Open Source Macroeconomics Laboratory Boot Camp
Business Cycle Moments and Time-Series Filtering

Kerk Phillips
Brigham Young University

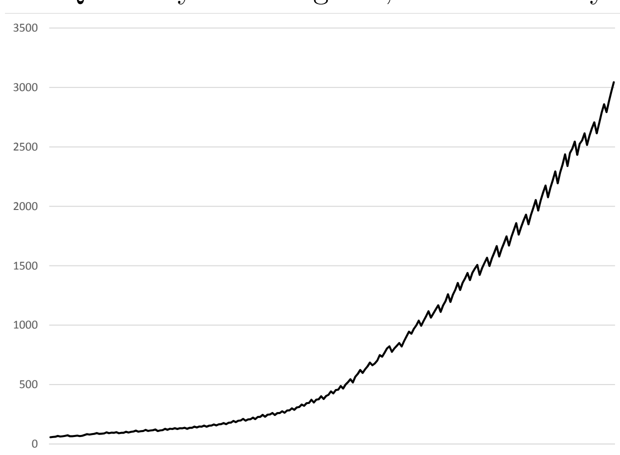
2017

1 Introduction

Any understanding of the macroeconomy must be rooted in data. Models that are otherwise elegant and well-crafted are useless if they do not accurately mimic the behavior of the real economy. We will focus a great deal on modeling, but must bear in mind that our models have targets which they are trying to hit. These targets are the stylized facts about the actual economy.

To establish these stylized facts requires that, in addition to collecting data, we must also analyze it and produce a set of robust facts. Economic data in its raw form is often difficult to interpret and may require adjustments. Consider the data plotted in figure 1. One of the most striking features of the time series for GDP is the sawtooth pattern due to seasonal variation in production. Over this sample period of 1947:1 to 2004:4 the highest level of production occurs in the fourth quarter. On average first quarter production is 6.74% below this, and second and third quarter production are 3.70% and 3.28% below the fourth quarter, respectively.

Figure 1: Quarterly GDP Figures, Not Seasonally Adjusted

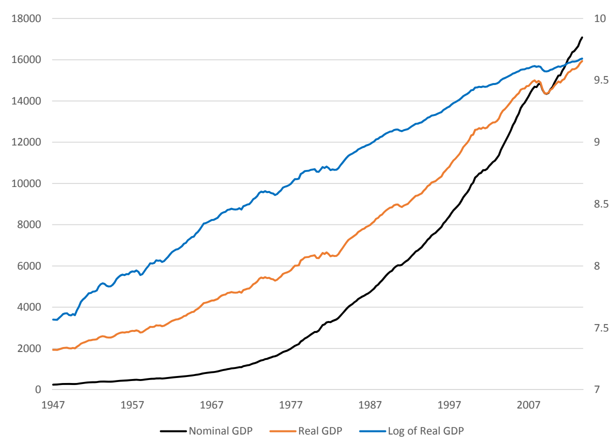


1947:1 - 2004:4, in billions of dollars

This seasonal variation in GDP dwarfs variation due to business cycles, making it difficult to determine when they begin and end. Often this seasonal variation is removed from the data before it is reported. The most common method is to use `X-13ARIMA-SEATS` which was developed and is available for download from the U.S. Bureau of the Census. Data with the seasonality removed are most frequently reported in the form of "seasonally adjusted at annual rates" or "s.a.a.r.", meaning they are interpretable as the value of goods and services the would be produced if the relative level of activity from that quarter had continued for a full year's time. This means the numbers are roughly four times bigger than the underlying data from which they are derived.

Figure 2 shows the seasonally adjusted numbers for U.S. GDP with the black line. These numbers are the nominal figures, meaning they are measured in current period dollars each quarter and are not adjusted for inflation. The real GDP numbers measured in constant dollars are shown with the orange line. As we will discuss below, it is often advantageous to transform macroeconomic data by taking the natural logarithm. This is plotted with the blue line and the scale is given on the right side of the graph.

Figure 2: Three Plots of Quarterly GDP, Seasonally Adjusted at Annual Rates



1947:1 - 2014:4, in billions of dollars, scale of the left

log series scale on the right

Raw economic data is rarely useful unless it has been broken down into series that can be analyzed for specific reasons. Usually this means dcomposing a set of observations over time into three series: 1) a trend or growth component, 2) a seasonal component and 3) a cyclical component. We use the first series to examine how economies behave in the long run where they tend to grow over time. We use the second to examine how economies behave in predictable ways over the course of a year. Finally, we use the third compoent to examine economic fluctuations or "business cycles".

We will write this decomposition using the notation in equation (1).

$$x_t = x_t^G + x_t^S + x_t^C \tag{1}$$

2 Finding Macroeconomic Data

There are many sources of data available for finding macroeconomic data. Below we list the most commonly used and discuss each briefly.

2.1 U.S. Data

For U.S. data national income and product accounts are calculated and reported by the Bureau of Economic Analysis, a part of the Department of Commerce. Labor market and price data are calculated and reported by the Bureau of Labor Statistics, which is part of the Department of Labor. The Federal Reserve Bank of St. Louis maintains the Federal Reserve Economic Database which pools data from these two sources an others in a single location. All of the sources below are free to the public.

2.1.1 Bureau of Economic Analysis (BEA)

The BEA website is found at http://bea.gov. Two very useful databases here are the NIPA accounts (accessed via http://bea.gov/iTable/index_nipa.cfm) and the Fixed Asset Tables (http://bea.gov/iTable/index_FA.cfm). The former is where you would find data on GDP and its various components. The later is a good source for data on the U.S. capital stock.

2.1.2 Bureau of Labor Statistics (BLS)

The BLS website is http://bls.gov/data. Here you will find detailed data on the labor market including employment, unemployment, labor force participation, hours worked, wages and earnings. You can also find detailed data on the CPI and consumer prices.

2.1.3 Federal Reserve Economic Database (FRED)

FRED is found at http://research.stlouisfed.org/fred2/. FRED collects data from both the BEA and BLS and from other sources. As a general rule, if you need to download many related series, such as GDP and all its components, you are better off going to the BEA or BLS websites where this can be accomplished in a single download. If, however, you need to download only a few series, FRED is probably the way to go. It has an easily searchable database and you can find the series you need very rapidly. One very nice feature of FRED is that when you download data you can choose to aggregate it from higher frequencies to lower ones. For example, the CPI numbers are reported once a month, but you may want to take the average of the monthly data over quarters so that the price data you download matches the GDP data you previously downloaded from the BEA. Getting the data directly from the BLS would require you to download the monthly data and then average on your own. FRED, however. allows you to average when you download so that you get a spreadsheet with quarterly data. This can be a real time saver.

2.2 International Data

FRED has a limited amount of international macroeconomic data. Often you will need a more extensive database to get the series you need. Some of the sources below are free or have limited free data.

2.2.1 Organization for Economic Cooperation and Development (OECD)

The OECD maintains a database of macroeconomic data for 34 relatively developed countries at http://stats.oecd.org/. There is a host of data here including NIPA statistics, labor data and a variety of other financial and social statistics.

2.2.2 International Monetary Fund (IMF)

The IMF maintains several databases, but the most useful one for macroeconomics is International Financial Statistics (IFS) at http://elibrary-data.imf.org/finddatareports.aspx?d=33061&e=169393. IFS has a smaller amount of data for each country, but it includes

NIPA data, prices, and other useful series. In addition the IMF covers a much larger set of countries, though some of these have limited or unreliable data.

2.2.3 International Labor Organization (ILO)

Finally, the ILO maintains a database on labor market data much like that maintained by the BLS in the U.S. This site is at http://laborsta.ilo.org/.

3 Natural Logarithms

The natural logarithm (written as ln to distinguish it from the base ten logarithm) of a number is the inverse function of the exponential function. That is, if $y = e^x$, then $x = \ln y$. A useful property of natural logarithms is that the difference between the logs of two numbers is approximately equal to the percent difference as long as the difference is small. More formally, this is written as:

$$\ln x_1 - \ln x_2 \doteq \frac{x_1 - x_2}{\frac{1}{2}(x_1 + x_2)} \tag{2}$$

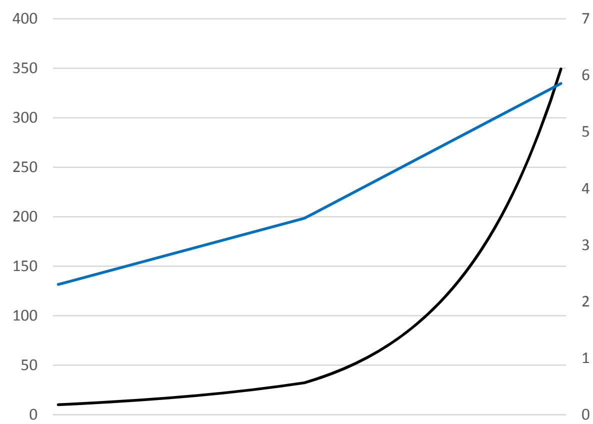
Table 1 shows the accuracy of the approximation for various "small" values of x. For values of x below 0.02 the approximation is quite good, but for values above .20 it is very bad.

