Low Frequency Econometrics

Ulrich Müller and Mark Watson
Presented by: Paul Sangrey

September 11, 2015

Really Technical. Really Fun. Really Useful.

- Use spectral theory and alternative asymptotics in sophisticated ways to understand questions that have not been thoroughly, rigorously addressed.
- Lots of decisions have very long time horizons, so we need reasonable estimators over long-time scales. But How?
- Clever way of using asymptotic approximations even in small samples.
- Cointegration.
- Devise new tests for various trend specifications.

What are we trying to forecast?

- The time-varying trend.
- How to tell the disentangle variation in the trend from business cycles?
- The last 5-10 years have seen very slow GDP growth. Is this the new normal, or have we simply had a long recession?

What do we know? Not Much

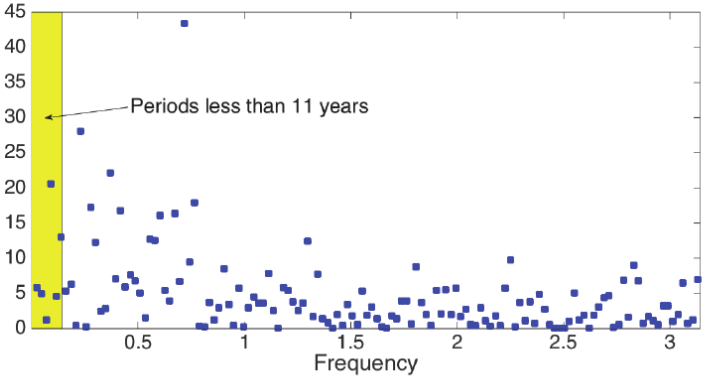

Figure: Periodogram of GDP Growth Rates

What don't we know?

- Essentially a small-sample inference problem .
  - With periods greater than 11 years, we essentially have only 6 datapoints.
  - So we can't really use traditional asymptotics.
- In any non I(0) model, the spectrum has non-zero curvature even within an $O(T^{-1})$ band around zero.

Cosine Transforms

- Use cosine transforms to extract the low-frequency information.
- $\Psi_j(s) = \sqrt{2}\cos(js\pi)$.
- Orthogonal to each other and to a constant.
- Thus the covariance matrix of the cosine transforms is an identity.
- Since the cosine transforms are orthogonal to a constant, in estimation procedures we probably want to include the sample mean, but this is not too hard to do. Although, this does mean that the covariance matrix is no longer an identity.

Asymptotic Approximations and Cosine Transforms

- Now we can do asymptotics!
- Hold the number (q) of low-frequency components fixed but let the time go to infinity.
- Better approximation of actual problem. As $T \to \infty$, we get more useless information but not more useful information.
- $T^{1-k} \sum_{t=1}^{T} \Psi_j \left(\frac{t-1/2}{T} \right) u_t \Longrightarrow \int_0^1 \Psi_j(s) dG(s) \sim \mathcal{N}(0, \Sigma)$.
- The k depends upon the particular parameterization of the trend we are using, and the G is a Gaussian process.

Why do it this way?

- Can elegantly consider a wide-variety of parametric trend models.
- This allows you to easily incorporate effects of low-frequency parameter-estimation uncertainty in your prediction.
- Robust to high-frequency misspecification.
- Can consider a wide-variety of tests concerning the different trend parameters.

Testing

- They consider several canonical trend models (Local-level, local-to-unity, fractional) and devise tests for their parameters.
- If you want to consider the persistence of the process, you probably want to be scale invariant.
- They consider several different techniques for deriving tests applied to this environment including point-optimal tests, weighted average power tests, and using least favorable distributions.
- One can also easily derive Bayesian tests and model selection procedures in this environment because you have an approximate Normal likelihood.

What Do We Know?

- The data is fairly informative about the business-cycle frequency components.
  - Use an ARMA model with stochastic volatility, or something simpler.
- Do we really have to be robust about this?
- We do not know much about the trend. Only 6 datapoints.

Structural Change

- The DGP of basically any model does change over time.
- You cannot forecast over long time horizons without worrying that the future will not be like the recent past.
- They do not even mention this.

Structural Change

- Maybe it's not as bad as it seems.
- Maybe we should think about structural change as the trend changing instead of the parameters changing. In some sense, it's just a question of interpretation if we consider a wide-enough class of DGP's since there's are severe identification issue.
- However, then we need to allow the trend's DGP to be even more general.
- We're not going to be able to estimate this. There is not enough information.

Dealing with Uninformative Data – Use Other Information

- In a Bayesian context, use informative priors.
- Bring in theory.
- Use additional data.
  - Give up some robustness and bring in higher frequency information. This can be done in the context of a parametric model, by using a greater number of cosine transformations, or considering other potential transformations.
  - Bring in other datasets. Consider using multiple series with a common trend such as GDP and TFP, stocks, etcetera. One could also consider series with the same trend structure but different actual trends to pick the correct parametric form.

Cointegration

- Work out asymptotic distribution for cointegrated processes and testing theory.
- Allow the trend to take on a number of different forms.
- "Low-frequency robust cointegration testing" by Müller & Watson (2013)

Biggest Problem

As an estimation procedure it does not work that well in practice. There just is not much information about the relationship between business cycle fluctuations and trend fluctuations once you throw out the business cycle fluctuations.

- Come to the lunch a week from Monday.

Contribution

- A very interesting – hard – question and considered using a great deal of robustness.
- Probably a good, rather atheortic way to make point predictions. Just cosine transform the data, take the first few transformations and run OLS.
- A very useful asymptotic framework, which should be able to be applied in a wide variety of settings.
- Lays the groundwork for a large and important literature.