Conference on

Forecasting and Monetary Policy

Berlin, 23-24 March 2009

Christian Schumacher
Deutsche Bundesbank

Discussion of „Forecasting Inflation with graduat regime shifts and exogenous information“

Comments on: Forecasting inflation with gradual regime shifts and exogenous information

Christian Schumacher
Deutsche Bundesbank

23 Mar 2009

1 Introduction

'Great Moderation': in the past decades, inflation rates have gone down in many industrialised countries

linear, time-constant models have difficulties matching this stylized fact

González, Hubrich and Teräsvirta (GHT henceforth) employ a **non-linear shifting-mean (SM) model to forecast inflation** in the Euro area (EA), UK and US

$\Longrightarrow$ high relevance for central banks

2 Model

Shifting-mean model (SM): inflation y_t is a function of time t plus noise

$$y_t = \delta(t) + \varepsilon_t \tag{1}$$

$$\delta(t) = \delta_0 + \sum_i \delta_i \frac{1}{1 + e^{-\gamma_i(t/T - c_i)}} \tag{2}$$

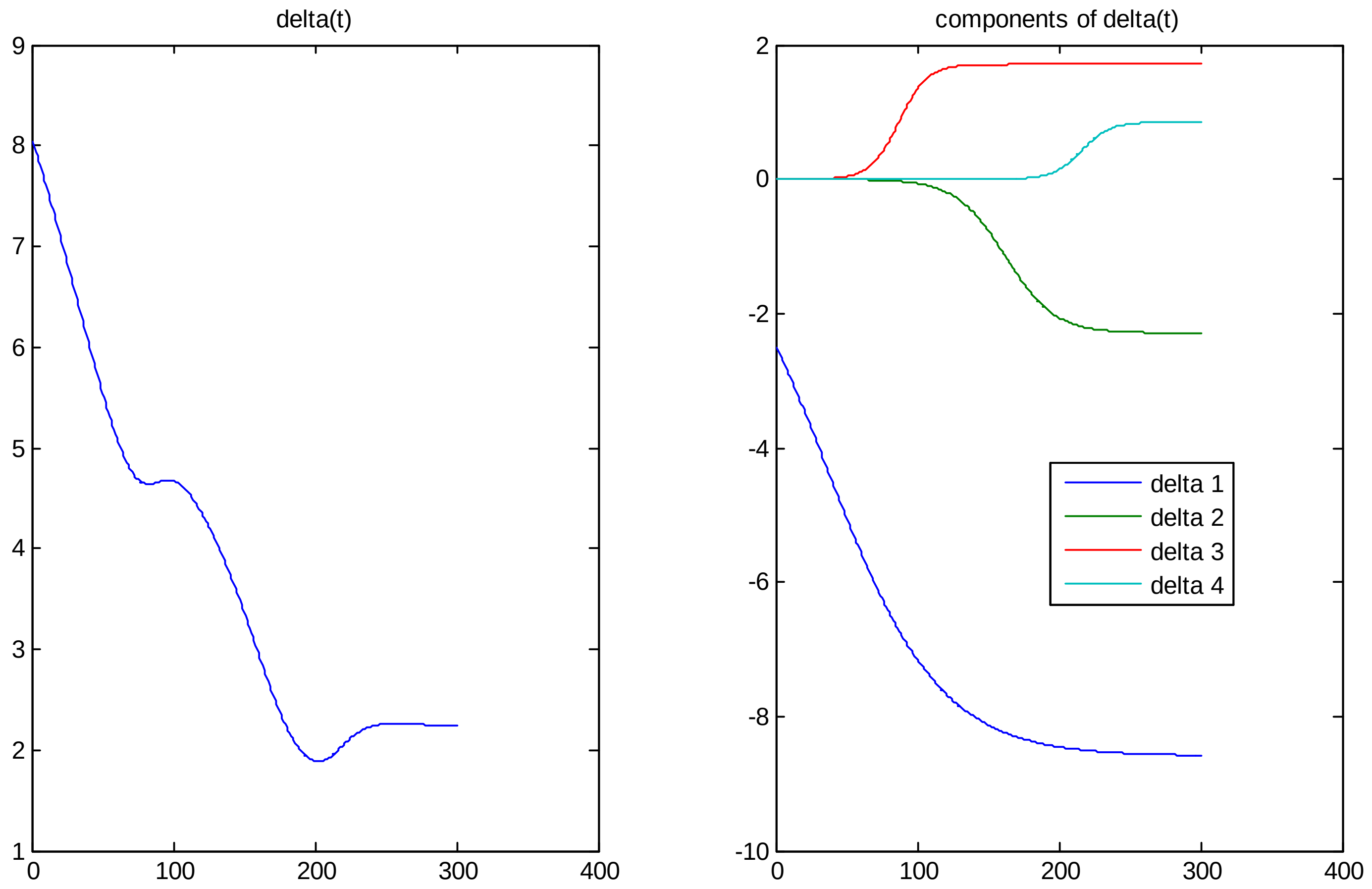
in the long run,

$$\lim_{t \to \infty} \delta(t) = \delta_0 + \sum_i \delta_i \tag{3}$$

holds, so model converges to steady state

$\Longrightarrow$ nice property compared to other non-linear functions (multiples of t e.g.)!

Figure 1: Euro area example



3 Forecasting with and without an inflation target

3.1 With inflation target

estimate model with exogenous information according to loss function

$$L = \sum_{t=1}^{T} l_t(\cdot|\cdot) - \lambda \sum_{t=T+1}^{T+\tau} \rho^{T+\tau-t} \{\delta(t) - x\}^2 \tag{4}$$

for trend forecast $\delta(t)$, horizons $t = T+1, \dots, T+\tau$ and $0 < \rho < 1$

x is a forecast equal to the central bank's **inflation target**

$\Longrightarrow$ model forecasts $\delta(T+\tau)$ are driven towards the exogenous forecast x in the long run as $\tau \to \infty$

$\Longrightarrow$ might be **useful in terms of credibility/transparency** if forecasts are published!

3.2 Forecasting with pure SM

long-run forecast of SM model without exogenous information

$$\lim_{\tau \to \infty} y_{T+\tau|T} = \lim_{\tau \to \infty} \delta(t) = \delta_0 + \sum_i \delta_i \tag{5}$$