Table 1: Approximation Error for $x = \ln(1 + x)$

x	$\ln(1+x)$	difference	percent difference
0	0	0	-
0.000001	1E-06	5E-13	0.0001%
0.00001	1E-05	5E-11	0.0005%
0.0001	1E-04	5E-09	0.0050%
0.001	0.001	5E-07	0.0500%
0.01	0.00995	4.97E-05	0.4992%
0.02	0.019803	0.000197	0.9967%
0.04	0.039221	0.000779	1.9869%
0.06	0.058269	0.001731	2.9709%
0.08	0.076961	0.003039	3.9487%
0.10	0.095310	0.004690	4.9206%
0.15	0.139762	0.010238	7.3254%
0.20	0.182322	0.017678	9.6963%
0.25	0.223144	0.026856	12.0355%
0.50	0.405465	0.094535	23.3152%

This a useful property when plotting the time series of macroeconomic variables. Since log differences are approximately percent differences, the slope of a log of a variable will be its growth rate over time. Figure 3 shows a time series that grows at rate of 5.0% for the first half of the sample and then growth at a rate of 10.0% thereafter. Since the slope of the log of the series is the growth rate, it is easy to spot the change in growth rates. The change is much less obvious when looking at the unadjusted series.

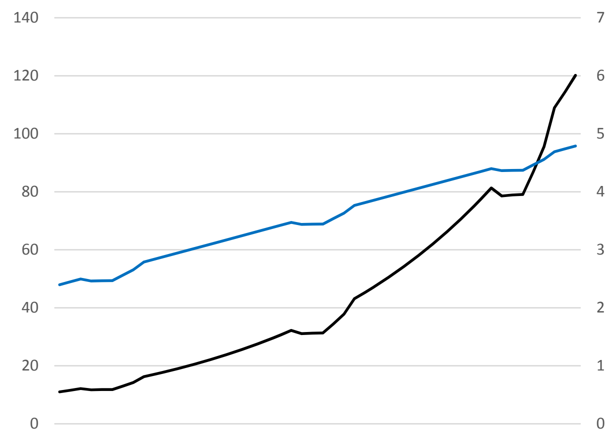
Figure 3: Natural Logarithms Growth Example



log series in blue with scale on the right

This property is also useful for determining the severity of recessions. Figure 4 shows an output series with three recessions. When examining the data in the original levels, the last recession appears the most severe and the first recession seems the most mild. Since log differences are percent differences the height from peak to trough in the log series will show the percent drop during the recession. As the log series shows, all three recession are of the same severity in percentage terms.

Figure 4: Natural Logarithms Recession Example



log series in blue with scale on the right

4 Filtering a Time-Series

In addition to removing seasonality, we also wish to remove the trend from many time series. The method we use below to measure volatility, persistence and cyclicality require that our data have no time trend. Since we are usually working with seasonally adjusted data at this point, the amounts to decomposing the data into a growth component and a cyclical component.

4.1 OLS Filter

The goal of any filter is to decompose the time series into several series with common frequencies. Let y_t be our data at time-period t. We want to decompose the data into growth component, τ_t, and the cyclical component, c_t

$$y_t = \tau_t + c_t \qquad \text{for } t = 1, \ldots, T \tag{3}$$

The simplest of filters involves fitting set τ to be the "polynomial of best fit" through the time series. The "polynomial of best fit" minimizes the following expression:

$$\min_{\beta} \|y - \beta x\|$$

For a polynomial of degree 1 (a straight line), the solution is given by the well known "normal equations" given by $\hat{\beta} = (X^T X)^{-1} X^T y$, where

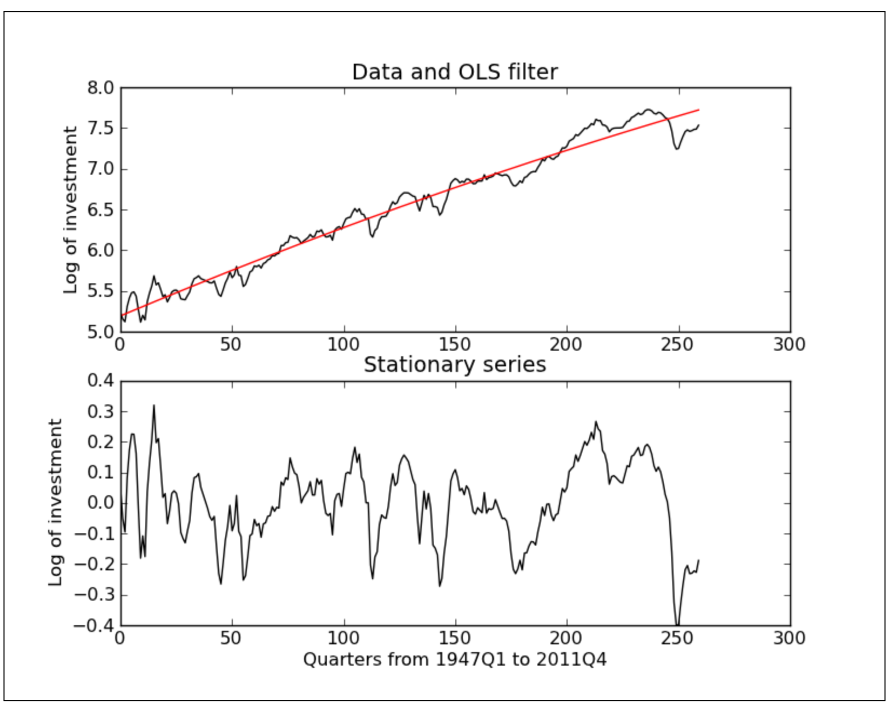
$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{pmatrix}$$

We can generalize this for a polynomial of degree k using the k^{th} degree Vandermonde matrix and the same set of normal equations.

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ & \vdots & \ddots & \vdots & \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{pmatrix}$$

Using a quadratic polynomial filter on logged investment data gives the following trend series (depicted in red) and cyclical series (depicted in the second subpanel)

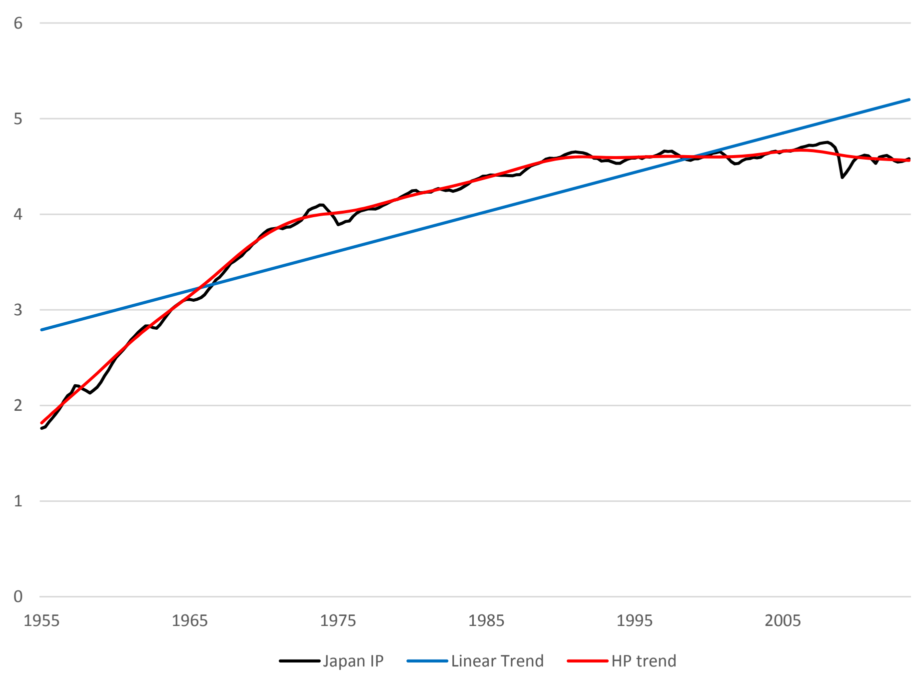
Figure 5: OLS Filter (quadratic)



As desired, the resulting series appears to have a constant mean and variance throughout the entire series.

