ECOM20001
Econometrics 1

Lecture Note 10
Time Series Regression

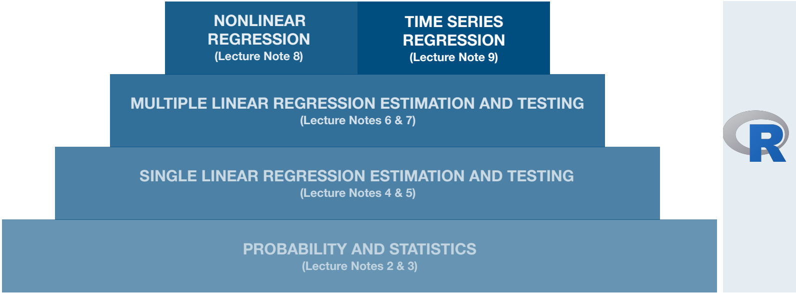
A/Prof David Byrne
Department of Economics
University of Melbourne

Stock and Watson: Chapter 15
sections 15.1-15.6 (inclusive)

Summary of Key Concepts

- Time series data
- Lags, first differences, growth rates
- Autocorrelation and autocovariance
- Autoregressions, AR(1) and more general AR(p) model
- Forecasting and forecast intervals
- Autoregressive distributed lag model AR(p, q)
- Generalised autoregressive distributed lag models AR($p, q_1, \ldots, q_k$)
- Key assumptions of ADL model
- Granger Causality to test for predictive content
- Model selection and determining model lag length
- Accounting for seasonality with dummy variables

Building our Econometric Toolkit

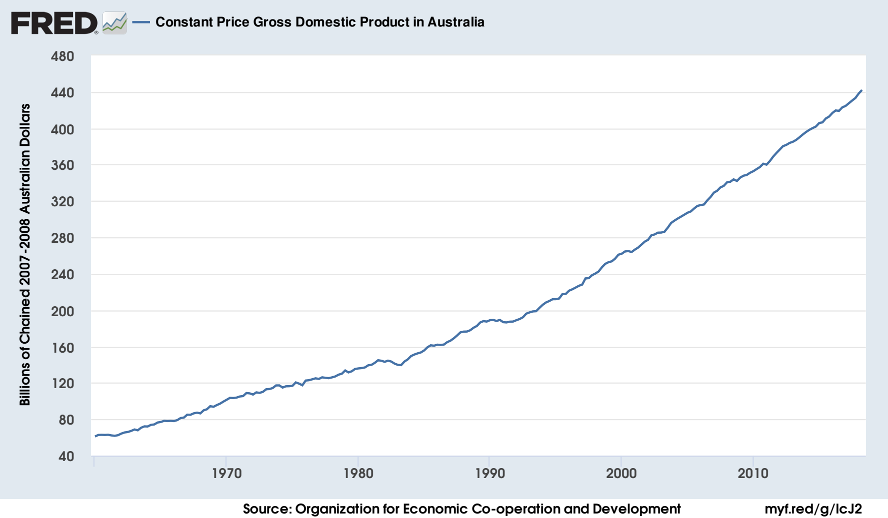


Cross-Sectional and Time Series Data

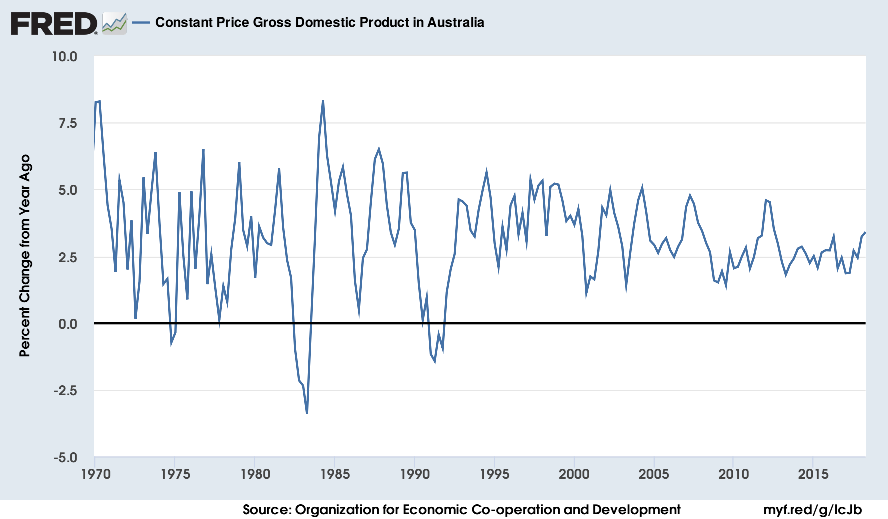
- For the entire subject, we have so far mainly worked with cross-sectional data based on a random sample across an observational unit (e.g., classroom, individual, country)
  - average test scores across classrooms in Victoria
  - earnings across immigrants and non-immigrants
  - GDP across countries
  - voting on women's issues across congress members
- In such studies, the cross-sectional unit index i and sample size of focus is $i = 1, \ldots, n$
- This lecture introduces a different data type: time series data where the time-series unit index and sample size of focus is in terms of time/periods: $t = 1, \ldots, T$

