Forecasting Euro Area Inflation with the Phillips Curve. An evaluation over 25 years

Marta Bańbura, Elena Bobeica*
European Central Bank

2nd Conference on Forecasting at Central Banks,
Bank of England
15-16 November 2018

* The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Purpose

Study the forecast performance of various Phillips curve specifications for the *euro area* starting the early-90s.

Analyse how it has changed over time (relative to naïve benchmarks)

Investigate the role of:

- Slack measures
- External variables
- Inflation trends
- Time variation in the coefficients
- Functional form

On the episodic performance of the Phillips curve

It would be extraordinarily useful to discover a specification of the Phillips curve that fits the data reliably ... as Stock and Watson (2010) observe, the history of the Phillips curve 'is one of apparently stable relationships falling apart upon publication.' Ball and Mazumder (2011) is a poignant example. Nonetheless, because of the practical importance of the Phillips curve, researchers must continue to search for better specifications.

from [Ball and Mazumder, 2014]

Motivation

After the Great Recession, there was a revival of interest in the relevance of the Phillips curves, as this key relationship seemed to break down

Puzzling nature of inflation and *Large forecast errors* led to:

A quest for the 'right' specification:

- Alternative measures of slack (short term unemployment [Gordon, 2013], unemployment recession gap [Stock and Watson, 2010], broad labour underutilisation measures [Bell and Blanchflower, 2018])
- Time variation in the PC slope ([IMF, 2013], [Blanchard et al., 2015])
- Oil via inflation expectations ([Coibion and Gorodnichenko, 2015], [Hasenzagl et al., 2018])
- Shifting inflation end point ([Kozicki and Tinsley, 2001], [Clark and Doh, 2014])

No systematic investigation of how all these proposals fare in forecasting.

Overview of the results

- Relative performance compared to naïve benchmarks

  (Slightly) better relative performance after the onset of the EMU but worse more recently.

- Slack measures

  Principal components and output gaps work OK compared to alternatives. Including financial variables in a composite measure of slack does not seem to help.

- External variables

  Do not seem to help much (out-of-sample)

- Inflation trends

  Important to account for changing inflation trend in the run-up to the EMU. More recently it appears less relevant. Although small improvement can be made by e.g. using long term survey expectations.

Overview of the results

- Time variation in the coefficients

  Occasionally helps

- Functional form

  Does not appear to be a key factor.

Literature

Vast literature for the US

See e.g. [Fuhrer et al., 2009]. On forecasting inflation:

[Stock and Watson, 2009], [Faust and Wright, 2013], [Chan et al., 2016],[Dotsey et al., 2018],[Hasenzagl et al., 2018]

Few studies on the euro area on the forecasting performance

[Rünstler, 2002], [Hubrich, 2005], [Canova, 2007], [Marcellino and Musso, 2010], [Buelens, 2012]. [Bereau et al., 2018], [Lenza and Jarocinski, 2016], [Moretti et al., 2018]

Often with a different focus, smaller evaluation sample.

Phillips curve, autoregressive distributed lag model
Benchmark model

We forecast annualised h-period inflation rate:

$$\pi_t^h = \frac{400}{h} \ln\left(\frac{P_t}{P_{t-h}}\right) = \frac{1}{h}\sum_{i=0}^{h-1} \pi_{t-i}$$

Direct forecast from:

$$\tilde{\pi}_{t+h}^h = \alpha(L)\tilde{\pi}_t + \beta_h(L)y_t + \gamma_h(L)z_t + \nu_{t+h}^h$$

- $\tilde{\pi}_t$: de-meaned (in the benchmark model) or de-trended inflation rate
- y_t: a measure of slack (de-meaned or gap)
- z_t: “external” variable (de-meaned)

Out-of-sample forecast evaluation

- Focus on *quarterly HICP excluding energy*
- Focus on 1-year-ahead forecast
- Disregarding data revisions
- Evaluations based on RMSFE
- Evaluation sample: 1994-2017

