Trends and Cycles, Potentials and Gaps, Permanent and Transitory decompositions

Fabio Canova
Norwegian Business School and CEPR
March 2020

Outline

- Intro and generics of the decompositions.
- Burns and Mitchell: turning point analysis.
- Lucas 1: LT, SEGM, FOD, Hamilton, UC, BN.
- Lucas 2: HP, BP, Wavelets, Butterworth.
- Economic model-based decompositions: BQ, KPSW.
- Collecting cyclical information: does it matter?
- Business and financial cycles.
- Fitting DSGE models to cyclical data.

References

Astrudillo, M.G. and J. Roberts, (2016). Can Trend and cycle decompositions be trusted? Finance and Economics Discussion series, Divisions of Finance, Statistics and Economic Affairs, Federal Reserve Board 2016-099

Baxter, M. and King, R., (1999), "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series", *Review of Economics and Statistics*, 81, 575-593.

Berge, T. and Jorda, O. (2011), "Evaluating the Classification of Economic Activity into Recessions and Expansions", *American Economic Journal: Macroeconomics*, 3, 246-277.

Beveridge, S. and Nelson, C., (1981), "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the Business Cycle", *Journal of Monetary Economics*, 7, 151-174.

Blanchard, O. and D. Quah (1989) The dynamic Effects of aggregate demand and supply disturbances. American economic Review, 79, 655-673.

Borio, C. (2012) The ifnancial cycles and the macroeconomics: what have we learned? BIS working paper 395.

Borio, C., Disyatat, P. and M. Juselius (2015). Rethinking potential output: embedding information about financial cycles. Oxford Economic Papers, 69(3), 655-677.

Borio, C., Drehemann, M. and D. Xia (2018). The financial cycle and recession risk. BIS quarterly review, December, 59-71.

Bry, G. and Boschen, C. (1971) *Cyclical analysis of time series: Selected Procedures and Computer Programs*, New York, NBER

Burns, A. and Mitchell, W. (1946) *Measuring Business Cycles*, New York, NBER.

Canova, F., (1998), "Detrending and Business Cycle Facts", *Journal of Monetary Economics*, 41, 475-540.

Canova, F., (1999), "Reference Cycle and Turning Points: A Sensitivity Analysis to Detrending and Dating Rules", *Economic Journal*, 109, 126-150.

Canova, F., (2014), "Bridging DSGE models and the raw data", *Journal of Monetary Economics*, 67, 1-15.

Canova, F. (2019a) Discussion to Fatima Pires speech at the Conference for the 140 anniversary of the Bulgarian Central Bank, https://sites.google.com/view/fabio-canova-homepage/home/policy-stuff.

Canova, F. (2019b). FAQ: How do I measure the output gap?, manuscript

Canova, F., and Ferroni, F. (2011), "Multiple filtering device for the estimation of DSGE models", *Quantitative Economics*, 2, 37-59.

Canova, F. and C. Matthes (2019). A new approach to deal with misspecification of structural econometric models, CEPR working paper

C. Chang, K. Chen, D. Waggoner and T. Zha (2015). Trend and cycles in China economy, NBER Macroeconomic annual, 30(1), 1-84.

Christiano, L. and J. Fitzgerard (2003) The Band Pass Filter, *International Economic Review*, 44, 435-465.

Cogley, T. and Nason, J., (1995), "The Effects of the Hodrick and Prescott Filter on Integrated Time Series", *Journal of Economic Dynamics and Control*, 19, 253-278.

Cornea Madeira, A. (2017), "The explicit formula for the HP filter in finite samples, *Review of Economics and Statistics*, 310-314

Cover, J., Enders, W. and Hueng, J. (2017), " Using and aggregated demand-aggregate supply model to identify structural demand-side and supply-side shocks: results using a bivariate VAR", *Journal of Money Credit and Banking*, 38, 777-790

De Jong, R. and N. Sakarya (2016) "The Econometrics of HP filter, *Review of Economics and Statistics*, 98(2), 310-317.

Grant, A. and Chan, J. (2017a). A Bayesian Model Comparison for Trend-Cycle Decompositions of Output. Journal of Money, Credit and Banking, 49, 525-552.

Grant, A. and Chan, J. (2017b). Reconciling output gaps: Unobservable component model and Hp filter. Journal of Economic Dynamics and Control, 75, 114-121.

Hamilton, J. (1989) "A New Approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica*, 57, 357-384

Hamilton J. (2018). Why you should never use the Hodrick and Prescott Filter, Review of Economics and Statistics, 100, 831-843.

Harvey, A. and Jeager, A., (1993), "Detrending, Stylized Facts and the Business Cycles", *Journal of Applied Econometrics*, 8, 231-247.

Hodrick, R. and Prescott, E., (1997), "Post-War US Business Cycles: An Empirical Investigation", *Journal of Money Banking and Credit*, 29, 1-16.

Kambler, G., Morley, J and B. Wong, (2018). Intuitive and reliable estimates of the output gap from Beveridge and Nelson filter. Review of Economics and Statistics, 100, 550-566.

Kim. S.J., Koh, K, Boyd, S. and Gorinevsky, D. (2009), L1-Trend filtering, *SIAM Review*, 51, 339-360

King, R. and Rebelo, S., (1993), "Low Frequency Filtering and Real Business Cycles", *Journal of Economic Dynamics and Control*, 17, 207-231.

King, R. Plosser, C., Stock, J. and Watson, M. (1991) "Stochastic Trends and Economic Fluctuations", *American Economic Review*, 81, 819-840.

O. Jorda, M. Shularick and A. Taylor (2016). The great mortgaging: housing finance, crises and business cycles. Economic policy, 107-152.

Lubik, T., Matthes, C. and Verona, F. (2019) Assessing U.S. Aggregate Fluctuations Across Time and Frequencies. Bank of Finland, manuscript.

Marcet, A. and Ravn, M. (2000) "The HP filter in Cross Country Comparisons", LBS manuscript.

Morley, J. (2002) A State-Space Approach to Calculating the Beveridge- Nelson Decomposition, Economics Letters 1, 123-127

Morley, J., Nelson, C. and Zivot, E. (2003) Why are Beveridge-Nelson and Unobservable Component Decompositions of GDP so Different? *Review of Economics and Statistics*, 86, 235-243.

Pagan, A. (2013) Patterns and their use. NCER working paper 96.

Pagan, A. (2019) Business cycle Issues: Some Reflections on a Literature. Manuscript.

Pagan, A. and D. Harding, (2002), Dissecting the Cycle: A Methodological Investigation, *Journal of Monetary Economics*, 49, 365-381.

Pagan, A. and D. Harding (2006). Synchronization of Cycles, Journal of Econometrics 132, 59-79.

Pagan, A.and M. Kulish (2018) Variety of cycles: Analysis and use. University of Sydney, working paper.

Phillips P. and O. Jin (2015). Business cycles, trend elimination and the HP filter, Cowles fundation working paper 2005.

Ravn, M and Uhlig, H. (2002), On adjusting the HP filter for the frequency of Observations, *Review of Economics and Statistics*, 84, 371-375.

Runstler, G. and M. Vlekke (2018). Business, housing, and credit cycles, Journal of Applied Econometrics, 33(2), 212-226.

Schuler, Y (2019). How should we filter economic time series? Bundesbank, manuscript.

Stock, J. and M. Watson. (1999), Forecasting Inflation, *Journal of Monetary Economics*, 44, 293-335.

Stock J. and M. Watson (2007). Why Has U.S. Inflation Become Harder to Forecast? Journal of Money, Banking and Credit, 39, 3-33.

Stock J. and M. Watson (2014). Estimating turning points using a large data set. Journal of Econometrics, 178, 368-381.

Stock, J. and M. Watson (2016). Core Inflation and Trend Inflation. Review of Economics and Statistics, 98, 770-784.

1 Introduction

• Why we care about business cycles? Why bout seasonal cycles?

• How do you measure business cycles? What are their features? Are they different from, say, financial cycles?

• Should we compute classical (level) or growth (detrended) cycles?

• Burns and Mitchell (BM) (1943):

"Business cycles are a type of fluctuations found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions **occurring at about the same time in many economic activities**, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; **this sequence of changes is recurrent but not periodic**; in duration business cycles vary from **more than one year to ten or twelve years**; they are not divisible into shorter cycles of similar characters with amplitudes approximating their own."

• Lucas (1977):

" **Movements about trend in gross national product** in any country can be well described by a stochastically disturbed difference equation of very low order. These movements do not exhibit uniformity of either period or amplitude, which is to say, **they do not resemble the deterministic wave motions which sometimes arise in the natural sciences**. Those regularities which are observed are **in the co-movements among different aggregative time series** (....). There is, as far as I know, **no need to qualify these observations by restricting them to particular countries or time periods**: they appear to be regularities common to all decentralized market economies. Though there is absolutely no theoretical reason to anticipate it, one is led by the facts to conclude that, with respect to the qualitative behavior of co-movements among series, **business cycles are all alike**.

• Burns and Mitchell consider level data. Characterize the business cycle of a nation by specifying interesting durations and looking at comovements across series.

• Lucas looks at detrended data (growth cycles). Seek commonality across series, time periods and, potentially, countries. Does not specify interesting periodicities.

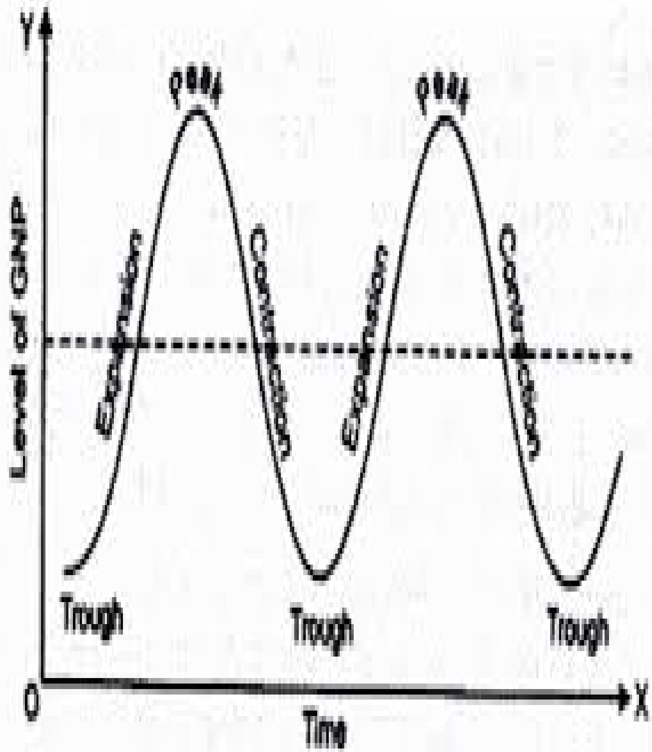


Fig. 13.1. *Four Phases of Business Cycles without Growth Trend*

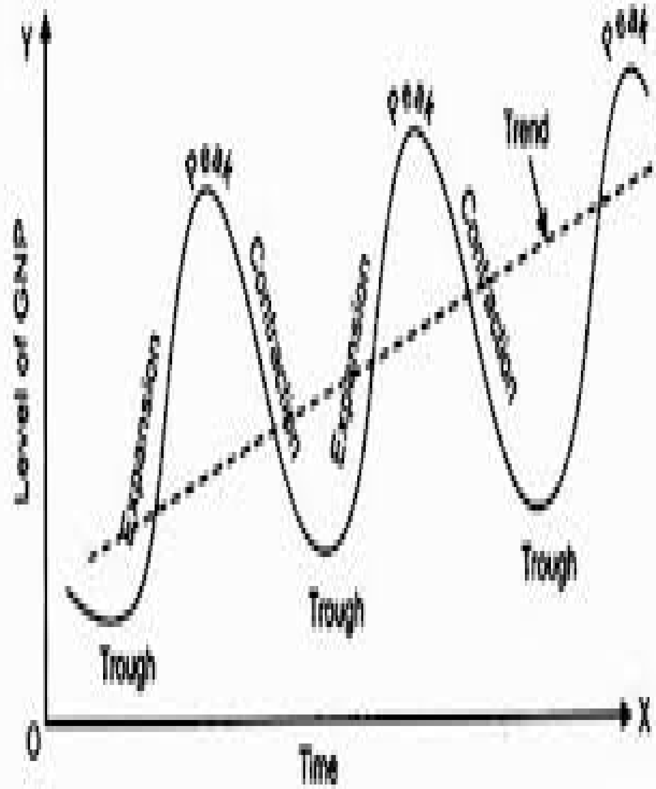


Fig. 13.2. *Cycles with Trend (i.e., Growth)*

Wikipedia (mixing and confusing): "The business cycle, also known as the economic cycle or trade cycle, **is the downward and upward movement of gross domestic product (GDP) around its long-term growth trend** The length of a business cycle is the period of time containing a single boom and contraction in sequence. These fluctuations typically involve shifts over time between periods of relatively rapid economic growth (expansions or booms) and periods of relative stagnation or decline (contractions or recessions).

Business cycles are usually measured by considering the growth rate of real gross domestic product. Despite the often-applied term cycles, these fluctuations in economic activity do not exhibit uniform or predictable periodicity. The common or popular usage boom-and-bust cycle refers to fluctuations in which the expansion is rapid and the contraction severe.

NBER (methodology):

“The NBER’s Business Cycle Dating Committee maintains a chronology of the U.S. business cycle. The chronology comprises alternating dates of peaks and troughs in economic activity. **A recession is a period between a peak and a trough, and an expansion is a period between a trough and a peak. During a recession, a significant decline in economic activity spreads across the economy and can last from a few months to more than a year.** Similarly, during an expansion, economic activity rises substantially, spreads across the economy, and usually lasts for several years. ... The Committee applies its judgment based on the above definitions of recessions and expansions and **has no fixed rule** to determine whether a contraction is only a short interruption of an expansion, or an expansion is only a short interruption of a contraction. .... The Committee **does not have a fixed definition of economic activity**. It examines and compares the behavior of various measures of broad activity: real GDP measured on the product and income sides, economy-wide employment, and real income. The Committee also may consider indicators that do not cover the entire economy, such as real sales and the Federal Reserve's index of industrial production (IP) ... a well-defined peak or trough in real sales or IP might help to determine the overall peak or trough dates, particularly if the economy-wide indicators are in conflict or do not have well-defined peaks or troughs".

CEPR (recessions):

"a significant decline in the level of economic activity, spread across the economy of the euro area, usually visible in two or more consecutive quarters of negative growth in GDP, employment and other measures of aggregate economic activity for the euro area as a whole"

Why we take deviations from trend?

- Data shows growth; and it has more than cyclical fluctuations.

- Economic models typically stationary and built to explain cyclical fluctuations.

- To collect cyclical facts or to match models and the data need to detrend/filter the data.

- Detrending and filtering are different operations!

Questions:

- If detrend: deterministic or stochastic trend? With breaks or without?

- If filtering: which filter? Which frequencies (cycles) to keep?

- Theories talk about permanent and transitory shocks. They discuss "potential" and "efficient" levels of the variables and "gaps" are deviations of actual from potential/efficient levels. How do they relate to statistical "trends" and "cycles"?.

- How do we link "neutrality" propositions (e.g. long run money neutrality) to "trend and cycle" decompositions?

General conundrums:

- What is the business cycle?

i) Burns-Mitchell/Harding-Pagan: the sequence of alternating, irregularly spaced turning points and repetition of expansion/recession phases or 2 quarters minimum duration.

ii) Majority of macroeconomists: the presence variability, serial and cross correlation in a vector of aggregate macroeconomic variables.

iii) Time series econometricians: spectral peak at cyclical frequencies in one or more time series.

iv) Policymakers: business cycle = output gap? (see Canova, 2019)

- How do one measures the cycle?

i) Use a statistical or an economic model?

ii) If a statistical model: use a univariate or a multivariate approach?

iii) If an economic model:

- Should it feature unit root shocks? What frictions should be in there?

- Should one try to measure the gap? Transitory fluctuations? Or cycles of a particular length?

2 Generics

Assume (for simplicity) that the "trend" is everything that it is not the "cycle", i.e., $y_t = y_t^x + y_t^c$.

- Trend and Cycles are unobservable.

- Nature of the decompositions depends:

i) Assumed properties (definition) of the properties of y_t^x.

ii) Correlation trend-cycle (call it ρ).

3 Burns-Mitchell/Pagan Approach

• Pattern recognition exercise: find cycles, expansions, contractions in the level of y_{it}.

• Use judgemental rules (NBER/CEPR dating committees): persistent periods (at least two quarters) of positive/negative growth. Arbitrary.

• Mechanical rules (Bry and Boschen (BB) algorithm): find peak and through dates (local max and min of the series).

• Example: Let $\mathsf{S}_t = 1$ if upturn occurs and zero otherwise (from some external information).Then $\mathsf{S}_t(1- \mathsf{S}_{t+1}) = 1$ if there is a peak and $(1-\mathsf{S}_t)\mathsf{S}_{t+1} = 1$ if there is a through.

• Measure durations and amplitudes of expansions and contraction phases.

- BB algorithm rules:

1. Peaks and throughs must alternate.

2. Each phase (peak to through or through to peak) must have a duration of at least six months (two quarters).

3. A cycle (peak to peak or through to through) must have a duration of at least 15 months (5 quarters).

4. Turning points within six months (2 quarters) of the beginning or end of the series are eliminated. Peaks or throughs within 24 months (8 quarters) of the beginning or end of the sample are eliminated if any of the points after or before are higher (or lower) than the peak (through).

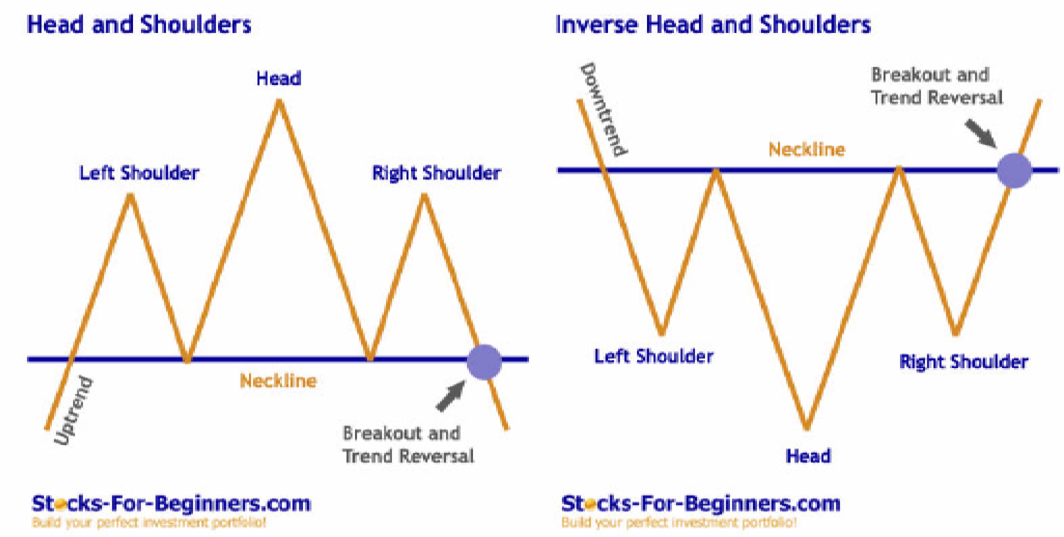


Figure 1: Head and Shoulders Pattern

- BB turning point dates may be different than NBER/CEPR turning point dates. Why?

- Peaks (throughs) may occur at negative (positive) values

- Recessions may be uniformly small (no sharp through).

Turning point dates: Euro area

Phase	CEPR	Length	BB (GDP)	Length
Peak	1974:3		1973:4	
Through	1975:1	2	1974:1	2
Peak	1980:1	20	1979:2	21
Through	1982:3	10	1979:4	2
Peak	1992:1	38	1991:2	46
Through	1993:3	6	1992:2	4
Peak	2008:1	58	2007:2	60
Through	2009:2	5	2008:3	5
Peak	2011:3	9	2010:4	9
Through	2013:1	6	2012:2	6

• How do you construct a synthetic BC indicator? Average-and-date or date-and-average? i.e. Would cycles in one indicator sufficient? Or is it better to date many series and take an average of turning points?

• Average-and-date. Take a standard coincident indicator (e.g. Conference Board (TCB) Indicator in US); or pick one relevant series (GDP, IP). Compute turning points. Compare with standard classification (NBER/CEPR) to check reasonableness of dates. Alternatives:

• Dynamic factor model (DFM):

$$\begin{aligned} y_{it} &= \lambda f_t + e_{it} \\ f_t &= a(L) f_{t-1} + u_t \quad u_t \sim (0, \sigma_u^2) \\ e_{it} &= b(L) e_{it-1} + v_{it} \quad v_{it} \sim iidN(0, \sigma_v^2) \end{aligned} \tag{1}$$

• ISD (Index standard deviation weighting)

$$I_t = \text{exp}(\sum_{i=1}^{N} \alpha_i y_{it}) \tag{2}$$

where $\alpha_i = \frac{s_i^{-1}}{\sum_{j=1}^{N} s_j^{-1}}$ and s_i is the standard deviation of y_{it}.

• In the US, the time paths of ISD and TCB similar (factor and ISD weights are very close).

Table 2
Average-then-date chronologies computed using three monthly coincident indexes and four measures of monthly GDP, as a lead (positive value) or lag (negative value) of the NBER turning point.

NBER		Coincident indexes			Monthly GDP			
		CI-TCB	CI-ISD	CI-DFM	GDP(E)	GDP(I)	GDP(Avg)	GDP-MA
1960:4	P	−2	0	-	−1	−2	−1	
1961:2	T	0	0	0	−2	−2	−2	
1969:12	P	−2	−2	−4	−4	-	−4	
1970:11	T	0	0	0	−10	-	0	
1973:11	P	0	0	0	1	0	1	
1975:3	T	1	1	1	0	−1	0	
1980:1	P	0	0	−10	-	0	-	
1980:7	T	0	0	0	-	−1	-	
1981:7	P	0	1	0	2	1	2	
1982:11	T	0	0	0	−6	0	−3	
1990:7	P	−1	−1	0	0	0	0	
1991:3	T	0	0	0	0	−2	−2	
2001:3	P	−6	−6	−6	-	0	-	-
2001:11	T	4	0	0	-	−1	-	-
2007:12	P	−1	0	0	1	−12	0	1
2009:6	T	0	0	0	0	1	0	0
Mean		−0.44	−0.44	−1.27	−1.58	−1.36	−0.75	0.50
MAE		1.06	0.69	1.40	2.25	1.64	1.25	0.50

Notes: Entries are the NBER turning point minus the series-specific Bry-Boschan turning point, in months. Episodes for which the series is available but does not have a Bry-Boschan turning point are denoted by "-". The GDP(E), GDP(I), and GDP(Avg) monthly GDP series are from Stock and Watson (2010a). The GDP-MA series is the Macroeconomic Advisors Monthly GDP series, which starts in 1992:4. The mean and mean absolute error (MAE) in the final two rows summarize the discrepancies of the chronology for the column series, relative to the NBER chronology; episodes in which a series does not have a Bry-Boschan recession are excluded from the summary statistics.

• Date-and-average. Compute turning points τ_{is} for series $i = 1, \ldots, n_s$ in episode s. Compute a location measure of the turning points distribution for each identified phase episode (e.g. NBER recession).

• If turning points are iid

$$n_s^{0.5}(\hat{\tau}_s^{mean} - \tau_s^{mean}) \xrightarrow{D} N(0, var(\tau_{is})) \quad (3)$$

$$n_s^{0.5}(\hat{\tau}_s^{median} - \tau_s^{median}) \xrightarrow{D} N(0, \frac{1}{4(g_s(\tau_s))^2}) \quad (4)$$

$$(h^3 n_s)^{0.5}(\hat{\tau}_s^{mode} - \tau_s^{mode}) \xrightarrow{D} N(0, \frac{g_s(\tau_s^{mode}) \int [K'(z)]^2 dz}{g_s''(\tau_s^{mode})}) \quad (5)$$

where $K(.)$ is a kernel, h the length of the kernel, $g_s(\tau)$ is the distribution of τ in episode s (see Stock and Watson, 2014).

• If certain types of series are over-represented in the sample relative to the population (e.g. there too many IP series and too few employment series) use weights; helps also if series do not have the same lengths.

• Weights

$$w_{i,s} = \frac{\pi_{m_i}}{p_{m_i,s}} \tag{6}$$

where π_m is the population probability of class m series (IPs, employments, interest rates, etc.) and $p_{m,s}$ is the sample probability of class m in business cycle episode s.

Table 3

Date-then-average chronologies and standard errors computed using turning points of 270 disaggregated series, as a lead (positive value) or lag (negative value) of the NBER turning point.