A linear trend will work fairly well if the trend is characterized by a constant growth rate. Consider figure 6. The black line plots the log of the index of Japanese industrial production from 1955:1 through 2014:4. The blue line is a linear trend fitted by ordinary least-squares (OLS) regression. This is a commonly-used statistical technique that fits a straight line to data. Clearly this is not the appropriate trend to remove if we are focusing on business cycles. Using a linear trend in this case would imply that industrial production was below trend from 1955:1 to 1966:3 and again from 1998:3 to the present. The notion that these two periods were recession periods and the time between was an economic boom is simply wrong. What we need to do is remove a trend that can vary over time. The red line in the figure shows the smoothed values for a time-varying trend computed using the Hodrick-Prescott filter. This method does a much better job of identifying recessions as period where the black line is below the red and booms as periods where the opposite is true.

Figure 6: Log of Japanese Industrial Production
Detrended using the HP-Filter and a Linear Trend



1955:1 - 2013:4, log scale

4.2 Moving Average and Differencing Filters

Suppose we create a filtered series, y_t, by taking a moving average of the raw data series x_t.

$$y_t = \sum_{j=-\infty}^{\infty} b_j x_{t-j} = b(L)x_t \tag{4}$$

A simple first-difference filter is sometimes used to render a series startionary. This is particularly applicable if the time series has a unit root.

$$y_t = x_t - x_{t-1} \tag{5}$$

Differences can be taken over longer periods as well. We will call these "d-difference" filters.

$$y_t = x_t - x_{t-d} \tag{6}$$

Finally, differences can be taken more than once. A second difference filter is given below.

$$y_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) \tag{7}$$

4.3 Hodrick-Prescott Filter

The Hodrick-Prescott (HP) filter was popularized by Hodrick and Prescott (1997). It generates a line of best fit very differently from polynomial interpolation or OLS. An HP is a kind of moving average (MA(k)) filter, meaning that the point in the filtered series at time t is a weighted average of the raw data over k points before/after t. As given in 3, it aims to decompose the raw data into a smooth "trend series" (sometimes referred to as the "growth series") and a stationary "cyclical series" such that $y_t = \tau_t + c_t$

To find the τ series that best fits the data, we minimize the following

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=1}^{T} [(\tau_t - \tau_{t-1}) - (\tau_{t-1} - \tau_{t-2})]^2 \right\} \tag{8}$$

with $c_t = y_t - \tau_t$ giving the deviation from the trend component (a mean zero process). The parameter λ penalizes changes in the trend component. A higher λ will result in a smoother trend component. In fact, as λ approaches infinity, the optimal $\tau_t - \tau_{t-1}$ tends towards a constant β, so $\tau_t = \tau_0 + \beta_t$, and the filtered series is simply the least squares solution. The proof of this is left as an exercise.

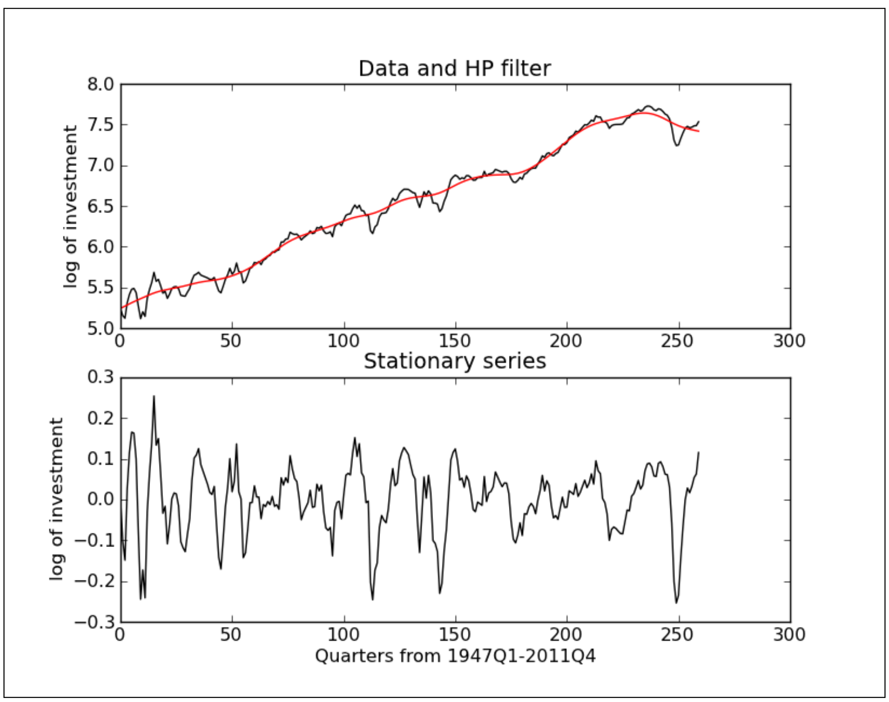
To compute the filter, we can write the first-order-conditions for 8, in matrix form as shown:

$$\Lambda T = Y$$
$$T = \Lambda^{-1} Y$$

where the $(n,m)^{th}$ entry of Λ is the λ coefficient for the i^{th} first-order-condition found by differentiating with respect to τ_i. The matrix Λ is symmetric in all cases and sparse for large values of T, which can allow us to compute Λ^{-1} very quickly.

An HP filter with $\lambda = 1600$ on quarterly data will generate a filter that moves along the data more flexibly than the quadratic polynomial, as shown in the figure below.

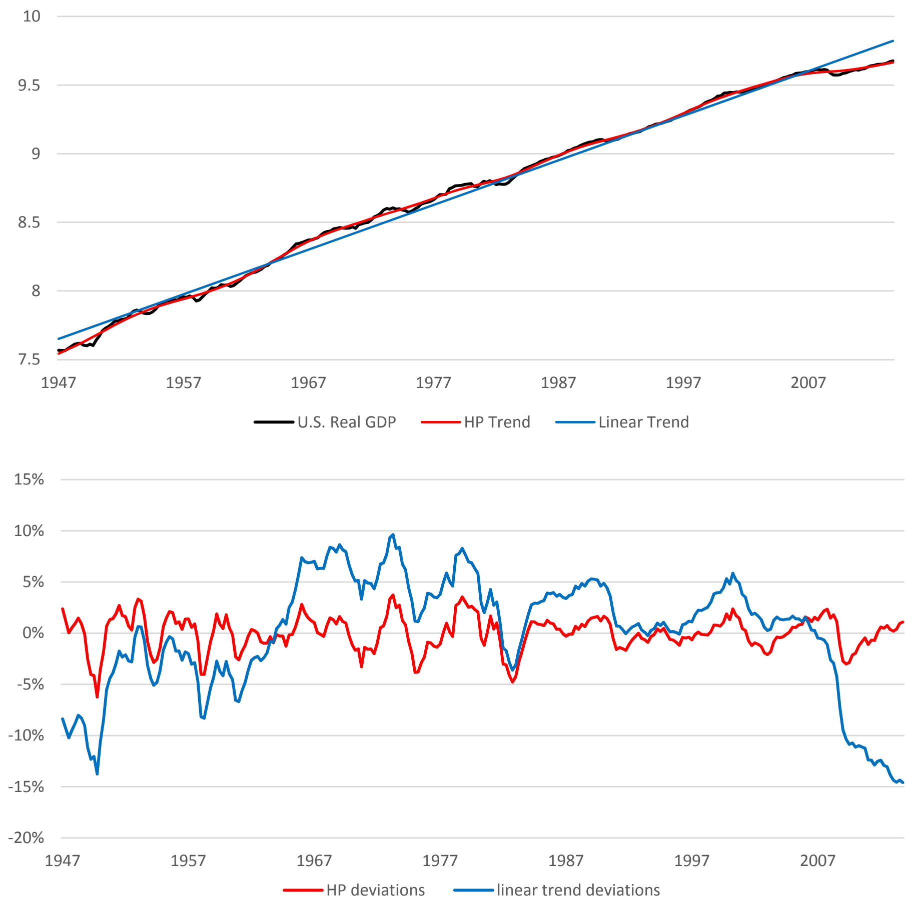
Figure 7: HP Filter ($\lambda = 1600$)



How does the stationary series in figure 7 (HP filter) compare to the stationary series from the OLS filter in figure 5? Does the HP generate a more cyclical series that is more or less stationary than OLS?