Australian GDP: 1960-2018, Quarterly

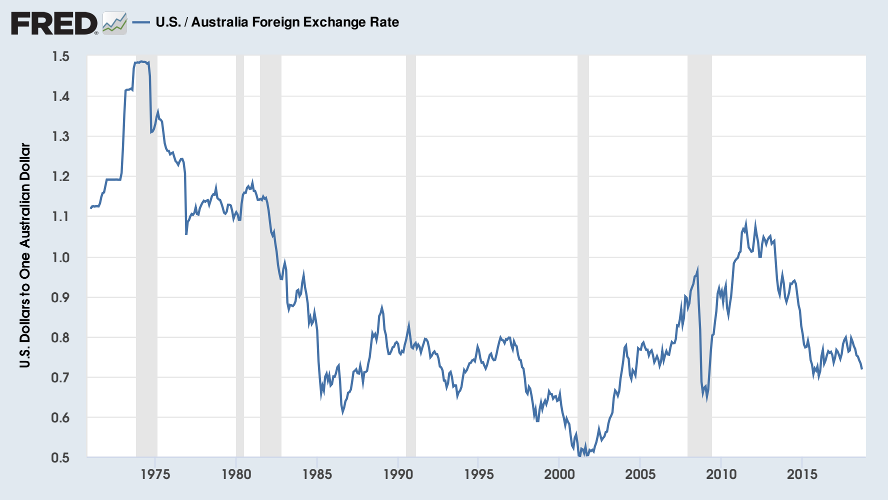
Source: Federal Reserve Bank of St Louis http://fred.stlouisfed.org



Australian GDP Growth: 1970-2018, Quarterly

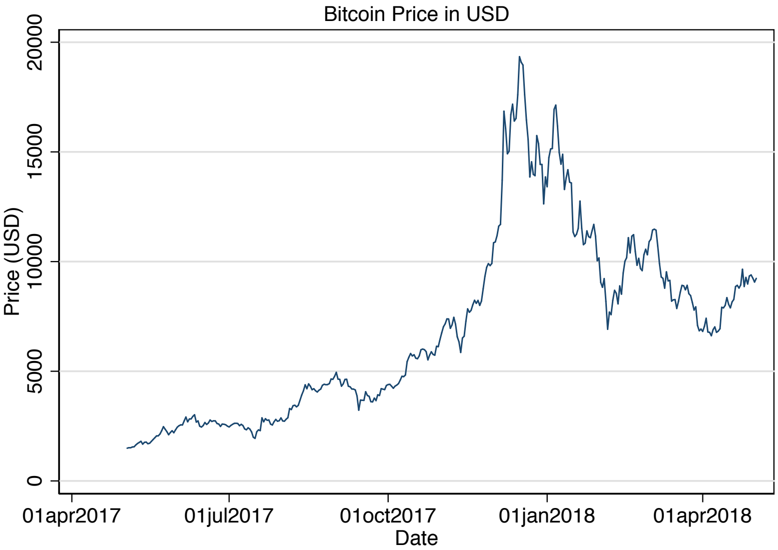


AUD/USD Exchange Rate: 1970-2018, Monthly

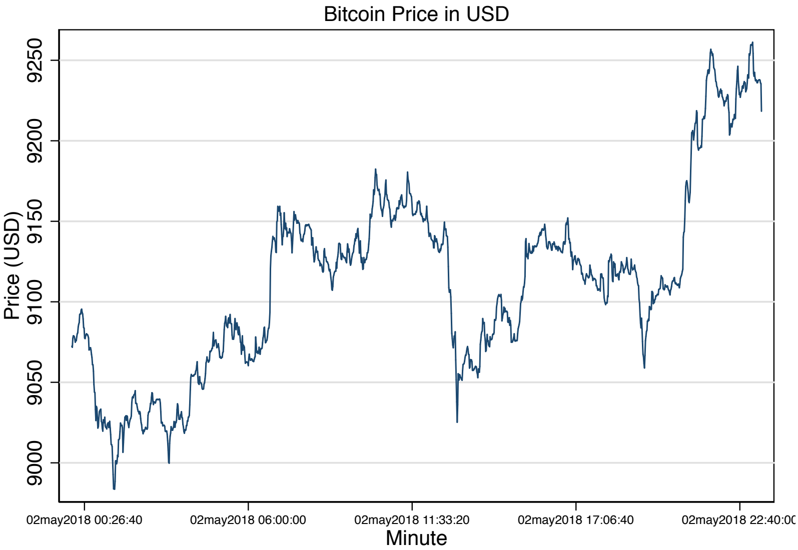


Shaded areas indicate U.S. recessions **Source: Board of Governors of the Federal Reserve System (US)** **myf.red/g/lcKs**