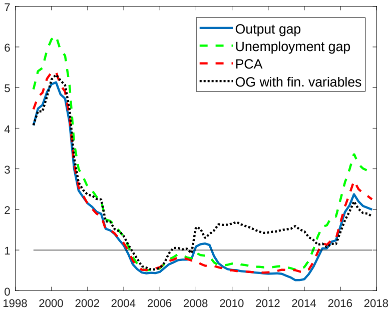
  Rolling RMSFEs and over sub-samples 1994-2000, 2001-08, 2009-17
- Rolling window estimation.

  Models allowing for time variation in the parameters have been estimated recursively.
- Benchmark models
  - Random walk (RW) of [Atkeson and Ohanian, 2001]: $\pi^h_{t+h|h} = \pi^4_t$
  - Unobserved component trend-cycle model with stochastic volatility (UCSV) of [Stock and Watson, 2007]

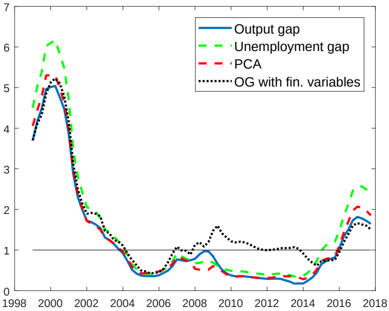
Benchmark results

Relative rolling RMSFE

Benchmark: UCSV



Benchmark: RW

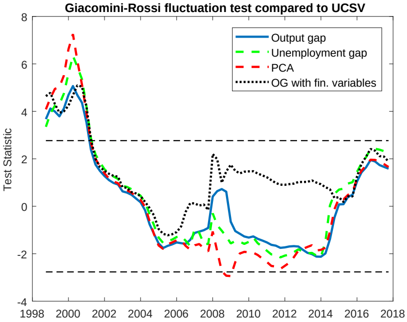


Note: The RMSFEs have been computed over a rolling window of 20 quarters.

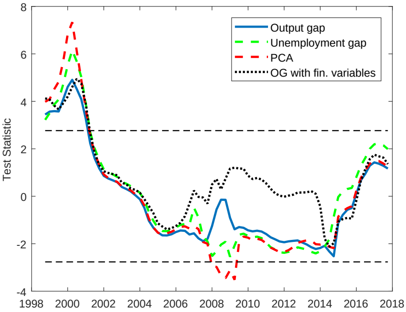
Benchmark results

The fluctuation test

Benchmark: UCSV



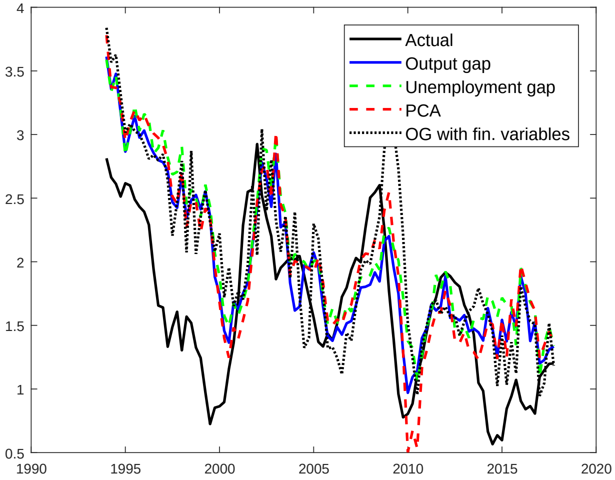
Benchmark: RW



Note: [Giacomini and Rossi, 2010] fluctuation test for a rolling window of 20 quarters. Dashed line shows the critical values.