in the empirical example for EA

$$\delta_0 + \sum_i \delta_i = 10.55 - 8.59 - 2.29 + 1.71 + 0.85 = 2.23 \tag{6}$$

$\Longrightarrow$ this is **not too far away from the target range of the Eurosystem** ($< 2.0\%$), thus shifting the pure model forecast to the target (as in GHT and section 3.1) doesn't matter much however, in case of relatively persistent deviations of inflation from target range, pinning down the forecast to target might deteriorate forecast a lot (GHT, p. 21)

this leads to a trade-off:

on the one hand: pinning down the forecast to target **increases credibility and transparency wrt the inflation target** of the central bank

on the other hand: in terms of mean-squared forecast error, **target can worsen the forecast record of the central bank, if target is too far away from the data**

3.3 Density forecasts with inflation target

depending on penalty parameter λ, optimising

$$L = \sum_{t=1}^{T} l_t(\cdot|\cdot) - \lambda \sum_{t=T+1}^{T+\tau} \rho^{T+\tau-t} \{\delta(t) - x\}^2 \quad (7)$$

can lead to **bimodal forecast densities**, as model forecast and inflation target forecast can be locally dominating (GHT, pp. 16-17, 19-20)

$\Longrightarrow$ due to communication difficulties, **policy makers could be reluctant to publish these forecast densities** (people perhaps think in a "mean-unimodal"-oriented way)

$\Longrightarrow$ what does bimodalitiy imply for optimal monetary policy? Should policy-makers accept the ambiguity in models like these?

4 Forecasting with SM and discounted and penalised regression

again, loss function used for estimating forecast version of SM

$$L = \sum_{t=1}^{T} l_t(\cdot|\cdot) - \lambda \sum_{t=T+1}^{T+\tau} \rho^{T+\tau-t} \{\delta(t) - x\}^2 \tag{8}$$

weights evolve according to

$$\begin{array}{l|lllll} t & T+1 & T+2 & \ldots & T+\tau-1 & T+\tau \\ \hline \rho^{T+\tau-t} & \rho^{\tau-1} & \rho^{\tau-2} & \ldots & \rho^{1} & \rho^{0}=1 \end{array} \tag{9}$$

given $0 < \rho < 1$, implying increasing weights $\rho^{T+\tau-t}$ if horizon τ increases: $\rho^{\tau-1} < \rho^{\tau-2} < \ldots < \rho^1 < \rho^0 = 1$