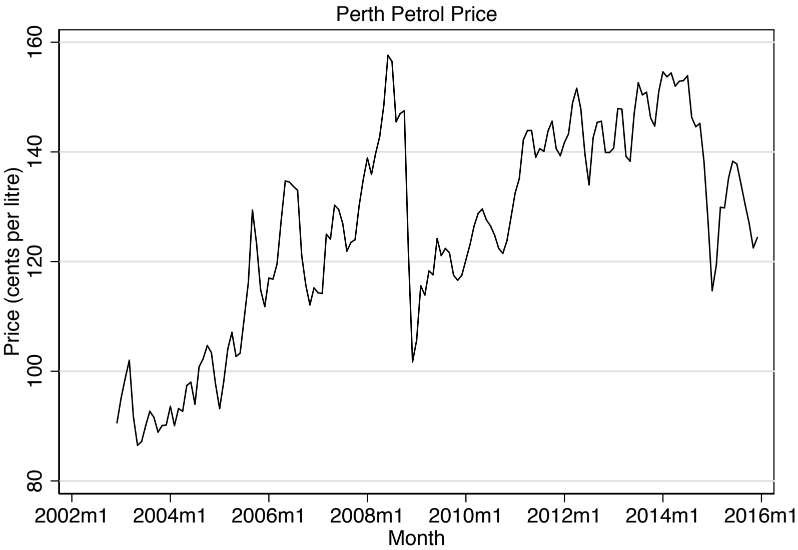
Bitcoin Price in USD: 2017-18, Daily



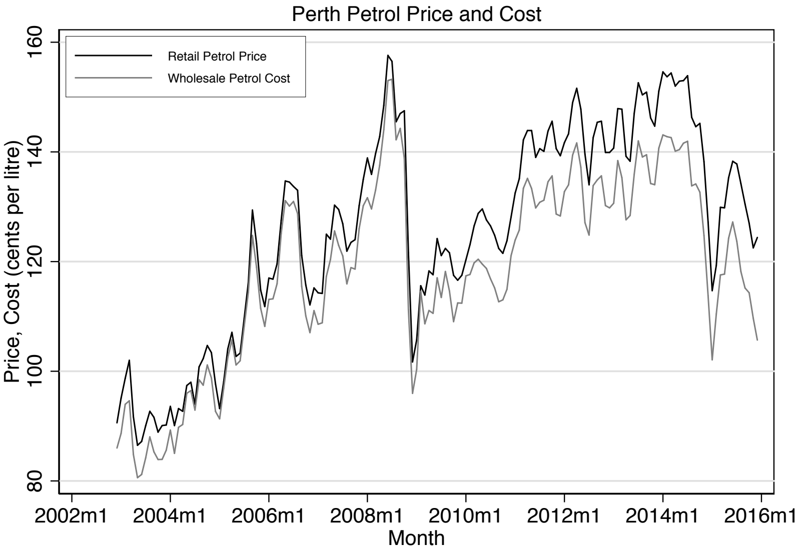
Bitcoin Price in USD: 2 May 2018, Every Minute