Benchmark results

Actual vs forecast



Measures of slack

- Product market:

  Output gap

  GDP growth, Industrial production growth, Industrial production gap, Capacity utilisation, Economic sentiment

- Labour market:

  Unemployment gap

  Unemployment rate, Unemployment recession gap [Stock and Watson, 2010], Employment growth, Employment gap, Short-term unemployment rate, U6 unemployment rate, U6 unemployment gap

- Composite:

  (First) principal component of all variables (PCA)

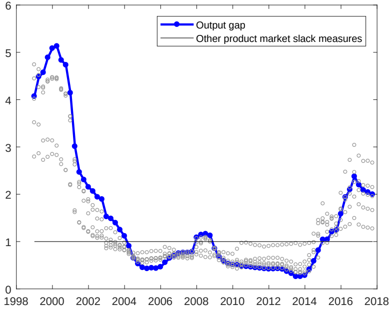
  Output gap with financial measures

Gaps have been computed recursively using the Christiano and Fitzgerald filter.

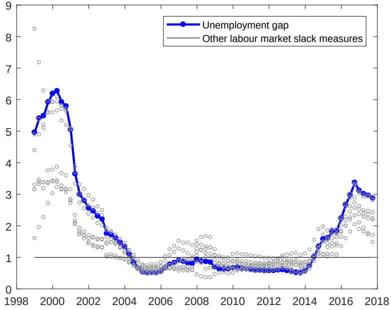
Measures of slack

Relative rolling RMSFE

Slack in product market



Slack in labour market

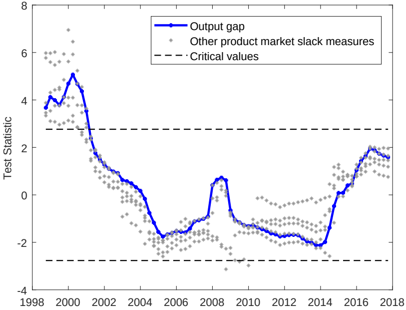


Note: The RMSFEs have been computed over a rolling window of 20 quarters. '*' represent the RMSFEs of models with other measures of slack from a given group. Benchmark: UCSV.

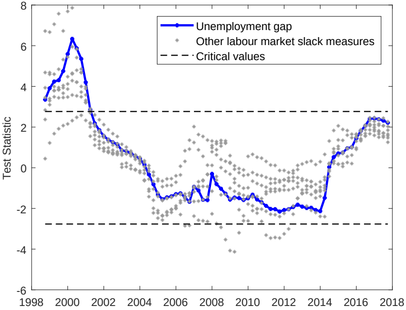
Measures of slack

The fluctuation test

Slack in the product market



Slack in the labour market



Note: Rolling window of 20 quarters. '*' represent the RMSFEs of models with other measures of slack from a given group. Dashed line shows the critical values. Benchmark: UCSV.

External variables

- Import prices and PPI:

  Import deflator
  Import deflator from outside the euro area
  Producer Price Index for total industry less construction
  US CPI

- Commodity prices:

  Oil price in EUR
  ECB commodity price index in USD (total non-energy)

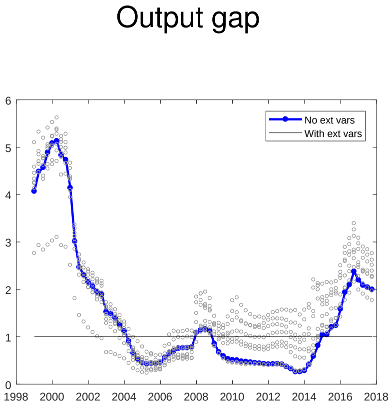
- Exchange rates:

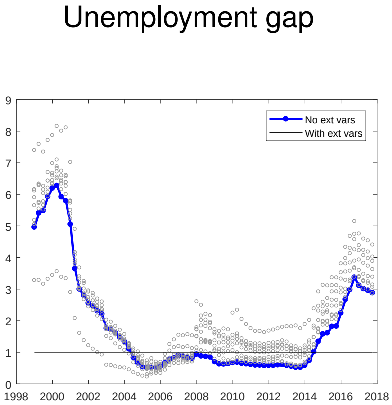
  EER12
  EUR/USD

[Bereau et al., 2018]: HICPX inflation forecast in the euro area cannot be improved by considering global factors

External variables

Relative rolling RMSFE





Note: The RMSFEs have been computed over a rolling window of 20 quarters. '*' represent the RMSFEs of models with different external variables. Benchmark: UCSV.

