Separating Yolk from White:
A Filter based on Economic Properties of Trend and Cycle

Peng Zhou[1]

Cardiff University, United Kingdom

Abstract

This paper proposes a new filter technique to separate trend and cycle based on stylised economic properties, rather than relying on *ad hoc* statistical properties such as frequency. Given the theoretical separation between economic growth and business cycle literature, it is necessary to make measures of trend and cycle match what the respective theories intend to explain. The proposed filter is applied to the long macroeconomic data collected by the Bank of England (1700-2015).

Key Words: Filter, Trend, Cycle

JEL Classification: C32

---

[1] D47, Aberconway Building, Cardiff University, Cardiff, CF10 3EU, UK. Tel: +44 (0)2920688778. Email: zhoup1@cardiff.ac.uk

Separating Yolk from White:
A Filter based on Economic Properties of Trend and Cycle

1. Introduction

The two fundamental issues in macroeconomic theory are long-run economic growth (driven by low-frequency changes) and short-run business cycle (driven by high-frequency changes). The literatures on these two issues are separated because the foci of the models are quite different. However, the data for economic growth and business cycle are not collected separately. A common practice to test/estimate an economic growth theory is simply to use the growth rate of the raw data, ignoring the possible role of business cycles[2]. Meanwhile, to test/estimate a business cycle theory, the raw data are usually filtered by some statistical procedures, ranging from de-meaning, linear de-trending and de-seasoning to sophisticated filters such as HP filter, band-pass filter and wavelet filter, but these procedures either fail to account for the possible variation in the trend component or tend to over-smooth the cycle component. A consequence of applying these statistical procedures before economic analysis is that the economic growth models are likely to over-state the importance of economic growth, while business cycle models are likely to under-state the importance of business cycle, because the filtered data may contain both trend and cycle. Without a proper measurement, the validity and reliability of the empirical inferences are questionable.

It very much resembles separating the yolk from the white, before a cook can use them for different recipes. The statistical procedures are like using a sharp knife to cut yolk out, but inevitably leaving some white behind and probably losing some yolk due to the inflexibility of the knife. A smart cook's trick is to make use of the physical properties of the yolk/white (different densities): simply squeeze a plastic bottle and press against the yolk, and the power of air suction will nicely separate the dense yolk from the thin white. The filter proposed in this paper follows a similar logic: we make use of some stylised economic properties of trend and cycle to separate them, rather than relying on some *ad hoc* assumptions on statistical properties such as frequency.

An earlier approach similar to the proposed filter can be found in Crafts, Leybourne and Mills (1989, CLM hereinafter) who use a flexible procedure to isolate the trend and cycle components using the Kalman filter. Their approach has the advantage over de-

---

[2] For example, Sala-i-Martin (2004), Badinger (2010) and Asimakopoulos and Karavias (2016).

meaning, de-trending and de-seasoning in that it allows for time-varying trend component. It also has the advantage over the sophisticated filters (e.g. HP filter, band-pass filter, wavelet filter) in that it uses a state-space model to incorporate some economic features into the filtering process. However, the CLM procedure only applies to $I(1)$ series. The present paper develops a more general filtering procedure which can be applied to both $I(0)$ and $I(1)$ processes. Any time series regardless of its stationarity property can be decomposed into a time-varying trend component and a cycle component without an assumption on frequency. A key advantage of this new filter over the traditional statistical procedures is that it incorporates some stylised economic properties of the trend and cycle into the filter design.

2. The Filter

A raw data series z_t, either $I(0)$ or $I(1)$, can be decomposed into a time-varying trend component (μ_t) driven by low-frequency changes and a cycle component (ψ_t) driven by high-frequency changes:

$$z_t = \mu_t + \psi_t \tag{1}$$

The cycle component ψ_t is modelled as a zero-mean stationary AR(2) process, which is a well-established stylised economic property in the business cycle literature[3].

$$\psi_t = \rho_1\psi_{t-1} + \rho_2\psi_{t-2} + \omega_t, \text{ where } \omega_t \sim WN(0,\sigma_\omega^2) \tag{2}$$

The AR(2) specification implies a pseudo-cyclical behavior (Harvey, 1985). The periodicity $\lambda \equiv 2\pi/\operatorname{acos}\left(|\rho_1|/2\sqrt{|\rho_2|}\right)$ can be used to measure the average duration of the cycle component and provide an intuitive measure of frequency. That is another reason why AR(2) is preferred to AR(1) in the specification of the cycle component.