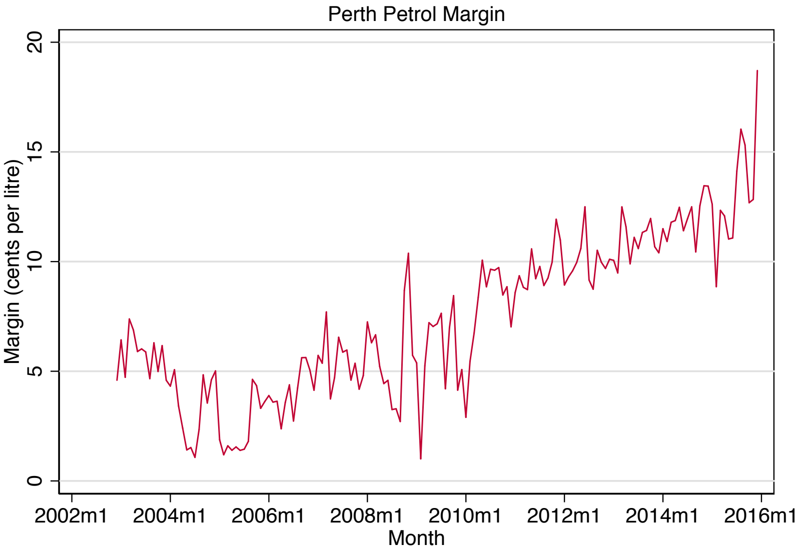
Perth Petrol Prices: 2003-2016, Monthly



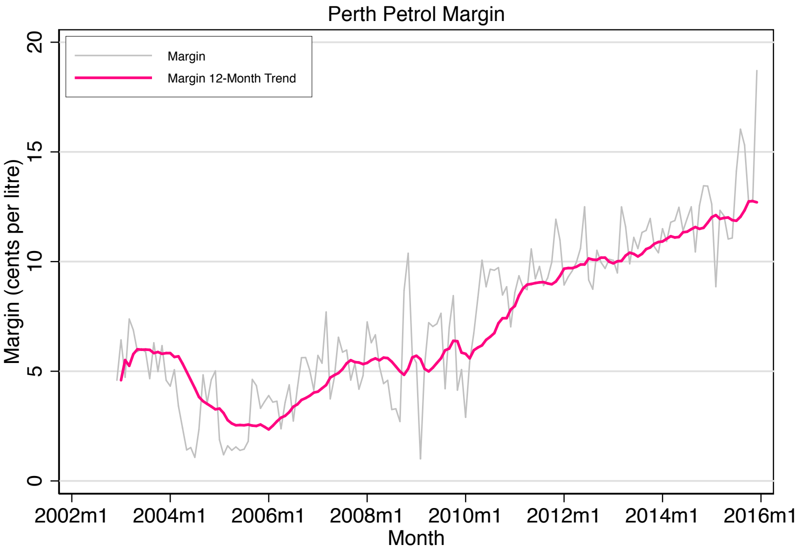
Perth Petrol Prices and Costs: 2003-2016, Monthly



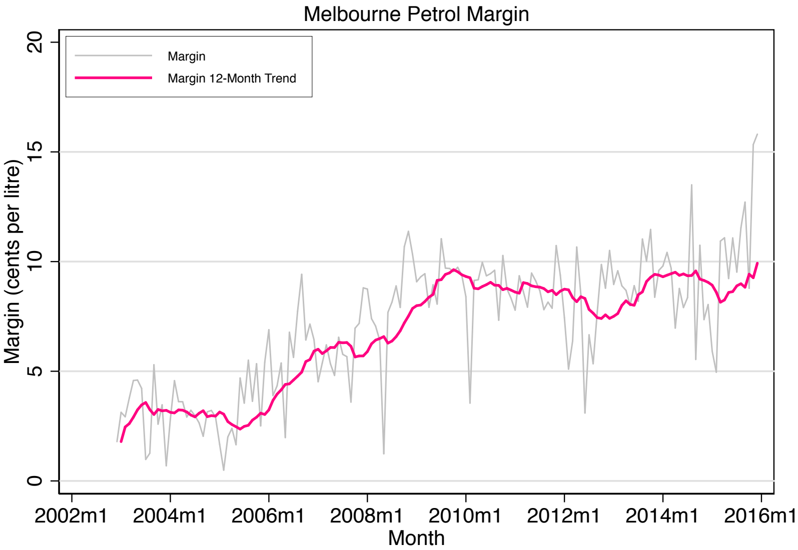
Perth Petrol Margins: 2003-2016, Monthly



Perth Petrol Margins: 2003-2016, Monthly



Melbourne Petrol Margins: 2003-2016, Monthly



Tacit Collusion

Herald Sun

How to save petrol money at the pump

VIC News

Study finds petrol stations colluding with each other to rip motorists off

Kara Irving, Herald Sun
February 14, 2017 8:00pm
Subscriber only

DRIVERS are paying more at the bowser each year because petrol stations have been colluding with each other to rip us off, a new study has found.

The sneaky profit-boosting strategy has seen petrol stations boost their prices by 15 to 20 cents per litre on Thursdays, before decreasing by two cents each day until the next hike

Research has found a motorist driving a 70L Toyota Camry could be paying up to $135 per year extra if they do not fill up on the cheapest day of the week, as seen between 2010 to 2015.

Tacit Collusion

The first of its kind study, which analysed 1.7 million petrol prices over 15 years, found all retailers including independent stations were involved in the price increases.

