely shifts in parameters of **deterministic** components.
Location shifts easy to detect: see figure 21.
Other breaks not so easy to detect:
impulse response analyses then unreliable.

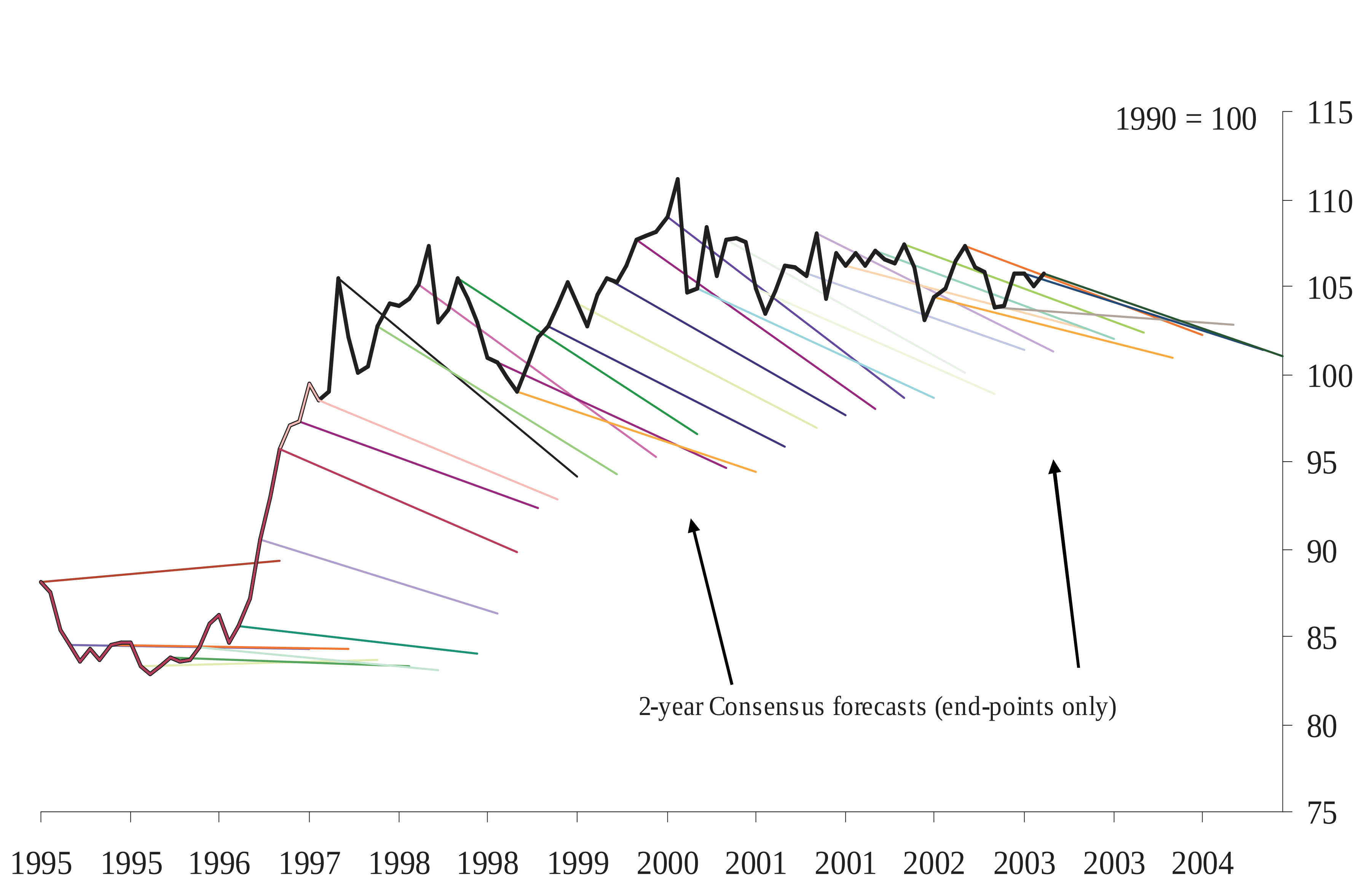
Many conventional results change radically
when parameter non-constancy:
non-causal models can outperform causal;
multi-step forecasts more accurate than 1-step;
intercept corrections can improve forecasts.