NBER Dates		No adjustments			Class lag-adjusted			Weighted estimation		
		Mean	Median	Mode	Mean	Median	Mode	Mean	Median	Mode
1960: 4	P	−1.8 (0.6)	−2.0 (0.7)	−1.4 (0.5)	−2.5 (0.7)	−2.3 (0.8)	−2.5 (0.3)	−2.0 (0.6)	−2.0 (0.3)	−1.4 (0.4)
1961: 2	T	−0.3 (0.4)	0.0 (0.6)	−0.5 (0.7)	−0.8 (0.3)	−1.1 (0.5)	−0.5 (0.2)	−0.3 (0.3)	0.0 (0.3)	−0.6 (0.2)
1969:12	P	−2.2 (0.7)	−2.0 (0.6)	−2.3 (0.4)	−1.7 (0.6)	−1.8 (0.7)	−1.3 (0.5)	−1.7 (0.8)	−2.0 (0.4)	−2.4 (5.9)
1970:11	T	1.2 (0.6)	0.0 (0.7)	−0.2 (0.4)	1.7 (0.6)	1.2 (0.7)	0.7 (0.3)	1.9 (0.7)	1.0 (0.6)	0.1 (2.7)
1973:11	P	1.3 (0.6)	2.0 (0.6)	1.6 (0.3)	1.9 (0.6)	3.0 (0.7)	2.2 (0.3)	2.4 (0.7)	3.0 (0.7)	1.7 (1.0)
1975: 3	T	1.0 (0.3)	0.0 (0.3)	0.4 (0.3)	1.6 (0.3)	1.2 (0.3)	1.0 (0.1)	1.3 (0.3)	1.0 (0.4)	0.8 (0.8)
1980: 1	P	−1.8 (0.7)	−1.0 (0.8)	−0.3 (0.4)	−1.3 (0.7)	−1.2 (0.9)	0.3 (0.2)	−1.8 (0.9)	−2.0 (0.8)	−0.1 (0.2)
1980: 7	T	−0.9 (0.5)	0.0 (0.4)	−0.5 (0.2)	−0.1 (0.5)	0.2 (0.3)	0.3 (0.1)	−0.5 (0.7)	0.0 (0.4)	0.0 (0.2)
1981: 7	P	−0.7 (0.5)	0.0 (0.5)	−0.1 (0.3)	−0.2 (0.5)	0.2 (0.5)	0.5 (0.1)	−0.1 (0.5)	0.0 (0.4)	0.1 (4.4)
1982:11	T	−0.6 (0.6)	0.0 (0.6)	1.1 (0.4)	−0.2 (0.6)	0.9 (0.6)	1.9 (0.2)	−0.5 (0.6)	0.0 (0.5)	0.9 (0.9)
1990: 7	P	−0.8 (0.6)	0.0 (0.7)	0.3 (0.5)	−0.3 (0.6)	−1.2 (0.8)	1.8 (0.4)	−1.1 (0.6)	−1.0 (0.5)	−0.3 (0.2)
1991: 3	T	2.1 (0.5)	1.0 (0.4)	0.4 (0.3)	2.1 (0.4)	1.1 (0.4)	0.4 (0.1)	2.0 (0.4)	1.0 (0.4)	0.2 (2.0)
2001: 3	P	−3.7 (0.5)	−3.0 (0.6)	−2.2 (0.3)	−4.1 (0.5)	−4.8 (0.6)	−3.2 (0.2)	−3.7 (0.6)	−3.0 (0.6)	−2.3 (4.4)
2001:11	T	0.2 (0.5)	1.0 (0.5)	0.6 (0.2)	0.5 (0.5)	1.2 (0.5)	1.5 (0.1)	0.6 (0.7)	1.0 (0.7)	0.6 (0.9)
2007:12	P	−1.0 (0.5)	−1.0 (0.9)	−6.1 (0.5)	−1.4 (0.5)	−1.8 (0.7)	−2.8 (0.8)	−1.4 (0.5)	−2.0 (0.9)	−6.0 (1.1)
2009:6	T	1.7 (0.3)	1.0 (0.5)	−0.1 (0.2)	1.5 (0.3)	1.7 (0.4)	1.4 (0.2)	1.6 (0.3)	1.0 (0.5)	−0.2 (0.2)
Mean		−0.39	−0.25	−0.59	−0.20	−0.20	0.10	−0.21	−0.25	−0.56
MAE		1.34	0.88	1.12	1.37	1.56	1.38	1.44	1.25	1.11

Notes: Entries are the NBER turning point minus the date-then-average chronology for that column, in months. Standard errors appear in parentheses. The mean and mean absolute error (MAE) in the final two rows summarize the discrepancies of the chronology for the column series, relative to the NBER chronology.

Alternatives:

• Pagan and Harding (2016). Construct a "reference phase": at least 50 per cent of the series are in a particular BC phase.

• Pagan (2019): Reference turning point minimizes the discrepancy among individual series turning points, i.e. if peaks are at 1973:1, 1973:5, 1973:9, reference peak is 1973:5.

• Construct a weighted average of turning points; weight depends on the (subjective) importance of individual series (GDP turning points have more weights than, say, labor productivity turning points).

Pagan and Harding (2002, 2006): compute useful statistics out of turning point classification, constructed following BM and BB.

Algorithm 3.1 1. *Smooth* y_t *to eliminate outliers, high frequency variations and other uninteresting fluctuations. Call* y_t^{sm} *the smoothed series.*

2. *Determine a potential set of turning points using a rule like, e.g.* $\Delta^2 y_t^{sm} > 0 (< 0), \Delta y_t^{sm} > 0 (< 0), \Delta y_{t+1}^{sm} < 0 (> 0), \Delta^2 y_{t+1}^{sm} < 0 (> 0)$.

3. *Add criteria to ensure that peaks and troughs alternate (may have consecutive peaks) and that the duration and the amplitude of phases are meaningful (minimum duration)*

Statistics

• Average durations (AD), i.e. the average length of time spent between throughs and peaks or peaks and throughs.

• Average amplitudes (AA), i.e. the average size of the drop between peaks and troughs or of the gain between troughs and peaks.

• Concordance index $CI_{j,j'} = n^{-1}[\sum \mathcal{I}_{jt}\mathcal{I}_{it} - (1 - \mathcal{I}_{jt})(1 - \mathcal{I}_{it})]$. Measures comovements over business cycle phases of two variables, where n is the number of complete cycles and $\mathcal{I}_{it} = 1$ in expansions and $\mathcal{I}_{it} = 0$ in contractions. $CI = 1 (= 0)$ if the two series are perfectly positively (negatively) correlated.

• Average cumulative changes over phases ($CM = 0.5 * (AD * AA)$) and excess average cumulative changes $ECM = ((CM - CM^A + 0.5 * AA)/AD)$, where CM^A is the actual average cumulative change.

Features

- No need to measure y_{it}^c.

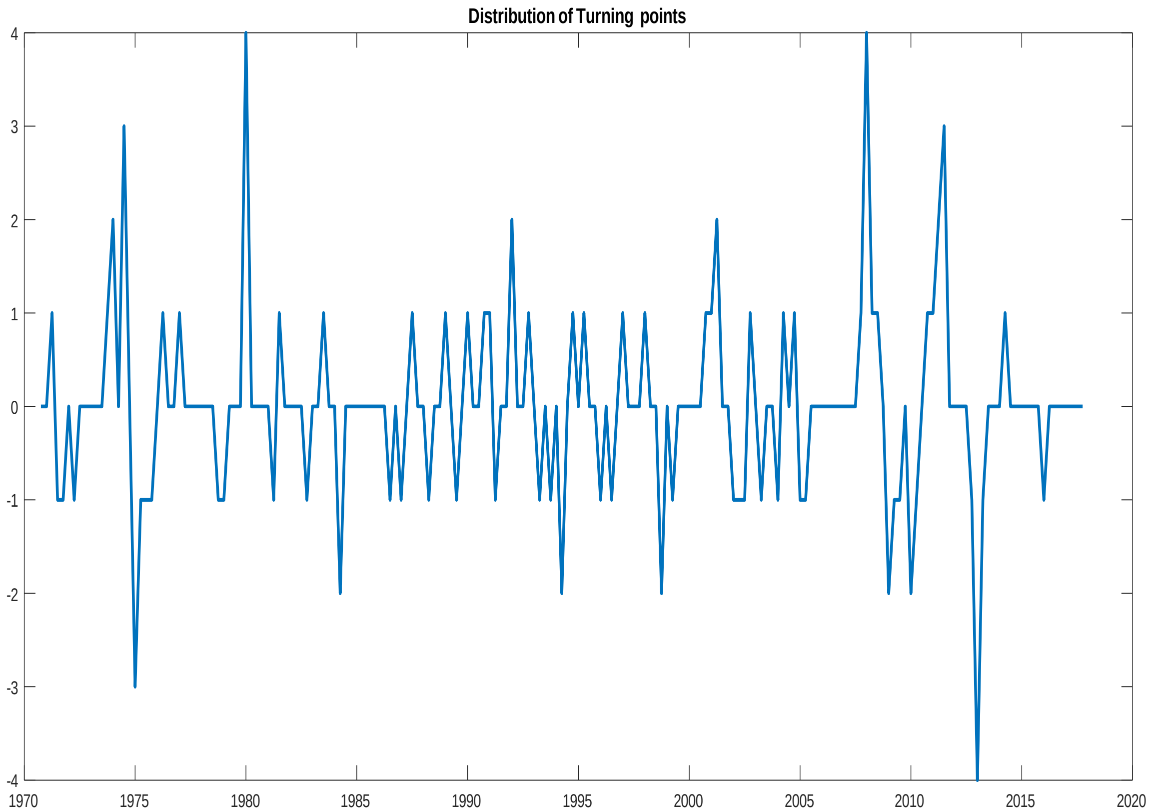
- Can collect statistics even if no econometrician cycles are present (good for DSGE models).

- Allows for asymmetries of cyclical phases.

- Results sensitive to dating rule [2.] and to minimum duration of phases (Typically: two or three quarters - so that complete cycles should be at least 5 to 7 quarters long) and to minimum amplitude restrictions (e.g. peaks to troughs drops of less than one percent should be excluded).

- How to adapt the procedure to international comparisons? How does it relates to the two-quarter negative/positive (NBER) rule?

- Euro data 1970:1-2017:4. Series: Y, C, Inv, Y/N, N, R, π.



- 1975:1 is it a though? Less than 50% of the series are in a downturn.

- 2008 is it a through?. Minimal distance through is 2009:2. (two series have minimum in 2008:1 and two in 2010:3).

- Important to have a good number of coincident series in the exercise.

Euro area Business Cycle Statistics

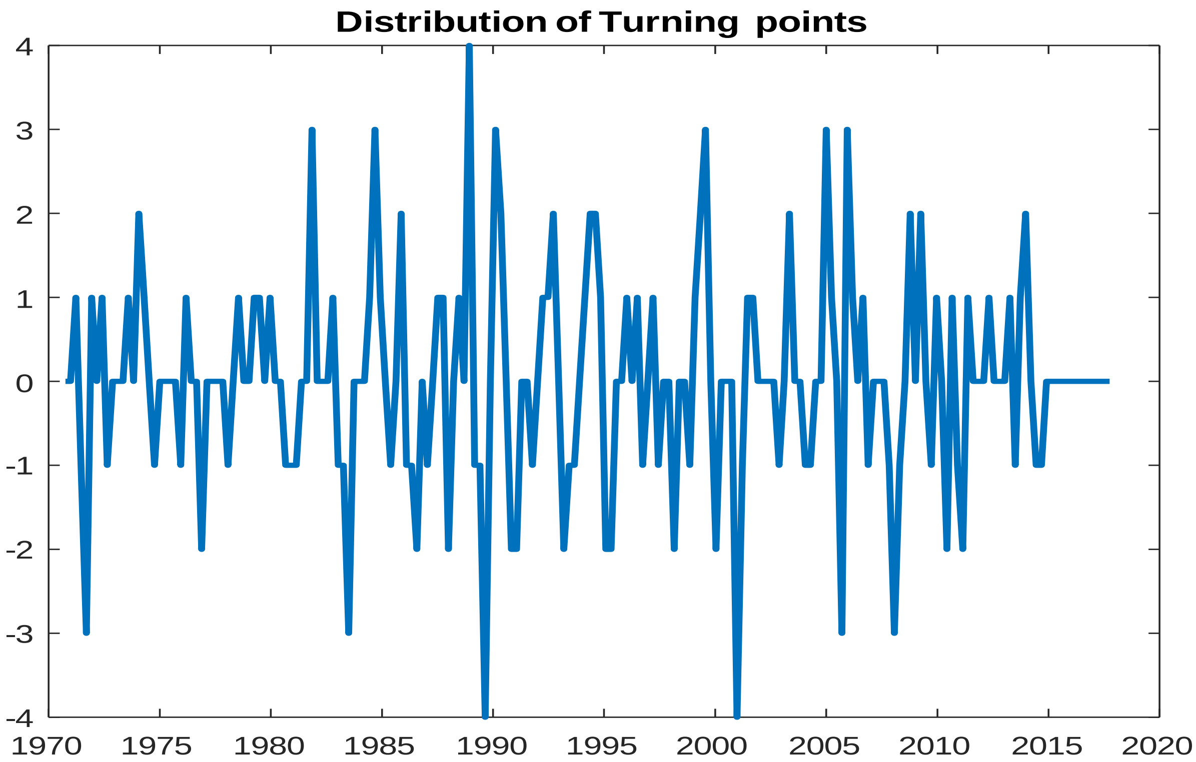
	AD (quarters)		AA (percentage)		ECM(percentage)		$CI_{j,j'}$ (phase)
	PT	TP	PT	TP	PT	TP	
GDP	3.8	33.7	-2.5	20.9	6.7	1.9	
C	5	36.6	-1.5	19.2	9.8	4.4	0.57
Inv	6.7	14.7	-7.2	14.7	14.9	1.1	0.52
Y/N	2.0	18.6	-1.2	8.9	1.7	10.15	0.61
N	9.0	22.8	-1.8	6.13	7.0	11.82	0.45
R	8.4	6.6	-3.1	2.69	10.5	7.80	0.04
π	9.0	6.9	-6.1	5.57	0.34	12.01	0.15

- Big asymmetries in durations and amplitudes.

- Output and consumption expansions longer and stronger than in other series.

- Low concordance of real and nominal series.

- US data 1970:1-2019:3. Series: Y, U, C, Inv, CapU, R1, R10, π, C/Y, I/Y, Term spread.



US Business Cycle Statistics

	AD (quarters)		AA (percentage)		ECM(percentage)		$CI_{j,j'}$ (phase)
	PT	TP	PT	TP	PT	TP	
GDP	3.4	27.4	-0.02	0.2	3.3	13.2	
C	3.7	42.6	-0.01	0.3	-15.7	7.5	0.41
Inv	4.9	10.2	-0.1	0.2	-15.8	8.4	-0.37
U	14.0	7.8	-2.2	2.8	19.3	5.6	0.29
capU	6.2	8.9	-6.6	6.0	-13.2	2.4	-0.02
π	5.3	6.8	-2.7	2.4	8.2	2.4	0.12
R	7.2	6.5	-3.6	2.8	15.2	2.8	-0.04

- Durations in U different than durations in C, I, Y, capU.
- Asymmetries large except for nominal variables.
- Concordance low (negative for I and capU).

3.1 Predicting Downturns

- Use probit/logit model: $\mathrm{P}(1 - S_t = 0|F_{t-1})$; F_{t-1} info available at $t-1$.
- Borio et al. (2018): F_{t-1} = financial cycle information.

Financial cycle proxies help in evaluating recession risk

Regression coefficients from panel probit models — Table 1

Horizon		Financial cycle[1]	DSR	Term spread	Financial cycle and term spread	DSR and term spread
		Advanced economies				
1 year	Financial cycle	0.69***			0.62***	
	DSR		0.61***			0.57***
	Spread			−0.35***	−0.21***	−0.28***
2 year	Financial cycle	0.63***			0.60***	
	DSR		0.38***			0.35***
	Spread			−0.23***	−0.09*	−0.17***
3 year	Financial cycle	0.43***			0.44***	
	DSR		0.16***			0.15***
	Spread			−0.08	0.03	−0.06

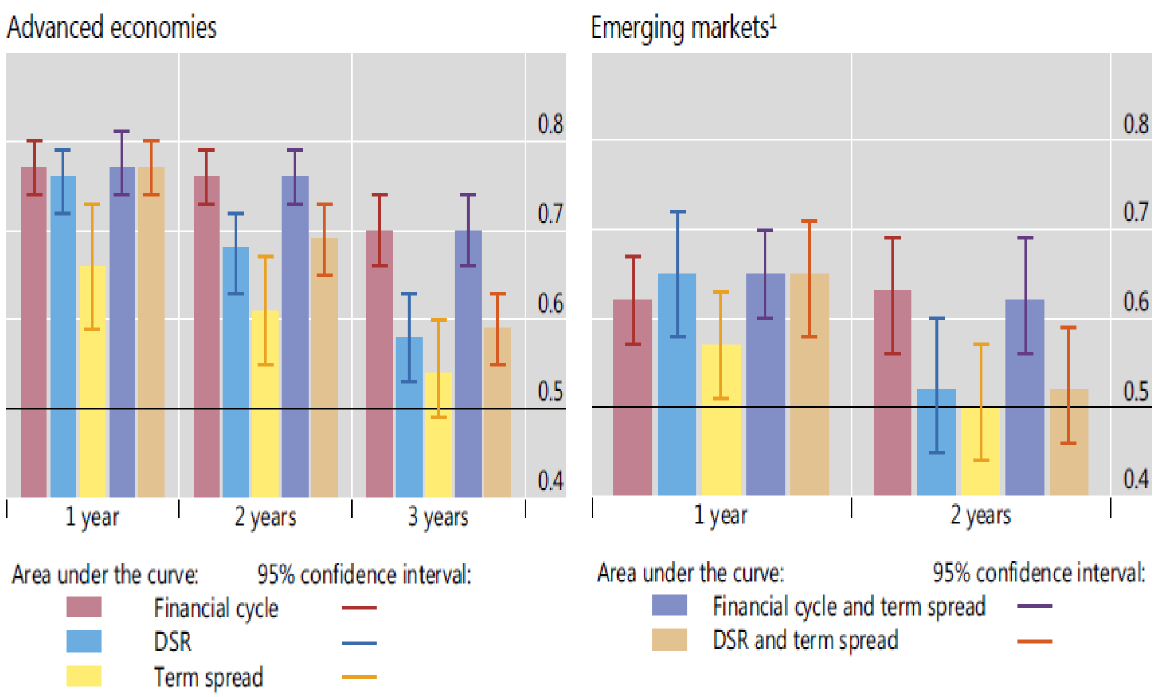
- Report area under the receiver operating characteristic (ROC) curve (Berge and Jorda, 2011).

- Curve maps out combinations of type I errors (missed recessions) and type II errors (false alarms). The area under the curve (AUC) measures the indicator's signalling quality.

- AUC=0.5 :Uninformative indicator; AUC=1.0: a perfect indicator. The AUC of an informative indicator is statistically different from 0.5.

Financial cycle measures are useful for assessing recession risk around the globe

AUCs for different forecast horizons Graph 3



The horizontal lines at 0.5 indicate the area under the curve (AUC) of an uninformative, random variable.

Problems:

- Predicting $1 - S_t$ different than predicting the sign of Δy_{t+1}.

- What are we measuring? P$(1 - S_t = 0|1 - S_{t-1} = 1)$= probability of entering a recession; P$(1 - S_t = 0|1 - S_{t-1} = 0)$= probability of staying in a recession. If just use F_{t-1} we are mixing these two probabilities.

- $1 - S_t$ is a generated variable that depends on $y_{t+k}, k = 1, 2$. Incorrect to use it as conditioning variable in a VAR to see if ,e,g responses differ in recession and expansions.

4 How do macroeconomists think about cycles?

- Use some procedure to remove y_{it}^x.

- Compute $var(y_{it}^c)$; $auto(y_{it}^c)$, i=1,2...,N; $corr(y_{it}^c, y_{1t}^c)$, $i = 2, ..., N$. where y_{1t}^c is output. What is the pattern across i?

- Fix a $t_0 < t < t_1$ (financial crisis, recession, etc.): compute variability, auto and cross correlations.

- Check if models can produce data 'patterns' (Pagan, 2013, 2019).

- What methods are available to estimate y_{it}^c?

Univariate (detrending) approaches

- Polynominal trend, $\rho = 0$.
- Segmented linear trend, $\rho = 0$.
- Differencing:RW trend, $\rho = 0$.
- Hamilton local projection, $\rho = 0$ (also multivariate)
- Unobservable components, ρ may be non-zero (also multivariate).
- Beveridge Nelson: $\rho = 1$ (also multivariate).

Univariate (filtering) approaches

- Hodrick and Prescott, $\rho = 0$.

- Band pass, $\rho = 0$.

- Wavelets, $\rho = 0$.

- Butterworth, $\rho = 0$.

Multivariate (economic) approches

- Blanchard and Quah; KPSW, $\rho \neq 0$ (structural shocks could be uncorrelated or correlated).

4.1 Deterministic Polynomial Trend

$$y_t^x = a + bt + ct^2 + ...$$

Estimate $a, b, c, \ldots$ in the regression

$$y_t = a + bt + ct^2 + ... + e_t$$

by OLS. Set $\hat{y}_t^c = y_t - a_{OLS} - b_{OLS}t - c_{OLS}t^2 - \ldots.$

• Can perfectly predict trend in the future.

• No acceleration/deceleration in the trend is possible.

• Unless a, b, c recursively estimated, timing of information in $\hat{y}_t^c$ and y_t differs.

• y_t^c is typically nearly non-stationary.

4.2 Deterministically Breaking Linear Trend

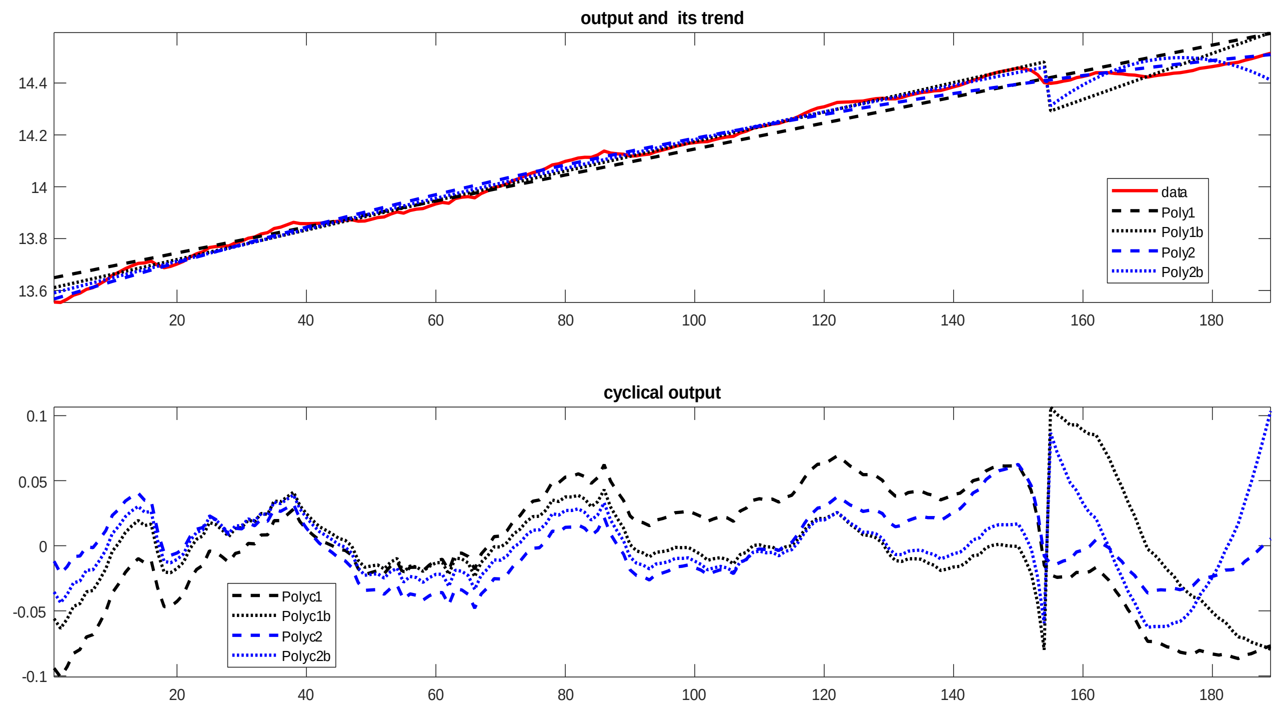
$$y_t^x = a_1 + b_1 t \quad if t \leq t_1 \tag{7}$$

$$y_t^x = a_2 + b_2 t \quad if t > t_1 \tag{8}$$

• Estimate a_i, b_i by OLS. Set $\hat{y}_t^c = y_t - a_{1OLS} - b_{1OLS}t,\ t \leq t_1$; $\hat{y}_t^c = y_t - a_{2OLS} - b_{2OLS}t,\ t > t_1$.

• What if t_1 unknown? Select $[t_a, t_b]$. Run OLS for every $t_1 \in [t_a, t_b]$. Use F-test to check $a_1 = a_2, b_1 = b_2$ each t_1. Break point is the t_1 producing $\max F(t_1) \rightarrow$ QML statistics (see Stock and Watson, 2002).

• Can still perfectly predict y_{t+h} after the break. Solution: Markov switching trend (Hamilton, 1989).



Trend and cycle: polynomial and break trends

- No break after 2008?