University of Melbourne's Dr David Byrne said the copycat tactic was never spoken about between retailers.

"It was around 2010 when we noticed petrol retailers were co-ordinating on price changes in an unspoken way," he said.

"They were achieving price co-ordination, and making higher profit margins, without speaking (to each other) at all."

It's been proven — we're being ripped off more each year at the petrol bowser. Picture: iStock

Algorithmic Collusion and Artificial Intelligence

The Economist

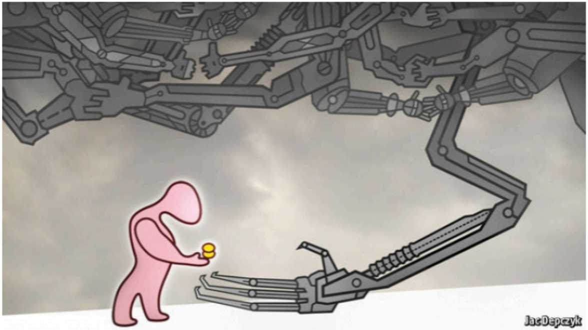
Topics Current edition More

Subscribe

Free exchange

Price-bots can collude against consumers

Trustbusters might have to fight algorithms with algorithms



Print edition | Finance and economics ›

May 6th 2017

Time Series Data

- All these datasets are example of time series data
- Repeatedly sample the <u>same</u> unit of observation over time (Australia, Bitcoin, Perth's Petrol Market)
- Our techniques for estimating and testing multiple linear regression models are the basis for analysing time series data
- Use and interpretation of these models are different than those we developed for cross-sectional data analysis in the subject thus far, with an emphasis on prediction and forecasting
- The next two lectures introduce time series analysis. ECOM30002: Econometrics 2 covers time series in depth.

Lags and First Differences

- Time series dataset: variable Y_t is observed for T periods: $Y_1, Y_2, \ldots, Y_T$ or just Y_t where t is the length of time between observations
  - examples of Y_t: GDP, GDP growth, exchange rate, bitcoin price, petrol margins
  - examples of t: year, quarter, month, day, hour, minute, second
- First lag of Y_t is Y_{t-1}; j^{th} lag of Y is Y_{t-j}

Lags and First Differences

t	gdp	lag1_gdp	lag2_gdp	lag3_gdp	diff_gdp
2003	607.48	489.60	442.72	487.94	117.88
2004	708.71	607.48	489.60	442.72	101.23
2005	755.45	708.71	607.48	489.60	46.74
2006	774.66	755.45	708.71	607.48	19.21
2007	894.78	774.66	755.45	708.71	120.11
2008	938.15	894.78	774.66	755.45	43.37
2009	894.82	938.15	894.78	774.66	-43.33
2010	1051.38	894.82	938.15	894.78	156.57
2011	1217.59	1051.38	894.82	938.15	166.21
2012	1260.11	1217.59	1051.38	894.82	42.52
2013	1193.98	1260.11	1217.59	1051.38	-66.13
2014	1147.24	1193.98	1260.11	1217.59	-46.74
2015	975.68	1147.24	1193.98	1260.11	-171.56
2016	998.53	975.68	1147.24	1193.98	22.85

First Lag

Lags and First Differences

t	gdp	lag1_gdp	lag2_gdp	lag3_gdp	diff_gdp
2003	607.48	489.60	442.72	487.94	117.88
2004	708.71	607.48	489.60	442.72	101.23
2005	755.45	708.71	607.48	489.60	46.74
2006	774.66	755.45	708.71	607.48	19.21
2007	894.78	774.66	755.45	708.71	120.11
2008	938.15	894.78	774.66	755.45	43.37
2009	894.82	938.15	894.78	774.66	-43.33
2010	1051.38	894.82	938.15	894.78	156.57
2011	1217.59	1051.38	894.82	938.15	166.21
2012	1260.11	1217.59	1051.38	894.82	42.52
2013	1193.98	1260.11	1217.59	1051.38	-66.13
2014	1147.24	1193.98	1260.11	1217.59	-46.74
2015	975.68	1147.24	1193.98	1260.11	-171.56
2016	998.53	975.68	1147.24	1193.98	22.85

