Discussion of Forecasting in the presence of recent and recurring structural breaks, by J. Eklund, S.Price and G.Kapetanios

F. Venditti

Banca d'Italia

January 18, 2010

Plan of the discussion

- Quick summary of the paper (aim and relevance)
- First environment (small frequent breaks)
  1. I suggest some exercises to further motivate the use of *naive* estimators
  2. I have a look at what happens if also the variance changes
- Second environment (large infrequent breaks)
  1. Raise a question on why Pesaran Timmermann (2007) does not apply here
- Some general comments
- Conclusions

What the paper is about

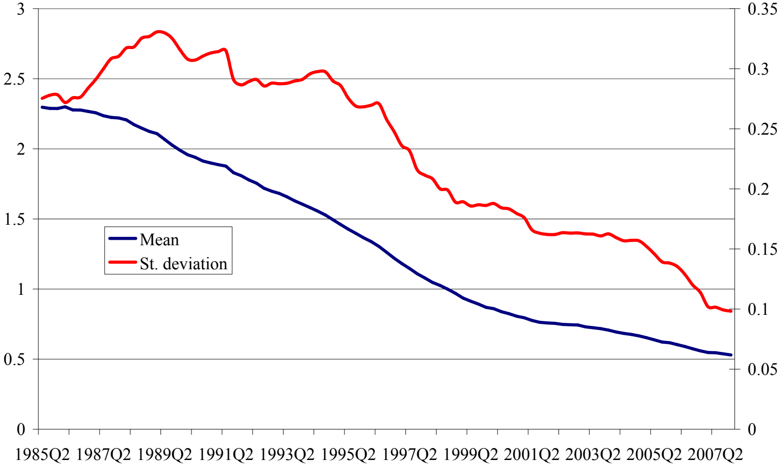
- It compares two approaches to breaks:
  1. Relaxed Guy: no need to be nervous, breaks happen all the time, I use simple strategies to discount past data.
     1. Rolling mean
     2. Pooled mean (Average of means over a shrinking window)
     3. EWMA
  2. Nervous Guy: large breaks might occur, I cannot relax, let me monitor all the time and if something happens I am ready to combine data...
- Issues particularly relevant if you update your forecasts frequently (Central Banks and alike)

Builds on Pesaran and Timmermann (2007)

- Intuition of PT (2007)
- You have a break that it is not too recent
- Does it pay off to use pre-break data?
- It might, especially if after the break the variable you want to forecast becomes noisier
- You have a bias from the slopes because you're using data from before the break (when the slopes were different)
- But you have a gain from the variance because you're mixing recent noisy data with old less noisy data
- In real life this might be of little use
  1. Break tests are worse than Alitalia flights: they're always late
  2. Small continuous breaks might never be caught by break tests (Benati, Drifts and breaks)

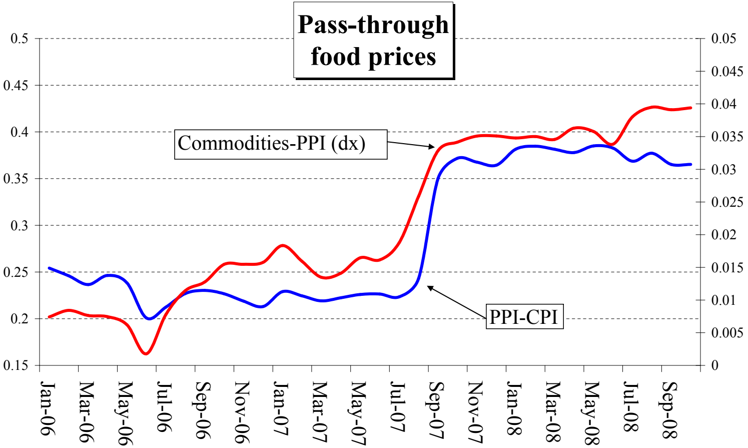
In real life we observe both small smooth breaks...

Inflation mean and volatility
(Euro area GDP deflator - q/q growth rates 15 years rolling window)



... and sudden large breaks

Recent spike in food prices pass-through in the euro area



First environment: small frequent structural breaks 1

- If breaks are small and frequent (i) tests don't catch them (ii) you're forced to use data across breaks due to data constraints
- Consensus emerged in the literature: this environment is well captured by slow moving drifting coefficients (SLOW-RW).

$$y_t = \mu_t + \epsilon_t$$
$$\mu_t = \mu_{t-1} + u_t$$

- If σ_u / σ_e is small: pile up problem (MUE, Bayesian methods etc...)