$\Longrightarrow$ trend forecast $\delta(t)$ receives a lot of weight for short horizons, x little however, GHT motivate the use of external information x by stating:

> "[...] SM-AR model may [...] also be used for forecasting. Nevertheless, it may suffer from the same problem as autoregressive models with a linear trend, namely, that extrapolating the deterministic component **may not yield satisfactory short-term forecasts**. [...] " (p. 11)

$\Longrightarrow$ this is a contradicton to the estimation procedure, where pure SM forecasts receive a lot of weight relative to the exogenous forecast in the short term

5 Recursive performance of SM at the end of the sample

GHT state:

> "[...] It should be pointed out, however, that the SM-AR model is **not a feasible tool for very short-term forecasting** because of its lack of adaptability. We provide a timely **measure of medium-term inflation** based on a single time series. [...] " (p. 6)

as forecasts are often serially dependent over the forecast horizons, it cannot be ruled out that **short-term and long-term forecastability are not independent** of each other question: how does selection and estimation of components (Quickshift with SM) work in real time? paper provides evidence for stable time periods in the EA and UK

to check the adaptability of SM, other interesting **samples chosen for forecast exercise** could include break, e.g. in the UK 1989-1992

5.1 Simulation exercise with recursive estimation of SM

stylized ‘Great Moderation’ DGP plus shift

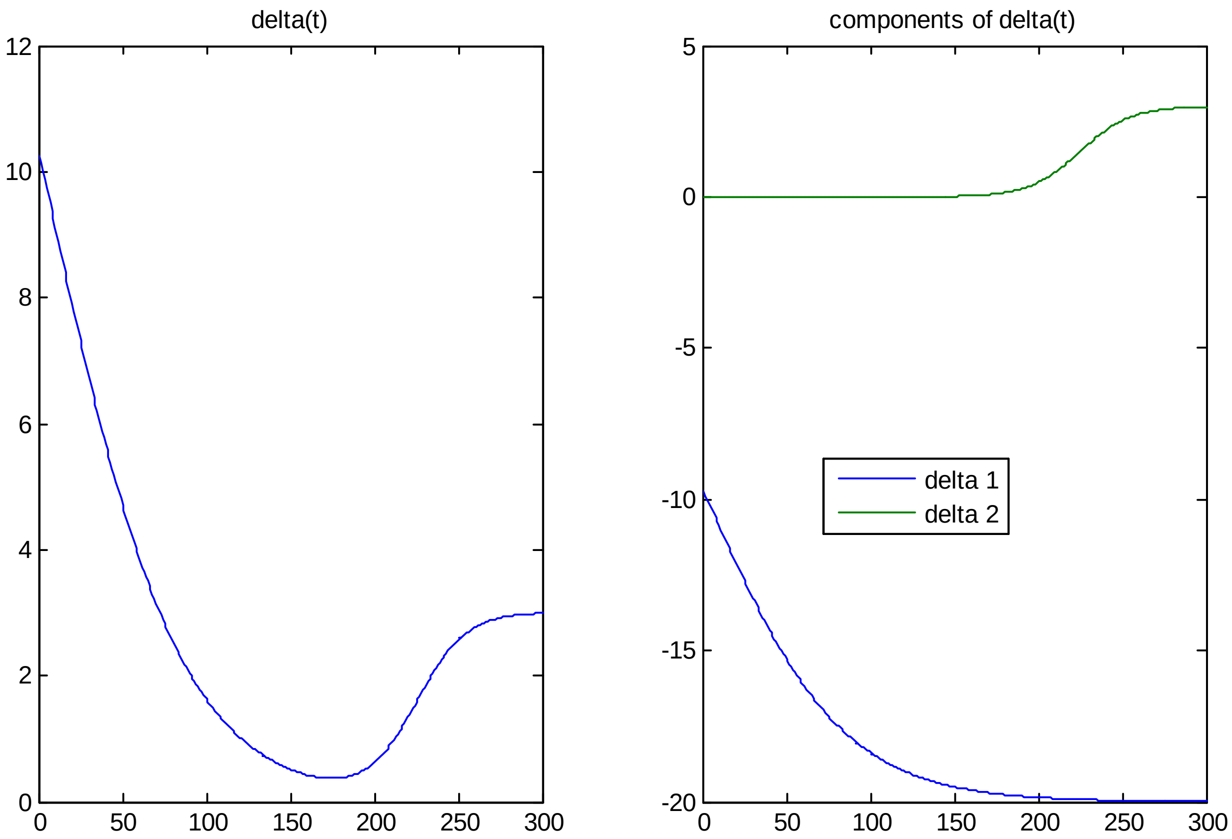
$$y_t = \delta(t) + \varepsilon_t \tag{10}$$