Second Lag

Lags and First Differences

t	gdp	lag1_gdp	lag2_gdp	lag3_gdp	diff_gdp
2003	607.48	489.60	442.72	487.94	117.88
2004	708.71	607.48	489.60	442.72	101.23
2005	755.45	708.71	607.48	489.60	46.74
2006	774.66	755.45	708.71	607.48	19.21
2007	894.78	774.66	755.45	708.71	120.11
2008	938.15	894.78	774.66	755.45	43.37
2009	894.82	938.15	894.78	774.66	-43.33
2010	1051.38	894.82	938.15	894.78	156.57
2011	1217.59	1051.38	894.82	938.15	166.21
2012	1260.11	1217.59	1051.38	894.82	42.52
2013	1193.98	1260.11	1217.59	1051.38	-66.13
2014	1147.24	1193.98	1260.11	1217.59	-46.74
2015	975.68	1147.24	1193.98	1260.11	-171.56
2016	998.53	975.68	1147.24	1193.98	22.85

Third Lag

Lags and First Differences

t	gdp	lag1_gdp	lag2_gdp	lag3_gdp	diff_gdp
2003	607.48	489.60	442.72	487.94	117.88
2004	708.71	607.48	489.60	442.72	101.23
2005	755.45	708.71	607.48	489.60	46.74
2006	774.66	755.45	708.71	607.48	19.21
2007	894.78	774.66	755.45	708.71	120.11
2008	938.15	894.78	774.66	755.45	43.37
2009	894.82	938.15	894.78	774.66	-43.33
2010	1051.38	894.82	938.15	894.78	156.57
2011	1217.59	1051.38	894.82	938.15	166.21
2012	1260.11	1217.59	1051.38	894.82	42.52
2013	1193.98	1260.11	1217.59	1051.38	-66.13
2014	1147.24	1193.98	1260.11	1217.59	-46.74
2015	975.68	1147.24	1193.98	1260.11	-171.56
2016	998.53	975.68	1147.24	1193.98	22.85

First Difference

Growth Rates

- Logarithms play an important role with time series analysis because they allow for direct calculation of growth rates:

$$\begin{aligned}\ln(Y_t) - \ln(Y_{t-1}) &\approx \frac{Y_t - Y_{t-1}}{Y_{t-1}} \\ &= \frac{\Delta Y_t}{Y_{t-1}}\end{aligned}$$

- This implies that $100 \times (\ln(Y_t) - \ln(Y_{t-1})) = 100 \times \frac{\Delta Y_t}{Y_{t-1}}$ is the growth rate of Y between periods t and $t-1$
- So we often transform a time series using a logarithm if we are looking to examine growth rates period-by-period

Autocorrelation

- How persistent is a time series? That is, how much correlation is there between Y_t and Y_{t-1}
- The correlation of a time series with its own lagged values is called autocorrelation or serial correlation
  - First autocorrelation is the correlation between Y_t and Y_{t-1}
  - Second autocorrelation is the correlation between Y_t and Y_{t-2}
  - j^{th} autocorrelation is the correlation between Y_t and Y_{t-j}
- In general, the j^{th} autocorrelation is computed as:

$$j^{th} \text{ autocorrelation} = \text{corr}(Y_t, Y_{t-j}) = \frac{\text{cov}(Y_t, Y_{t-j})}{\sqrt{\text{var}(Y_t)\text{var}(Y_{t-j})}}$$

- The p^{th} autocovariance is the covariance between Y_t and Y_{t-p} is:

$$j^{th} \text{ autocovariance } = \text{cov}(Y_t, Y_{t-j})$$

Autocorrelation

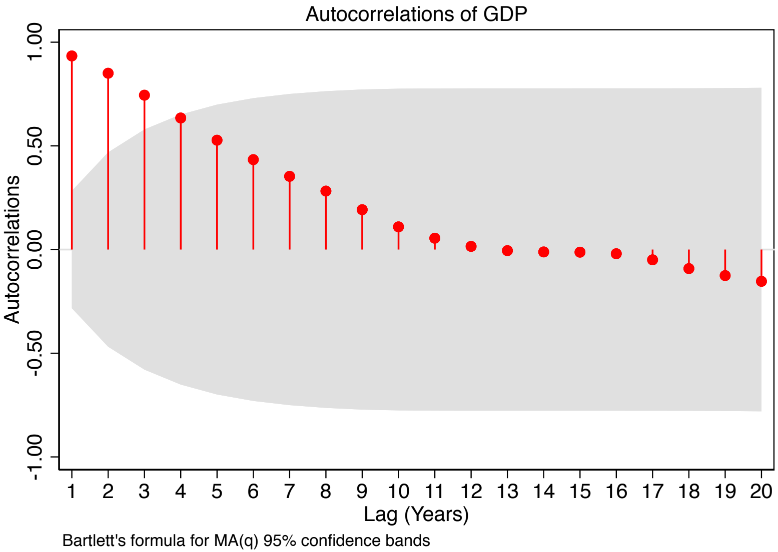
- Using a time series data set $Y_1, \ldots, Y_T$, we can estimate the j^{th} autocovariances ($\widehat{\text{cov}(Y_t, Y_{t-j})}$) and autocorrelations ($\hat{\rho}$) as follows:

$$\widehat{\text{cov}(Y_t, Y_{t-j})} = \frac{1}{T} \sum_{t=j+1}^{T} (Y_t - \bar{Y}_{j+1:T})(Y_{t-j} - \bar{Y}_{1:T-j})$$