The trend component μ_t depends on whether the series is $I(1)$ or $I(0)$:

$$\mu_t = \mu_{t-1} + \beta_t + \eta_t, \text{ where } \eta_t \sim WN\left(0,\sigma_\eta^2\right) \tag{3A}$$

$$\mu_t = \rho\mu_{t-1} + (1-\rho)\beta_t + \eta_t, \text{ where } \eta_t \sim WN\left(0,\sigma_\eta^2\right) \tag{3B}$$

The specification of the trend component again makes use of some well-established

---

[3] Most business cycle models assume AR(1) or AR(2) in the structural equations and shock processes. Very few models have longer than 2-period lag. See for example Smets and Wouters (2007).

stylised economic properties[4]. In the $I(1)$ case (e.g. log GDP), equation (3A) can also be written in the first difference form: $\Delta\mu_t = \mu_t - \mu_{t-1} = \beta_t + \eta_t$ to see that the trend growth ($\Delta\mu_t$) contains a time-varying deterministic component (β_t) and a stochastic component (η_t). $\Delta\mu_t$, rather than the raw growth Δy_t, is what most economic growth models actually try to explain. In the $I(0)$ case (e.g. unemployment rate, inflation and interest rate), equation (3B) is a mean-reverting process, allowing for a time-varying mean β_t. Lastly, the time-varying mean/trend is modelled as a random walk:

$$\beta_t = \beta_{t-1} + \xi_t, \text{ where } \xi_t \sim WN(0, \sigma_\xi^2) \tag{4}$$

The equation system (1), (2), (3A) and (4) defines the dynamic behavior of an $I(1)$ process, while (1), (2), (3B) and (4) defines the counterpart of an $I(0)$ process. If we denote $\mathbf{x}_t \equiv [\mu_t; \beta_t; \psi_t; \psi_{t-1}]$ as the unobserved state vector, z_t as the observed raw data and v_t as a white noise process with a standard normal distribution, the equation system can be written in the state space form:

- State Equation: $\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}v_t$;
- Measurement Equation: $z_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}v_t$.

The matrices are defined as:

$$\mathbf{A} = \mathbf{A}_0^{-1}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \rho_1 & \rho_2\\ 0 & 0 & 1 & 0\end{bmatrix},\ \mathbf{B} = \mathbf{A}_0^{-1}\begin{bmatrix}\sigma_\eta\\ \sigma_\xi\\ \sigma_\omega\\ 0\end{bmatrix},\ \mathbf{C} = [1 \quad 0 \quad 1 \quad 0],\ \mathbf{D} = 0, \text{ where:}$$

$$\mathbf{A_0} \equiv \begin{bmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \text{ if } z_t \sim I(1), \text{ and } \mathbf{A_0} \equiv \begin{bmatrix}1 & 1-\rho & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \text{ if } z_t \sim I(0).$$

Note that the measurement error is assumed to be equal to zero ($\mathbf{D} = 0$) here, but in general, it can be non-zero. The above state space model can be estimated numerically using the Kalman filter. This procedure of decomposing a time series z_t into trend and cycle based on the stylised economic properties of trend and cycle is termed as the Z-filter for the convenience of comparison with other procedures[5].