4.3 Differencing

- Trend estimate

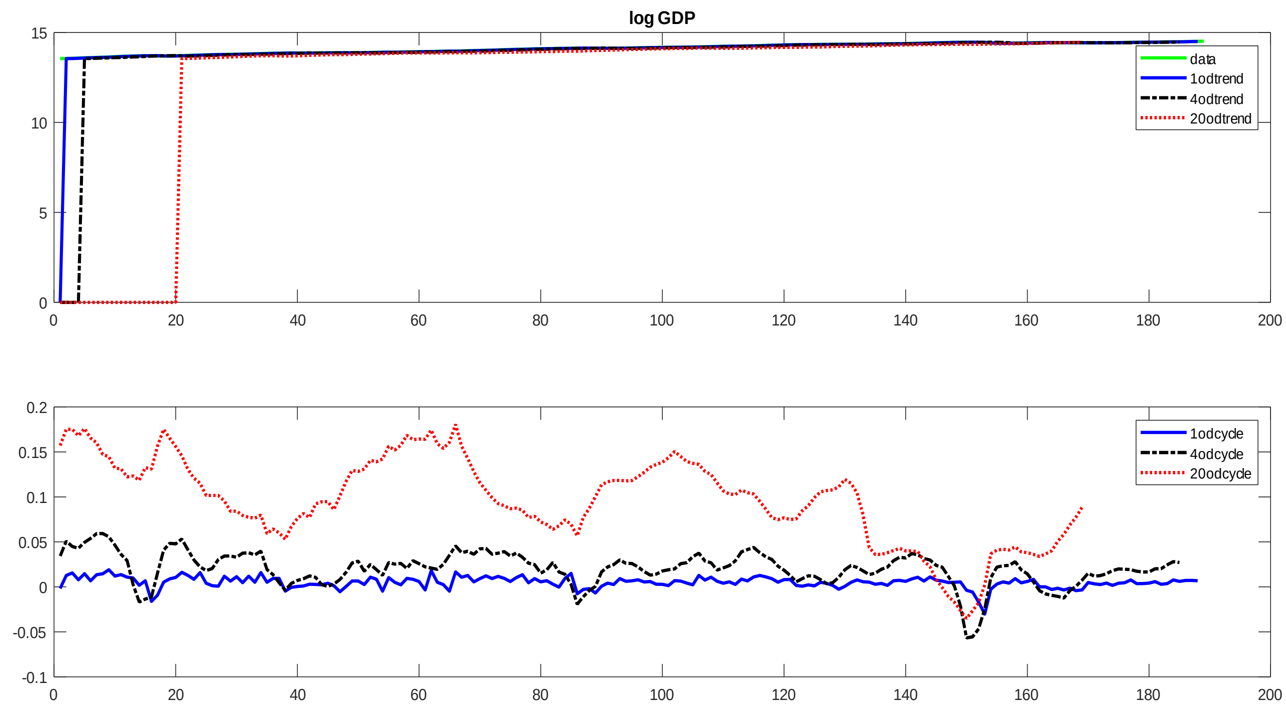
$$y_t^x = y_{t-d}^x \quad d = 1, 4, 8, 24, ... \tag{9}$$

- Cycle estimate: $y_t^c = \Delta^d y_t$,

- Long or short differencing? How do you choose d?

- For d=1 (quarter-on-quarter growth rates) cycles very volatile. Difficult to have models to explain them.

- If $d > 1$ artificial MA(d-1) components in y_t^c.



Trend and cycle: differencing

- Long differencing leaves a downward trend in filtered data

4.4 Hamilton: local projection technique

- Same ideas used to compute impulse responses/ direct forecasts:

$$y_{t+h} = \kappa_{1h}\Delta y_t + \kappa_{2h}\Delta y_{t-1} + \ldots + \kappa_{dh}\Delta^{d-1} y_t + w_{t+h} \tag{10}$$

where, typically, $h = 8$ and $d = 4$. In practice run:

$$y_{t+h} = \alpha_{1h} y_t + \alpha_{2h} y_{t-1} + \alpha_{3h} y_{t-2} + \alpha_{4h} y_{t-3} + w_{t+h} \tag{11}$$

- w_{t+h} is an estimate of y^c_{t+h}.

- w_{t+h} is a function of h, d.

Properties:

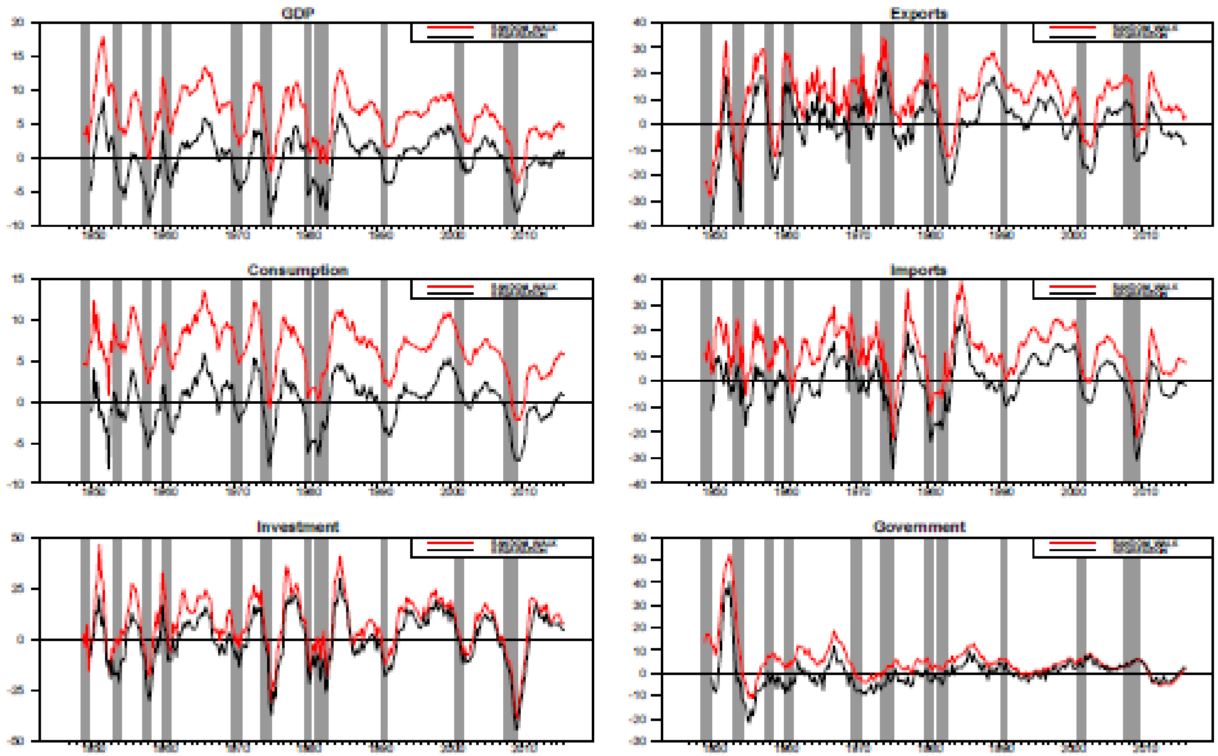
• w_{t+h} is model free. Robust to misspecification of the trend process.

• w_{t+h} is stationary if y_t has up to d unit roots.

• Can be applied to seasonally non-adjusted data; to data of any frequency (quarter, month, week: adjust d and h).

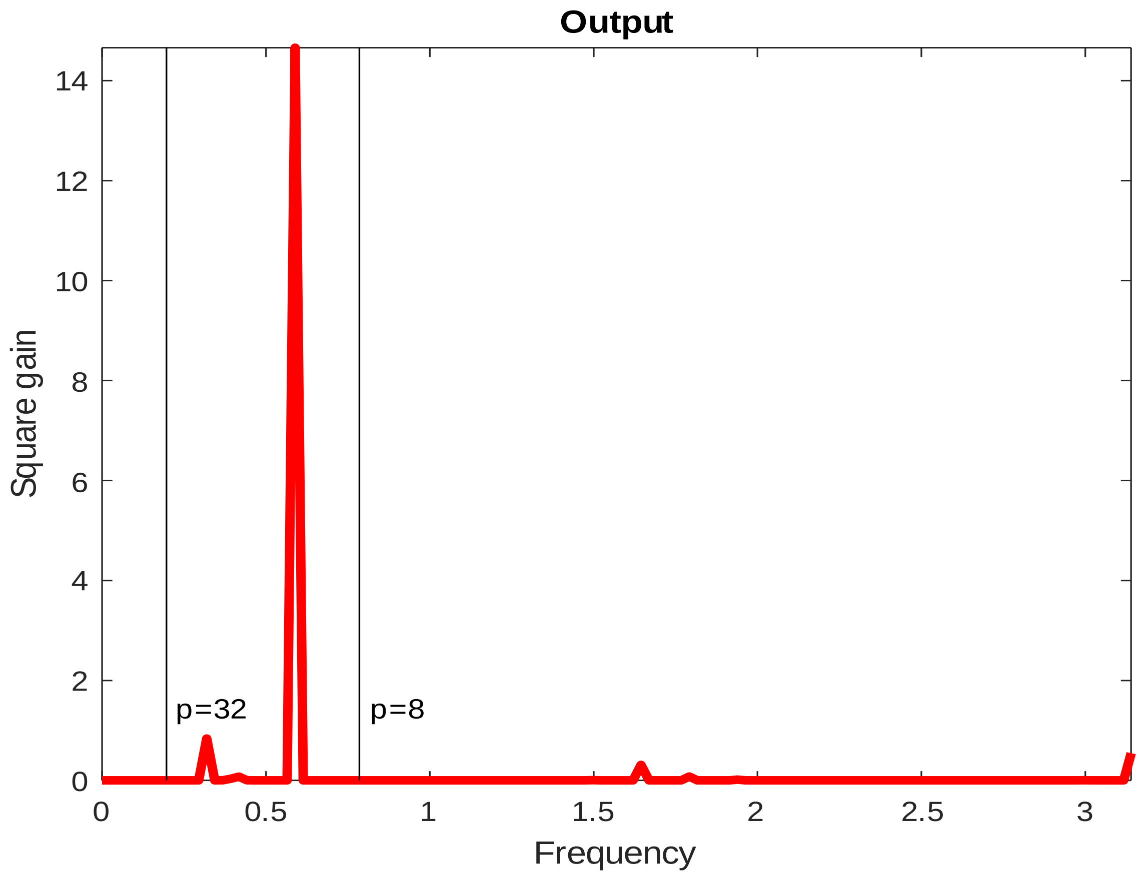
• w_{t+h} similar to those obtained with h differencing.

Figure 6. Results of applying regression (black) and 8-quarter-change (red) filters to 100 times the log of components of U.S. national income and product accounts.



Questions

- Do cycles in w_{t+h} have standard durations and amplitude?
- What kind of comovement does the procedure generates?
- Are their features dependent on h, d?
- What are the properties of the Hamilton trend? (Schuler, 2019)



- Hamilton filter is not a business cycle filter. Peak is at 10.66 quarters.

4.5 Unobservable component methods

• State space model-based.

• Assume certain time series properties for the trend and cycle, e.g. trend is a RW, cycle is an AR(2).

• Can be boosted up with observable regressors or additional features for the error process, see e.g. Stock and Watson, 2016.

• Can be made multivariate, see e.g. Astrudillo and Roberts, 2016; Grant and Chan, 2017a, 2017b. Can restrict trends to be common.

Two setups:

$$
\begin{aligned}
y_t &= \tau_t + c_t + u_t \\
\tau_t &= \tau_{t-1} + \mu + \eta_t \\
c_t &= \theta_1 c_{t-1} + \theta_2 c_{t-2} + \epsilon_t
\end{aligned} \tag{12}
$$

Estimate $(\theta_1, \theta_2, \mu, \sigma_u^2, \sigma_\eta^2, \sigma_\epsilon^2), \rho = corr(\eta_t, \epsilon_t)$ by KF-ML approach or by MCMC with flat prior.

$$
\begin{aligned}
y_t &= \tau_t + c_t + u_t \\
\tau_t &= \tau_{t-1} + \mu + \eta_t \\
c_{1t} &= \theta((\cos\omega)c_{1t-1} + (\sin\omega)c_{2t}) + \epsilon_{1t} \\
c_{2t} &= \theta(-(\sin\omega)c_{1t-1} + (\cos\omega)c_{2t}) + \epsilon_{2t}
\end{aligned} \tag{13}
$$

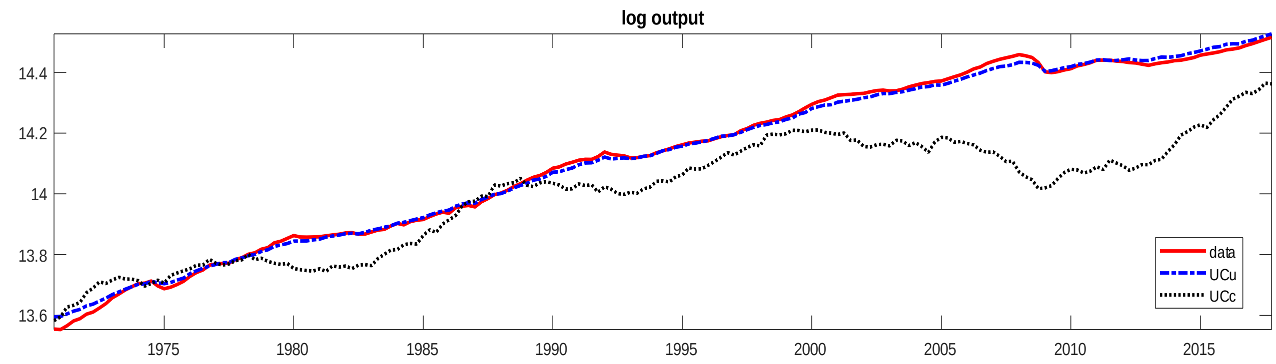
$c_t = [c_{1t}, c_{2t}]$. Fix $0 < \omega < \pi$, estimate $(\theta, \mu, \sigma_u^2, \sigma_\eta^2, \sigma_{\epsilon_1}^2, \sigma_{\epsilon_1}^2)$. (see Runstler and Vlekke, 2018).

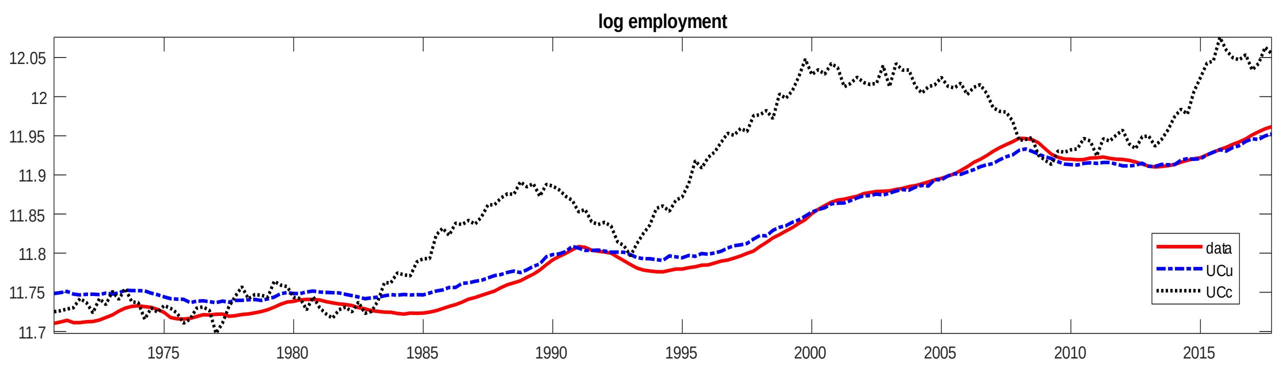
- Can use a more flexible local linear trend specification (see next page)

- Can pick up more than one ω in the cycles in (13).

- Often omit u_t (measurement error).

- Can allow breaks in the trend $(\mu_1(t < t_0), \mu_2(t \geq t_0))$, Markov switching in μ, rare events (jumps in σ_η), and stochastic volatility in σ_ϵ^2.





Data and Trends, UC $\rho = 0$, and $\rho \neq 0$

- Contemporaneous correlation of cyclical outputs: 0.19579
- AR(1) of cyclical outputs: 0.96784 ; 0.99363
- Variabilities of cyclical output: 0.00015967; 0.015188
- Quite a lot of differences! Which one to choose?

Multivariate UC

$$
\begin{aligned}
y_t &= \tau_t + c_t + u_t \\
\tau_t &= \tau_{t-1} + \mu_t + \eta_t \\
\mu_t &= \mu_{t-1} + \nu_t \\
c_{it} &= \sum_{j_i} \rho_j c_{it-j} + \epsilon_{it}, \quad \text{i=1,2,...N}
\end{aligned} \tag{14}
$$

where y_t is $N \times 1$, τ_t is a scalar. Here there is a common (stochastic) local-linear trend. The model for the cycle is allowed to be series specific.

$$
\begin{aligned}
y_t &= \tau_t + c_t + u_t \\
\tau_{it} &= \tau_{it-1} + \eta_{it}, \quad \text{i=1,2,...N} \\
c_t &= \sum_{j} \rho_j c_{t-j} + \epsilon_t
\end{aligned} \tag{15}
$$

where y_t is $N \times 1$, c_t is a scalar. Here there is common cyclical components but there are variable-specific trends.

4.6 Beveridge-Nelson decomposition

• Trend is the long run forecastable component of y_t

• It assumes y_t features a unit root (otherwise long run forecastable component is the mean of y_t).

• Features of estimated y_t^c depend on lag length of the estimating model and sample size.

• Univariate setup :$(\Delta y_t - \bar{y}) = A(\ell)\Delta y_{t-1} + e_t$, where $e_t \sim iid(0, \Sigma_e)$ and all the roots of $det(A(\ell))$ are less than one.

• MA: $(\Delta y_t - \bar{y}) \equiv \Delta y_t^* = D(\ell)e_t$, where $D(\ell) = (1 - A(\ell))^{-1}, D_0 = I$. If $D(1) \neq 0$,

$$\Delta y_t^* = D(1)e_t + (1-\ell)D^\dagger(\ell)\Delta e_t \tag{16}$$

where $D^\dagger(\ell) = \frac{D(\ell)-D(1)}{1-L}$. Cumulating

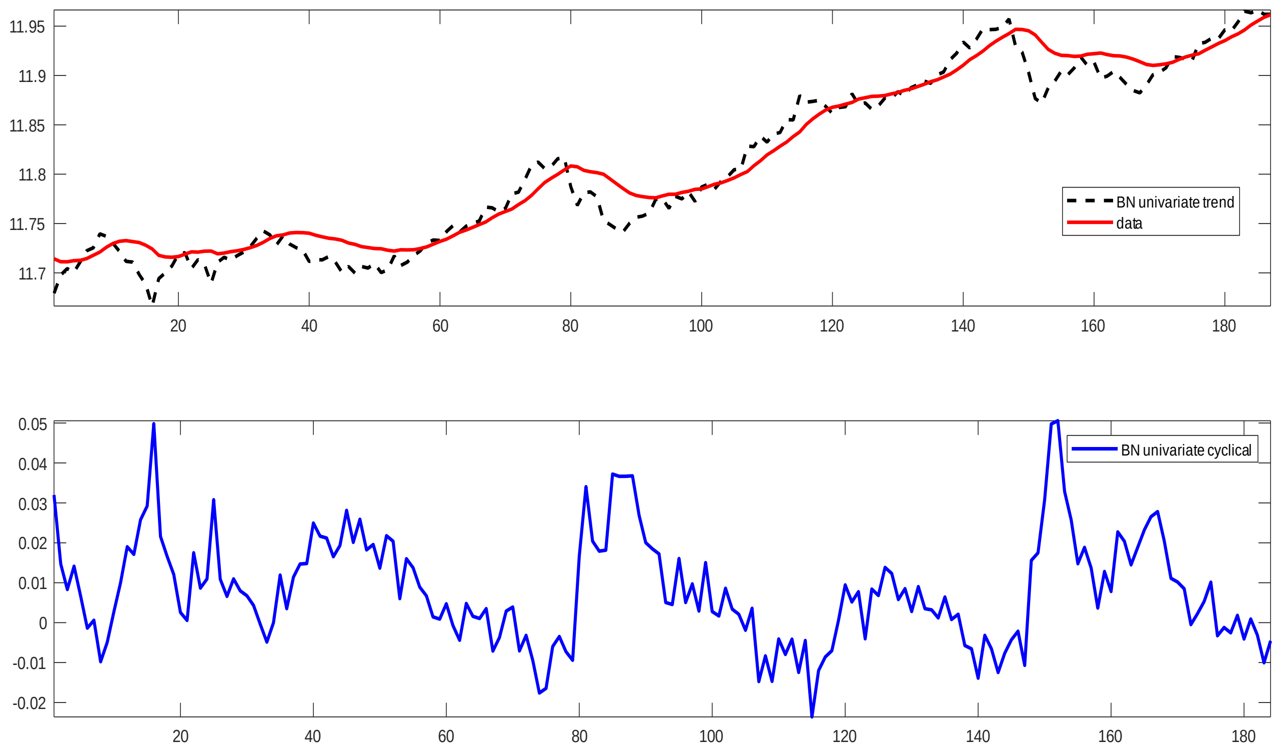
$$y_t = (\bar{y} + D(1)\sum_{j=1}^{t} e_j) + D^\dagger(\ell)e_t = y_t^x + y_t^c \tag{17}$$

- Trend and cycle perfectly correlated, $\rho = 1$.

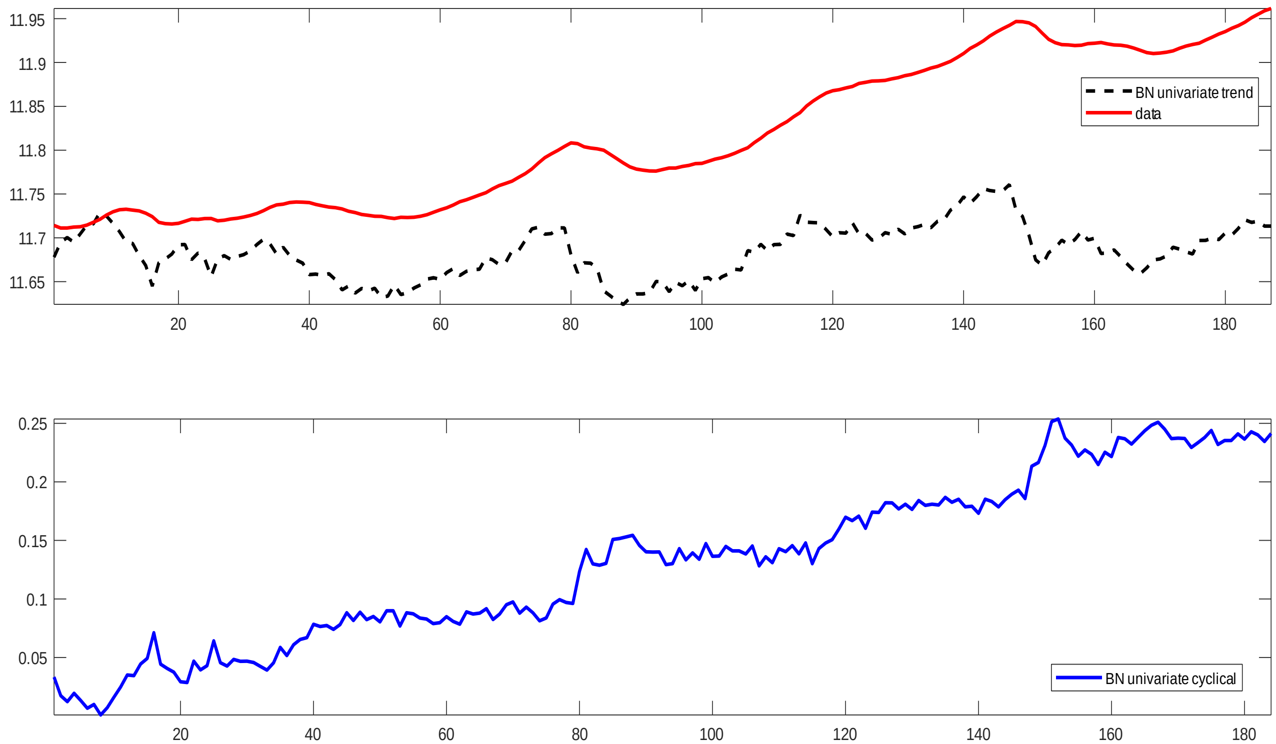
- Trend is a random walk with drift.

- Can be cast into a state space framework (see Morley, et al., 2003).

- Quality of the decomposition depends on the estimate of $\bar{y}$, the drift in the random walk.



Log Employment BN sample mean of Δy_t



Log Employment BN Long run mean $c/(1 - A(\ell))$

Multivariate Beveridge-Nelson

• Let $y_t = [\Delta y_{1t}, y_{2t}]$ $(m \times 1)$; where y_{1t} are I(1); and y_{2t} are I(0);

• Suppose $y_t = \bar{y} + D(\ell)e_t$, where $e_t \sim iid(0, \Sigma_e)$ and $D_0 = I$, the roots of $det(D(\ell))$ are equal or greater than one; and that $D_1(1) \neq 0$, where $D_1(\ell)$ is $m_1 \times 1$ (first m_1 rows of $D(\ell)$). Then

$$\begin{pmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{pmatrix} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} + \begin{pmatrix} D_1(1) \\ 0 \end{pmatrix} e_t + \begin{pmatrix} (1-\ell)D_1^{\dagger}(\ell) \\ (1-\ell)D_2^{\dagger}(\ell) \end{pmatrix} \Delta e_t \qquad (18)$$

$D_1^{\dagger}(\ell) = \frac{D_1(\ell)-D_1(1)}{1-\ell}$ $D_2^{\dagger}(\ell) = \frac{D_2(\ell)}{1-\ell}$, $0 <$ rank$[D_1(1)] \leq m_1$ and $y_t^x = [\bar{y}_1 + D_1(1)\sum_s e_s, \bar{y}_2]'$ is the permanent component of y_t.