De-trending

Forecast gains when accounting for a long-term inflation trend

(see e.g., [Clark and McCracken, 2010], [Faust and Wright, 2013], [Clark and Doh, 2014], [Zaman, 2013])

In the Phillips curve inflation we take $\tilde{\pi}_t = \pi_t - \pi_t^{TR}$ with π_t^{TR}:*

- Long-term inflation expectations from surveys

  6-10 year inflation forecasts from Consensus Economics.

  Available for the euro area as of 2003; backdated to 1990 using the forecast for the largest euro area countries (see [Castelnuovo et al., 2003])

- UCSV
- Exponentially weighted moving average

  Equivalent to UCSV with constant ratio of variances of temporary and permanent shocks.
- [Stock and Watson, 1999] unit root formulation

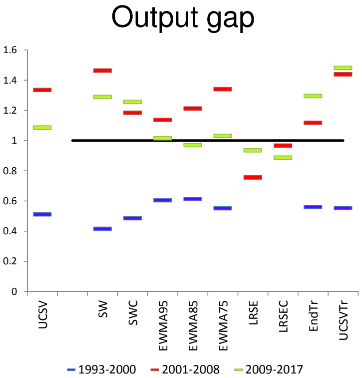
  For one-quarter-ahead horizon equivalent to taking $\pi_t^{TR} = \pi_{t-1}$
- Endogenous trend

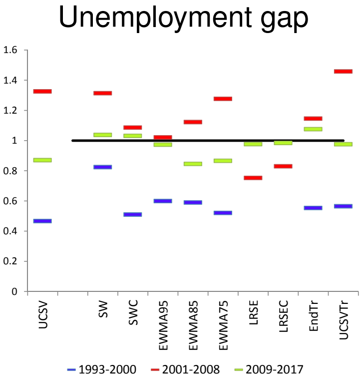
  $\pi_t^{TR} = \pi_{t-1}^{TR} + u_t$ estimated simultaneously with the coefficients of the Phillips curve

* related to the shifting endpoint concept of [Kozicki and Tinsley, 2001]; see also [Clark and Doh, 2014].

De-trending

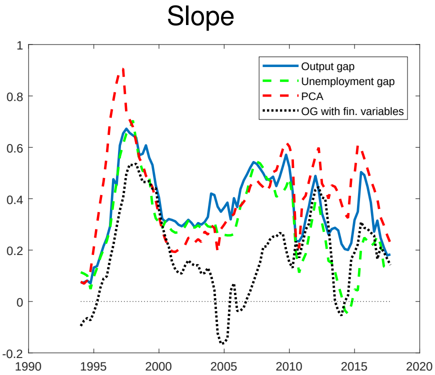
Relative RMSFE





Note: The RMSFEs are relative to the model with constant mean (trend).

Slope



Note: The slope chart shows the sum of slack coefficients with a rolling window of 40Q

Considered TVP models

- á la [Primiceri, 2005]

  $\tilde{\pi}_t = \pi_t$, intercept is included in the equation and all coefficients follow random walks

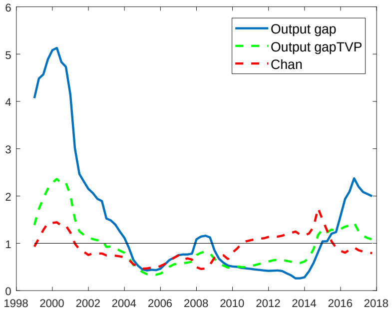
- [Chan et al., 2016]