Figure 8 shows the effect of applying the HP filter to US GDP data and compares this with removing a linear trend.

Figure 8: Log of US GDP Detrended using the HP-Filter and a Linear Trend



1947:1 - 2013:4

top panel - log scale, bottom panel - percent deviation from trend

When we remove a trend from data that has been transformed to logs, the interpretation of the remaining deviations from trend is the percent the value of the series is below its trend value. For example figure 8 shows that during the lowest point of the most recent recession in the second quarter of 2009, GDP was 3.0% below its trend value as calculated by the HP filter ($14.36 trillion versus $14.80).

4.4 Band-Pass Filters

A band-pass filter works in a way similar to the HP filter in the way that it decomposes the data into a trend series and a cyclical series. The major difference between the band-pass filter and an HP filter is that the band-pass filter is two-sided. That is, it removes frequencies outside the chosen band that are higher than the upper cutoff and lower than the lower cutoff. By contrast, the HP filter is one-sided and removes only frequencies below a lower cutoff. As a result, the term "band-pass filter" refers to a whole family of filtering techniques.

4.5 Filtering in Python

A good source for time-series filter, including the three discussed in this chapter is the `scikits.statsmodels` package. The filters are called as:

- `tsa.filters.hpfilter`
- `tsa.filters.bkfilter`
- `tsa.filters.cffilter`.

The internal documentation follows.

Hodrick-Prescott filter

```
   hpfilter(X, lamb=1600)
Parameters:
X : array-like -- The 1d ndarray timeseries to filter of length (nobs,)
or (nobs,1)
lamb :  float -- The Hodrick-Prescott smoothing parameter.  A value of
1600 is suggested for quarterly data.  Ravn and Uhlig suggest using a
value of 6.25 (1600/4**4) for annual data and 129600 (1600*3**4) for
monthly data.
Returns:
```

```
cycle :  array -- The estimated cycle in the data given lamb.
trend :  array -- The estimated trend in the data given lamb.
```

Baxter-King bandpass filter

```
   bkfilter(X, low=6, high=32, K=12)
Parameters:
X : array-like -- A 1 or 2d ndarray.  If 2d, variables are assumed to be
in columns.
low :  float -- Minimum period for oscillations, i.e., Baxter and King
suggest that the Burns-Mitchell U.S. business cycle has 6 for quarterly
data and 1.5 for annual data.
high :  float -- Maximum period for oscillations BK suggest that the U.S.
business cycle has 32 for quarterly data and 8 for annual data.
K : int -- Lead-lag length of the filter.  Baxter and King propose a
truncation length of 12 for quarterly data and 3 for annual data.
Returns:
Y : array -- Cyclical component of X
```

Christiano Fitzgerald asymmetric, random walk filter

```
   cffilter(X, low=6, high=32, drift=True)
Parameters:
X : array-like -- 1 or 2d array to filter.  If 2d, variables are assumed
to be in columns.
low :  float -- Minimum period of oscillations.  Features below low
periodicity are filtered out.  Default is 6 for quarterly data, giving
a 1.5 year periodicity.
high :  float -- Maximum period of oscillations.  Features above high
```

```
periodicity are filtered out.  Default is 32 for quarterly data, giving
an 8 year periodicity.
drift :  bool -- Whether or not to remove a trend from the data.  The
trend is estimated as np.arange(nobs)*(X[-1] - X[0])/(len(X)-1).
Returns:  cycle :  array -- The features of ‘X‘ between periodicities
given by low and high.
trend :  array -- The trend in the data with the cycles removed.
```

4.6 Filtering Series that are not Bounded by Zero

Some economic variables can be either positive or negative. Since the log of a value less than or equal to zero is undefined, we cannot transform a series like this using natural logs. Even if the series does not exhibit an upward or downward trend over time it may still need to be adjusted for increased variability over time.

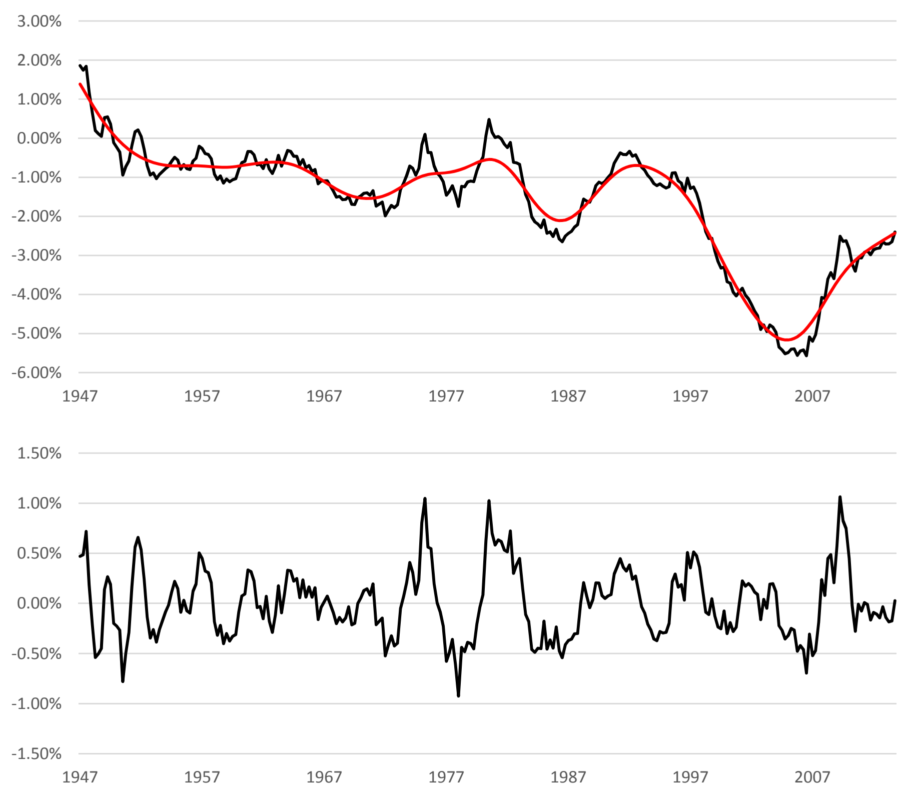
A good example of this is net exports (NX). While both exports and imports must be positive and we can take their natural log, the difference between them can be either positive or negative. See figure 9 below which shows the value of real net exports for the U.S. from 1947:1 to 2013:4. The values are mostly negative but their are periods (1947:1-1949:2, 1951:3-1952:1, 1975:2 and 1980:2-1981:2) when the values are positive.

A common way to adjust this data is to examine NX as a percent of GDP. Figure 10 shows this time series. The swings in net exports still rise over time, but not as much. This series can then be detrended using the HP filter to yield the trend deviations shown in the bottom panel. Since NX/GDP is already a percent, the interpretation of these deviation is the number of percentage **points** NX/GDP is above or below its trend value. So in figure 10 we see that in the second quarter of 2009, the ratio of net exports to GDP peaked a value 1.06 percentage points above its trend (-2.51% versus -3.57%).

Figure 9: U.S. Real Net Exports

1947:1 - 2013:4, billions of constant (2009) dollars

Figure 10: NX/GDP Detrended using the HP-Filter



1947:1 - 2013:4

top panel - log scale, bottom panel - percentage point deviation from trend

5 Spectral Analysis and the Effects of Filters

Any time-series can be fitted with an appropriately high-order polynomial. Suppose we have a time series $X = \{x_1, x_2, \ldots, x_T$. With T observations we can specify a polynomial of order $T-1$ that will pass through each point.

Similarly any time series can be fitted with a weighted sum of an appropriate number of sinusoidal wave functions of various wavelengths. Again, with T observations we can sum $T-1$ wave functions and the resulting time-path will pass through each point. The mapping from observations at various points in time to wave functions of various frequencies is called a Fourier transformation.

When removing trends from data it is useful to think about the effects the filter has on the various constituent sine waves. In order to do this we need to first go though the basics of spectral analysis. We will only be scratching the surface here.