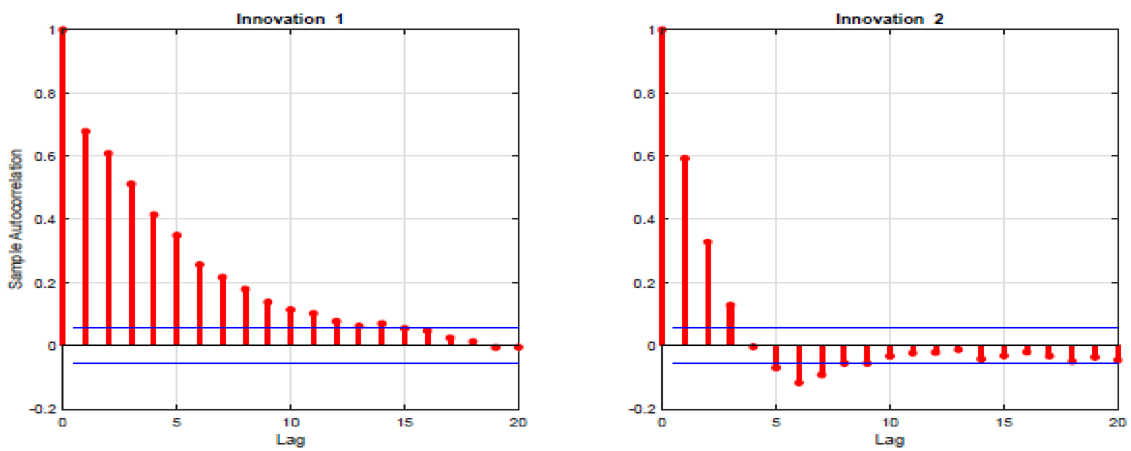
• Kambler et al., 2018: smooth BN decomposition (add penalty in the estimation)

5 How do econometricians think about cycles?

• Stationary data summarized with the autocovariance function (ACF):

$$ACF(\tau) = E_t(y_t - E_t y_t)(y_{t-\tau} - E_t y_{t-\tau}) \tag{19}$$

• ACF is symmetric, has correlated elements $(E(ACF(\tau), ACF(\tau')) \neq 0, \tau \neq \tau')$.



• Alternatively, stationary data can be summarized with spectral density:

$$\mathcal{S}(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{-\infty} ACF(\tau) e^{-i\omega\tau}, \qquad (20)$$

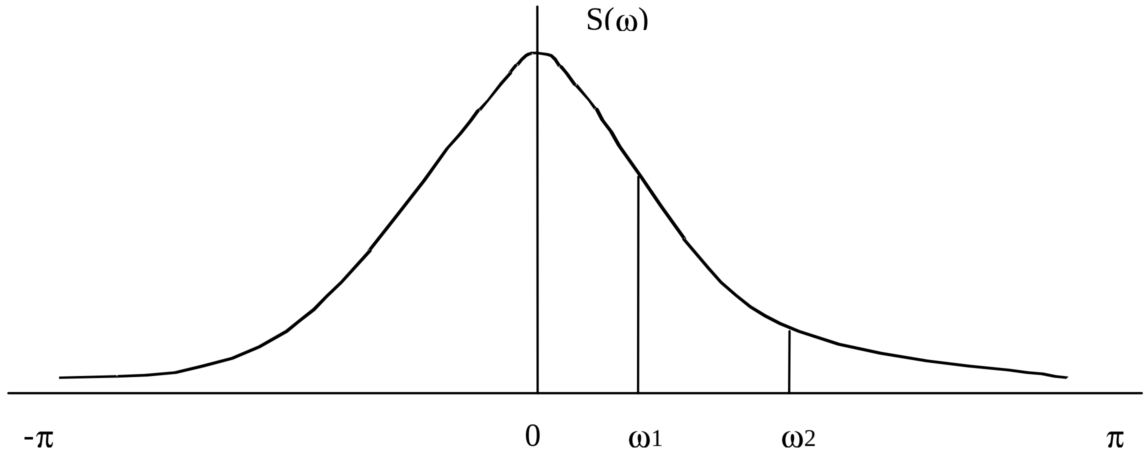
$\omega \in [0, 2\pi]$, $i = (-1)^{0.5}$, $e^{-i\omega_\tau} = \cos(\omega_\tau) - i \sin(\omega_\tau)$.

• Spectral density changes coordinates relative to ACF.

• If $S(\omega)$ is evaluated at $\omega_\tau = \frac{2\pi\tau}{T}$, $\tau = 0, \ldots, T-1$ (Fourier frequencies):

i) $\mathcal{S}(\omega_\tau) = \mathcal{S}(\omega_{-\tau})$ (symmetry around $\omega_\tau = 0$).

ii) $E(\mathcal{S}(\omega_\tau)\mathcal{S}(\omega_{\tau'})) = 0$ (uncorrelatedness at two different ω_τ's)



• Area under the spectral density ($\sum_{\omega} S(\omega)$) is the variance of the process. Given orthogonality (by i. of previous slide), can perform variance decomposition by frequencies.

• $\mathcal{S}(\omega = 0) = \sum_{\tau=-\infty}^{-\infty} ACF(\tau)$ measures of the persistence of y_t.

• If y_t has a unit root, $S_y(\omega = 0) \uparrow \infty$ and for $x_t = \Delta y_t$ $S_x(\omega = 0) = 0$.

• How do I associate a frequency ω_τ with the length of the fluctuations? The length of the fluctuations at Fourier frequency ω_τ is $p = \frac{2\pi}{\omega_\tau} = \frac{T}{\tau}$.

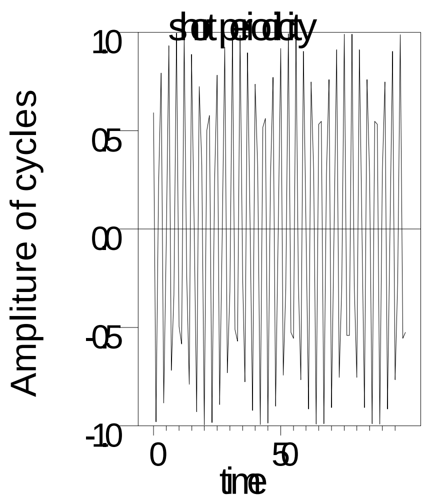
Example 5.1 $\omega_\tau = \frac{\pi}{16} \rightarrow p = 32$; $\omega_\tau = \frac{\pi}{2} \rightarrow p = 4$ *(quarters, years,etc.)*

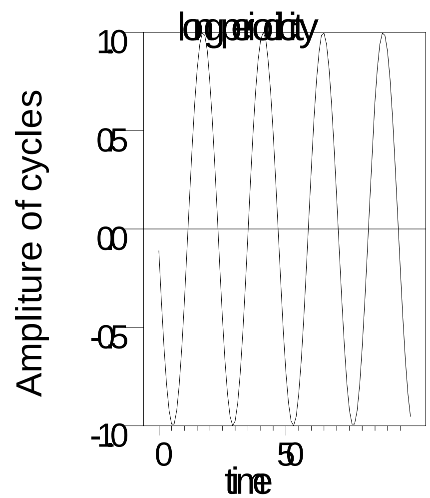
• Components of spectral density:

(1) Trends: $\omega_\tau \in (0, \omega_1)$ (low frequencies) (Not just $S(0)$).

(2) Cycles: $\omega_\tau \in (\omega_1, \omega_2)$ (cyclical frequencies)

(3) Seasonals, irregulars: $\omega_\tau \in (\omega_2, \pi)$ (high frequencies)





• Low frequencies (trends) associated with cycles featuring long periods of oscillations (time series moves infrequently from peaks to throughs).

• High frequencies (irregulars) are associated with short cycles (time series move frequently from peaks to throughs).

Multivariate analysis

• The spectral density matrix of a stationary $N \times 1$ vector $\{y_t\}_{t=-\infty}^{\infty}$ is $\mathcal{S}(\omega) = \frac{1}{2\pi} \sum_\tau ACF(\tau) \exp(-i\omega\tau)$ where

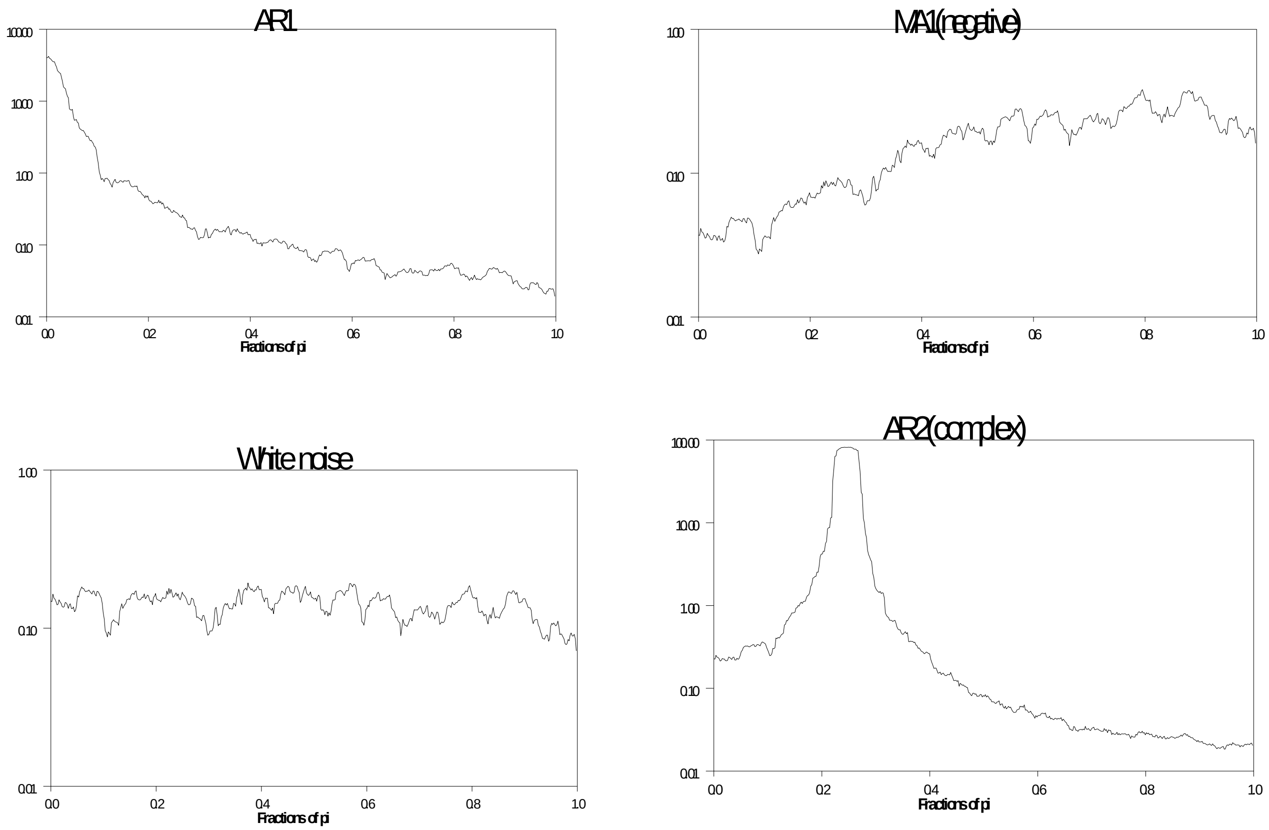
$$\mathcal{S}(\omega) = \left[\begin{array}{cccc} \mathcal{S}_{y_1 y_1}(\omega) & \mathcal{S}_{y_1 y_2}(\omega) & \dots & \mathcal{S}_{y_1 y_m}(\omega) \\ \mathcal{S}_{y_2 y_1}(\omega) & \mathcal{S}_{y_2 y_2}(\omega) & \dots & \mathcal{S}_{y_2 y_m}(\omega) \\ \dots & \dots & \dots & \dots \\ \mathcal{S}_{y_N y_1}(\omega) & \mathcal{S}_{y_N y_2}(\omega) & \dots & \mathcal{S}_{y_N y_N}(\omega) \end{array} \right]$$

• Diagonal of the spectral density matrix real; off-diagonal complex.

• The coherence between y_{it} and y_{jt} is $Co_{y_i,y_j}(\omega) = \frac{|\mathcal{S}_{y_i,y_j}(\omega)|}{(\mathcal{S}_{y_i,y_i}(\omega)\mathcal{S}_{y_j,y_j}(\omega))^{0.5}}$.

• It measures the strength of the association between y_{it}, y_{jt} at frequency ω. $\int Co(\omega) d\omega = \rho_{y_1,y_2}$: decomposition of correlation by frequency. $Co(\omega)$ is real since $|y| =$ real part of complex number y.

Examples of univariate spectral densities



Conclusions

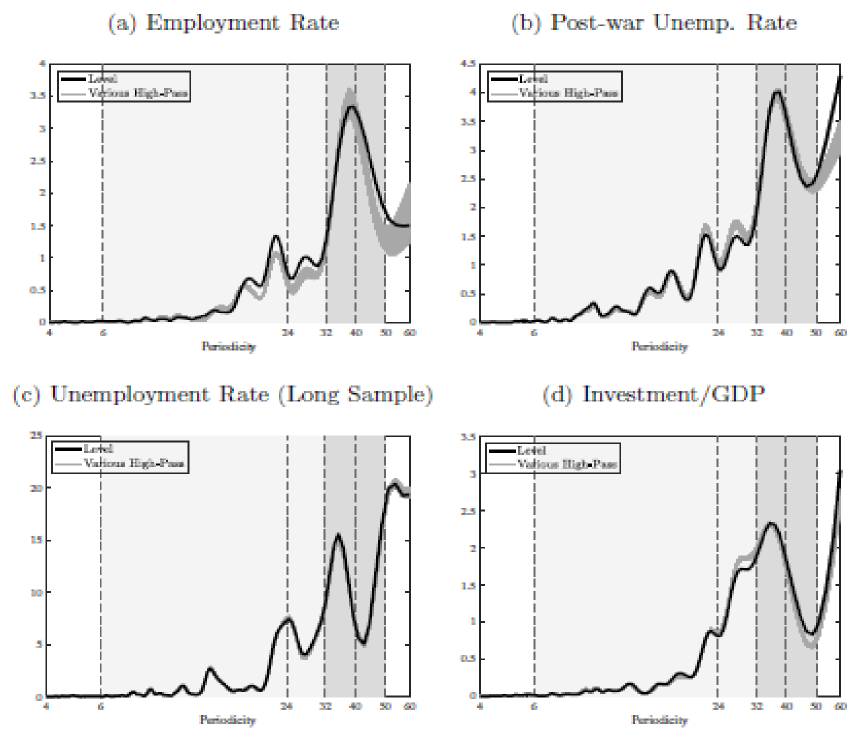
- Displaying variability and serial correlation (e.g. AR(1) or MA(1)) does not generate cycles for econometricians.

- Alternating sequence of irregularly sparse turning points does not necessarily generate cycles for econometricians.

- Need, at least, an AR(2) with complex roots to have econometrician cycles.

- Need large coherence at certain frequencies to have y_{it} and y_{jt} comoving.

• Beaudry et al. (2019): labor market variables have econometrician cycles.



Note: cycle length reversed (left short cycle, right long cycles).

Filters

- Spectral densities defined only for stationary series.

- Interested in variability at certain frequencies (Why? Electrical engineers arguments?)

- Filters may make y_t stationary under certain assumptions.

- Filters eliminate variability at certain frequencies.

- Two birds with one stone!

• A filter is a linear transformation of a primitive stochastic process y_t.

$$y_t^f = \sum_{-J}^{J} \mathcal{B}_j y_{t-j} = \mathcal{B}(\ell) y_t \tag{21}$$

• The filter is symmetric if $\mathcal{B}_j = \mathcal{B}_{-j}$. Symmetric filters have the property that the timing of the cycles in y_t and y_t^f is the same (zero phase shift).

• If $\sum_{-J}^{J} \mathcal{B}_j = 0$ and y_t is non-stationary, y_t^f is stationary (filtering de-trends/stationarize time series with unit roots).

• Two MA filters

1) $y_t^f = y_t + D_1 y_{t-1}$.

2) $y_t^f = \sum_{j=-J}^{J} y_{t-j}$. The larger is J the smoother is y_t^f.

• If $CGF_y(z)$ is the covariance generating function of y_t, and $y_t^f = \mathcal{B}(\ell) y_t$, then $CGF_{y^f}(z) = \mathcal{B}(z)\mathcal{B}(z^{-1}) CGF_y(z)$. When univariate $CGF_{y^f}(z) = |\mathcal{B}(z)|^2 CGF_y(z)$, where $|\mathcal{B}(z)|$ is the real part (modulus) of $\mathcal{B}(z)$.

Example 5.2 *Let e_t be a white noise. Its spectrum is $\mathcal{S}_e(\omega) = \frac{\sigma^2}{2\pi}$ (this is the CGF for $z = e^{i\omega}$). Let $y_t = a(\ell) e_t$ where $a(\ell) = a_0 + a_1 \ell + a_2 \ell^2 + \ldots$. The spectrum of y_t is $\mathcal{S}_y(\omega) = |a(e^{-i\omega})|^2 \mathcal{S}_e(\omega)$, where $|a(e^{-i\omega})|^2 = a(e^{-i\omega}) a(e^{i\omega})$.*

Terminology

• The frequency response function of the filter is $\mathcal{B}(\omega) = \mathcal{B}_0 + 2\sum_j \mathcal{B}_j \cos(\omega j)$ (i.e. set $\ell^j = e^{i\omega j}$); it measures the effect of a shock in y_t on y_t^f at frequency ω (IRF in frequency domain).

• $|\mathcal{B}(\omega)|$ is the gain (transfer) function; it measures how much the **amplitude** of the fluctuations y_t^f changes relative to the amplitude of y_t at frequency ω.

• $|\mathcal{B}(\omega)|^2$ is the **squared gain**; it measures how much the **variance** of y_t^f changes relative to the variance of y_t at frequency ω.

5.1 The Hodrick and Prescott (HP) Filter

• Trends are smooth (variations are small; could be almost deterministic or stochastic). Assumption formalized in the constrained problem:

$$\min_{y_t^x}\{\sum_{t=1}^{T}(y_t - y_t^x)^2 + \lambda \sum_{t=2}^{T}((y_{t+1}^x - y_t^x) - (y_t^x - y_{t-1}^x))^2\} \tag{22}$$

If $\lambda = 0$, the solution is $y_t^x = y_t$. As $\lambda \uparrow, y_t^x$ becomes smoother. If $\lambda \to \infty$, y_t^x becomes linear (no variations). Typically: $\lambda = 1600$ for quarterly data.

• Ravn and Uhlig (2002): if $\lambda = 129000$ for monthly data and $\lambda = 6.25$ for annual data, HP filters picks cycles with similar periodicity for monthly, quarterly and annual data.

Solution to the constrained optimization:

$$\hat{y}^x = Ay = (H'H + \lambda Q'Q)^{-1} H y \qquad (23)$$
$$\hat{y}^c = y - \hat{y}^x = (I - A) y \qquad (24)$$

where $y = [y_T, \ldots, y_1]'$ is a $T \times 1$ vector, $y^x = [y^x_T, \ldots, y^x_1, y^x_0, y^x_{-1}]'$ is a $(T+2) \times 1$ vector, $H = [I,\ 0]$ where I is a $T \times T$ identity matrix and 0 a $T \times 2$ matrix of zeros and

$$Q = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & -2 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & -0 & 1 & -2 & 1 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 & 1 \end{bmatrix}$$

- (23) is a "ridge" estimator (typically used for multicollinearity problems).

- Bayesian interpretation: $\Delta^2 y^x_{t-1} = \epsilon_t$ is a prior with $\epsilon_t \sim N(0, \lambda * \sigma^2_c)$.

• Alternative (UC) setup:

$$\begin{aligned} y_t &= y_t^x + y_t^c \\ \Delta y_t^x &= \epsilon_t \end{aligned} \tag{25}$$

where both ϵ_t and y_t^c are white noise, uncorrelated with y_0^x, y_{-1}^x. Two solutions (see literature on curve fitting, e.g. Wabha, 1980).

i) If $C_0^{-1} = var(y_0^x, y_{-1}^x)^{-1} \longrightarrow 0$, find a_t such that $\tilde{y}_t^x = a_t y_t$ by min $E(y_t^x - a_t y_t)^2$. Solution: $\tilde{y}^x = E(y^x y')E(yy')^{-1}y = \tilde{A}y$. If $\lambda = \frac{\sigma_c^2}{\sigma_\epsilon^2}$, then $A = \tilde{A}$.

ii) (25)) is a state space system. Can use the Kalman smoother to solve the signal extraction problem (still assuming large C_0).

• $\lambda = 1600$ means σ_c, the standard deviation of the cycle, is 40 times larger than σ_ϵ the standard deviation of the second difference of the trend.

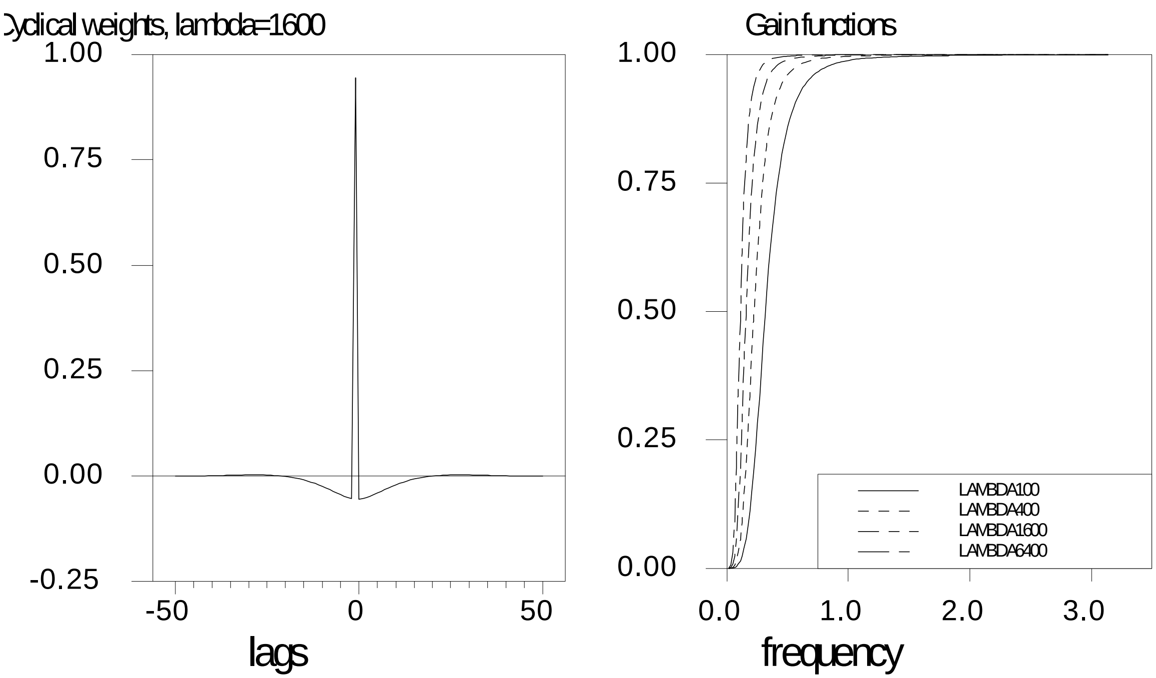
• HP Solution is optimal when the cycle is a white noise.

• HP Solution: is time dependent (the cycle at t depends on how large is T). Beginning and end-of-sample problems.

• Premultiplying (23) by $(H'H + \lambda Q'Q)^{-1}$ and letting T grow to infinity one can show that $y_t^c = \mathcal{B}^c(\ell) y_t$, where

$$\mathcal{B}^c(\ell) \simeq \frac{(1-\ell)^2(1-\ell^{-1})^2}{\frac{1}{\lambda} + (1-\ell)^2(1-\ell^{-1})^2} \tag{26}$$

- When $\lambda = 1600$, $\mathcal{B}_j^c$ and the $|\mathcal{B}^c(\omega)|^2$ looks like in the picture below.



- Properties of HP filter:

(i) It eliminates linear and quadratic trends from y_t.

(ii) Stationarize y_t with up to 4 unit roots (King and Rebelo, 1993).

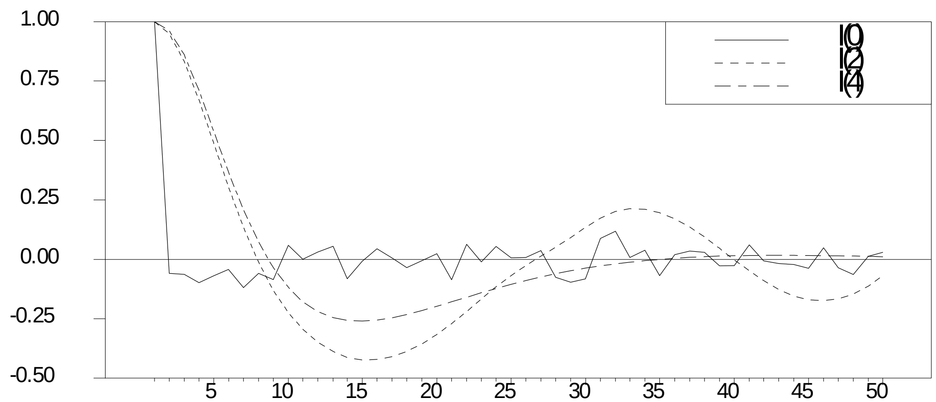
- What happens if y_t has less than 4 unit roots? Overdifferencing.
- **HP filter may create spurious autocorrelation in** y_t^f (Slutzky effect).