First environment: small frequent structural breaks 2

- Relaxed guy doesn't buy this view + he doesn't like sophisticated methods
- He lives in a world in which :

$$
\begin{aligned}
y_t &= \mu_t + \epsilon_t \\
\mu_t &= \mu_{t-1} && \text{with probability 1-p} \\
\mu_t &= \eta_t && \text{with probability} \quad \text{p}
\end{aligned}
$$

where η_t is a random uniform shock that can have large or small variance.

- Problem 1: are his estimators really robust?
- If the world were to be a SLOW-RW how would they perform?
- I would like to see some robustness checks
- Change the DGP and see how naive estimators perform compared to more sophisticated ones (TVP, MS and so on)

An aside

- I have a problem with the formula for EWMA in the paper
- The weights for observation j are written like: $1/T\lambda(1-\lambda)^{T-j}$
- Yet they should be: $\lambda(1-\lambda)^{T-j}$
- EWMA errors look far too large to me
- In my simulations EWMA actually performs quite well

Simulate a random walk plus noise model

High Signal/Noise			Low Signal/Noise		
ROLL	POOL	EWMA	ROLL	POOL	EWMA
0.92	1.04	0.93	0.96	1.02	0.98

First environment: small frequent structural breaks 3

- Problem 2: the noise variance never changes in this paper
- Even if you had enough data (and you don't in this environment as breaks are frequent) you could never exploit any trade off by using pre-break data
- What if the world became less or more noisy at some point?
- Would naive estimators perform well compared to the full sample? Simulate the model in the paper:

$$y_t = \mu_t + \epsilon_t \tag{1}$$

$$\mu_t = \mu_{t-1} \qquad \text{with probability 1-p} \tag{2}$$

$$\mu_t = \eta_t \qquad \text{with probability p} \tag{3}$$

- Now I let σ_e (noise variance) switch **only once** randomly within the sample with probability 1/T
- σ_e can double (from low to high volatility environments)
- σ_e can halve (from high to low volatility environments)

	From High to Low variance					
	High Signal/Noise			Low Signal/Noise		
	ROLL	POOL	EWMA	ROLL	POOL	EWMA
p=.1	0.98	0.98	0.80	1.01	1.00	1.02
p=.01	0.86	1.01	0.80	1.00	1.00	1.03

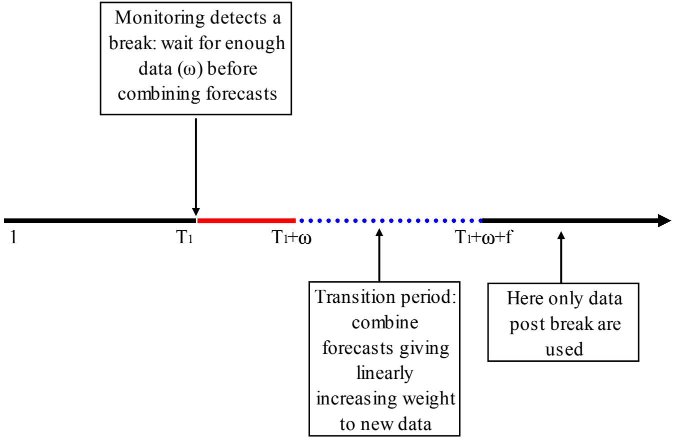
	From Low to High variance					
	High Signal/Noise			Low Signal/Noise		
	ROLL	POOL	EWMA	ROLL	POOL	EWMA
p=.1	0.99	0.99	0.90	1.01	1.00	1.04
p=.01	0.93	1.00	0.93	1.01	1.00	1.04

Summary on this

- The rationale for using naive estimators seems rather weak
- Relaxed guy risks to be confused for Lazy guy
- Some robustness checks could ‘sell’ the story that he’s using models that are robust to misspecification of the underlying DGP

Second environment: rare large breaks

- After $T_1 + \omega$ you have enough data to combine: PT (2007) here becomes relevant
- Why doesn't Nervous guy combine data even after $T_1 + \omega + f$?
- The issue of the constant variance is even more important



More general points

- The theoretical part is not very informative: you still have to simulate even for the simple local level model
- As a reader I'd prefer to have a small paragraph on monitoring
- The paper tries to tackle a lot of points, maybe too many
- Small frequent breaks: with robustness checks and comparisons with more sophisticated models could be a paper on its own

Conclusions

- Issue very relevant for people that actually forecast frequently
- The choice not to model time variation with some unobserved components model raises eyebrows
- Robustness to misspecification of the underlying process could be a way to go
- Playing around with the variance could give further insights
- I learned a lot from reading this paper, which is always good!

Thanks for listening