  Inflation trend, natural rate of unemployment and coefficients follow random walks (bounded for inflation trend)

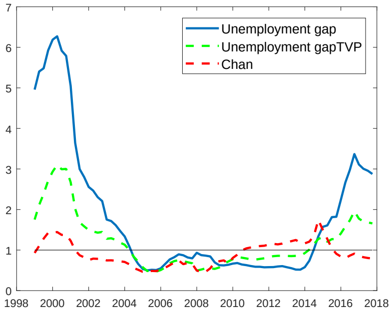
Fixed versus TV parameters

Relative rolling RMSFE

Slack in the product market



Slack in the labour market



Note: The RMSFEs have been computed over a rolling window of 20 quarters.

Functional form

We compare the direct forecasts from the ADL model presented above with indirect (iterated) forecasts from:

- ADL model for one-quarter-ahead inflation

$$\tilde{\pi}_{t+1} = \alpha(L)\tilde{\pi}_t + \beta(L)y_t + \gamma(L)z_t + \nu_{t+1}$$

  Forecasts for the explanatory variables are obtained from univariate AR models.

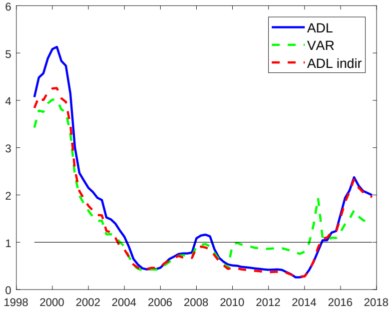
- VAR

$$X_t = \mu + \Phi(L)X_{t-1} + \nu_t\,, \qquad X_t = [\tilde{\pi}_t\ y_t\ z_t]'$$

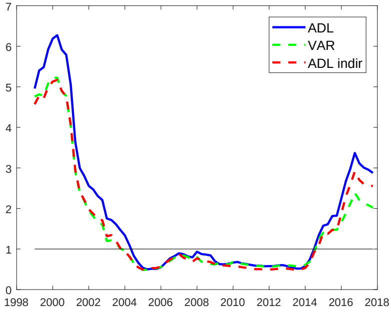
See e.g. [Hubrich, 2005], [Garratt et al., 2010], [Benkovskis et al., 2011], [Clark and Doh, 2014].

Functional form

Slack in the product market



Slack in the labour market



Note: The RMSFEs have been computed over a rolling window of 20 quarters.

Further work

- Parameter shrinkage

  AIC sometimes selects many lags and the coefficients are not reliably estimated on a relatively short estimation window. Particularly problematic for the VAR versions.

- Forecast combinations

  Relative performance changes over time.

- Density forecasts (combinations)

  Could also improve point forecasts (e.g. [Aastveit et al., 2014])

- Conditional forecasts

  Scenarios often considered at central banks and other institutions (e.g. inflation forecasts conditional on certain path for output gap or commodity prices).

Summary of the results

- In general, univariate benchmarks are hard to beat
- Euro area inflation was particularly hard to forecast in the run-up to the EMU and in the aftermath of the twin recessions
- The performance of the Phillips curve is episodic also for the euro area. Some specifications manage to improve upon the UCSV but the relative performance of various models changes.
- Among all “dimensions” (slack / external variables / de-trending / time variation / functional form), the most important seems to be de-trending/time variation (at the beginning of the sample). The latter deserves further research.
- Among all the exogenous slack measures, the output gap and the PCA perform well.

Finding the Phillips curve is like finding a needle in a haystack. But it is hidden there somewhere!

[Reichlin, 2018]

THANK YOU

Aastveit, K. A., Gerdrup, K. R., Jore, A. S., and Thorsrud, L. A. (2014).
Nowcasting gdp in real-time: A density combination approach.
Journal of Business and Economic Statistics, 32(1):48–68.