5.1 Discrete Fourier Transform

Consider a sequence of (possibly complex) numbers, $X = \{x_1, x_2, \ldots, x_T\}$. The discrete Fourier transform (DFT) of this series is defined as:

$$\hat{x}_k = \sum_{t=1}^{T} x_t e^{-i2\pi kt/T} \tag{9}$$

Each $\hat{x}_k$ is a complex number that contains information on the amplitude and phase of a sine wave. The sum of these sine waves exactly recreates the time-series X. The frequency for $\hat{x}_k$ is k/T cycles per sample. The amplitude of the sine wave is given by (10) and the phase by (11).

$$\frac{|\hat{x}_k|}{T} = \frac{\sqrt{Re(\hat{x}_k)^2 + Im(\hat{x}_k)^2}}{N} \tag{10}$$

$$arg(\hat{x}_k) = atan2\{Im(\hat{x}_k), Re(\hat{x}_k)\} \tag{11}$$

The inverse discrete Fourier transform (IDFT) is:

$$x_t = \frac{1}{T} \sum_{k=1}^{T} y_k e^{i2\pi kt/T} \tag{12}$$

The DFT is often written as $\mathcal{F}$ so that the series $\hat{X} = \{\hat{x}_1, \hat{X}x_2, \ldots, \hat{x}_T\}$ can be expressed as the DFT of X by writing $\hat{X} = \mathcal{F}(X)$. The IDFT, not surprisingly, is written $X = \mathcal{F}^{-1}(\hat{X})$

X is a representation of the series in the "time domain" and Y is a representation of the same series in the "frequency domain".

5.2 Spectral Density

Consider an discrete infinite time series denoted X. We will define the power of X as in (13).

$$P \equiv \lim_{t\to\infty} \frac{1}{2T} \sum_{-T}^{T} x_t^2 \tag{13}$$

Consider now a subsample of the series running from 1 to T and denote this as $X^T(t)$

If we define ω as the "angular frequency" measured in radians, so that $\omega = 2\pi k$, then the DFT of $X^T(t)$ can be denoted $\hat{X}^T(\omega) = \mathcal{F}\{X^T(t)\}$. The power spectral density (PSD), spectral density function (SDF) or simply "spectral density" of X is defined in (14).

$$S_x(\omega) \equiv \lim_{T\to\infty} E\{|\hat{X}^T(\omega)|^2\} \tag{14}$$

Since ω is the angular frequency all the information about the spectral density is contained on the interval $[0, 2\pi]$.

Some useful properties of the SDF are: First, $S_x(-\omega) = S_x(\omega)$, the SDF is symmetric about $\omega = 0$. Second, the SDF describes the distribution of the variance of X^T over various frequencies. $\mathrm{Var}(X) = 2\int_0^{1/2} S_x(\omega)d\omega$.

The cross-spectral density between two time series X and Y is defined in (15)

$$S_{xy}(\omega) = \lim_{T\to\infty} E\left\{|\hat{X}_T(\omega)\hat{Y}_T(\omega)|\right\}. \tag{15}$$

The cross-spectral density describes the distribution of the covariance of X^T and Y^T over various frequencies. $\text{Cov}(X,Y) = 2\int_0^{1/2} S_{xy}(\omega)d\omega$.

5.3 Periodogram

In practice, we cannot get an infinite series of observations and we need an approximation or estimate of based on finite data. The classic estimate is the periodogram. The periodogram is defined by (16).

$$\begin{aligned} a(k) &= \frac{2}{T}\sum_{t=1}^{T} x_t \cos(kt)dt \\ b(k) &= \frac{2}{T}\sum_{t=1}^{T} x_t \sin(kt)dt \\ P_x(k) &= \sqrt{a(k)^2 + b(k)^2} \end{aligned} \tag{16}$$

The periodogram is normally plotted with ω on the horizontal axis and P_x on the vertical axis.

There are a host of additional issues involved with properly estimating a the spectral density, and other transformations exist that are more appropriate for different contexts. We will stick with the periodogram.

In Python there are several repackaged functions that will generate the periodogram. The most accessible is found in the Scipy package – `scipy.signal.periodogram`

5.4 Effects of Filters

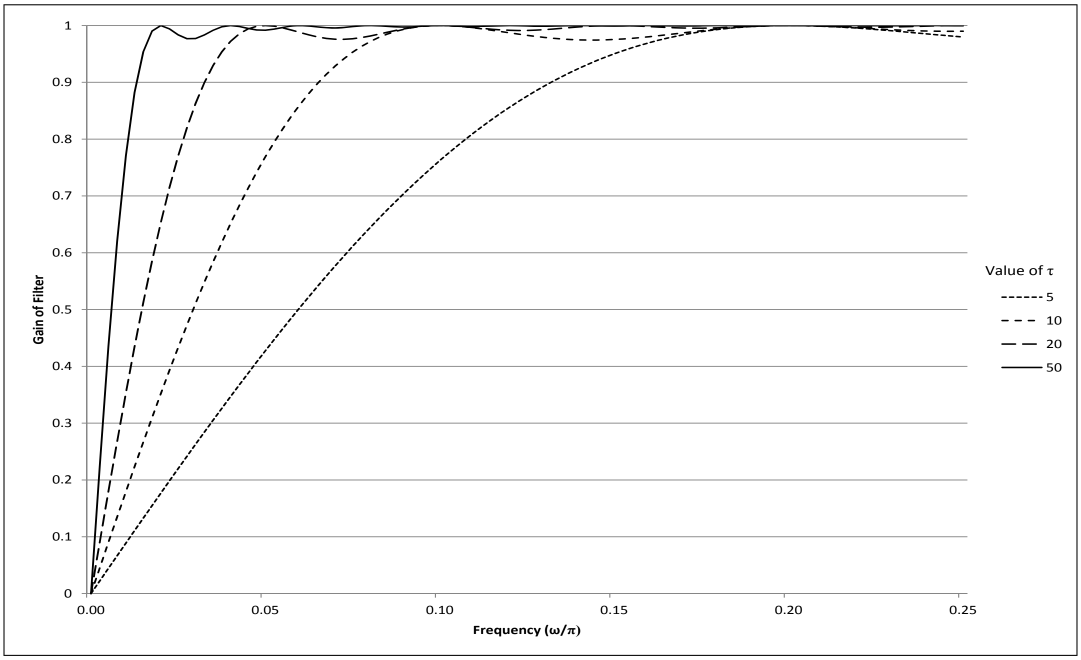
5.4.1 Moving Average Filters

The Fourier transform of an MA(1) process, $x_t = (1+\theta L)\epsilon_t$ is given by $S_x(\omega) = (1+\theta e^{-\omega})(1+\theta e^{i\omega})\sigma_\epsilon^2 = (1+\theta(e^{i\omega}+e^{-i\omega})+\theta^2)\sigma_\epsilon^2 = (1+2\theta cos(\omega)+\theta^2)\sigma_\epsilon^2$

Pedersen (2001) shows that the power transfer function for a moving average filter is $|H_{MA}(\omega)|^2 = \left(\frac{\sin(\tau\omega/2)}{\tau\sin(\omega/2)}\right)^2$, where τ is the number of periods over which the moving-average is taken. The gain of any filter is related to the transfer function by the following formula.

$$G = \sqrt{1 - |H(\omega)|^2}. \tag{17}$$

Figure 11: Gain of the Moving-Average Filter



5.4.2 Hodrick-Prescott Filter

We can also use it on an HP filter. Start by rewriting 8 using the lag operator L to put all τ_i variables into the same period t. Differentiate with respect to τ_t to get the following:

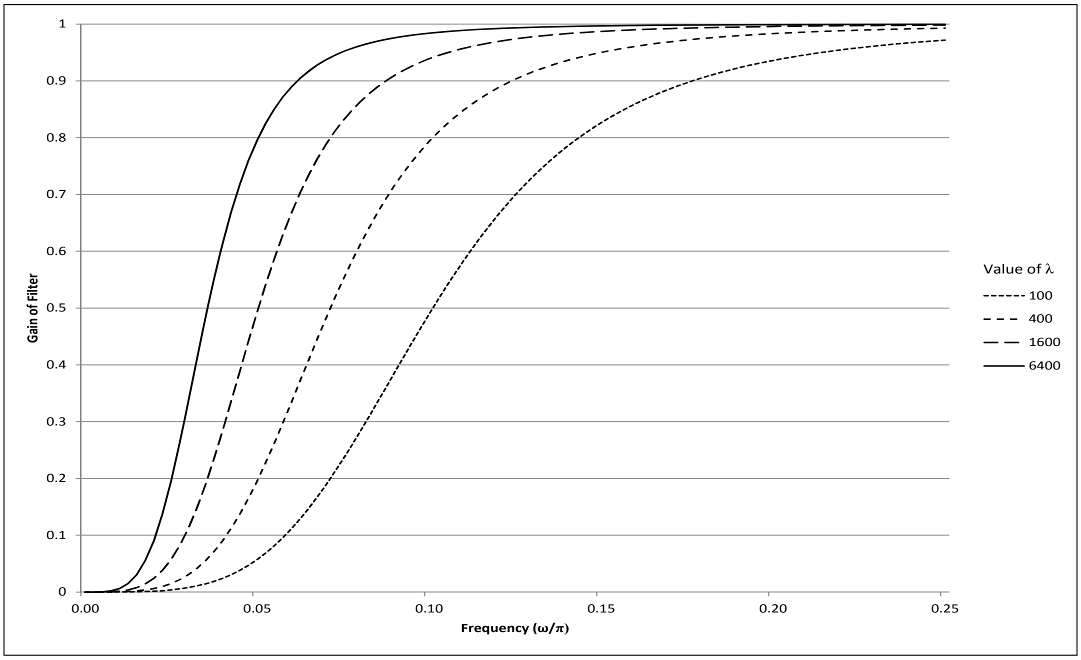
$$C(L) \equiv \frac{\lambda(1-L)^2(1-L^{-1})^2}{\lambda(1-L)^2(1-L^{-1})^2+1} = \frac{\lambda L^{-2}(1-L)^4}{\lambda L^{-2}(1-L)^4+1} \tag{18}$$