Location shifts in UK unemployment

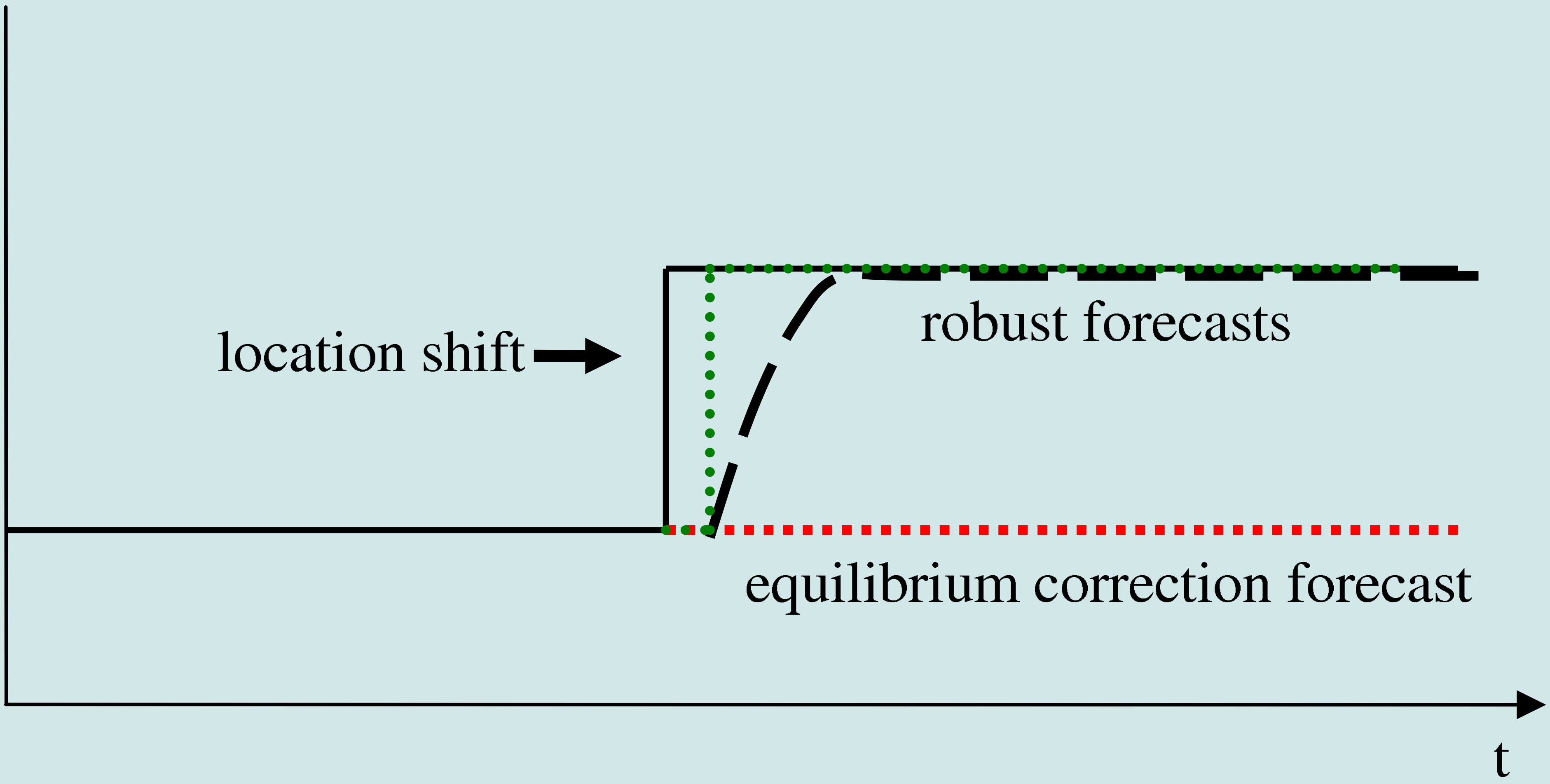


Forecast failure for UK unemployment

£ERI outturns & 2-year consensus forecasts

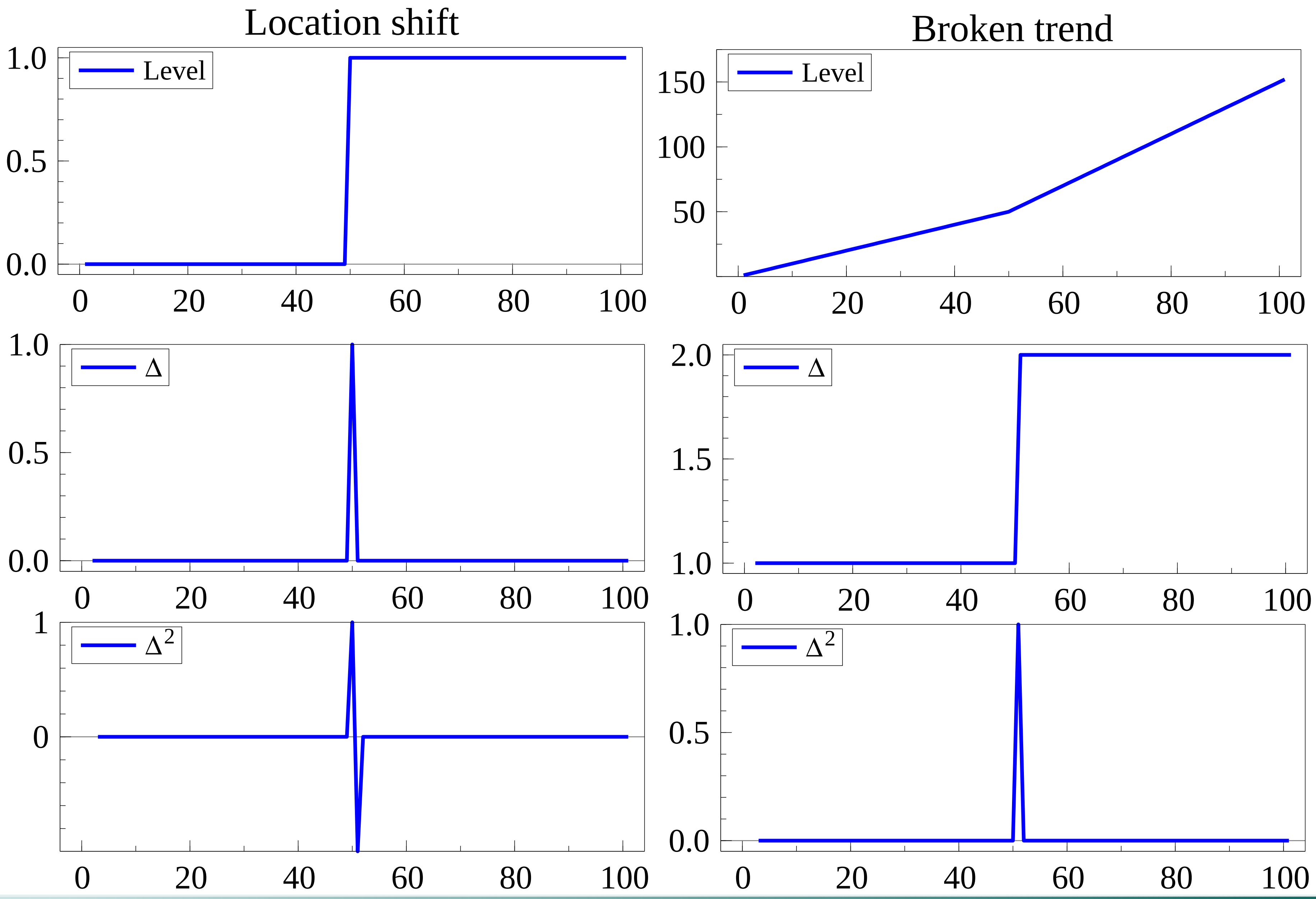


Robust Forecasts



When $f_{t+1}(\cdot) \neq f_t(\cdot)$, forecasting devices robust to location shifts win forecasting competitions.

Location shifts and broken trends



Using $\Delta \mathrm{x}_T$ to forecast

Consider the in-sample DGP:

$$\Delta \mathbf{x}_T = \boldsymbol{\gamma} + \boldsymbol{\alpha}\left(\boldsymbol{\beta}'\mathbf{x}_{T-1} - \boldsymbol{\mu}\right) + \boldsymbol{\Psi}\mathbf{z}_T + \mathbf{v}_T, \tag{1}$$

where $\mathbf{z}_t$ denotes many omitted effects, with:

$$\Delta \mathbf{x}_{T+i} = \boldsymbol{\gamma}^* + \boldsymbol{\alpha}^*\left((\boldsymbol{\beta}^*)'\mathbf{x}_{T+i-1} - \boldsymbol{\mu}^*\right) + \boldsymbol{\Psi}^*\mathbf{z}_{T+i} + \mathbf{v}_{T+i}. \tag{2}$$