Atkeson, A. and Ohanian, L. E. (2001).
Are Phillips Curves Useful for Forecasting Inflation.
Federal Reserve Bank of Minneapolis Quarterly Review, 25(1):2–11.

Ball, L. and Mazumder, S. (2014).
A Phillips Curve with Anchored Expectations and Short-Term Unemployment.
NBER Working Papers 20715, National Bureau of Economic Research, Inc.

Bell, D. N. and Blanchflower, D. G. (2018).
Underemployment in the US and Europe.
NBER Working Papers 24927, National Bureau of Economic Research, Inc.

Benkovskis, K., Caivano, M., D'Agostino, A., Dieppe, A., Hurtado, S., Karlsson, T., Ortega, E., and Várnai, T. (2011).
Assessing the Sensitivity of Inflation to Economic Activity.
Working Paper Series 1357, European Central Bank.

Bereau, S., Faubert, V., and Schmidt, K. (2018).
Explaining and forecasting euro area inflation: the role of domestic and global factors.
Working Paper 663, Banque de France.

Blanchard, O., Cerutti, E., and Summers, L. H. (2015).
Inflation and activity: Two explorations and their monetary policy implications.
Peterson Institute for International Economics, Working Paper WP15-19.

Buelens, C. (2012).
Inflation Forecasting and the Crisis: Assessing the Performance of Different Forecasting Models and Methods.
European Comission Economic Papers, (451).

Canova, F. (2007).
G-7 Inflation Forecasts: Random Walk, Phillips Curve, or What Else?
Macroeconomic Dynamics, 11(1):1–30.

Castelnuovo, E., Nicoletti-Altimari, S., and Rodríguez-Palenzuela, D. (2003).
Definition of Price Stability, Range and Point Inflation Targets: The Anchoring of Long-Term Inflation Expectations.
Working Paper Series 273, European Central Bank.

Chan, J. C. C., Koop, G., and Potter, S. M. (2016).
A Bounded Model of Time Variation in Trend Inflation, Nairu and the Phillips Curve.
Journal of Applied Econometrics, 31(3):551–565.

Clark, T. E. and Doh, T. (2014).
Evaluating alternative models of trend inflation.
International Journal of Forecasting, 30(3):426–448.

Clark, T. E. and McCracken, M. W. (2010).
Averaging Forecasts from VARs with Uncertain Instabilities.
Journal of Applied Econometrics, 25(1):5–29.

Coibion, O. and Gorodnichenko, Y. (2015).
Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation.
American Economic Journal: Macroeconomics, 7(1):197–232.

Dotsey, M., Fujita, S., and Stark, T. (2018).
Do Phillips Curves Conditionally Help to Forecast Inflation?
International Journal of Central Banking, 14(4):43–92.

Faust, J. and Wright, J. H. (2013).
Forecasting Inflation, volume 2 of *Handbook of Economic Forecasting*.

Elsevier.
pages 2-56.

Fuhrer, J., Kodrzycki, Y., Little, J. S., and Olivei, G. (2009).
Understanding Inflation and the Implications for Monetary Policy, a Phillips Curve Retrospective.
The MIT Press.

Garratt, A., Mitchell, J., and Vahey, S. (2010).
Measuring Output Gap Uncertainty.
CEPR Discussion Papers 7742.

Giacomini, R. and Rossi, B. (2010).
Forecast comparisons in unstable environments.
Journal of Applied Econometrics, 25(4):595–620.

Gordon, R. (2013).
The phillips curve is alive and well: Inflation and the nairu during the slow recovery.
Working Paper Series 19390, National Bureau of Economic Research.

Hasenzagl, T., Pellegrino, F., Reichlin, L., and Ricco, G. (2018).
A Model of the Fed's View on Inflation.
CEPR Discussion Papers 12564, C.E.P.R. Discussion Papers.