The filtering formula gives us the following:

$$|H_{HP}(\omega)|^2 = \left|\frac{\lambda(1-e^{-i\omega})^2(1-e^{i\omega})^2}{\lambda(1-e^{-i\omega})^2(1-e^{i\omega})^2+1}\right|^2 = \left|\frac{4\lambda[1-\cos(\omega)]^2}{4\lambda[1-\cos(\omega)]^2+1}\right|^2 \tag{19}$$

Plotting the gain over the domain of $\omega \in [0, \pi)$ gives us an idea of what frequencies are filtered out by the HP series. Figure 12 gives plots for several values of λ .

Figure 12: Gain of the HP Filter



Hodrick and Prescott (1997) argue that the appropriate value of λ when filtering quarterly data for business cycle frequencies is $\lambda = 1600$.

5.4.3 Differencing Filters

Other simple filters have been proposed and used. It is helpful to examine their effect on the spectral density of a time series. Pedersen (2001) shows the power transfer functions for some of these filters. One of these that is commonly used is the First-Difference filter. The power transfer function for a first-difference filter is $|H_{FD}(\omega)|^2 = 2[1 - cos(\omega)]$. From this we can calculate the gain using (17).

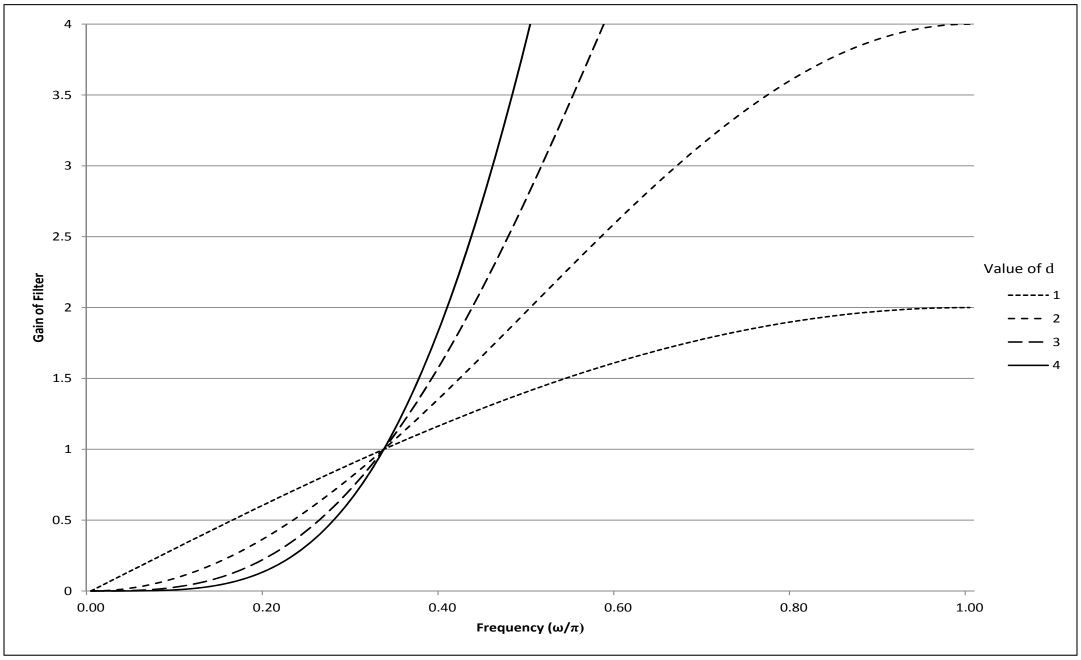
Differencing the series d times is a d-difference filter and it's power transfer function is given by (20)

$$|H_{DD}(\omega)|^2 = 2^d[1 - cos(\omega)^d] \tag{20}$$

5.4.4 Band-Pass Filters

A band-pass filter works in a way similar to the HP filter in the way that it decomposes the data into a trend series and a cyclical series. The major difference between the band-

Figure 13: Gain of the d-Difference Filter

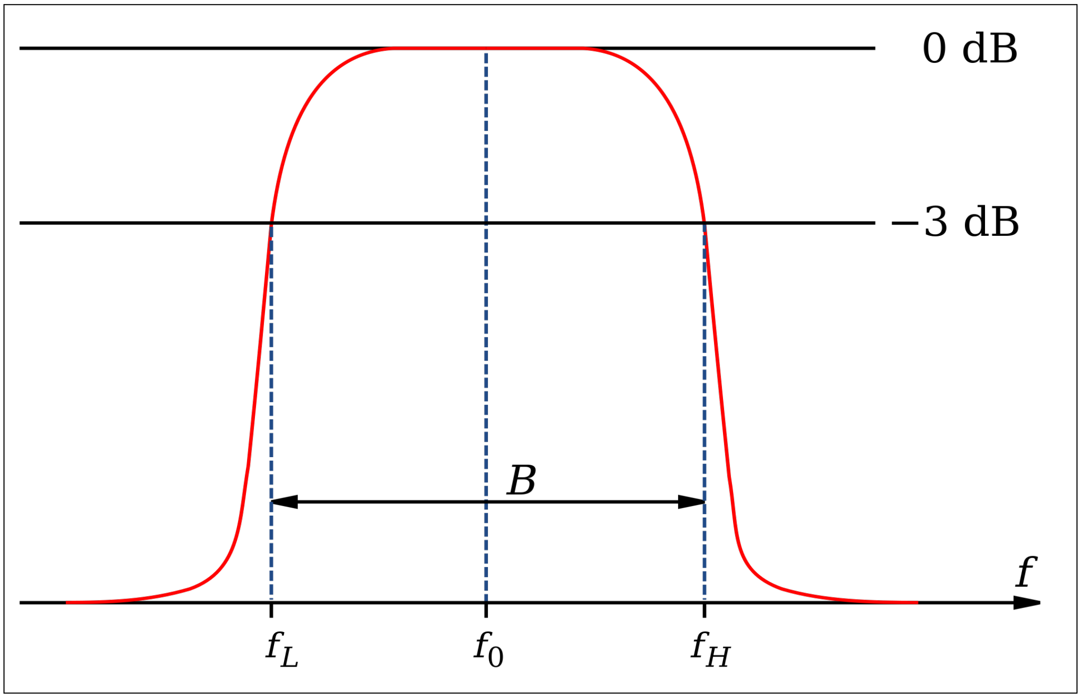


pass filter and an HP filter is that the band-pass filter is two-sided. That is, it removes frequencies outside the chosen band that are higher than the upper cutoff and lower than the lower cutoff. By contrast, the HP filter is one-sided and removes only frequencies below a lower cutoff. As a result, the term "band-pass filter" refers to a whole family of filtering techniques.

Figure 14, which is shamelessly downloaded from Wikipedia, illustrates a typical band-pass filter, in this case for use with audio equipment. The gain of the filter is the red line. The cutoffs for the bandwidth (B) are f_L and f_H. An ideal filter would follow the vertical dashed lines and be zero outside the bandwidth and one inside. However, with finite samples it is impossible to actually implement an ideal filter. As the red line shows, some frequencies outside the band will be passed, and some frequencies inside the band will be attenuated. A good filter is one that comes close to the step function of the ideal filter.

There are a variety of band-pass filters available for use with economic time series. The most widely used are the Baxter-King filter (Baxter and King, 1999) and the Christiano-Fitzgerald filter (Christiano and Fitzgerald, 2003). Band-pass filters specify a range of frequencies that are passed through the filter. For business cycle analysis frequencies corresponding to a band between six and thirty-two quarters are most often used.