---

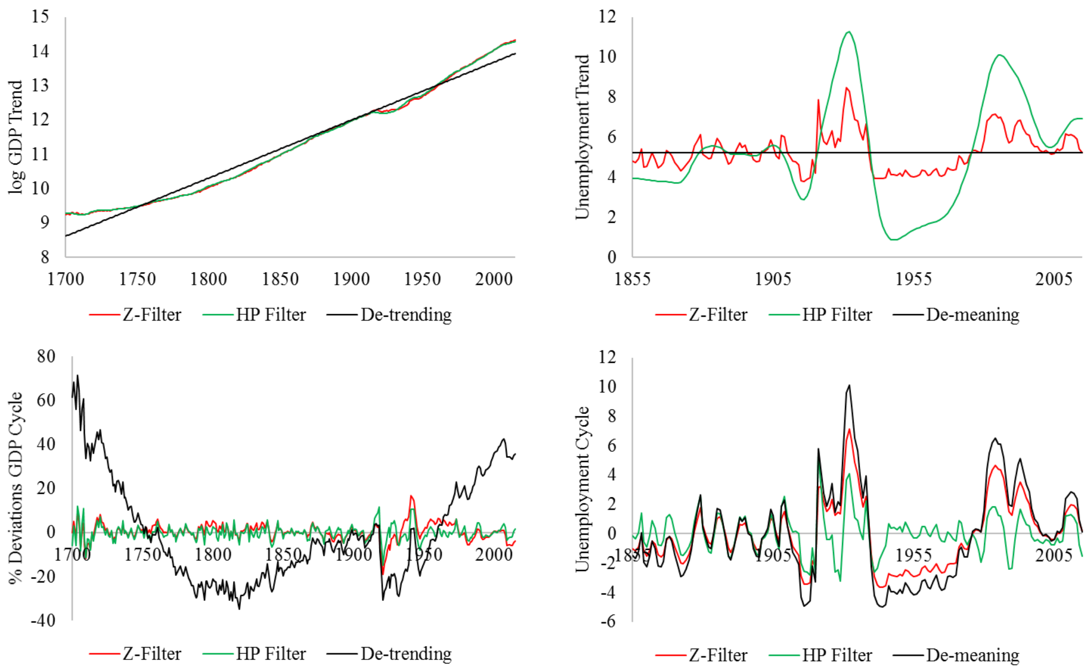
[4] For example, see Nelson and Plosser (1982) for an early discussion of the model specification of trend.

[5] Another reason is because of the initial of the author's last name.

3. Application

To demonstrate, we apply the Z-filter to probably the longest macroeconomic data series in the world accumulated by the Bank of England[6]. We compare the filtered trend and cycle components of (log) GDP (an $I(1)$ process) and unemployment rate (an $I(0)$ process) with other statistical procedures in Figure 1. As expected, the volatility of trend component in Z-filter is greater than that in HP-filter, and the volatility of cyclical component in Z-filter is smaller. For example, the standard deviation of Z-filtered GDP cycles is 2.8% while that of HP-filter is 3.8%.

Figure 1 The Comparison of De-trend, De-mean, HP-filter and Z-filter



Notes: De-trending is based on a simple linear regression with an intercept and a deterministic time trend.

One expected feature of the Z-filtered trend component is its high volatility (both for $I(1)$ and $I(0)$ processes). It is because the Z-filter is more flexible and does not restrict the frequency of changes in the trend component. Arguably, there is no good reason and no good basis to restrict the frequency of changes in the trend. The resulting trend shows that the fluctuations of the economy can well come from the fluctuations in the trend as much as in the cycle. However, if de-meaning, de-trending or HP filter are used, the fluctuations will be mainly assigned to cycle components.

---

[6] Bank of England, November, 2016. *Three Centuries of Macroeconomic Data* (version 2.3). Available at: http://www.bankofengland.co.uk/research/pages/onebank/threecenturies.aspx.

The state-space models of (log) GDP and unemployment rate are estimated using the Kalman filter and the results are shown in Table 1. The most informative conclusion is the periodicity (λ) of the two series. A complete cycle in the output market takes about 5 years (similar to the findings in CLM), while that in the labour market is about 20 years. Different cycle lengths imply that recessions of different severity may be due to composition of different cycles. If only the output market experiences the trough, then the recession is mild (minor trough). If more than one markets are at the trough simultaneously, then it may form a Great Recession (medium trough) or even a Great Depression (major trough). Combined with financial cycles (which can be estimated using interest rate), the composition of cycles with different lengths may be more complicated, but has more potential to explain the observed patterns of business cycles.