A VEqCM in $\mathbf{x}_t$ is used for forecasting:

$$\Delta \widehat{\mathbf{x}}_{T+i|T+i-1} = \widehat{\boldsymbol{\gamma}} + \widehat{\boldsymbol{\alpha}}\left(\widehat{\boldsymbol{\beta}}'\mathbf{x}_{T+i-1} - \widehat{\boldsymbol{\mu}}\right). \tag{3}$$

All main sources of forecast error occur given (2):
stochastic and deterministic breaks;
omitted variables;
inconsistent parameters;
estimation uncertainty;
innovation errors.

DDV avoids failure

Contrast using sequence of $\Delta \mathbf{x}_{T+i-1}$ to forecast:

$$\Delta \widetilde{\mathbf{x}}_{T+i|T+i-1} = \Delta \mathbf{x}_{T+i-1}. \tag{4}$$

But because of (2), $\Delta \mathbf{x}_{T+i-1}$ is $(i > 1)$:

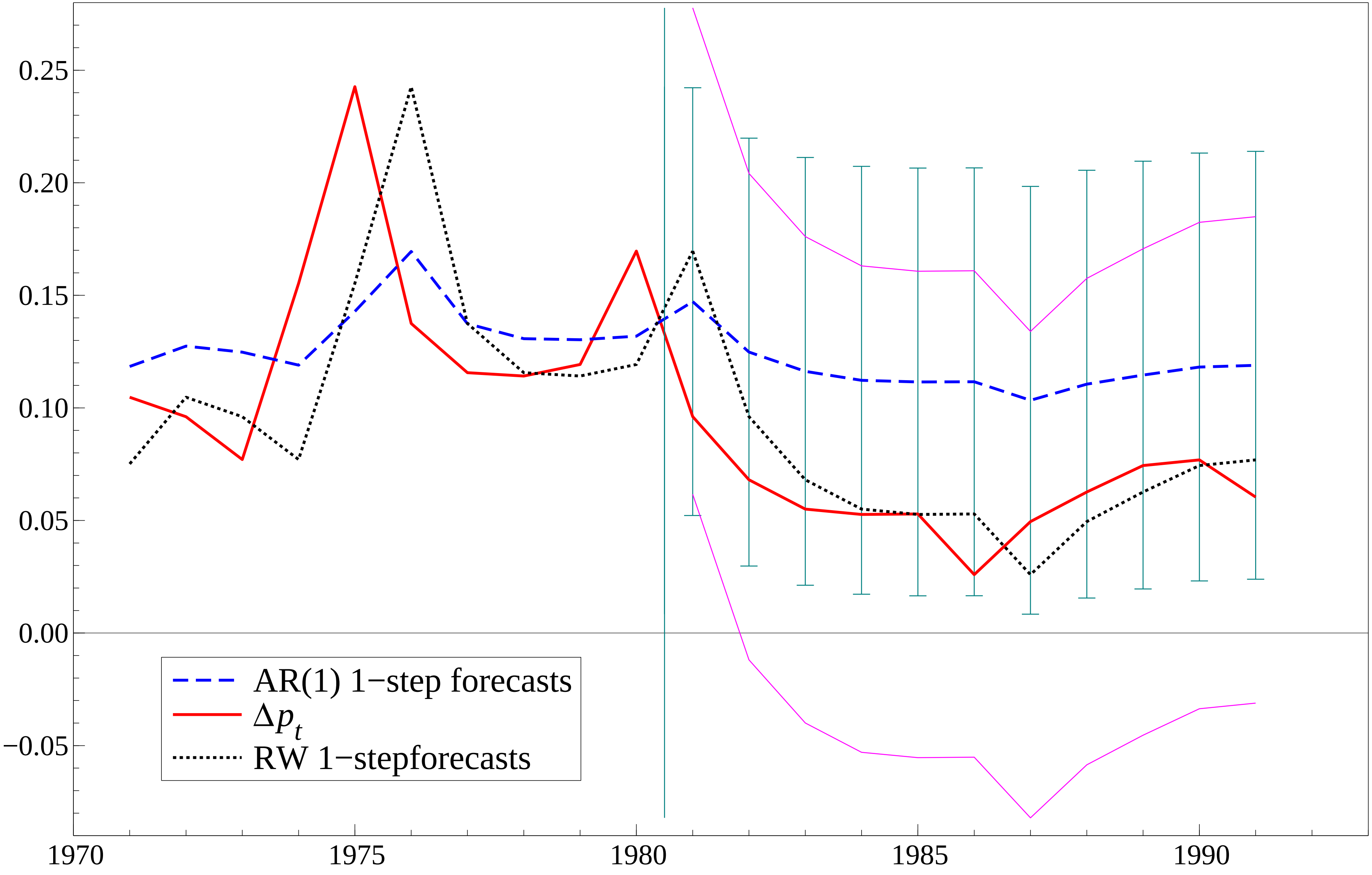
$$\Delta \mathbf{x}_{T+i-1} = \boldsymbol{\gamma}^* + \boldsymbol{\alpha}^* \left((\boldsymbol{\beta}^*)' \mathbf{x}_{T+i-2} - \boldsymbol{\mu}^* \right) + \boldsymbol{\Psi}^* \mathbf{z}_{T+i-1} + \mathbf{v}_{T+i-1}. \tag{5}$$