- Intuition: $y_t = e_t \sim iid(0, \sigma^2)$. Then

$$\begin{aligned} \Delta y_t &= e_t - e_{t-1} \quad \text{correlation of order 1} \\ \Delta^2 y_t &= e_t - e_{t-1} - (e_{t-1} - e_{t-2}) \quad \text{correlation of order } 2, \text{etc.} \end{aligned}$$

- Differencing a stationary y_t induces spurious serial correlation.

Example 5.3 *Let y_t be I(2) or I(4). Pass them through a HP filter. The figure plots the ACF of y_t^f. The serial correlation in filtered I(2) higher then in the filtered I(4).*



- **It can create spurious variability in the filtered data**.

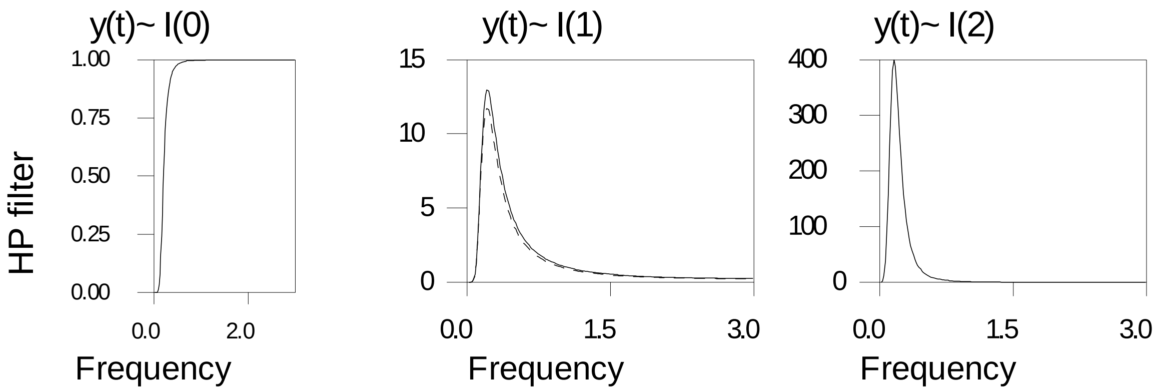
- If y_t is stationary, the squared gain function is:

$$\mathcal{B}^c(\omega) \simeq \frac{16\sin^4(\frac{\omega}{2})}{\frac{1}{\lambda} + 16\sin^4(\frac{\omega}{2})} = \frac{4(1-\cos(\omega))^2}{\frac{1}{\lambda} + 4(1-\cos(\omega))^2}$$

:

- It damps fluctuations with periodicity $\geq$ 24-32 quarters per cycle, it passes short cycles without changes.

- If y_t is I(1) $\mathcal{B}^c(\ell)$ is a combination of two filters: $(1-\ell)$ makes y_t stationary, $\frac{\mathcal{B}(\ell)}{1-\ell}$ filters Δy_t. When $\lambda = 1600$ the gain function of $\frac{\mathcal{B}(\ell)}{1-\ell}$ is $\simeq 2(1-\cos(\omega))B(\omega)$, which peaks at $\omega^* = \arccos[1 - (\frac{0.75}{1600})^{0.5}] \simeq 30$ periods:



• If y_t is I(1) HP damps long and short run growth cycles and amplifies business cycle frequencies (e.g. the variance of the cycles with average duration of 7.6 years is multiplied by 13).

• Problem even larger if y_t is I(2).

• Same problem if y_t nearly integrated ($\rho_y = 0.95$)? (see dotted line)

- What is the intuition for the increased variability?

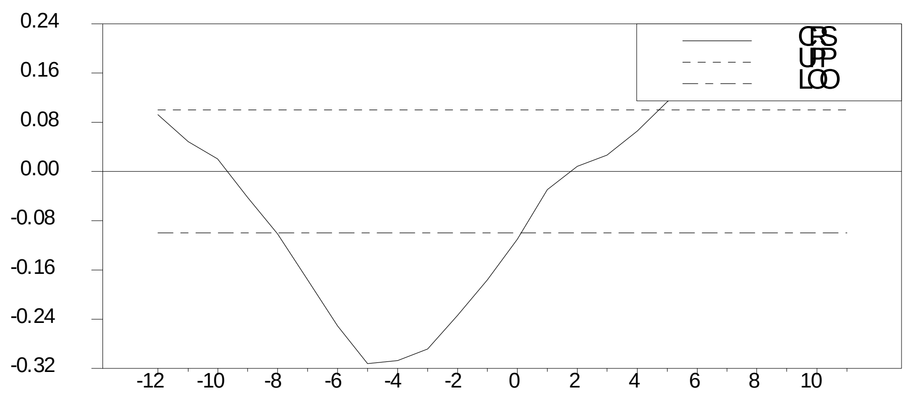
- Suppose $\Delta y_t = e_t \sim iid(0, \sigma^2)$. Then

$$\begin{aligned} var(\Delta^2 y_t) &= var(e_t - e_{t-1}) = var(e_t) + var(e_{t-1}) = 2\sigma^2 \\ var(\Delta^3 y_t) &= var(e_t - 2e_{t-1} + e_{t-2}) = 4\sigma^2 \end{aligned}$$

etc.. So the filter $\frac{\mathcal{B}(\ell)}{1-\ell}$ can augment the variability of Δy_t.

- **It can produce spurious comovements among series.**

Example 5.4 *y_{1t} and y_{2t} are two uncorrelated random walks. Pass them through a HP filter. The figure plots the cross correlation function of y_{1t}^c, y_{2t}^c and a 95 percent asymptotic tunnel for the hypothesis of no correlation.*



- What is the intuition for this result?

The two filtered series have similar spectrum. Therefore, it is possible that they go up and down together (Note: this does not happen all the times).

- **Conclusions: The HP filter has the potential to generate spurious variability, spurious serial and cross variable correlations**

Properties of HP filter (continue)

iii) It leaves high frequency variability unchanged (high pass filter).

iv) HP cyclical component predicts the future. Alternative to (26):

$$y_t^c = \frac{\lambda(1-\ell)^4}{1+\lambda(1-\ell)^2(1-\ell^{-1})^2} y_{t+2} \tag{27}$$

v) $\lambda = 1600$ inconsistent with KF estimates of $\sigma_c^2, \sigma_\epsilon^2$ and UC setups.

Table 1. Maximum likelihood estimates of parameters of state-space formalization of the HP filter for assorted quarterly macroeconomic series.

	σ^2_c	σ^2_v	λ
GDP	0.115	0.468	0.245
Consumption	0.163	0.174	0.940
Investment	4.187	12.196	0.343
Exports	5.818	3.341	1.741
Imports	4.423	4.769	0.927
Government spending	0.221	1.160	0.191
Employment	0.006	0.250	0.023
Unemployment rate	0.014	0.092	0.152
GDP Deflator	0.018	0.081	0.216
S&P 500	21.284	15.186	1.402
10-year Treasury yield	0.135	0.054	2.486
Fed Funds Rate	0.633	0.116	5.458
Real Rate	0.875	0.091	9.596

vi) Two-sided filter (do not use y_t^c in VARs!).

vii) Cross county comparisons difficult because cycles may have different length. Marcet-Ravn (2000) solve

$$\min_{y_t^x} \quad \sum_{t=1}^{T} (y_t - y_t^x)^2 \tag{28}$$

$$\mathcal{V} \geq \frac{\sum_{t=1}^{T-2} (y_{t+1}^x - 2y_t^x + y_{t-1}^x)^2}{\sum_{t=1}^{T} (y_t - y_t^x)^2} \tag{29}$$

where $\mathcal{V} \geq 0$ is a constant to be chosen by the researcher, $\mathcal{V}$ measures the relative variability of the acceleration in the trend and the cycle, and may be country specific.

Example 5.5 *200 data points from a stationary RBC model with utility* $U(c_t, c_{t-1}, N_t) = \frac{c_t^{1-\varphi}}{1-\varphi} + log(1 - N_t)$ *assuming* $\beta = 0.99, \varphi_c = 2.0, \delta = 0.025, \eta = 0.64$*, steady state hours equal to 0.3,* $\rho_\zeta = 0.9, \rho_g = 0.8, \sigma_\zeta = 0.0066, \sigma_g = 0.0146$*. Table reports average unconditional moments across 100 simulations, before and after HP filtering.*

Simulated statistics

	Raw			*HP filtered*		
	K	*W*	*LP*	*K*	*W*	*LP*
cross (GDP_t, x_t)	*0.49*	*0.65*	*0.09*	*0.84*	*0.95*	*-0.20*
cross (GDP_{t+1}, x_t)	*0.43*	*0.57*	*0.05*	*0.60*	*0.67*	*-0.38*
St. Dev	*1.00*	*1.25*	*1.12*	*1.50*	*0.87*	*0.50*

5.2 One sided HP filter

• The HP-filter is two-sided and thus not very useful for real analysis and forecasting. In addition, by construction, y_t^c artificially predicts the future.

• There is a version of the HP filter which is one-sided and does not feature future predictability.

• The trend and the cycle can be estimated with standard Kalman filter/ EM algorithm iterations, MCMC, or by serial implementation.

- The model is:

$$y_t = y_t^x + y_t^c \tag{30}$$

$$y_t^x = 2y_{t-1}^x - y_{t-2}^x + \epsilon_t \tag{31}$$

where ϵ_t, y_t^c are white noise sequences.

- State space representation (see Stock and Watson, 1999):

1. State Equation

$$\begin{bmatrix} y^x_{t|t} \\ y^x_{t-1|t} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y^x_{t-1|t-1} \\ y^x_{t-2|t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ 0 \end{bmatrix} \tag{32}$$

2. Observation Equation

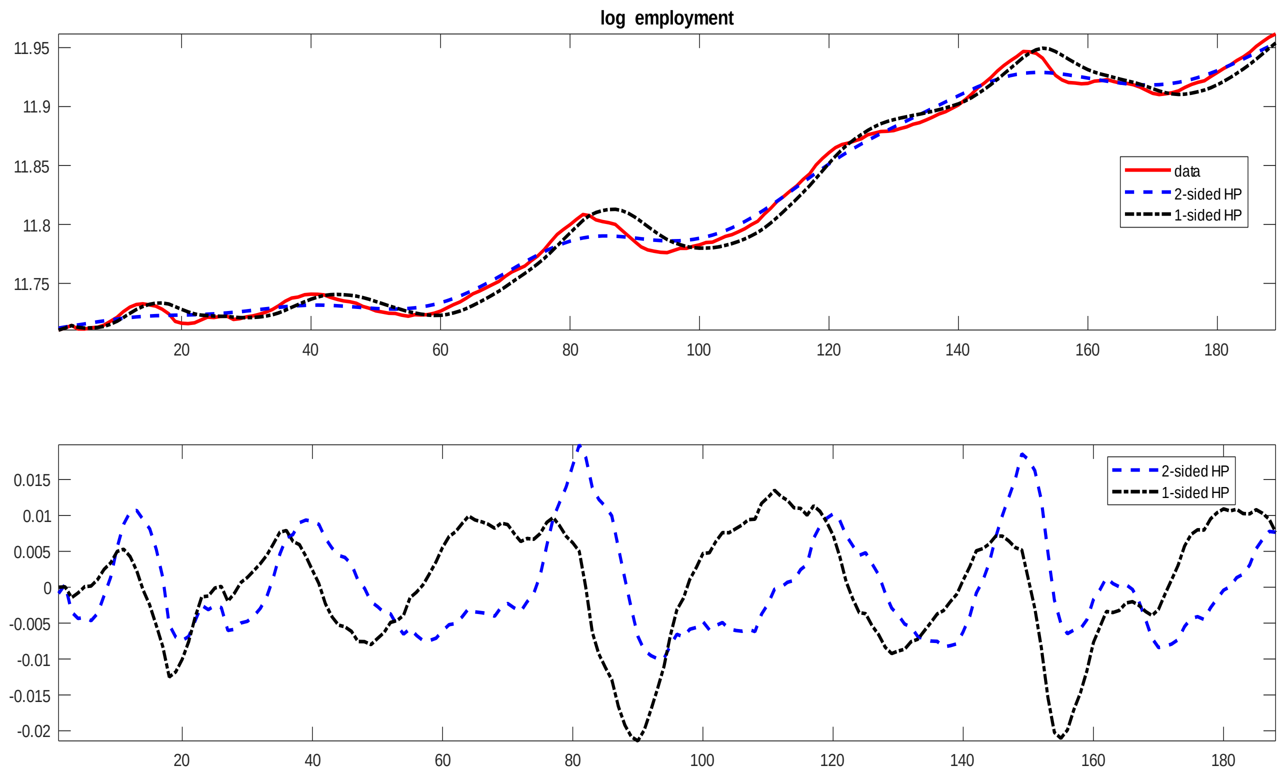
$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} y^x_{t|t} \\ y^x_{t-1|t} \end{bmatrix} + \begin{bmatrix} y^c_t \\ 0 \end{bmatrix} \tag{33}$$

- Can restrict $\lambda = \frac{\sigma_c^2}{\sigma_\epsilon^2}$ with a prior, e.g. $\lambda \sim N(1600, 10)$.

- Serial implementation (Meyer-Gohde, 2010).

- Much faster than KF; gives almost identical results.

- $\{y_t^x\}_{t=1}^T$ is obtained calculating for each t the standard HP filtered trend using data up to that t and equating y_t^x with trend value for period t (i.e. compute T two-sided HP filters trends).



Log Employment: one and two sided HP

5.3 L1-HP filter

- Standard problem:

$$\min_{y_t^x}\{\sum_{t=1}^{T}(y_t - y_t^x)^2 + \lambda \sum_{t=1}^{T}((y_{t+1}^x - y_t^x) - (y_t^x - y_{t-1}^x))^2\} \quad (34)$$

- L1 problem (Kim et al.,2009):

$$\min_{y_t^x}\{\sum_{t=1}^{T}(y_t - y_t^x)^2 + \lambda \sum_{t=1}^{T}|(y_{t+1}^x - y_t^x) - (y_t^x - y_{t-1}^x)|\} \quad (35)$$

- Same features as standard HP.

- Non-linear filter.

- Gives rise to piecewise linear segments:

$$y_t^x = a_k + b_k t, \quad t_k \leq t \leq t_{k+1}, \quad k = 1, \ldots, p-1 \tag{36}$$

and

$$a_k + b_k t_{k+1} = a_{k+1} + b_{k+1} t_{k+1} \quad k = 1, \ldots, p-1 \tag{37}$$

- p is the number of break points where the estimated trend changes slope.

- The number of break points in y_t^x typically decreases as λ increases.

- Used in (business) finance to signal "changes in market trends".

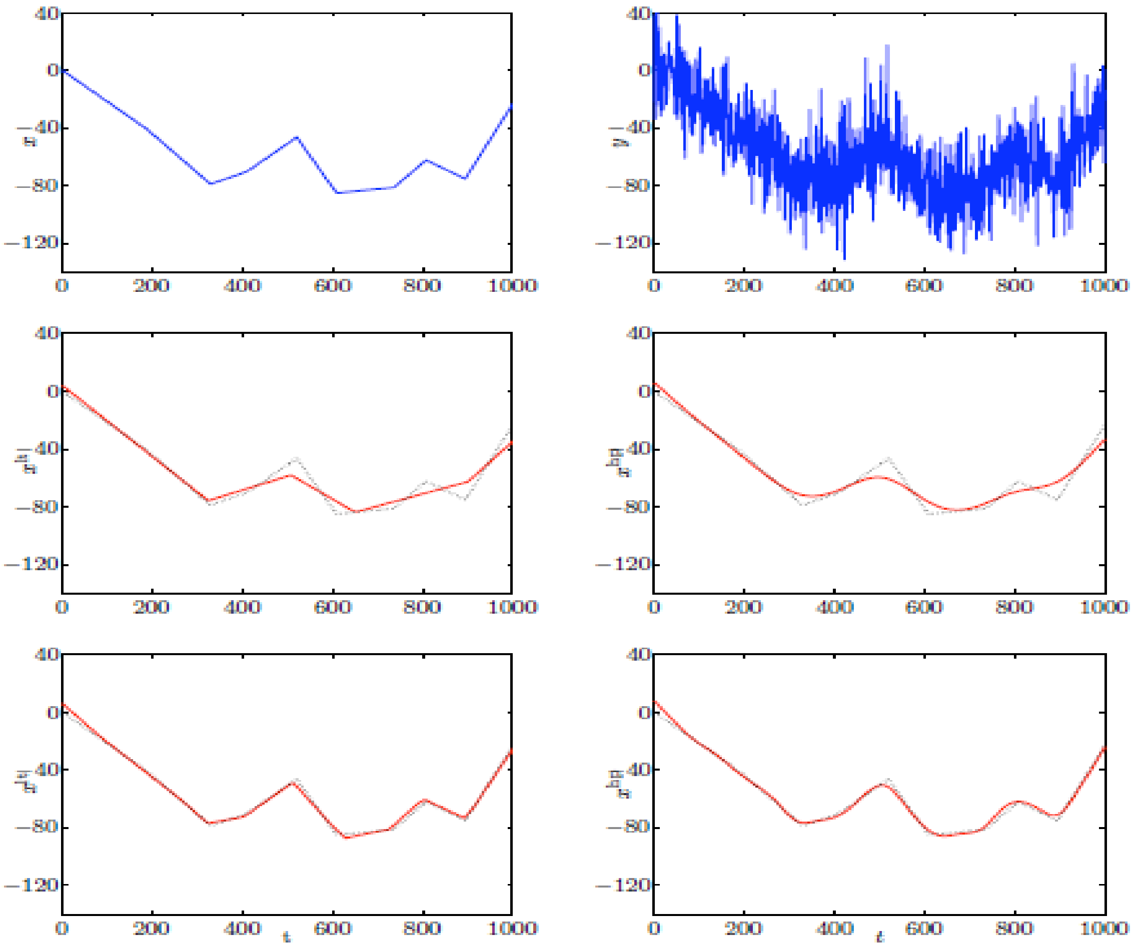


Fig. 1 *Trend estimation on synthetic data. Top left: The true trend x_t. Top right: Observed time series data y_t. Middle left: ℓ_1 trend estimate x^{lt} with four total kinks ($\lambda = 35000$). Middle right: H-P trend estimate x^{hp} with same fitting error. Bottom left: x^{lt} with seven total kinks ($\lambda = 5000$). Bottom right: H-P trend estimate x^{hp} with same fitting error.*

5.4 Other MA filters.

$$y_t^f = \sum_{-J}^{J} \mathcal{B}_j y_{t-j} = \mathcal{B}(\ell) y_t \tag{38}$$

• Symmetric MA filters ($\mathcal{B}_j = \mathcal{B}_{-j}$) with $\lim_{J\to\infty} \sum_{-J}^{J} \mathcal{B}_j = 0$ preferred because they maintain lead/lag relationships and eliminate unit roots.

• HP is a symmetric, truncated MA filter. Other filters?

Example 5.6 *A symmetric (truncated) MA filter:* $\mathcal{B}_j = \frac{1}{2J+1}$, $0 \leq j \leq |J|$ *and* $\mathcal{B}_j = 0, j > |J|$. *If* $y_t^c = (1 - \mathcal{B}(\ell))y_t \equiv \mathcal{B}^c(\ell)y_t$ *the cyclical weights are* $\mathcal{B}_0^c = 1 - \frac{1}{2J+1}$ *and* $\mathcal{B}_j^c = \mathcal{B}_{-j}^c = -\frac{1}{2J+1}$, $j = 1, 2 \ldots, J$.

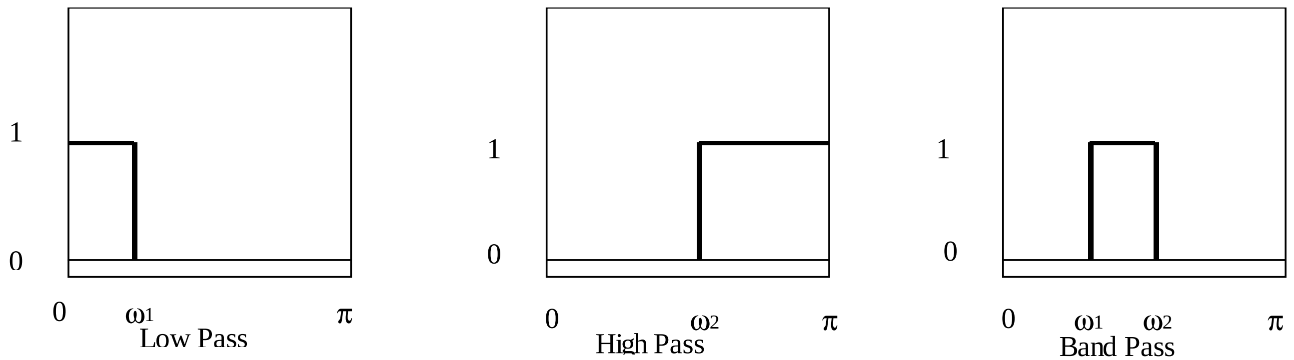
Band Pass (BP) Filters

- Combination of high pass and low pass MA filters.

- Low pass filter: $\mathcal{B}(\omega) = 1$ for $|\omega| \leq \omega_1$ and 0 otherwise.

- High pass filter: $\mathcal{B}(\omega) = 0$ for $|\omega| \leq \omega_1$ and 1 otherwise.

- Band pass filter: $\mathcal{B}(\omega) = 1$ for $\omega_1 \leq |\omega| \leq \omega_2$ and 0 otherwise.



Time series representation of the weights of the filters:

Low pass: $\mathcal{B}_0^{lp} = \frac{\omega_1}{\pi}$; $\mathcal{B}_j^{lp} = \frac{sin(j\omega_1)}{j\pi}$; $0 < j < \infty$, some ω_1.

High pass: $\mathcal{B}_0^{hp} = 1 - \mathcal{B}_0^{lp}$; $\mathcal{B}_j^{hp} = -\mathcal{B}_j^{lp}$; $0 < j < \infty$.

Band pass: $\mathcal{B}_0^{bp} = \mathcal{B}_j^{lp}(\omega_2) - \mathcal{B}_j^{lp}(\omega_1)$; $0 < j < \infty, \ \omega_2 > \omega_1$.

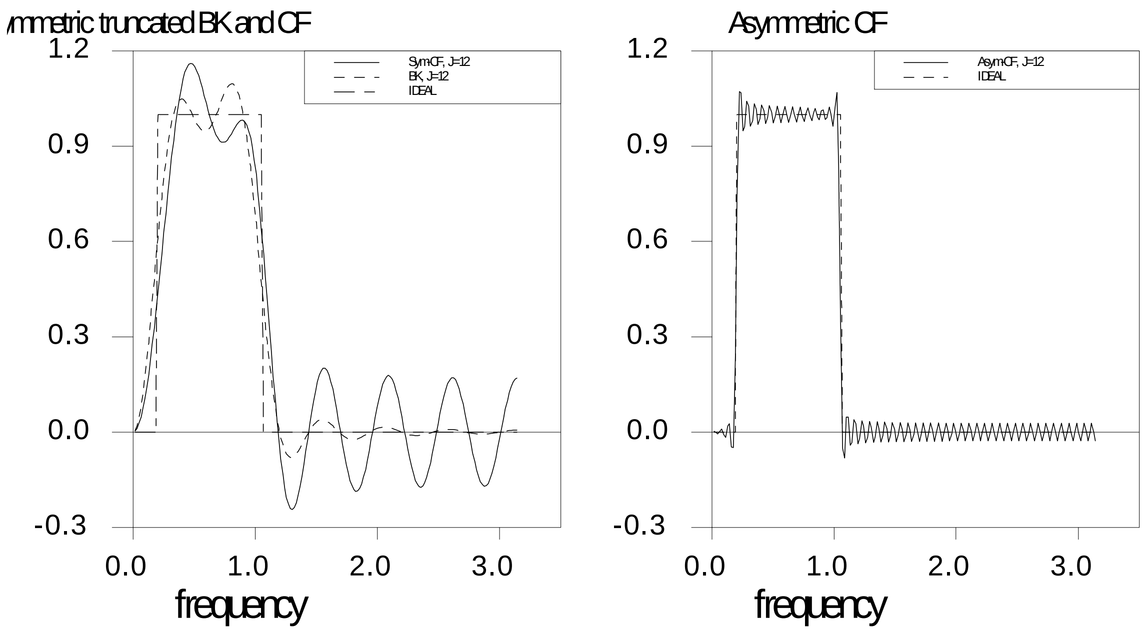
- j must go to infinity. Hence, these filters are **not realizable** for T finite.

• Baxter and King (1994): for finite T, cut at some $\bar{J} < \infty$.