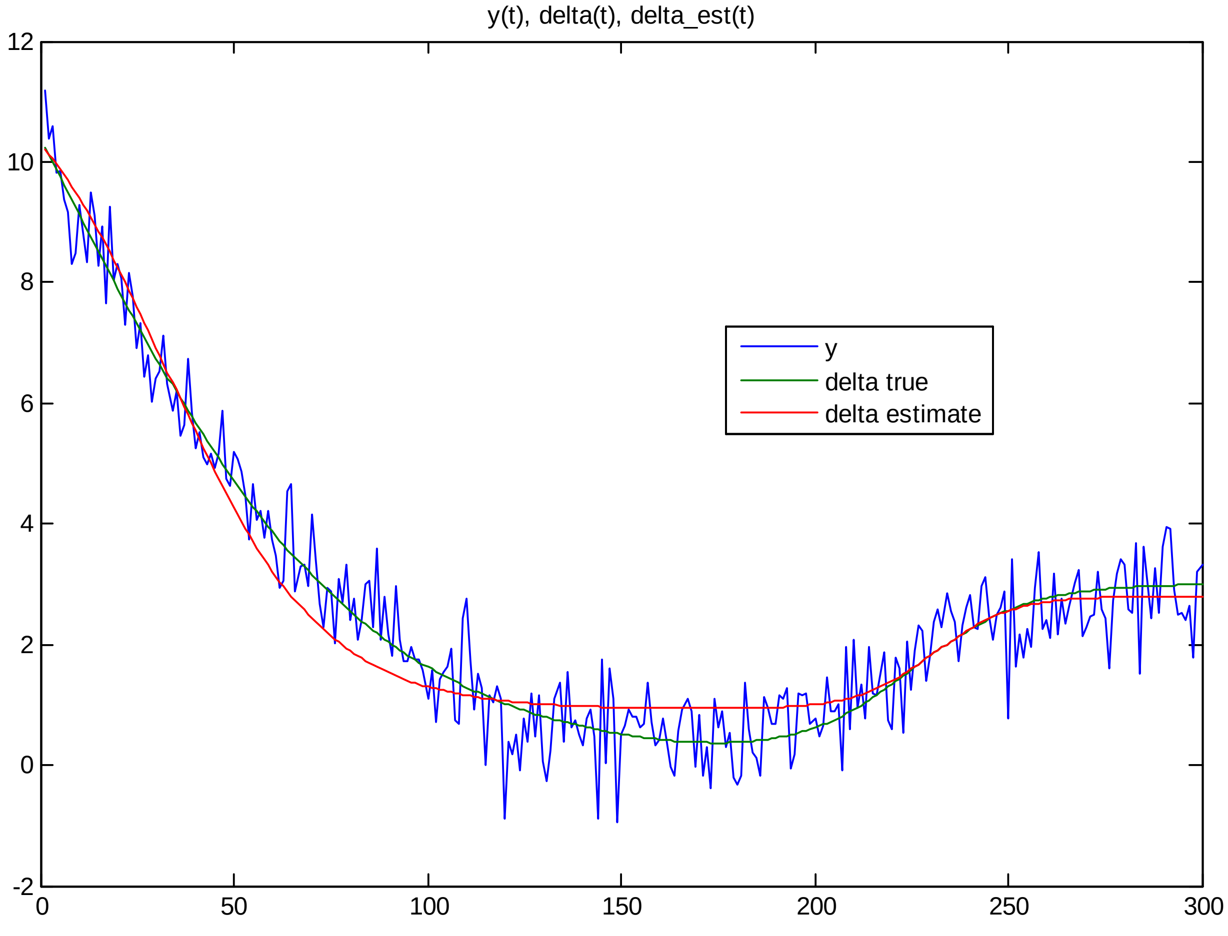
$$\delta(t) = 20.0 - 20.0 \times \frac{1}{1 + e^{-7.5(t/T - 0.01)}} + 3.0 \times \frac{1}{1 + e^{-20.0(t/T - 0.75)}} \tag{11}$$

$$\varepsilon_t \sim N(0, 0.3) \tag{12}$$

sample: $T = 300$ (months, similar to GHT)