Thus, $\Delta \mathbf{x}_{T+i-1}$ reflects all the effects needed:
parameter changes; no omitted variables;
with no estimation issues at all.

Two drawbacks:
$\mathbf{v}_{T+i-1}$ doubles innovation error variance;
variables lagged one extra period – adds ‘noise’. **Trade-off**.

Forecast error is $\mathsf{I}(-1)$, so very ‘noisy’,
but no systematic failure.

RW and AR(1) forecasts for inflation



Does not refute causal models

Analogy: rocket to moon predicted to land 4th July but oxygen tank exploded and mission was aborted: forecast is systematically and badly wrong.

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

Macro-economic forecast failure occurs regularly.
Forecast failure depends on forecast-period events:
need not invalidate theory or model,
nor be predictable from in-sample tests;
neither avoided, nor induced, by congruence.

Forecasting breaks

Objectives:

develop methods for forecasting breaks
with
robust strategies if breaks incorrectly predicted

First requires that:

(1) breaks are predictable

(2) we have information relevant to that predictability

(3) such information is available at the 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

Robust strategies

Second builds on considerable recent research:

(7) robust forecasting devices

(8) improved intercept corrections

(9) pooling of forecasts

(10) Also need accurate forecast-error uncertainty measures

(1) Unpredictability of breaks

Role of information analyzed in **Clements and Hendry (2005)**

New formulation with two information sets

which potentially might be very different–

one economics: regular forces from agents' behaviour

other could be politics (say): causes of sudden shifts

No claim that such information actually exists in any given instance, but key to model both if it does

Classic example:
one set of forces that lead to outbreak of civil war

other factors facilitate its continuation–
see (e.g.) **Collier and Hoeffler (2007)**

(2) Relevant information

Depends on which breaks matter

Location shifts are most pernicious:
induce non-stationarity & systematic forecast failure

Theory in **Clements and Hendry (1998, 2006)**

Other breaks of less relevance for forecasting

So seek information relevant to location shifts:

or relevant to ongoing effects as a shift occurs

Forecast failure remains common–and systematic

(3) Available information

Several possibilities:

'leading indicators'–but historical record unimpressive
non-linear functions of variables already in models–**same**

Rapid information updates at forecast origin–
higher frequency data should help

Forecast-error taxonomy for time disaggregation:
higher frequency does not reduce impacts of breaks
see **Castle and Hendry (2008)**
But may detect breaks sooner, so adapt better

So consider information outside usual subject matter

(4) Dectecting non-linearity

Appropriate model form entails non-linear reactions

Portmanteau test for general form of non-linearity

Castle and Hendry (2006)

Low-dimensional, orthogonalized-representation of polynomial functions

Test only needs $2n$ functions for n linear regressors

Power against up to quintics and inverses thereof

Provides basis for general-to-simple approach:
linear model embedded in non-linear general model

(5) Modelling non-linearity

Non-linear model selection–many sub-problems:

(A) specify general form of non-linearity;
– polynomials, exponentials in orthogonalized regressors;

(B) collinearity between non-linear functions;
– double demeaning to remove key collinearity;

(C) non-normality: non-linear functions capturing outliers;
– remove outliers by impulse saturation;

(D) excess numbers of irrelevant variables;
– super-conservative *Gets* strategy;

(E) potentially more variables than observations;
– multi-stage ‘combinatorial selection’;

(F) determine specific form of non-linearity.
– encompassing tests against specific non-linear forms (e.g., ‘ogive’, LSTAR, bilinear, . . .).