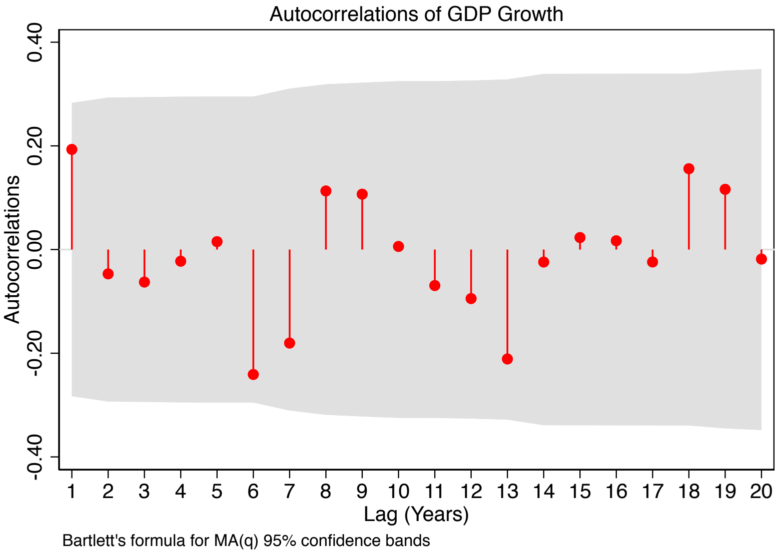
$$\hat{\rho} = \frac{\widehat{\text{cov}(Y_t, Y_{t-j})}}{\widehat{\text{var}(Y_t)}}$$

  where:
  - $\bar{Y}_{j+1:T}$ is the sample average of Y_t using observations from $t = j+1, \ldots, T$
  - $\bar{Y}_{1:T-j}$ is the sample average of Y_t using observations from $t = 1, \ldots, T-j$
  - $\widehat{\text{var}(Y_t)}$ is the sample variance of Y_t
- Autocorrelations are the analogue to time series analyses as scatter plots are to cross-sectional analyses

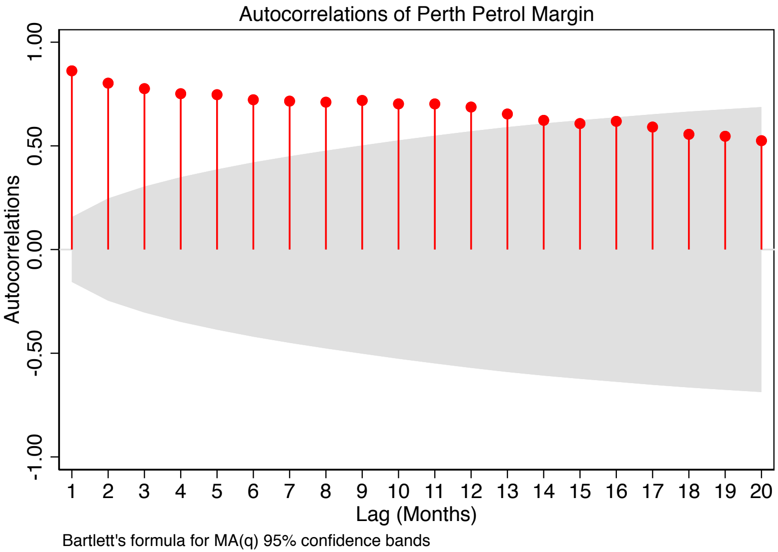
Autocorrelations ($\hat{\rho}$'s) for GDP



Autocorrelations ($\hat{\rho}$'s) for GDP Growth



Autocorrelations for Perths' Petrol Margin



Autoregressions

- Autoregressions are regression models that relates a time series variable to its past (lagged) values
- The first-order autoregression or the AR(1) model (where the (1) indicates 1 lag) for a time series variable Y_t is given by:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$$

  where u_t is an error term
- Examples of AR(1) models estimated by OLS:

$$\widehat{GDP_t} = \underset{(26.463)}{23.300} + \underset{(0.052)}{0.984}\, GDP_{t-1}; \;\; \bar{R}^2 = 0.924$$

$$\widehat{Growth_t} = \underset{(0.694)}{2.746} + \underset{(0.180)}{0.169}\, Growth_{t-1}; \;\; \bar{R}^2 = 0.032$$