Figure 14: Stylized Band-Pass Filter



The Hodrick-Prescott filter, Baxter-King filter and Christiano-Fitzgarald filter are all available through the Statsmodels package. You must install this and import it using "`import statsmodels.api as sm`". See details at http://www.statsmodels.org/. The filters are called with:

- `sm.tsa.filters.hpfilter(data, lambda)`
- `sm.tsa.filters.bkfilter(data, low, high, K)`
- `sm.tsa.filters.cffilter(data, low, high, drift`

6 Calculating Moments

Once we have rendered a time-series stationary by removing its growth trend, we want to analyze its behavior. Since we have taken natural logs of our original data and have removed a growth trend, the portion remaining is interpretable as the percent deviation of the variable from its growth trend. We are particularly interested in the following: volatility, cyclicality, and persistence.

Volatility is how much the variable in question moves around over time, in our case this is in percent terms. We will use the standard deviation of a filtered series as our measure of volatility.

Cyclicality measures how the variable moves in relation to the business cycle. A variable such as consumption, which rises when the business cycle is on an upswing, is considered **procyclical**, while one which falls when the business cycle improves, such as unemployment, is **countercyclical**. If a variable moves in ways that are unrelated to the business cycle, or remains constant, it is considered **acyclical**. We will use the correlation coefficient of a filtered series with filtered GDP as our measure of cyclicality.

Persistence tells us how likely it is that when a series is above trend today it will remain above trend next period. We will use the autocorrelation coefficient of a filtered series as our measure of persistence.

Let us denote the filtered time series we are considering as $\{x_t^C\}_{t=1}^T$ and the filtered time series for GDP as $\{y_t^C\}_{t=1}^T$.

6.1 Means

Our filtered time series can be thought of as a random variable. The mean of a random variable is defined as $\mu_x \equiv E\{x\}$, where E stands for the expected value and x is the random variable in question.

The sample mean of x is calculated using the formula in equation (21).

$$\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t^C \tag{21}$$

It turns out that by removing the trend from our time series we guarantee that the sample mean ($\bar{x}$) is zero. We will continue to carry the term $\bar{x}$ in our formulas below, however, to make them as general as possible.

6.2 Variance and Standard Deviation

The variance of a variable is also known as its second moment about the mean. This is formally defined as $\sigma_x^2 = E\{(x - \bar{x})^2\}$.

A sample variance can be calculated using the following formula.

$$s_x^2 = \frac{1}{T-1} \sum_{t=1}^{T} (x_t^C - \bar{x})^2 \tag{22}$$

Note that since each term of the sum in equation (22) is squared, the variance must be a number that is greater than or equal to zero. It cannot be negative. The larger the deviations of the series from its mean (of zero in our case) the bigger the terms in the sum will be on average. Hence, a more volatile series will have a larger variance.

The variance is just the average value of the squared deviations from the mean. However the squaring of the deviations can lead to difficulties in interpretation. A simpler concept to interpret is the standard deviation. The standard deviation is interpretable as the average distance of the variable from its mean over the sample. The standard deviation is simply the square root of the variance, or s_x

6.3 Covariance and Correlation

A covariance is also a second moment about means, but unlike the variance, it considers two variables. Formally the covariance between two variables x and y is defined as $\sigma_{xy} = E\{(x - \bar{x})(y - \bar{y})\}$

The sample covariance between two variables is calculated as in equation (23).

$$s_{xy} = \frac{1}{T-1} \sum_{t=1}^{T} (x_t^C - \bar{x})(y_t^C - \bar{y}) \tag{23}$$

Note now that the terms in the sum can be either positive or negative. If the two series tend to move together, then when x is above its mean y will usually be above its mean and the term for that observation will be a positive number. When x is below its mean y will also be below its mean and again term will be a positive number from multiplying two negative numbers. Hence the sample covariance will contain mostly positive terms and will yield a positive sum. If the two variables move in opposite directions however, then when one is above the mean the other will tend to be below its mean and the terms will be mostly negative yielding a negative sample covariance.

Note that the variance is a special case of the covariance. That is, the variance is the covariance of a variable with itself.

Covariances have interpretation problems just as variances do. A covariance can be large for two reasons. First, the two variables could be highly correlated. Second, they could have high variances leading to big terms in the sum. To control for this second effect we divide the covariance by the two standard deviations to get the correlation coefficient. This is calculated as shown below.

$$\rho_{xy} = \frac{s_{xy}}{s_x s_y} \tag{24}$$

By construction, this number will always lie between -1 and +1. Hence highly correlated series will have correlation coefficients close to one. For our purposes a series is considered procylical if its correlation coefficient with GDP is positive. If the correlation coefficient is negative, the variable is countercyclical. If the correlation coefficient is close to zero, the series is considered acyclical.

6.4 Autocorrelations

An autocorrelation is the correlation coefficient of a variable with the lagged value of itself. A sample autocorrelation coefficient is calculated as in equation (25).

$$\rho_x = \frac{a_x}{s_x^2};\ a_x = \frac{1}{T-2}\sum_{t=2}^{T}(x_t^C - \bar{x})(x_{t-1}^C - \bar{x}) \tag{25}$$

The more persistent a series is, the more likely it will remain above (below) its mean next period when it is above (below) its mean this period. Hence a highly persistent series will have an autocorrelation coefficient that is close to one, while a series with little persistence will have one close to zero.

Note that we can calculate autocorrelation coefficients for lags longer than one period if we wish. A more general formula for (25) is given by equation (26)

$$\rho_{xi} = \frac{a_{xi}}{s_x^2};\ a_{xi} = \frac{1}{T-i-1}\sum_{t=i+1}^{T}(x_t^C - \bar{x})(x_{t-i}^C - \bar{x}) \tag{26}$$

6.5 Measures of Business Cycle Asymmetry

The onset of recessions is more sudden and characterized by greater drops in GDP than the more gradual recovery from recessions during an economic boom. Sichel (1993) proposed a couple of measures designed to quantify this asymmetry.

Sichel recommends using the sample skewness as one measure. He calls this “deepness”. Skewness is calculated using formula in (27).

$$d_x = \frac{T\sum_{t=1}^{T}(x_t^C - \bar{x})^3}{(T-2)s_x^3} \tag{27}$$

where s_x is defined in (22).

Sichel’s second measure is called “steepness” and is the skewness of the difference in a series between two successive periods. For lack of a better alternative we will denote this as g_x.

$$g_x = \frac{(T-1)\sum_{t=2}^{T}(\Delta x_t^C - \Delta\bar{x})^3}{(T-3)s_{\Delta x}^3};\ \Delta x_t^C \equiv x_t^C - x_{t-1}^C \tag{28}$$

Note that because we take differences the sample size for this statistic will be $T-1$.

7 U.S. Moments

Applying the above methodology to a wide variety of data from the United States we find the moments reported in table 2.

There are several key facts to glean from table 2. We discuss a few of the most important ones below.

Table 2: Stylized Facts for the U.S. Economy - HP Filter

	Volatility		Persistence	Cyclicality		Asymmetry	
variable	standard deviation	relative to GDP	auto-correlation	correlation -w- GDP		deepness	steepness
GDP	1.66%	1.0000	0.8527	1.0000	strongly procyclical	0.6062	0.2055
Consumption	1.29%	0.7737	0.8186	0.7763	strongly procyclical	-0.0689	0.5372
Investment	7.54%	4.5411	0.7960	0.8431	strongly procyclical	0.6083	0.2423
Government Spending	3.35%	2.0178	0.9043	0.1521	mildly procyclical	0.0763	-1.3746
Net Exports	0.34%	0.2054	0.8287	-0.3272	countercyclical	-0.4268	-0.1434
Exports	5.47%	3.2940	0.7228	0.4166	procyclical	-0.0739	-0.1236
Imports	5.15%	3.1032	0.7304	0.7117	procyclical	0.7426	-0.3485
Employment	1.08%	0.6501	0.8950	0.7920	strongly procyclical	0.2520	0.2302
Hours	1.02%	0.6123	0.8006	0.7800	strongly procyclical	0.9120	0.1540
Total Labor	1.83%	1.1001	0.8571	0.8990	strongly procyclical	0.5633	0.4712
Nominal Wages	0.90%	0.5444	0.7966	0.0114	acyclical	0.0936	-0.2438
Real Wages	1.06%	0.6408	0.7901	0.2578	procyclical	-0.0761	-0.4862
Nominal Interest Rate	1.12%	0.6732	0.8261	0.3993	procyclical	-0.6874	0.6155
Real Interest Rate	1.08%	0.6485	0.7141	0.0528	acyclical	-0.5591	-0.6989
Money Supply	2.53%	1.5221	0.9239	0.0616	acyclical	-0.1132	-1.2376
Real Money Supply	3.16%	1.9029	0.9241	0.2045	mildly procyclical	0.0678	-1.5241
GDP Delfator	0.93%	0.5612	0.9031	-0.1696	mildly countercyclical	0.0727	-1.6188
CPI - All	1.35%	0.8124	0.9082	-0.1957	mildly countercyclical	-0.0095	-0.2552
Solow Residual 1	0.83%	0.4987	0.6762	0.5682	procyclical	0.3517	0.1109
Labor Productivity	0.79%	0.4734	0.5702	0.0134	acyclical	0.2157	0.1028
Unemployment Rate	14.49%	8.7260	0.8933	-0.8633	countercyclical	-0.4849	-1.1987
Participation Rate	0.39%	0.2327	0.6947	0.2484	mildly procyclical	-0.0477	-0.1903
Budget Deficit	1.35%	0.8114	0.7929	-0.7277	countercyclical	0.1233	1.5156