• If the filter is symmetric and $\sum_{-\bar{J}}^{\bar{J}} \mathcal{B}_J = 0$ a truncated BP makes stationary series with quadratic trends and with up to two unit roots.

• BK approximation has the same problems of HP filter if y_t is (nearly) integrated.

• J needs to be large for the approximation to be good, otherwise leakage and compression.



mmetric truncated BK and CF
SymCF, J=12
BK, J=12
IDEAL
1.2
0.9
0.6
0.3
0.0
-0.3
0.0
1.0
2.0
3.0
frequency
Asymmetric CF
AsymCF, J=12
IDEAL
1.2
0.9
0.6
0.3
0.0
-0.3
0.0
1.0
2.0
3.0
frequency

• Christiano and Fitzgerald (2003): use a **non-stationary, asymmetric** approximation which is optimal in the sense of making the approximation error as small as possible.

• MA coefficients depend on t and change magnitude and even sign.

• Better spectral properties (see picture) but:

a) Need to know the properties of time series before taking the approximation (need to know if it is a I(0) or I(1)).

b) Phase shifts may occur.

• Christiano and Fitzgerald approximation is the same as Baxter and King if y_t is a white noise. In general, they will differ at the beginning and end of the sample.

5.5 Wavelets filter

Similar idea as BP filters but:

- Implementation is in time domain and one-sided MA.

- Size of the MA window adjusted depending on the cycles one wants to extract.

- Can be used on stationary and non-stationary series.

- Implementation: Haar wavelet filter (see Lubik et al., 2019).

$$y_t = \sum_{j=1}^{J} D_{jt} + S_{J,t} \tag{39}$$

$$D_{jt} = 1/(2^j) * \left(\sum_{i=0}^{2^{j-1}-1} y_{t-i} - \sum_{i=2^{j-1}}^{2^j-1} y_{t-i} \right) \tag{40}$$

$$S_{J,t} = 1/(2^J) * \left(\sum_{i=0}^{2^J-1} y_{t-i} \right) \tag{41}$$

- Typically $J = 6$. Low j's capture high frequency; j=3,4 business cycles and j=5 low frequencies.

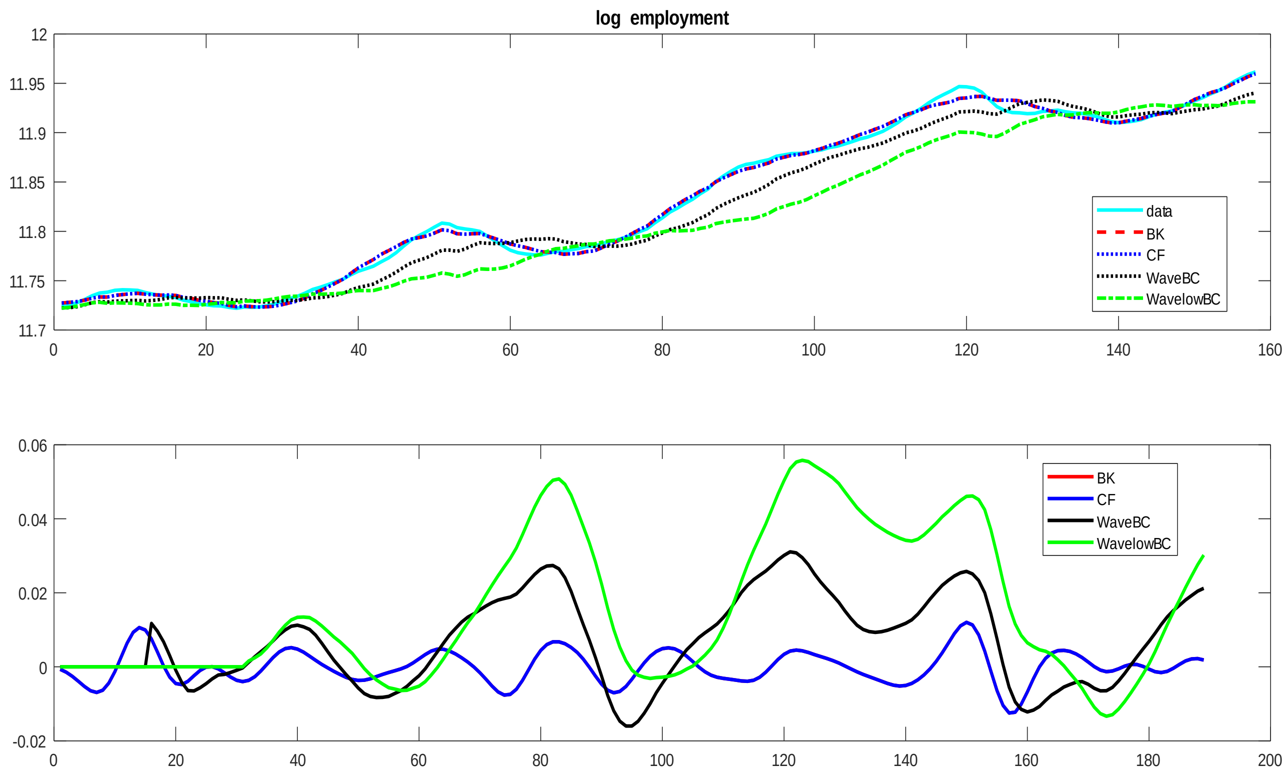
- S_{Jt} captures the long run component.

- 8-16 quarters cycles $D_{3t} = (1/8) * (y_t + y_{t-1} + y_{t-2} + y_{t-3} - y_{t-4} - y_{t-5} - y_{t-6} - y_{t-7})$.

- 16-32 quarters cycles $D_{4t} = (1/16) * (y_t + y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5} + y_{t-6} + y_{t-7} - y_{t-8} - y_{t-9} - y_{t-10} - y_{t-11} - y_{t-12} - y_{t-13} - y_{t-14} - y_{t-15})$.

- 32-64 quarters cycles $D_{5t} = (1/32) * (y_t + y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5} + y_{t-6} + y_{t-7} + y_{t-8} + y_{t-9} + y_{t-10} + y_{t-11} + y_{t-12} + y_{t-13} - y_{t-14} - y_{t-15} - y_{t-16} - y_{t-17} - y_{t-18} - y_{t-19} + y_{t-20} - y_{t-21} - y_{t-22} - y_{t-23} + y_{t-24} - y_{t-25} - y_{t-26} - y_{t-27} + y_{t-28} - y_{t-29} - y_{t-30} - y_{t-31})$.

- Window changes with the components.



Log Employment; Wavelet and BP filters

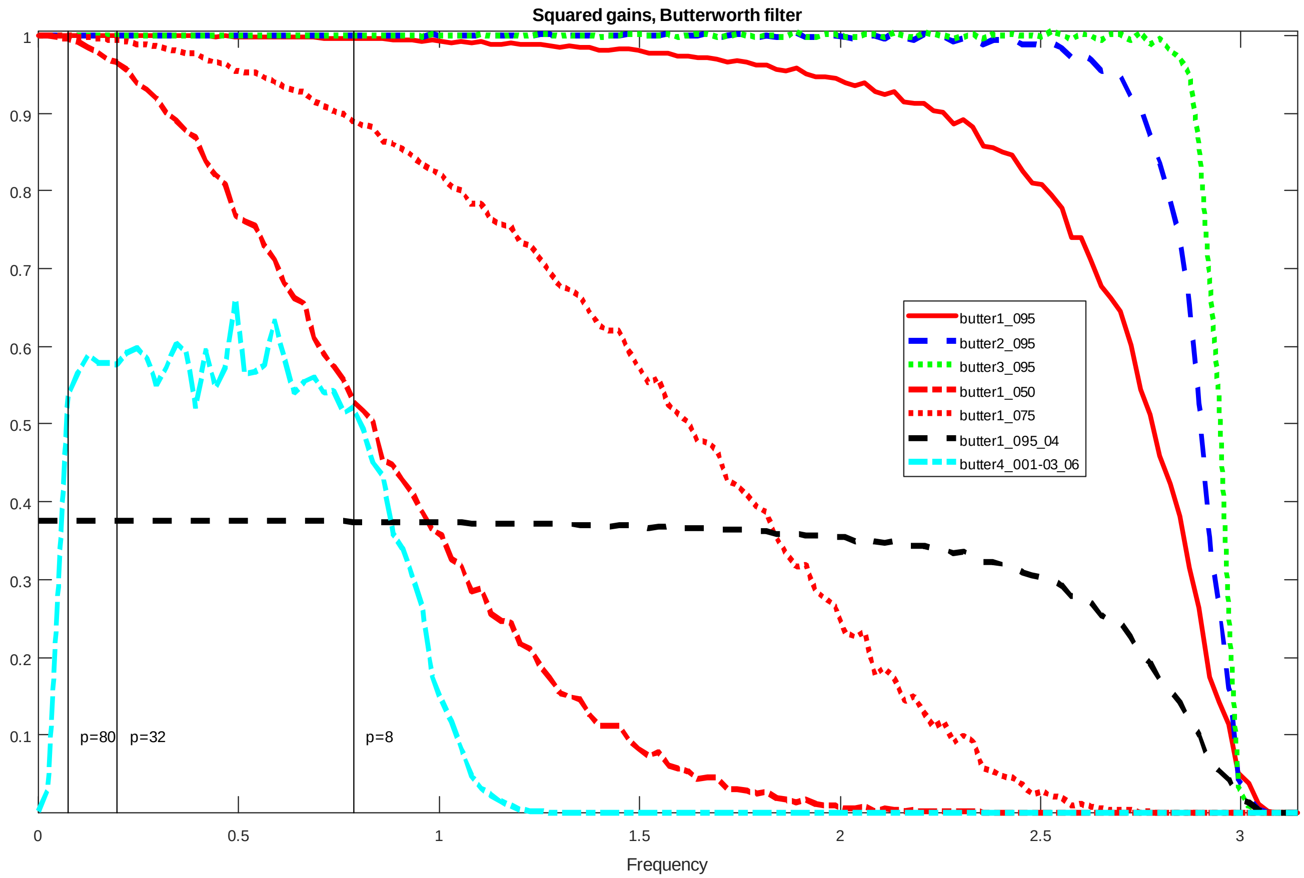
5.6 Butterworth filters

• Designed as low pass; can be adapted to high pass, band pass and even stop pass.

• Butterworth (1937): 'An ideal electrical filter should not only completely reject the unwanted frequencies but should also have uniform sensitivity for the wanted frequencies'

• Squared gain function: $G(\omega) = \frac{G_0}{1+(\frac{\omega}{\omega_c})^{2n}}$ where G_0 is the gain at the zero frequency, n is the (polynomial) order of the filter and ω is a selected frequency and ω_c a reference point (typically $\omega_c = 1$).

• Flexible. Can be designed to capture medium and low frequency variations. Can be designed to eliminate unit roots without affecting medium frequencies.



Squared gains, Butterworth filter
butter1_095
butter2_095
butter3_095
butter1_050
butter1_075
butter1_095_04
butter4_001-03_06
p=80
p=32
p=8
0
0.5
1
1.5
2
2.5
3
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Frequency

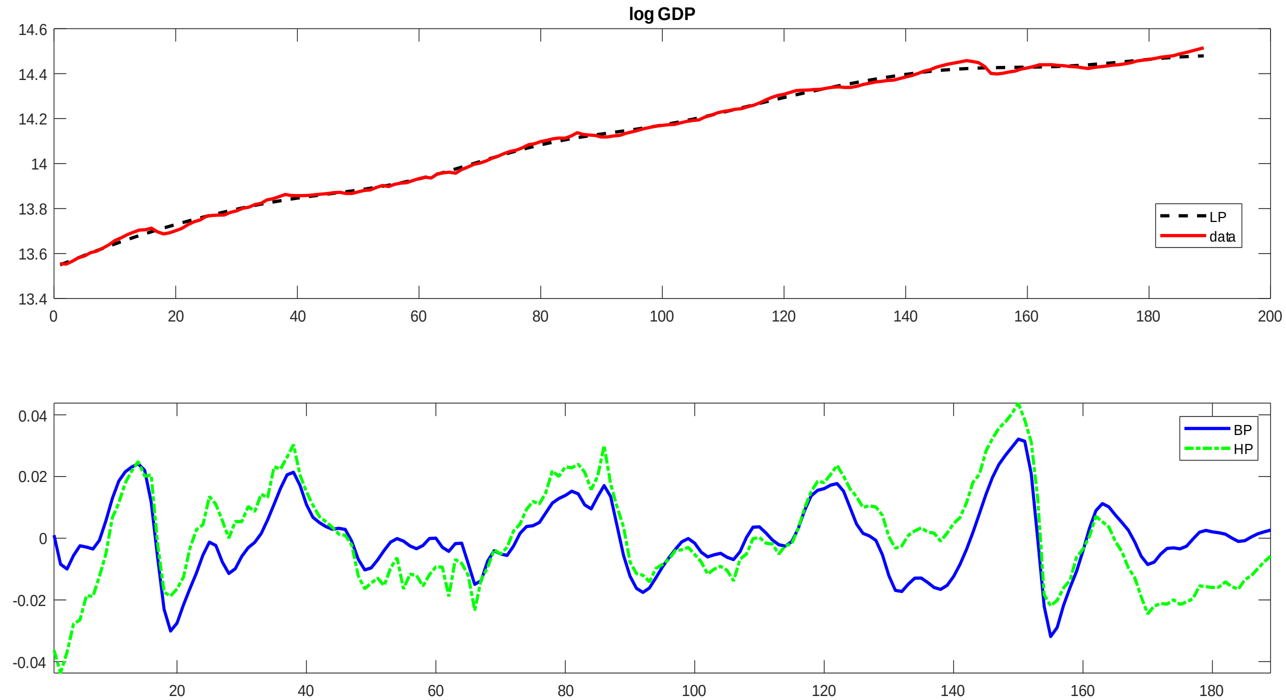
- Different decays are possible depending on n.
- Scale depends on G_0.
- Starting of decay depends on ω
- Useful to extract components wiht power at all frequencies.

Matlab commands to build a Butterworth filter

- [a,b]=butter(n,cutoff,type), where n is the degree of the polynomial, cutoff is where the squared gain falls and type could be low, high, stop-pass. If cutoff is a vector with two values, butter computes band pass weights.

- y1=filtfilt(a,b,y). Creates the filtered series using an ARMA(a,b) with y as input.

- Normalized to have $G_0 = 1$. Rescale the b coefficients to chnage G_0 (up if coefficients up you get lower squared gain).



Log output: Low pass, Band pass, High pass Butterworth filters

• Canova (2019b) BW good to extract gaps produced by economic models. How does it perform on real data?

Analytical computation of statistics

- How to compute ACF of filtered data?

- If $y_t^c = \mathcal{B}^c(\ell) y_t$ and $\mathcal{B}^c(\ell)$ known: $ACF_{y^c}(\tau) = ACF_y(0) \sum_{i=-\infty}^{\infty} \mathcal{B}_i^c \mathcal{B}_{i-\tau}^c + \sum_{\tau'=1}^{\infty} ACF_y(\tau') \sum_{i=-\infty}^{\infty} \mathcal{B}_i^c \mathcal{B}_{i-\tau'-\tau}^c + \sum_{\tau'=1}^{\infty} ACF_y(\tau')' \sum_{i=-\infty}^{\infty} \mathcal{B}_i^c \mathcal{B}_{i-\tau'-\tau}^c$.

- We need to truncate the sums at some $\bar{\imath}$, except in some special cases.

6 Economic Decompositions

- Use economic models to split y_t into unobservable components.

- Leading examples: Blanchard and Quah (1989), trend random walk, $\rho \neq 0$. King, Plosser, Stock and Watson (KPSW) (1991), cointegrated trend, $\rho \neq 0$.

- Recover permanent-transitory components (not trend/cycle: permanent may have cyclical features; not potential/gap: gap may have permanent features).

- Results sensitive to model specification and sample size.

- Example of a BQ decomposition: Fisher's model

$$gdp_t = gdp_{t-1} + a(\epsilon_t^s - \epsilon_{t-1}^s) + \epsilon_t^s + \epsilon_t^d - \epsilon_{t-1}^d \quad (42)$$

$$un_t = N_t - N^{fe} = -\epsilon_t^d - a\epsilon_t^s \quad (43)$$

$d =$ demand, $s =$ supply. This model implies that un_t has no trend; the trend in gdp_t is $gdp_t^x = gdp_{t-1}^x + a(\epsilon_t^s - \epsilon_{t-1}^s)$; and the cycle is $gdp_t^c = \epsilon_t^d - \epsilon_{t-1}^d + \epsilon_t^s$.

- Only supply shocks have long run effects on gdp_t.

- Both supply and demand shocks have cyclical effects on gdp_t.

- gdp_t^x and gdp_t^c correlated (ϵ^s drives both).

• Example of KPSW decomposition: RBC model. $y_t = [gdp_t, inv_i, C_t]$.

$$y_t = y_t^x + y_t^c \tag{44}$$

y_t^x a scalar, y_t^c a 3×1 vector. Δy_t has a MA representation

$$\Delta y_t \;=\; \bar{y} + D(\ell) e_t \tag{45}$$

• Trend component of y_t identified using $D(1)e_t = [1, 1, 1]' e_t^x$, where e_t^x is a permanent innovation (use Cholesky decomposition of $D(1)\Sigma_e D(1)'$).

• Cyclical component $y_t - y_t^x$.

• Implementation is like in multivariate BN but e_t^x is a supply (technology) disturbance (not a reduced form shock)

Alternative identification assumptions

- The BQ decomposition implicitly normalizes the variance of structural shocks to one and assumes that structural shocks are uncorrelated.

- Evidence suggests that long run and short run disturbances may be correlated, e.g. Morley et et. (2003), Grant and Chan (2017a).

- Normalization chosen my matter (Waggoner and Zha, 2003).

- Cover et al. (2003): use alternative normalization plus identification assumptions that allow demand and supply shocks to be correlated.

- Structural model ($\alpha > 0$, unitary slope AD)

$$
\begin{aligned}
gdp_t &= E_{t-1} gdp_t + \alpha (p_t - E_{t-1} p_t) + \epsilon_{1t} && (46) \\
gdp_t &= p_t + E_{t-1}(gdp_t + p_t) + \epsilon_{2t} && (47)
\end{aligned}
$$

$\epsilon_{1t}, \epsilon_{2t}$ potentially correlated; (46) is AS; (47) is AD.

- VAR : $y_t = a_0 + a(L) y_{t-1} + e_t$, $y_t = [gdp_t, p_t]'$.

- Relationship VAR-structural model

$$
\begin{aligned}
e_{1t} &= \frac{1}{1+\alpha}\epsilon_{1t} + \frac{\alpha}{1+\alpha}\epsilon_{2t} && (48) \\
e_{2t} &= -\frac{1}{1+\alpha}\epsilon_{1t} + \frac{1}{1+\alpha}\epsilon_{2t} && (49)
\end{aligned}
$$

or $e_t = B\epsilon_t$.

• Identification: i) Normalization: $\epsilon_{it}, i = 1, 2$ has a unitary effect on y_t; ii) slope of aggregate demand is unit (demand shocks may be persistent); iii) long run demand shock neutrality. i)-ii)-iii) imply:

$$\alpha = -\frac{a_{12}(1)}{1 - a_{22}(1)} \tag{50}$$

• Given (50) use (48) and (49), to recover structural shocks.

• Permanent/transitory components correlated. Permanent component:

$$y_t = a_0 + a(L)y_{t-1} + B_1\epsilon_{1t} \tag{51}$$

where B_1 is the first column of B.

• BQ setup:

$$e_{1t} = c_{11}\epsilon_{1t} + c_{12}\epsilon_{2t} \tag{52}$$

$$e_{2t} = c_{21}\epsilon_{1t} + c_{22}\epsilon_{2t} \tag{53}$$

Identification: $\sigma^2_{\epsilon_i} = 1; \sigma_{\epsilon_1,\epsilon_2} = 0$; long run demand shock neutrality

$$c_{12}(1 - a_{22}(1)) + c_{22}a_{12}(1) = 0 \tag{54}$$

• Given (54), use (52), (53) to get the structural shocks (3 unknowns in 3 moments).

• Permanent and transitory components correlated because supply shocks drive both even if supply and demand shocks are uncorrelated.

6.1 Are BQ estimates robust?

• Coibion et al (2018): BQ estimates of output gaps only depend on supply shocks. Traditional estimates depend on both supply and demand shocks.

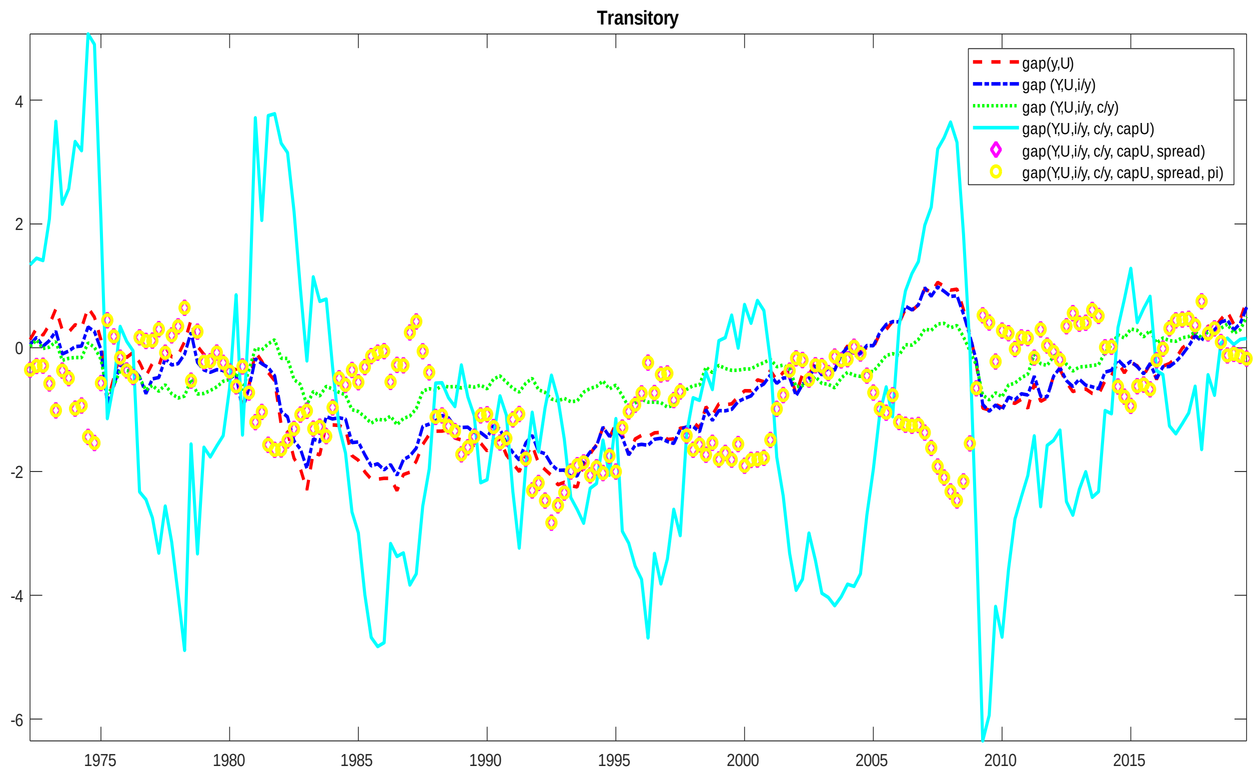
• Canova and Ferroni (2019): VAR estimates subject to *deformation*. Estimates and inference about latent variables may depend on the VAR model used.

• Deformation occurs if the DGP has more shocks than the variables of the VAR.

• Cross sectional and time deformation could be present.

• Run a VAR with US output growth and unemployment. Compute the permanent and the transitory components of output.

• Add to the VAR:
- investment/output ratio
- consumption/output ratio
- capacity utilization
- term spread (10 years bond rate -call rate)
- inflation

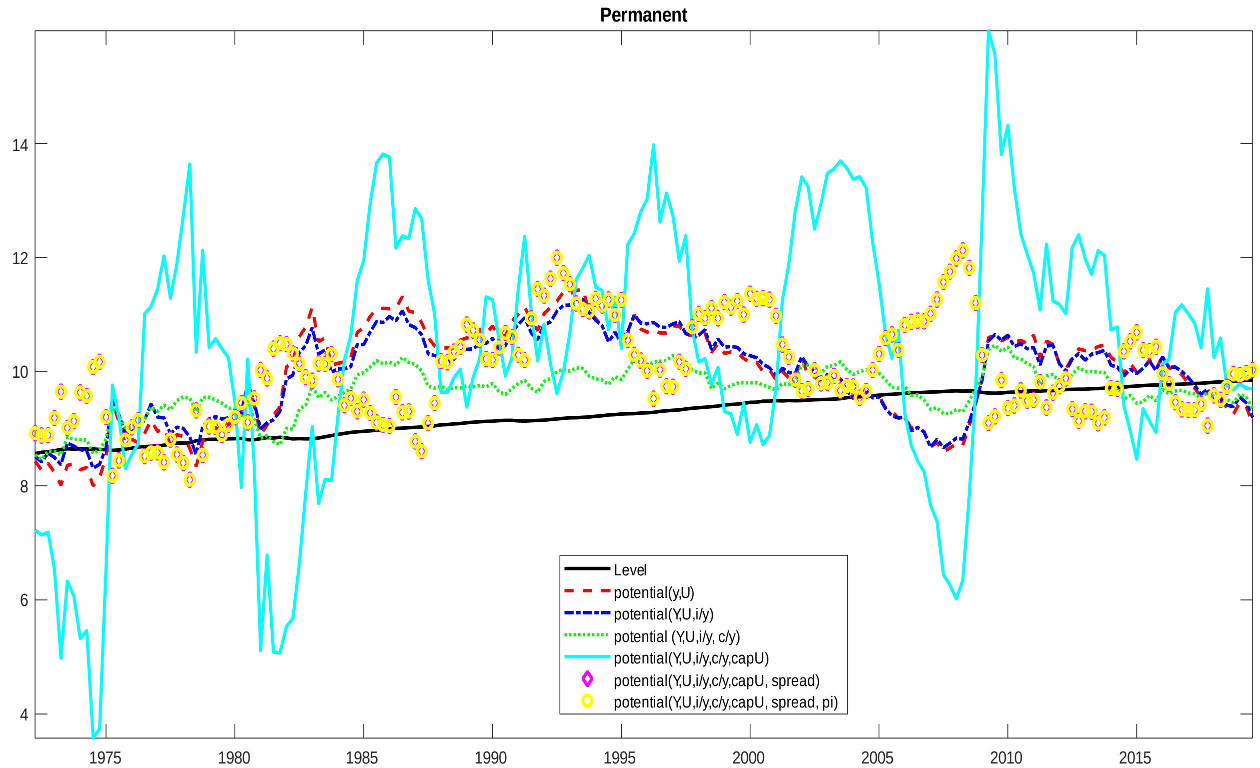


Transitory BQ estimates

- Timing of peaks and throughs changes.