(6) Forecast accuracy

Even if location shift is predictable
by available information
embodied in a well-selected non-linear model
problems remain

a] **Breaks alter collinearities** between variables

c] **adverse impact on MSFE if collinearity changes**
despite large increase in information content of data

d] **unavoidable**–deleting collinear variables does not help:
unless they are actually irrelevant;

e] hence **immediate updating** can be crucial

Changing collinearity

Simplest conditional regression DGP:

$$y_t = \boldsymbol{\beta}'\mathbf{z}_t + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{6}$$

with $\mathbf{z}_t$ independent of $\{\epsilon_t\}$:

$$\mathbf{z}_t \sim \mathsf{IN}_n\left[\mathbf{0}, \boldsymbol{\Sigma}\right] \tag{7}$$

for $\boldsymbol{\Sigma} = \mathbf{H}'\boldsymbol{\Lambda}\mathbf{H}$ with $\mathbf{H}'\mathbf{H} = \mathbf{I}_n$.

1-step MSFE for known regressors from (6):

$$\mathsf{E}\left[\widehat{\epsilon}^2_{T+1|T}\right] = \sigma_\epsilon^2\left(1 + \sum_{i=1}^{n}\frac{\lambda_i^*}{T\lambda_i}\right) \tag{8}$$

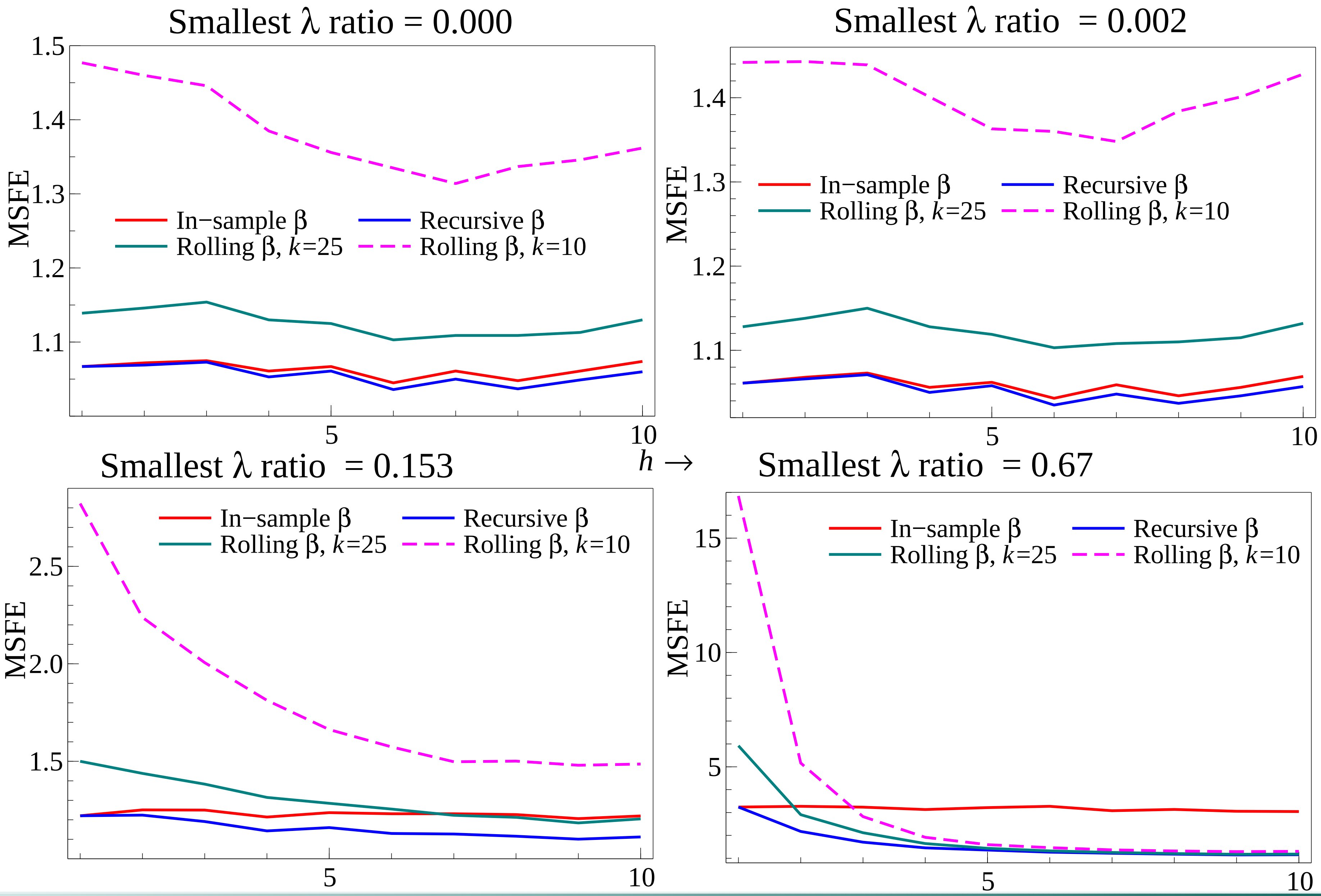
where $\mathsf{E}\left[\mathbf{z}_{T+1}\mathbf{z}'_{T+1}\right] = \boldsymbol{\Sigma}^* = \mathbf{H}'\boldsymbol{\Lambda}^*\mathbf{H}$.