7.1 Volatility

- First, we have quantified the volatility of GDP at 1.66%, meaning that on average GDP is a little under two percent away from its trend value. This turns out to be a fairly easy fact to replicate in most macroeconomic models, but it is a key one nonetheless.

- Consumption is less volatile than output with about 80% of the volatility we observe in GDP. We will discuss this phenomenon when we talk about savings and consumption behavior in a dynamic framework in later chapters. However, we point out now that this is consistent with the notion of "consumption smoothing".

- Investment is several times more volatile than output. In our sample it is four and a half times as volatile in percentage terms. Investment is the component of GDP that is most sensitive to the business cycle.

- Both exports and imports are more than three times as volatile as GDP. Comparing net exports is problematic, however, since the measures of volatility are different due to the different filtering techniques.

- Government spending is twice as volatile as GDP.
- Employment and hours per worker are less volatile than GDP, but their product, total labor, has roughly the same volatility as GDP (1.83% vs 1.66%).
- Real and nominal wages and interest rates are less volatile than GDP, while real and nominal money supplies are more volatile.
- Prices are less volatile than GDP.
- Total factor productivity and labor productivity (GDP per unit of labor) are both less volatile than GDP.

7.2 Cyclicality

- Consumption, investment, employment, hours per worker, and total labor are all highly procyclical.
- While exports and imports are both procyclical, imports are much more highly correlated with GDP. Net exports are countercyclical. This means as GDP rises exports and imports both rise, but imports rise more or more often and net exports therefore fall on average.
- Government purchases of goods and services is essentially acyclical.
- Real wages are mildly procyclical, while nominal wages are acyclical.
- Nominal interest rates are mildly procyclical, and real interests rates are acyclical.
- The real money supply is mildly procyclical, nominal money is acyclical.
- Prices are mildly countercyclical.
- Total factor productivity is procyclical, but labor productivity is acyclical.
- The unemployment rate and the budget deficit are both strongly countercyclical.
- The labor force participation rate is mildly procyclical.

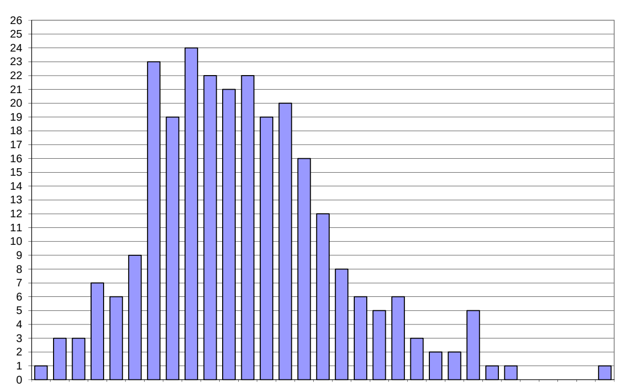
7.3 Persistence

All of the variables listed in our table have a fairly high degree of persistence. The lowest is labor productivity at 0.5702 which implies a half-life[1] of just over one quarter (1.23 quarters). That is, when labor productivity jumps up above trend it will take it just over one quarter to return halfway to trend. The highest persistence is for the nominal money supply at 0.9239, implying a half-life of just over two years (8.74 quarters).

7.4 Asymmetry

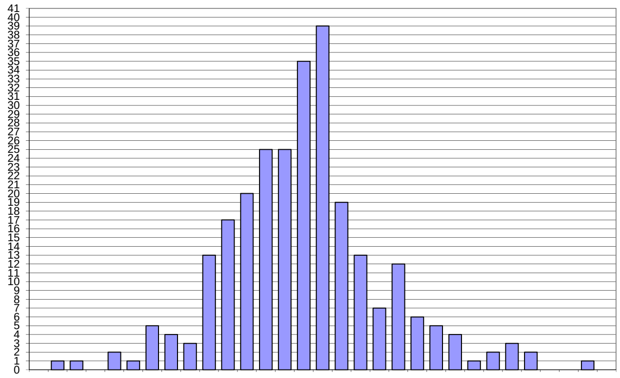
Table 2 shows that no variables have deepness or steepness of exactly zero, indicating there is at least some symmetry. To get a feel for how much asymmetry there is we plot histograms for the values of HP-filtered GDP and its difference in figures 15 and 16. The deepness and steepness numbers, which are skewness coefficients for these two time series are 0.6062 and 0.2055, respectively. The former value is one of the largest magnitudes in the table. Figure 15 shows that there is a noticeable positive skewness in the data, but it is not overwhelming, and might be very close to zero were it not for a few outliers. The skewness in figure 16 is much less noticeable. We conclude that while asymmetries exist, they are not of critical importance in understanding business cycles. The one caveat is that the budget deficit has a great deal of positive deepness, indicating that it has more large increases than large decreases.

Figure 15: Distribution of HP-Filtered GDP Values



[1]The time it takes a variable to return halfway from its current value to its trend value.

Figure 16: Distribution of Differenced HP-Filtered GDP Values



Exercises

Exercise 1. Prove that the spectral density of $X + Y$ is given by:

$$S_X(\omega) + S_Y(\omega) + 2Re(S_{XY}(\omega))$$

Exercise 2. Recall the minimization function for an HP filter is

$$\min_{g_t}\{\sum_{t=1}^{T}(y_t - g_t)^2 + \lambda \sum_{t=2}^{T}[(g_{t+1} - g_t) - (g_t - g_{t-1})]^2\} \tag{29}$$

Prove that as $\lambda \to \infty$, the filtered series g_t is linear ($g_t = g_0 + \beta t$ for some β)

Let $0 < \lambda < \infty$. What are the first order conditions for 29 for $t = 1, 2$ and the general case k?

Exercise 3. Find and plot the (non-smoothed) spectral density estimates for the following quarterly US time series: real GDP, real consumption, real nvestment, and the GDP deflator.

Exercise 4. Use the HP filter to filter the time series from Exercise 3 taking $\lambda = 1600$. Find and plot the spectral density estimates for both the trend and cyclical parts of the filtered time series.

Exercise 5. Using the data from Exercise 3, filter each series using the HP filter with values of $\lambda = \{100, 400, 1600, 6400, 25600\}$. For each series and each value of λ, calculate the standard deviation, autocorrelation, and correlation with GDP. Report these values and comment on the results. Plot the time-series for investment along with its 5 smoothed trend series.

Exercise 6. Replicate Table 2 using data collected from the sites mentioned in Section 2. Use quarterly data going back as far as 1947, if possible. Do this for each of the following filters:

- A linear trend filter
- the first difference ($y_t - y_{t-1}$) filter
- HP($\lambda = 1600$)
- BP(6,32,$K = 8$)

Do not calculate the deepness and steepness measures.

References

Baxter, Marianne and Robert G. King, "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series," *The Review of Economics and Statistics*, November 1999, *81* (4), 575–593.

Christiano, Lawrence J. and Terry J. Fitzgerald, "The Band Pass Filter," *International Economic Review*, 05 2003, *44* (2), 435–465.

Hodrick, Robert and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 1997, *9*, 1–16.

Pedersen, Torben Mark, "The Hodrick-Prescott filter, the Slutzky effect, and the distortionary e!ect of filters," *Journal of Economic Dynamics and Control*, 2001, *25*, 1081–1101.

Sichel, D., "Business Cycles Asymmetry: A Deeper Look," *Economic Inquiry*, 1993, *31*, 224–236.