- Amplitude of cycles changes.

- End of the sample: transitory component is positive or negative ?



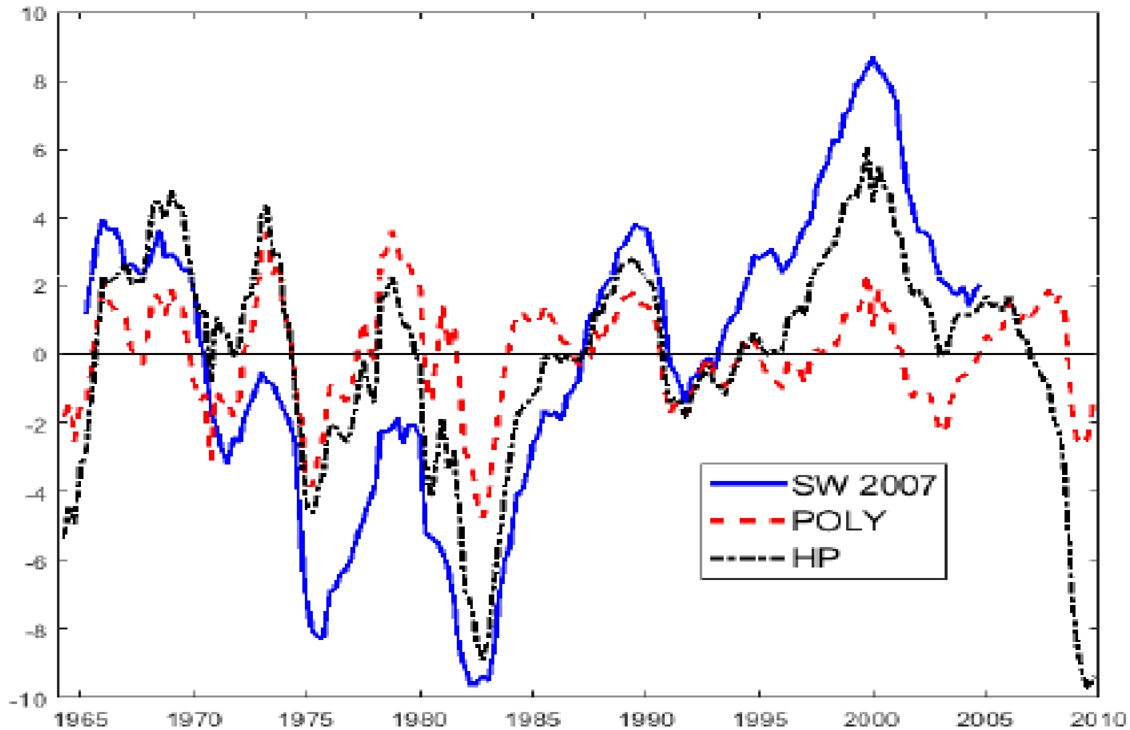
Permanent BQ estimates

7 How policymakers think about cycles?

• Policymakers interested in gaps. Very loosely defined.

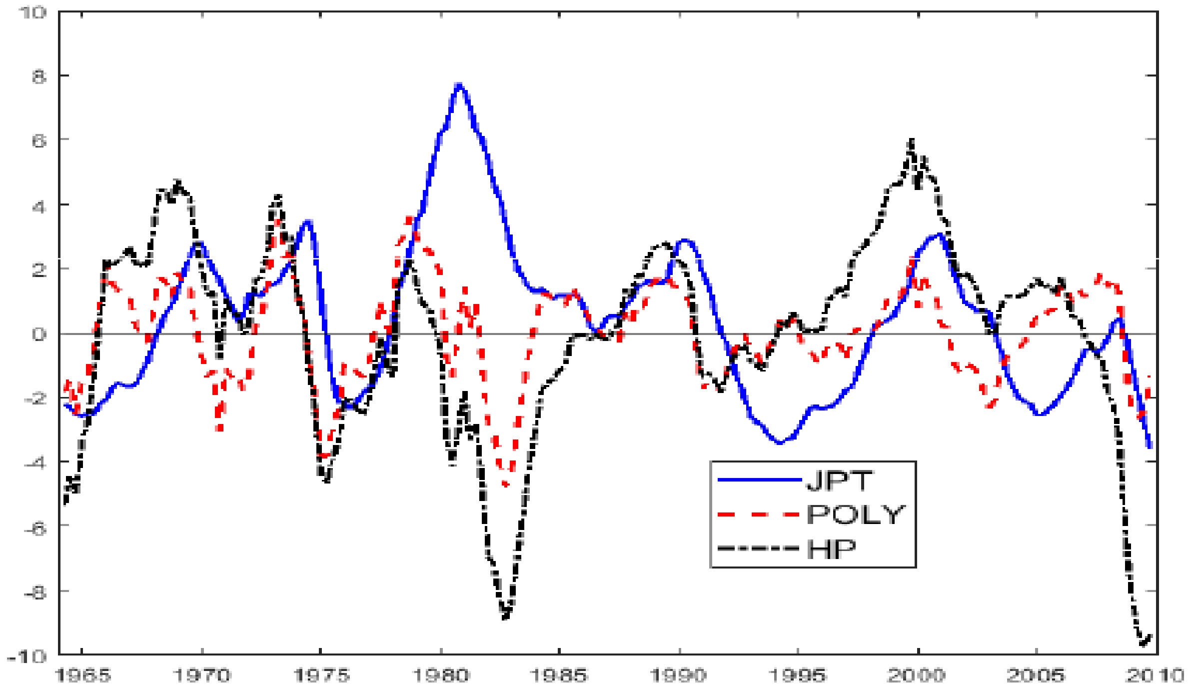
• Gaps are meaningful only in terms of a model. Potential is the path of the variables when nominal frictions are eliminated. Gap is the difference of actual from potential level.

• How do DSGE-based estimates of gaps look like?

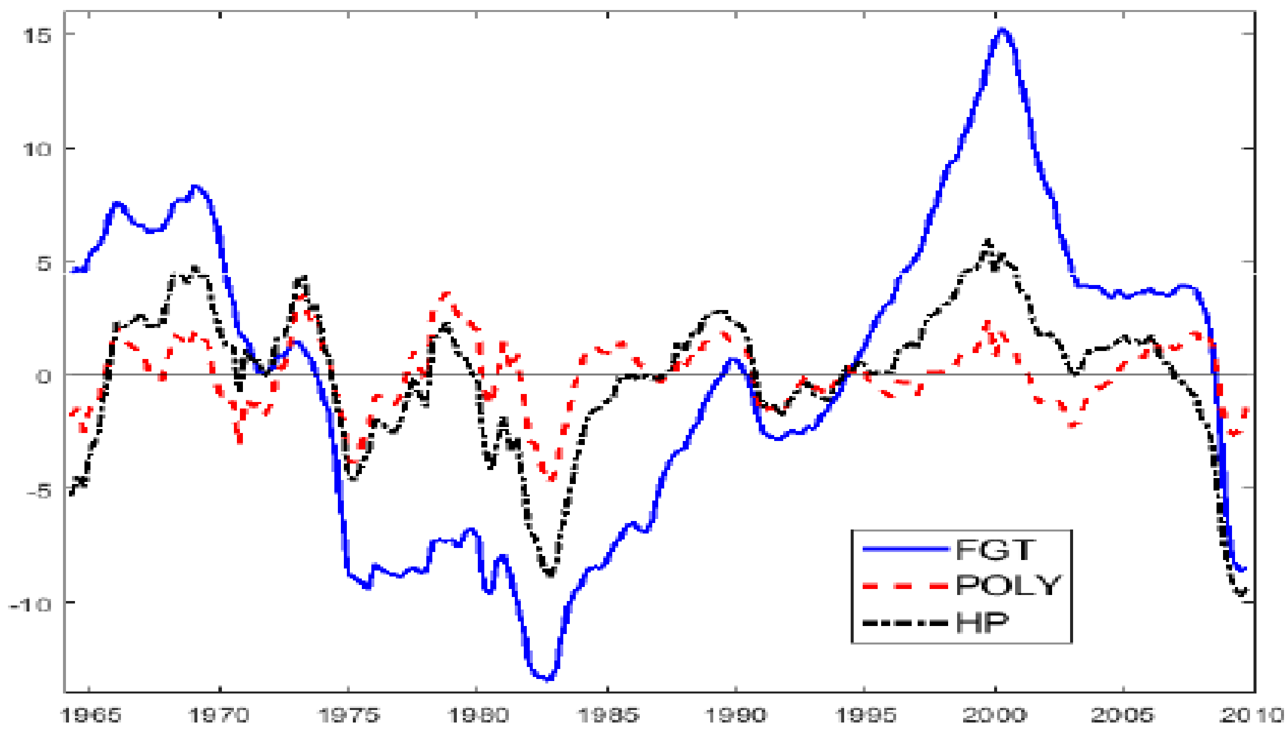


10
8
6
4
2
0
-2
-4
-6
-8
-10
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
SW 2007
POLY
HP

SW (2007) gap



JPT (2013) gap: model with two observable hours series



FGT (2020) gap: model with financial frictions

- Model chosen matters.

- Tend to have larger/longer swings than traditional statistical estimates: amplitude and duration of phases change.

- If you do not trust a model, what do you do? Canova and Matthes (2019) robust CL approach.

• What are the features of gaps? Are they similar to cycles obtained detrending/filtering the data? Are they similar to transitory fluctuations? Canova (2019b). NO

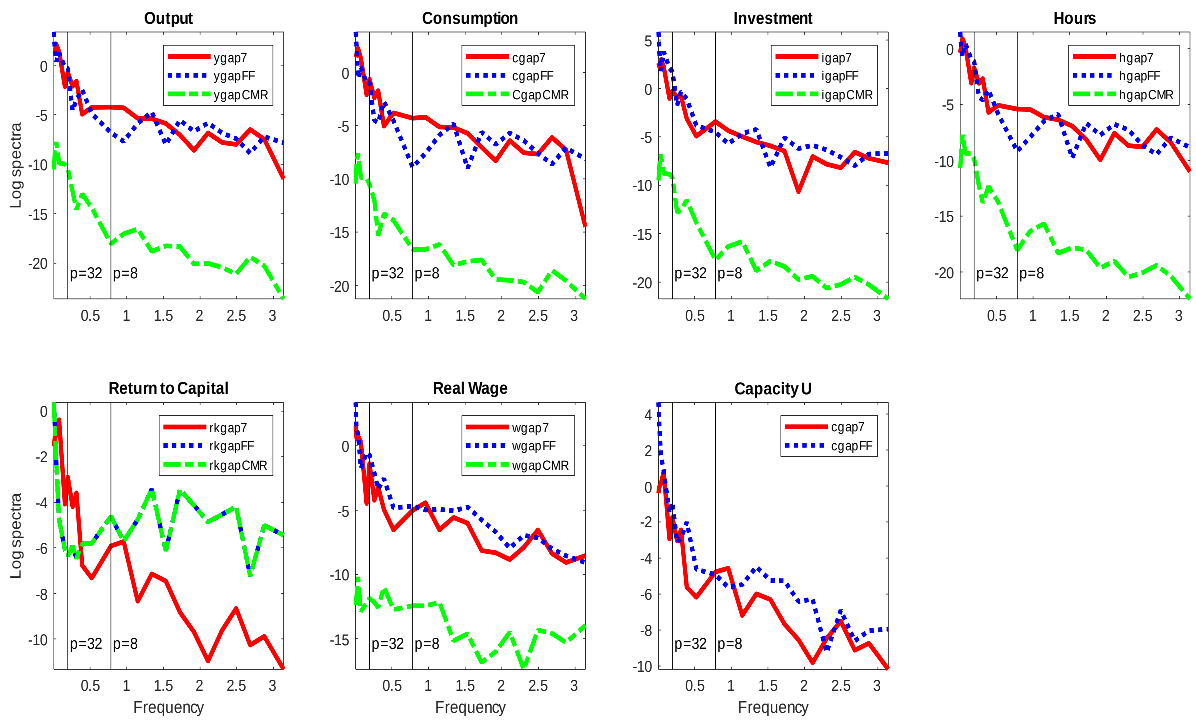
• Gaps depend on what frictions are included in the model but

i) Generally persistent and feature important low frequency variations.

ii) Have little power (variability) at business cycle frequencies.

iii) Are correlated with potentials.

iv) Do not look like standard gap measures (CBO, Fed measures (y-potential)).



Log spectra gaps: SW, SWFF, CMR models

8 Collecting cyclical information

• Approach one uses would not matter if cyclical statistics would be more or less the same. Are they?

• When filtering why do we concentrate on 8-32 quarter cycles? In developing countries the trend may be cycle: shocks may have permanent features (Aguiar and Gopinath, 2007). Cycles in labor market data may be longer than 8-32 quarters (Beaudry, et al. 2019).

• How do you compare countries different cyclical features (length)?

• Filtering and detrending subject to specification errors and small sample or truncation biases.

Canova (1998)-(1999) Business cycle facts depend:

- Assumptions about the trend and procedures used to remove it.
- Whether decompositions are univariate vs multivariate.
- Whether components are orthogonal vs. non-orthogonal.
- What portions of spectrum are emphasized.
- Sample size (in small samples cyclical coefficients poorly estimated)

Summary statistics

	Variability	Relative Variability		Contemporaneous Correlations			Periodicity
Method	GDP	Consumption	Real wage	(GDP,C)	(GDP,Inv)	(GDP, W)	(quarters)
HP1600	1.76	0.49	0.70	0.75	0.91	0.81	24
HP4	0.55	0.48	0.65	0.31	0.65	0.49	7
BN	0.43	0.75	2.18	0.42	0.45	0.52	5
BP	1.14	0.44	1.16	0.69	0.85	0.81	28
KPSW	4.15	0.71	1.68	0.83	0.30	0.89	6

• Differences present also in other statistics, e.g. dating of cyclical turning points or measuring business cycle phases.

Conclusion

• Extraction of growth cycles and calculation of statistics problematic.

• Empirical facts should be collected without growth removal and should be conditional (rather than unconditional).

• If you care about gaps, use models. If you do not trust models use composite methods or BW filters.

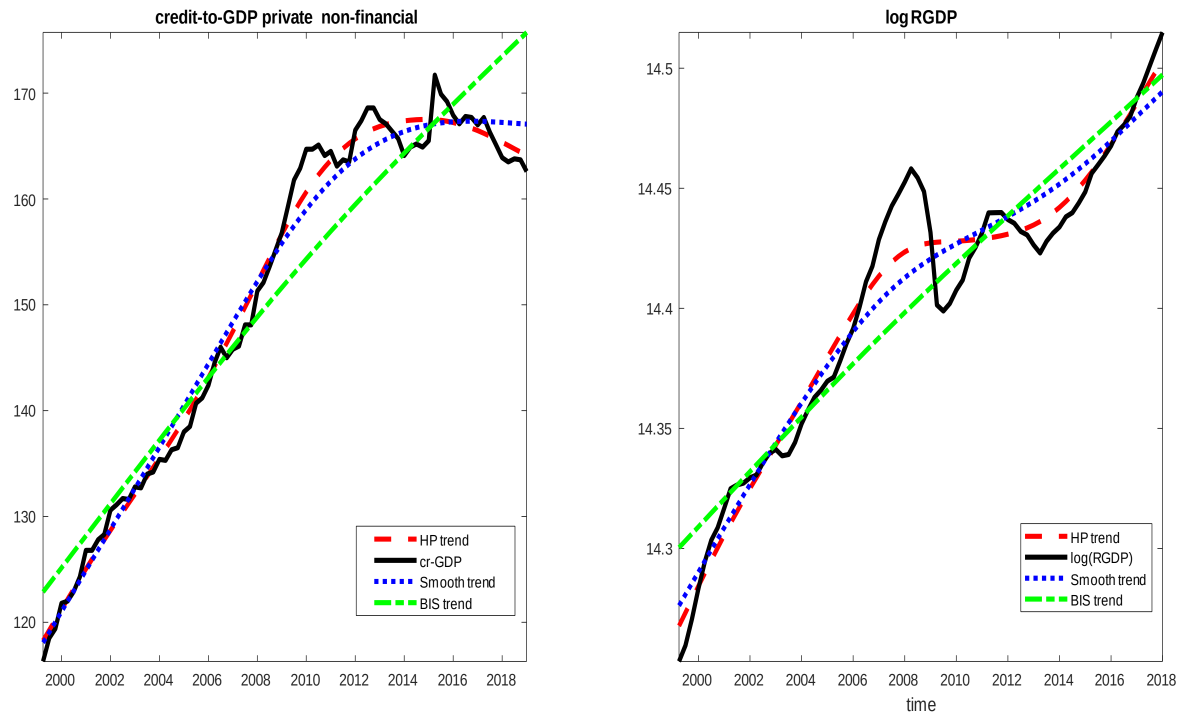
9 Business and financial cycles. Are they different?

- Are they different? BIS: Financial cycles are longer than business cycles. see Borio (2012)

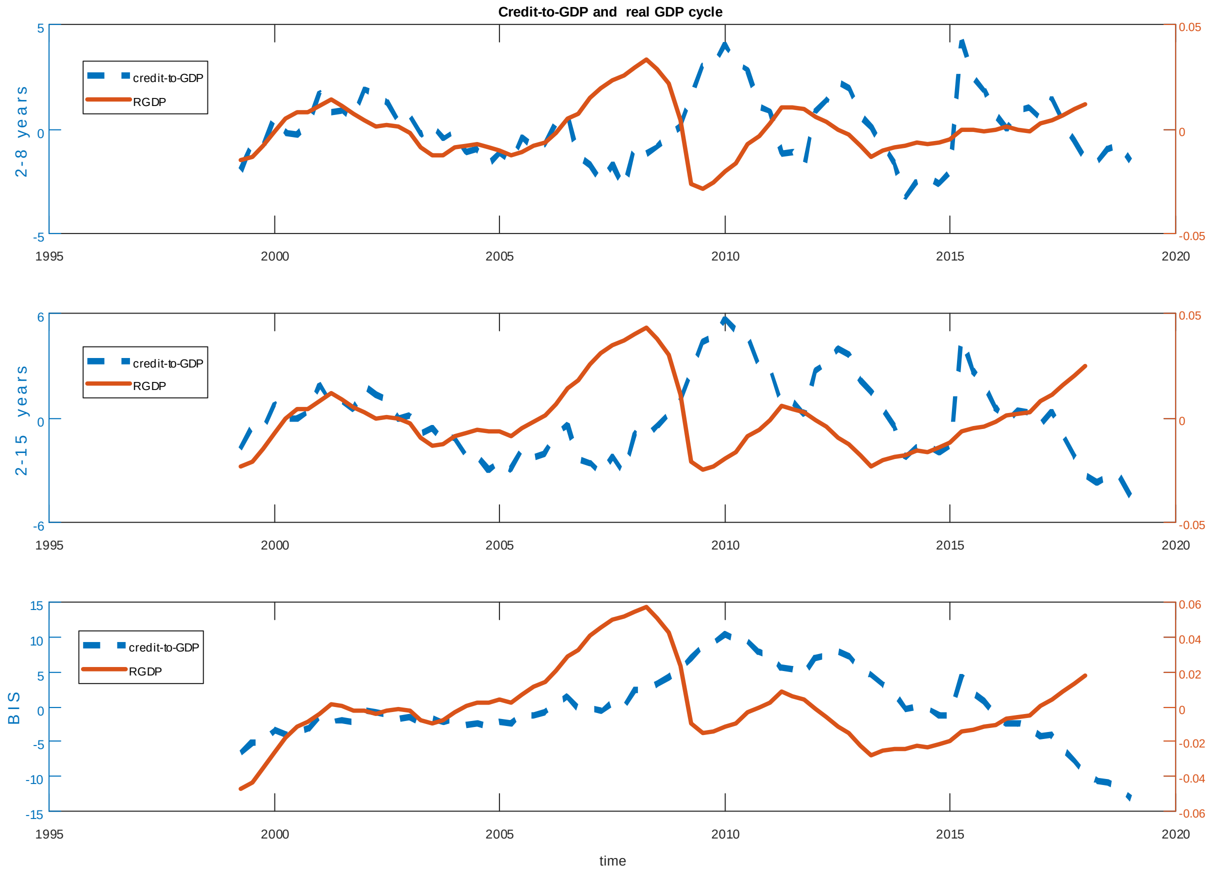
- Lots of literature on the topic, see e.g. Runstler and Vlekke (2018).

- Compare credit to GDP to nonfinancial corporations and output for illustration.