Figure 2: DGP with great moderation plus shift





y(t), delta(t), delta_est(t)
y
delta true
delta estimate
12
10
8
6
4
2
0
-2
0
50
100
150
200
250
300

given the simulated trend $\delta(t)$ and data, I investigate **how well recursive estimation works**

I **estimate with one or two components** of the trend

in the DGP, we have two components: 1) the long-run trend over the whole sample, 2) the shift active in the final quarter of the sample period

5.2 Results

Figure 3: SM estimation with one component

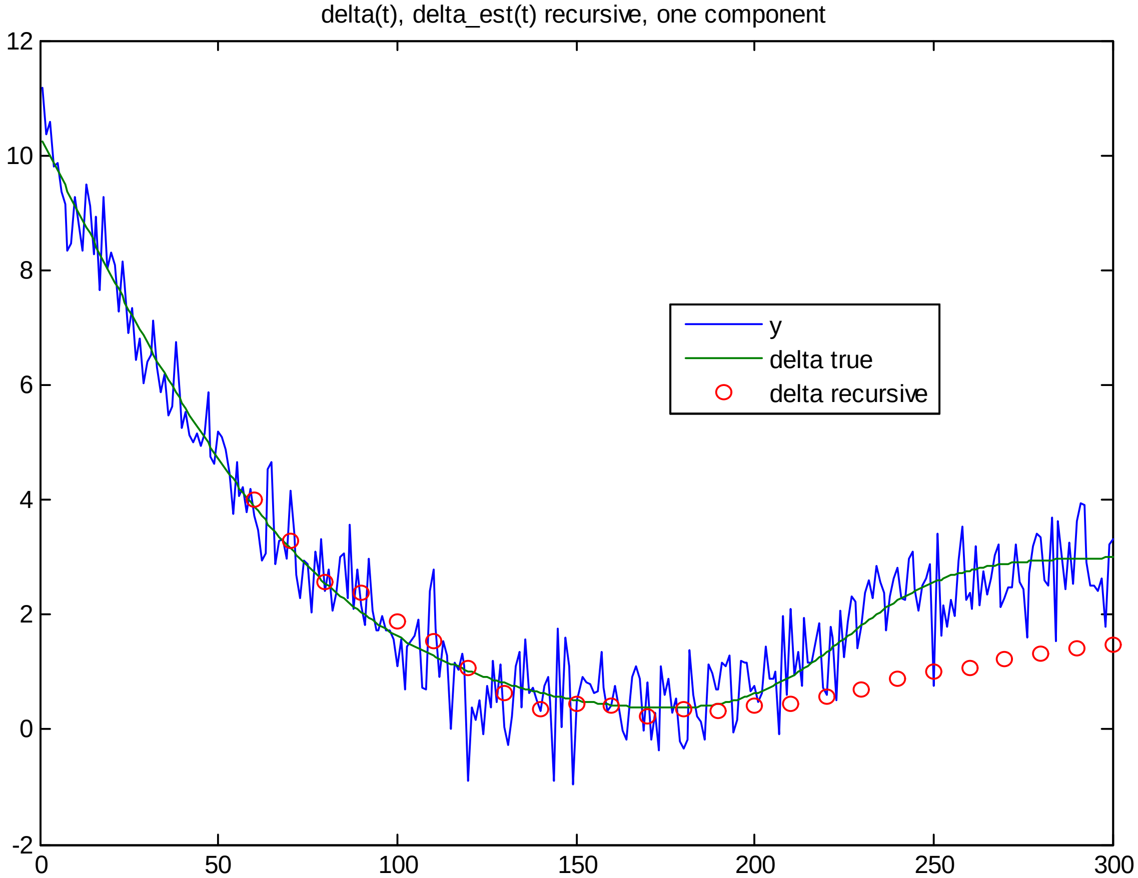
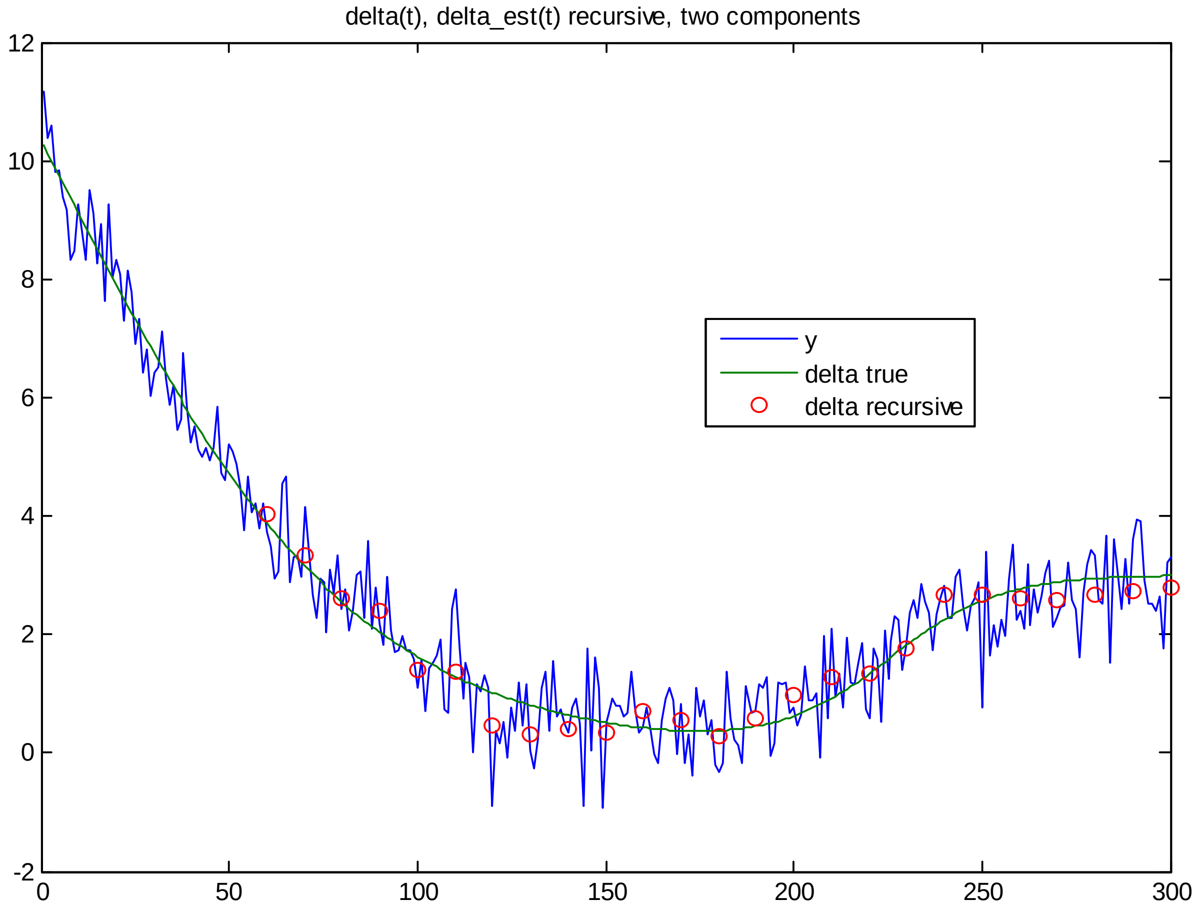