Hubrich, K. (2005).
Forecasting Euro Area Inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy?
International Journal of Forecasting, 21(1):119–136.

IMF (2013).
Chapter 3, The dog that didn't bark: has inflation been muzzled or was it just sleeping?
In *World Economic Outlook, April*, pages 1–17. International Monetary Fund, Washington.

Kozicki, S. and Tinsley, P. A. (2001).
Shifting endpoints in the term structure of interest rates.
Journal of monetary economics, 47(3):613–652.

Lenza, M. and Jarocinski, M. (2016).
An inflation-predicting measure of the output gap in the euro area.
Working Paper Series 1966, European Central Bank.

Marcellino, M. and Musso, A. (2010).
The Forecasting Performance of Real Time Estimates of the Euro Area Output Gap.
CEPR Discussion Papers 7763.

Moretti, L., Onorante, L., and Zakipour-Saber, S. (2018).
Phillips curves in the euro area.
mimeo, European Central Bank.

Primiceri, G. E. (2005).
Time Varying Structural Vector Autoregressions and Monetary Policy.
Review of Economic Studies, 72(3):821–852.

Reichlin, L. (2018).
Discussion of slack and cyclically sensitive inflation by Stock and Watson.
Sintra 18-20 June, ECB Annual Forum.

Rünstler, G. (2002).
The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area.
Working Paper Series 182, European Central Bank.

Stock, J. H. and Watson, M. M. (1999).
Forecasting Inflation.
Journal of Monetary Economics, 44(2):293–335.

Stock, J. H. and Watson, M. M. (2007).
Why has U.S. Inflation Become Harder to Forecast?

Journal of Money, Credit, and Banking, 39(7):3–34.

Stock, J. H. and Watson, M. M. (2009).

Phillips Curve Inflation Forecasts.

In Fuhrer, J., Kodrzycki, Y., Little, J. S., and Olivei, G., editors, *Understanding Inflation and the Implications for Monetary Policy, a Phillips Curve Retrospective*, pages 101–186. The MIT Press.

Stock, J. H. and Watson, M. M. (2010).

Modelling Inflation After the Crisis.

mimeo.

Zaman, S. (2013).

Improving Inflation Forecasts in the Medium to Long Term.

Economic commentary, Federal Reserve Bank of Cleveland.

De-trending - Stock and Watson (1991) formulation

- Stock and Watson (1999,2009) unit root formulation:

$$\pi^h_{t+h} - \pi_t = \mu_h + \alpha(L)\Delta\pi_t + \beta_h(L)y_t + \gamma_h(L)z_t + \nu^h_{t+h}$$

For $h = 1$ this is equivalent to formulation above with $\pi^{TR}_t = \pi_{t-1}$.

De-trending - EWMA

Local level model:

$$\pi_t = \pi_t^{TR} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2)$$
$$\pi_t^{TR} = \pi_{t-1}^{TR} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

then

$$\pi_{t|t}^{TR} = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i \pi_{t-i}$$

where $\lambda = f(\sigma_\eta^2 / \sigma_\epsilon^2)$.

Alternatively, π_t has an IMA(1,1) representation:

$$\Delta \pi_t = (1 - \theta L) u_t$$

and $\lambda = \theta$.

We consider $\lambda = 0.95$, 0.85 and 0.75.

De-trending - Endogenous trend

$$
\begin{aligned}
\pi^h_{t+h} - \bar{\pi}^h_{t+h} &= \alpha(L)(\pi_t - \bar{\pi}_t) + \beta(L)x_t + \epsilon^h_{t+h} \\
\bar{\pi}_t &= \bar{\pi}_{t-1} + u_t
\end{aligned}
$$

$\pi^h_{t+h} = \frac{1}{h}\sum_{i=0}^{h-1} \pi_{t+h-i}$ and $\bar{\pi}^h_{t+h} = \frac{1}{h}\sum_{i=0}^{h-1} \bar{\pi}_{t+h-i}$