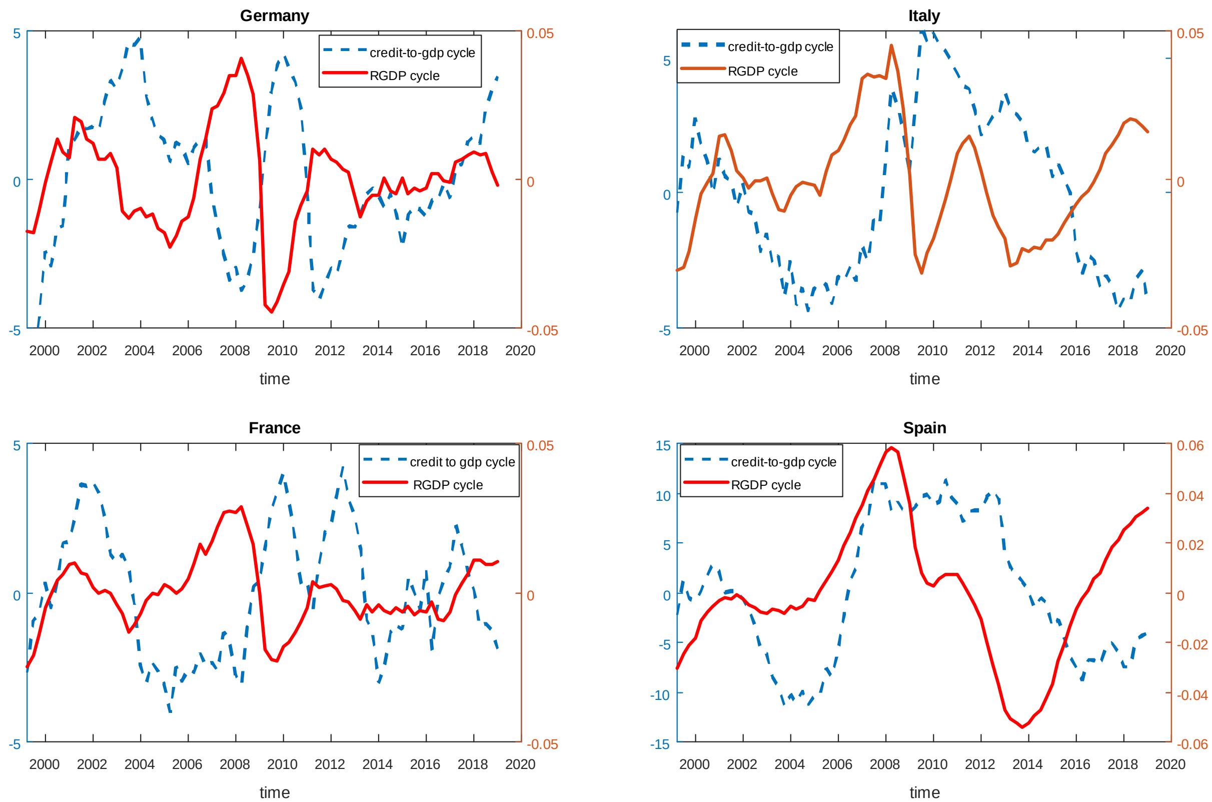
Canova (2019a): Euro area data



Variable	% of variance 2-8 years cycles	% of variance 8-15 years cycles	Persistence AR1
Credit/GDP total	1.5	18.3	0.99
Credit/GDP households	1.6	19.0	0.99
Credit/GDP private non financial	1.7	19.1	0.99
log(real GDP)	2.1	20.3	0.99
Labor Productivity	2.2	20.4	0.99
Unemployment rate	1.6	18.1	0.98

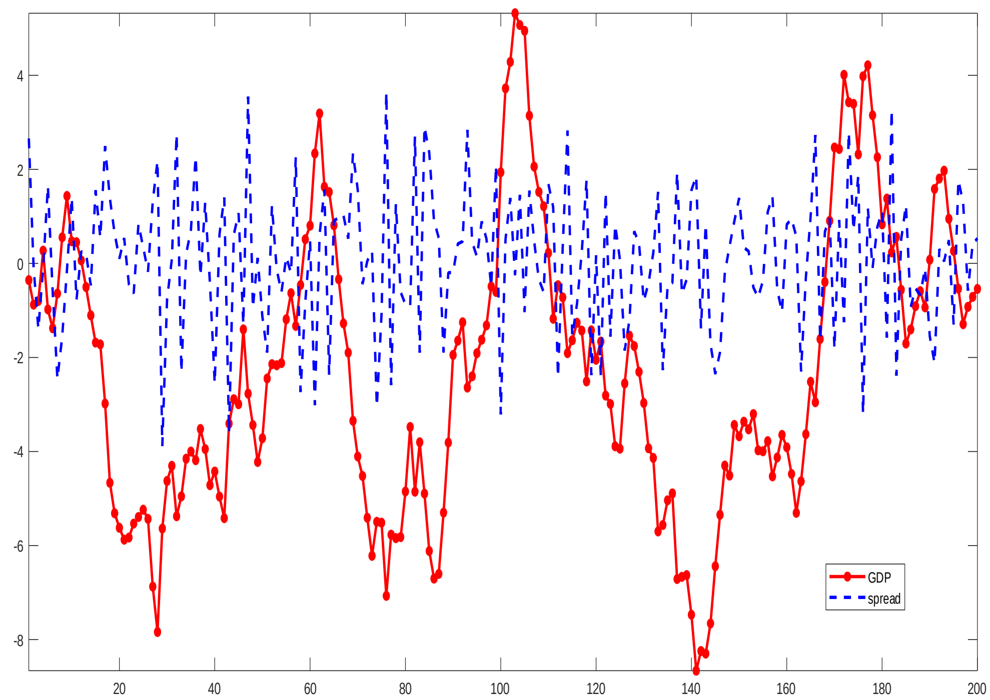


Credit-to-GDP and real GDP cycle
credit-to-GDP
RGDP
2-8 years
2-15 years
BIS
time

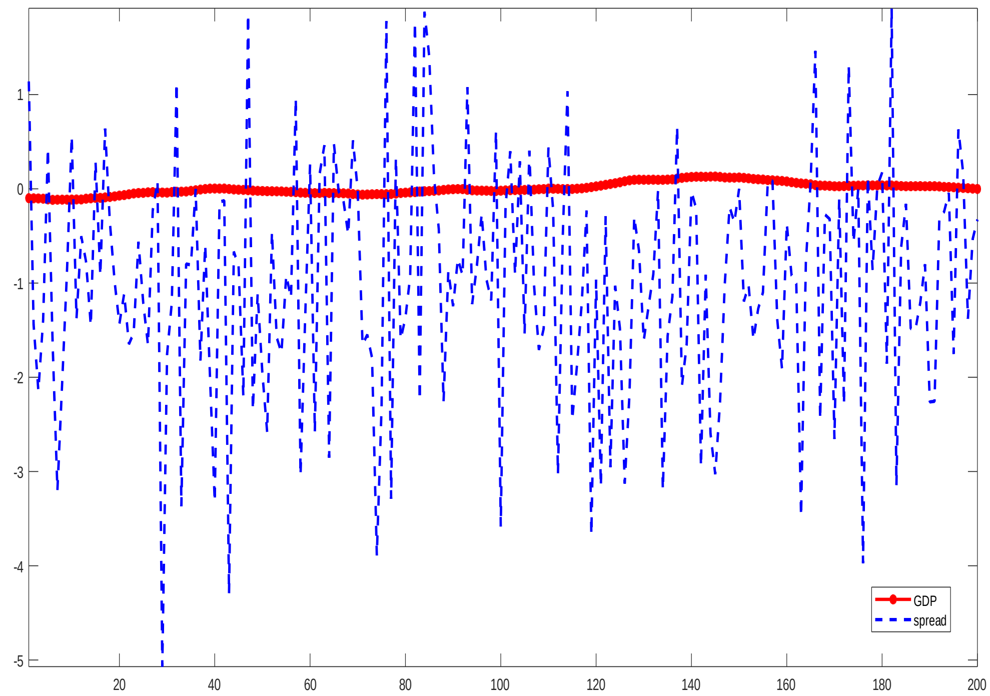

Germany
credit-to-gdp cycle
RGDP cycle
Italy
credit-to-gdp cycle
RGDP cycle
France
credit to gdp cycle
RGDP cycle
Spain
credit-to-gdp cycle
RGDP cycle
time

Real and Financial cycles in models

- Use SW-FF (Del Negro et al, 2015) and CMR (Cristiano, et al. 2011)
- Do cycles look like those of the data?



Output and spread gaps, SWFF



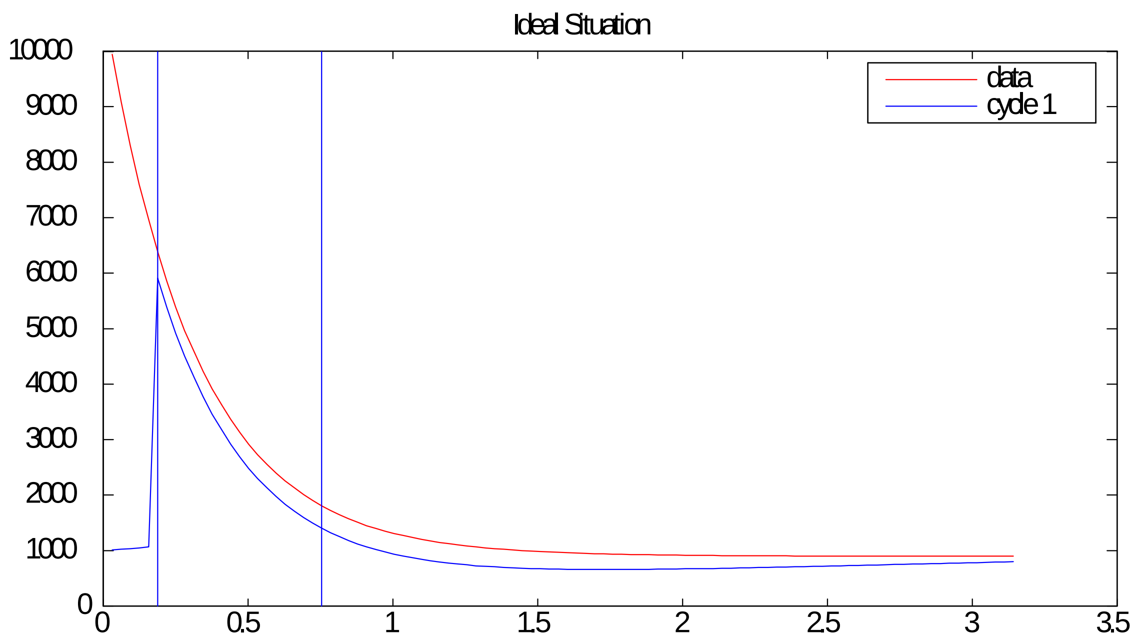
Output and spread gaps, CMR

10 Fitting structural models to filtered data

• Statistical filtering: Find B_j such that y_t^f has $\mathcal{S}(\omega_\tau) \neq 0$ only for certain $\omega_\tau \in (\omega_1, \omega_2)$.

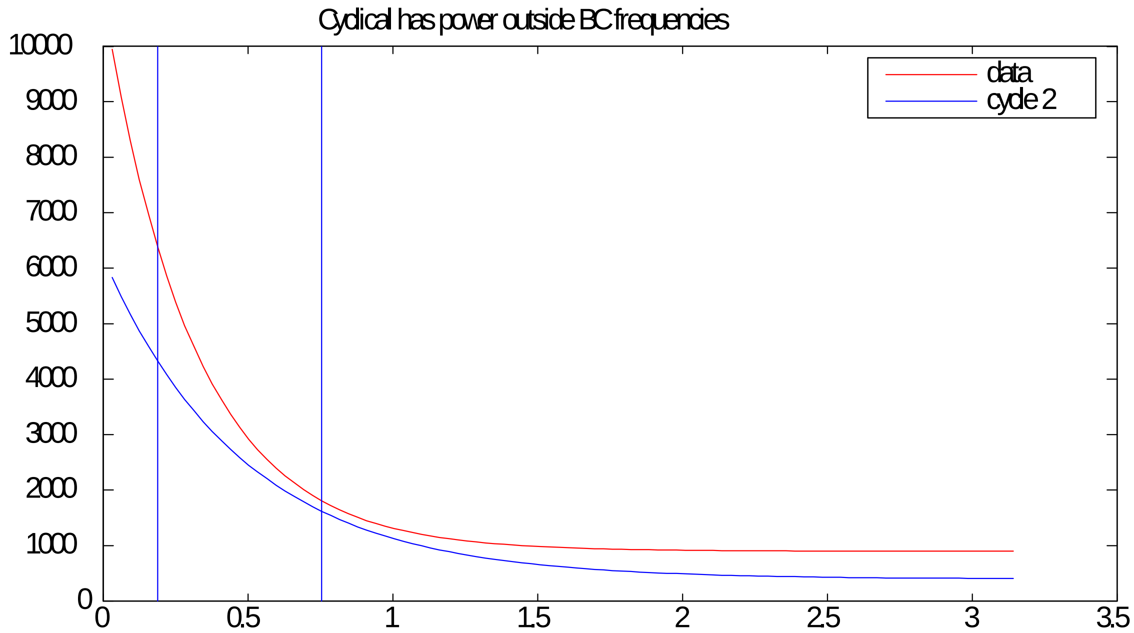
• Economic filtering: $y_t = y_{1t} + y_{2t} = A(L)e_t + B(L)u_t$, where e_t are permanent shocks, u_t are transitory shocks or e_t are disturbances entering the potential and u_t disturbances entering the gap. Note u_t and e_t may overlap.

• In general, $y_{2t} \neq y_t^f$ since y_{1t}, y_{2t} have $\mathcal{S}(\omega_\tau) \neq 0$ for all $\omega_\tau \in (0, \pi)$.



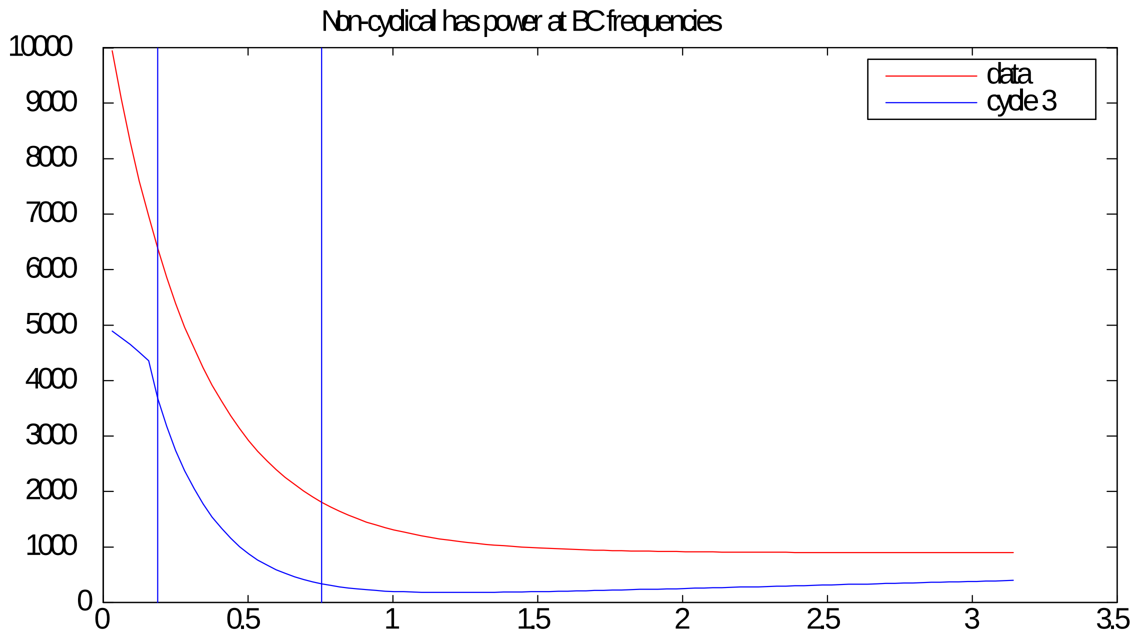
Ideal case

• (Cyclical) model has most of the variability located at business cycle frequencies. Statistical filtering would ok.



Realistic case

- If (cyclical) model is driven by persistent AR(1) shocks, lots of variability in the low frequencies. Filtering throws away information.



General case

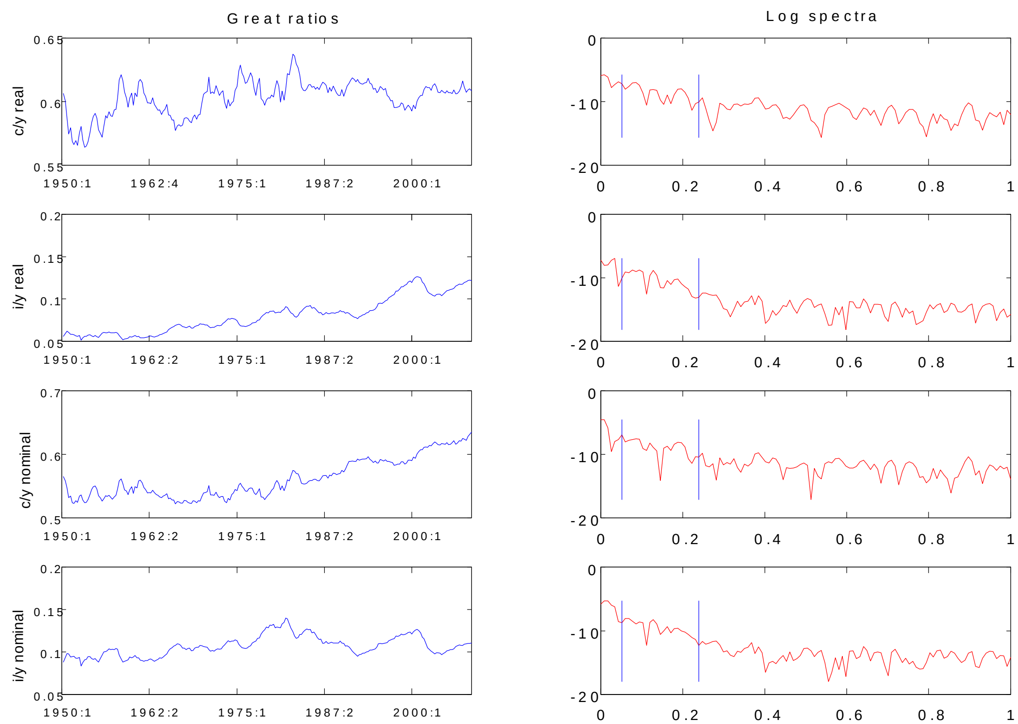
• If (cyclical) model is driven by persistent AR(1) shocks, and permanent shocks are cyclical, filtering is distortive. Different filters will give different results.

- Typical solution: Build in a trend in a (cyclical) model. Transform the data with a model consistent approach. Problems:

- Models with trends (in technology) imply balanced growth path. Typically violated in the data.

- Where do we put the trend (e.g. technology or preferences) matters for estimates of the structural parameters - nuisance parameter problem.

- Should we use a unit root or trend stationary specification?



Real and nominal Great ratios in US, 1950-2008.

Filter	LT		HP		FOD		BP		Ratio	
	Median	(s.d.)	Median	(s.d.)	Median	(s.d.)	Median	(s.d.)	Median	(s.d.)
σ_c	2.19	(0.10)	2.25	(0.12)	2.54	(0.16)	2.21	(0.10)	1.69	(0.11)
σ_n	1.79	(0.08)	1.57	(0.10)	1.90	(0.19)	1.78	(0.08)	2.16	(0.10)
h	0.67	(0.01)	0.59	(0.03)	0.44	(0.03)	0.66	(0.02)	0.64	(0.02)
α	0.17	(0.03)	0.12	(0.02)	0.12	(0.03)	0.16	(0.02)	0.13	(0.02)
ϵ	3.90	(0.12)	4.27	(0.14)	2.92	(0.11)	3.72	(0.05)	4.09	(0.12)
ρ_r	0.16	(0.04)	0.52	(0.04)	0.22	(0.06)	0.49	(0.04)	0.22	(0.04)
ρ_π	1.36	(0.08)	1.67	(0.04)	1.74	(0.05)	1.77	(0.08)	1.71	(0.05)
ρ_y	-0.15	(0.02)	0.35	(0.06)	0.13	(0.07)	0.44	(0.05)	-0.02	(0.01)
ζ_p	0.81	(0.01)	0.60	(0.03)	0.33	(0.03)	0.56	(0.03)	0.81	(0.01)
ρ_χ	0.76	(0.02)	0.59	(0.04)	0.29	(0.04)	0.82	(0.03)	0.82	(0.02)
ρ_z	0.96	(0.01)	0.54	(0.05)	0.87	(0.05)	0.46	(0.05)	0.92	(0.01)
σ_χ	0.23	(0.04)	0.37	(0.05)	0.23	(0.04)	0.20	(0.03)	0.95	(0.16)
σ_z	0.12	(0.02)	0.08	(0.01)	0.09	(0.01)	0.09	(0.01)	0.08	(0.01)
σ_{mp}	0.11	(0.01)	0.08	(0.01)	0.12	(0.02)	0.08	(0.01)	0.12	(0.01)
σ_μ	30.54	(1.17)	1.01	(0.)	0.16	(0.03)	0.63	(0.21)	34.70	(1.04)

Posterior estimates NK model. For LT, HP, FOD and BP real variables detrended, nominal demeaned. For Ratio, real variables are in terms of hours,all variables demeaned.

Which column should be trusted?

Alternatives:

- Use a data rich environment (Canova and Ferroni, 2011).

Let y_t^i be the actual data filtered with method $i = 1, 2, ..., I$ and $y_t^d = [y_t^1, y_t^2, \ldots]$. Assume:

$$y_t^d = \lambda_0 + \lambda_1 y_t(\theta) + u_t \qquad (55)$$

where $\lambda_j, j = 0, 1$ are matrices of parameters, measuring bias and correlation between data and model based quantities, u_t measurement errors and θ the structural parameters.

- Factor model setup a-la Boivin and Giannoni (2005).

- Can jointly estimate θ and λ's.

- Same interpretation as GMM with many instruments.

- Bridge cyclical model and the raw data with a flexible specification (Canova, 2014).

$$y_t^d = c + y_t^T + y_t^m(\theta) + u_t \tag{56}$$

where $y_t^d \equiv \tilde{y}_t^d - E(\tilde{y}_t^d)$ the log demeaned vector of observables, $c = \bar{y} - E(\tilde{y}_t^d)$, y_t^T is the non-cyclical component, $y_t^m(\theta) \equiv S[y_t, x_t]'$, where S is a selection matrix, is the model based- cyclical component (the solution of a DSGE model), u_t is a iid $(0, \Sigma_u)$ (measurement) noise, $y_t^T, y_t^m(\theta)$ and u_t are mutually orthogonal.

- Non cyclical component

$$y_t^T = y_{t-1}^T + \bar{y}_{t-1} + e_t \qquad e_t \sim \; iid \; (0, \Sigma_e^2) \tag{57}$$

$$\bar{y}_t = \bar{y}_{t-1} + v_t \qquad v_t \sim \; iid \; (0, \Sigma_v^2) \tag{58}$$

- $\Sigma_v^2 > 0$ and $\Sigma_e^2 = 0$, y_t^T is a vector of I(2)processes.

- $\Sigma_v^2 = 0$, and $\Sigma_e^2 > 0$, y_t^T is a vector of I(1) processes.

- $\Sigma_v^2 = \Sigma_e^2 = 0$, y_t^T is deterministic.

- $\Sigma_v^2 > 0$ and $\Sigma_e^2 > 0$ and $\frac{\sigma^2_{v_i}}{\sigma^2_{e_i}}$ is large, y_{it}^T is "smooth" and nonlinear (as in HP).

- Jointly estimate structural θ and non-structural parameters (joint estimation and filtering)

- Equivalent to assume a rich measurement error structure.

How does the procedure do in a simple experimental design?

	True	Small variance Median	(s.e)	True	Large variance Median	(s.e)
σ_c	3.00	3.68	(0.40)	3.00	3.26	(0.29)
σ_n	0.70	0.54	(0.14)	0.70	0.80	(0.13)
h	0.70	0.55	(0.04)	0.70	0.77	(0.04)
α	0.60	0.19	(0.03)	0.60	0.41	(0.04)
ϵ	7.00	6.19	(0.07)	7.00	6.95	(0.09)
ρ_r	0.20	0.16	(0.04)	0.24	0.31	(0.04)
ρ_π	1.30	1.30	(0.04)	1.30	1.25	(0.03)
ρ_y	0.05	0.07	(0.03)	0.05	0.08	(0.10)
ζ_p	0.80	0.78	(0.04)	0.80	0.72	(0.02)
ρ_χ	0.50	0.53	(0.04)	0.50	0.69	(0.05)
ρ_z	0.80	0.71	(0.03)	0.80	0.90	(0.03)
σ_χ	0.011	0.012	(0.0003)	0.011	0.012	(0.0003)
σ_z	0.005	0.006	(0.0001)	0.005	0.007	(0.0001)
σ_{mp}	0.001	0.002	(0.0004)	0.001	0.002	(0.0004)
σ_μ	0.206	0.158	(0.0006)	0.206	0.1273	(0.0004)
σ_χ^{nc}	0.02			0.23		

Parameters estimates using flexible specification. σ_χ^{nc} is the standard error of the shock to the non-cyclical component.

Appendix: Other elements of Spectral Analysis

• The periodogram of y_t is $Pe(\omega) = \sum_\tau \widehat{ACF}(\tau)e^{-i\omega\tau}$ where $\widehat{ACF}(\tau) = \frac{1}{T}\sum_t (y_t - \bar{y})(y_{t-\tau} - \bar{y})$ and $\bar{y} = \frac{1}{T}\sum_t y_t$.

• Periodogram is inconsistent estimator of the spectrum. Periodogram consistently estimate only an average of the frequencies of the spectrum. For consistency need to "smooth" periodogram with a filter (kernel).

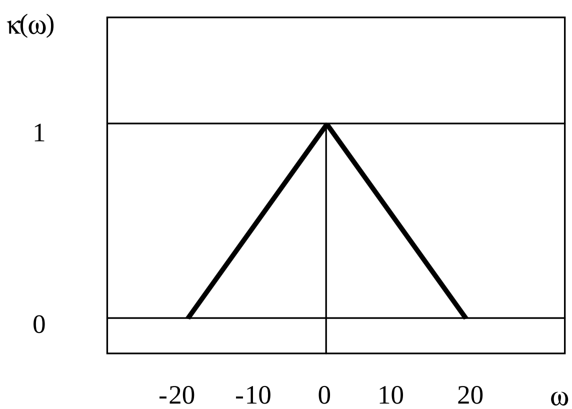
• A filter is a kernel (denoted by $\mathcal{K}_T(\omega)$) if, as $T \to \infty$, $\mathcal{K}(\omega_\tau) = 1, \quad$ for $\omega_\tau = \omega$ and $\mathcal{K}(\omega_\tau) = 0$ otherwise.

• Kernels eliminate bias in $\widehat{ACF}(\tau)$. Since as $T \to \infty$ bias disappears, wants kernels to converge to $\delta-$function as $T \to \infty$.

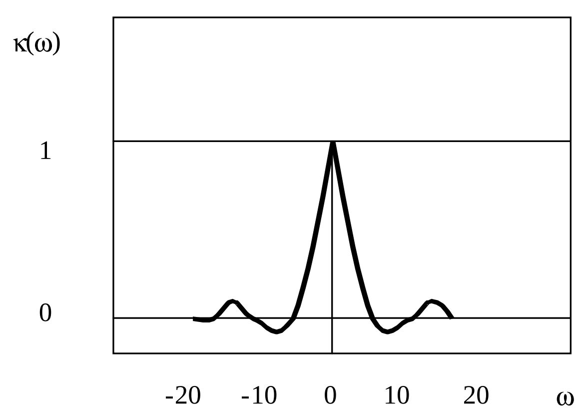
Two useful Kernels.

• Bartlett kernel: tent shaped, width $2J(T)$; $\mathcal{K}(\omega_j) = 1 - \frac{|\omega|}{J(T)}$. $J(T)$ chosen so that $\frac{J(T)}{T} \to 0$ as $T \to \infty$.

• Quadratic spectral kernel: wave with infinite loops;
$\mathcal{K}(\omega_j) = \frac{25}{12\pi^2 j^2}\frac{\sin(6\pi j/5)}{(6\pi j)/5} - cos(\frac{6\pi j}{5})$.



Bartlett Kernel



Quadratic Spectral Kernel