Figure 4: SM estimation with two components



summary of results: **two components work relatively well** also in periods, where only one component is active in the DGP (the trend)

one component alone captures only the long-run component, but fails in periods where two components are active in the DGP

$\Longrightarrow$ conditional on the proper choice of the number of components, **adaptability is not bad in the SM model**

$\Longrightarrow$ as the end of the sample is very important for policy-makers, this could be investigated in more depth, also relative to other methods

6 Minor issues

taking into account long-run information with **penalised regression is applicable to any other regression approaches**, not only to SM; might help to pin down long-run forecasts of other models currently used in central banks

example: in Bayesian VAR framework, Sveriges Riksbank centers forecasts around inflation target by choosing priors appropriately

the **intercept model** $y_t = \delta_0 + \varepsilon_t$, estimated over a short rolling window, can be a strong competitor in forecast comparisons in the Great Moderation era, and works better than the AR model (due to instabilities in the AR coefs and variance of ε_t)

there is a **corner solution** for γ_3, γ_4: intitial values (grids) are in the range $0.01 \leq \gamma_3, \gamma_4 \leq 30$, estimators for EA are $\widehat{\gamma}_3, \widehat{\gamma}_4 = 30$, and more speed of convergence ($\gamma_3, \gamma_4 > 30$) could improve the fit

Thank you!