If $\widehat{\boldsymbol{\beta}}_{(T)}$ retained, (8) continues to hold. **But with updating**:

$$\mathsf{E}\left[\widehat{\epsilon}^2_{T+2|T+1}\right] = \sigma_\epsilon^2\left(1 + \sum_{i=1}^{n}\frac{\lambda_i^*}{T\lambda_i + \lambda_i^*}\right) \tag{9}$$

Reduction depends on smallest eigenvalue ratio

Impact of breaks in collinearity



(7) Insurance policies

Use robust forecasting devices
'Insurance' after a break to mitigate systematic failure
Hendry (2006): **explanation for success of naive devices**

(8) **Improved intercept corrections**
'Set on track' at the forecast origin, while smoothing recent corrections: **Hendry and Reade (2006)**

(9) **Pooling of forecasts**
'Model averaging' can go seriously wrong, but improved by Gets model selection : **Hendry and Reade (2004)**

(10) Accurate forecast-error uncertainty measures:
on-going research

Conclusions

Despite weak assumptions of non-stationary economy,
subject to unanticipated structural breaks,
model differs from DGP in unknown ways,
selected and estimated from unreliable data,
can derive many useful insights.

Econometric systems should outperform—but do not.
Causal information swamped by unmodelled breaks.
Strategy: retain former yet avoid systematic failure.

Surprisingly:
poor methods; bad models; inaccurate data; and data-based selection
not primary causes of systematic mistakes.
Main causes are unanticipated large changes
affecting forecast period.

Conclusions

Whether breaks are predictable from relevant information available at the forecast origin remains unknown as yet.
But progress in developing forecasting models;
and methods of testing for and selecting such models.

Predictability theory: 2 information sets, regular and shifts; model latter as non-linear ogive.

Considerable progress since:
“The only function of economic forecasting is to make astrology look respectable”
John Kenneth Galbraith

Bibliography

References

Castle, J. L., and Hendry, D. F. (2005). Extending the boundaries of automatic Gets to non-linear models. Mimeo, Economics Department, Oxford University.

Castle, J. L., and Hendry, D. F. (2006). A low-dimension, collinearity-robust test for non-linearity. Discussion paper 326, Economics Department, Oxford University.

Castle, J. L., and Hendry, D. F. (2008). Forecasting UK inflation: Empirical evidence on robust forecasting devices. In Rapach, D. E., and Wohar, M. E. (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, pp. 41–92. Amsterdam: Elsevier. Forthcoming.

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

Clements, M. P., and Hendry, D. F. (eds.)(2002). *A Companion to Economic Forecasting*. Oxford: Blackwells.

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

Clements, M. P., and Hendry, D. F. (2006). Forecasting with breaks in data processes. In Elliott, G., Granger, C. W. J., and Timmermann, A. (eds.), *Handbook of Econometrics on Forecasting*, pp. 605–657. Amsterdam: Elsevier.

Collier, P., and Hoeffler, A. (2007). Civil war. Chapter 22, *Handbook of Defense Economics*, T. Sandler and K. Hartley (eds), Elsevier.

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

Hendry, D. F., and Reade, J. J. (2004). Problems with model averaging, and a better alternativ