Table 1 Summary of the Estimated State Space Models

	Log (GDP)	Unemployment Rate
ρ_1	0.8338	1.2749
	(0.060)	(0.481)
ρ_2	-0.0451	-0.3674
	(0.039)	(0.344)
ρ		0.8624
		(0.135)
σ_η	0.0114	0.4917
	(0.008)	(0.930)
σ_ξ	0.0013	0.0027
	(0.001)	(35.232)
σ_ω	0.0242	0.7548
	(0.007)	(2.391)
λ	4.8507	19.6279

Notes: Standard deviations are in parentheses. $\lambda \equiv 2\pi/\operatorname{acos}\left(|\rho_1|/2\sqrt{|\rho_2|}\right)$ measures the periodicity.

4. Conclusion and Discussion

This paper introduces a new filter, the Z-filter, to decompose any time series into the trend and cycle consistent to the stylised economic properties of trend and cycle in the theoretical literature. In contrast, the traditional filters mainly rely on statistical properties such as frequency. Arguably, economic properties are more relevant than statistical properties in separating trend and cycle. The Z-filter can improve the consistency in empirical inference by a better measure of what the theories are intended to explain. In other words, it can greatly reduce the measurement error in empirical studies. As an analogy, to bake a good cake, the cook needs to separate yolk from white, while to verify a specific model, the economist needs to isolate trend from cycle.

There are two close substitutes in existing literature to the proposed filter. One is Auto-Regressive-Moving-Average (ARMAX) model, which is a well-understood statistical tool for stationary data, so it can only capture the dynamic properties of the cyclical component with no consideration in the trend component. Thus, the proposed filter includes ARMAX model as a special case, where the trend component is already filtered out. Another popular technique close to the proposed filter is Linear Gaussian State Space (LGSS) model, which is a general model for any time series data. However, the LGSS model is essentially a statistical model without using any economic properties of the data. The Z-filter combines the powerfulness of ARMAX model in characterizing stationary dynamics and the generality of LGSS model in capturing nonstationary dynamics, together with explicit use of economic properties of the filtered series.

Some may argue for a different order of lags rather than AR(2) in the model. Note that for a stationary series, the only difference between models with different lag orders is how many peaks (or turning points) there are in the impulse response functions. If it is AR(1), there is usually only one peak, while AR(2) can have two and AR(p) can have p peaks. However, in our context, it is well documented in the empirical literature that there are only one or two peaks in macroeconomic data (in the case of two peaks, it is called overshooting). The fundamental point of this paper is to distinguish "economic models" from "statistical models". Therefore we make use of the economic properties of the macroeconomic data to streamline the specification, rather than compiling statistical possibilities. This assumption of a lag order of 2 is also adopted by CLM.

References

Asimakopoulos, S. and Karavias, Y. 2016. “The impact of government size on economic growth: A threshold analysis.” *Economics Letters*, 139 (2): 65-68.

Badinger, H. 2010. “Output volatility and economic growth.” *Economics Letters*, 106, (1): 15-18.

Crafts, N. Leybourne, S. and Mills, T.1989. “Trends and Cylcs in British Industrial Production, 1700-1913.” *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 152 (1): 43-60.

Harvey, A. 1985. “Trends and cycles in macroeconomic time series.” *Journal of Business Economics and Statistics*, 3: 216-227.

Nelson, R. and Plosser, C. 1982. “Trends and random walks in macroeconomic time series.” *Journal of Monetary Economics*, 10 (2): 139-162.

Smets, F. and R. Wouters. 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *The American Economic Review*, 97 (3):586-606.

Sala-i Martin, X., G. Doppelhofer, and R. Miller. 2004. “Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach.” *American Economic Review*, 94 (4): 813-835.