$$\widehat{Margin_t} = \underset{(0.450)}{1.866} + \underset{(0.059)}{0.747}\, Margin_{t-1}; \;\; \bar{R}^2 = 0.539$$

  where $Growth_t = 100 \times \Delta \ln(GDP_t)$

Prediction/Forecasting

- A key use for time series models is forecasting
- For example, the Reserve Bank of Australia (RBA) is forecasts GDP growth, inflation, unemployment, housing prices, population, exchange rates, and other macro variables
- These forecasts are critical to informing current monetary policy decisions with interest rates (which the RBA sets), and also informs fiscal policy with government taxation and expenditure

Forecasting: Reserve Bank of Australia

RESERVE BANK OF AUSTRALIA

Search RBA website

Media Releases | Speeches | Publications | Statistics | Chart Pack

About Us | Monetary Policy | Market Operations | Financial Stability | Payments & Infrastructure | Financial Services | Banknotes

Home › Publications › Statement on Monetary Policy › 2018 › Statement on Monetary Policy – February 2018

In Publications

- Minutes
- Statement on Monetary Policy
  - 2018
    - February
    - Boxes
  - 2017
  - 2016
  - 2015
  - 2014
  - 2013
  - 2012
  - 2011
  - 2001–2010
  - 1997–2000
- Financial Stability Review
- Bulletin
- Annual Reports
- Corporate Plan
- Research Discussion Papers
- Conferences

Statement on Monetary Policy

February 2018

The *Statement on Monetary Policy* sets out the Bank's assessment of current economic conditions, both domestic and international, along with the outlook for Australian inflation and output growth. A number of boxes on topics of special interest are also published. The *Statement* is issued four times a year.

Download the complete *Statement* PDF 5.9MB

- Overview — Download PDF 162KB
- International Economic Developments — Download PDF 1.1MB
- International and Foreign Exchange Markets — Download PDF 1.3MB
- Box A: The Period of Low Volatility in Financial Markets — Download PDF 476KB
- Domestic Economic Conditions — Download PDF 1.1MB
- Box B: The Recent Economic Performance of the States — Download PDF 342KB
- Box C: Spillovers from Public Investment — Download PDF 186KB
- Domestic Financial Markets — Download PDF 1.2MB
- Box D: Measures of Investor and Owner-occupier Housing Credit — Download PDF 192KB
- Inflation — Download PDF 686KB
- Economic Outlook — Download PDF 407KB
- Copyright and Disclaimer Notices — Download PDF 833KB

Forecasting

- Suppose you estimated an AR(1) model using a time series $Y_1, \ldots, Y_{T-1}$ and obtained regression estimates $\hat{\beta}_0$ and $\hat{\beta}_1$
- Using the estimated model, the forecast of Y in period $T+1$ (e.g,. the period after the estimation sample ends) based on the AR(1) model is:

$$\hat{Y}_{T+1|T} = \hat{\beta}_0 + \hat{\beta}_1 Y_T$$

- The forecast error is the mistake made by the forecast

$$\text{Forecast error} = Y_{T+1} - \hat{Y}_{T+1|T}$$

where Y_{T+1} is the realized/actual value of Y in period $T+1$

Forecasting Example with GDP Growth

- Growth AR(1) model estimated on periods $t = 1969, \ldots, 2010$

$$\widehat{Growth}_t = \underset{(0.680)}{2.672} + \underset{(0.183)}{0.178}\, Growth_{t-1}; \quad \bar{R}^2 = 0.037$$

- Realized growth in 2010 was $Growth_{2010} = 2.01$
- Forecasted growth in 2011 is:

$$\widehat{Growth}_{2011} = 2.672 + 0.178 \times 2.01 = 3.03$$

- Realized growth in 2011 was $Growth_{2011} = 2.37$
- So the AR(1) forecast error is:

$$\text{Forecast Error} = \text{Actual-Forecast} = 2.37\text{-}3.03 = \text{-}0.66$$

In words, the model overpredicts growth in 2011

Forecasts and Predicted Values

- Forecasts are <u>not</u> simply predicted values from the regression
- In-sample forecasts: predicted values are what the model predicts Y_t is for $t = 1, \ldots, T$ ()
- Out-of-sample forecasts: what the model predicts for Y_t in the future after the sample period $t = T + 1, \ldots$

Root Mean Forecast Error

- We measure the size of a model's forecast error, or the magnitude of a typical mistake made using a forecasting model, using the root mean squared forecast error (RMSFE):

$$\text{RMSFE} = \sqrt{E[(Y_{T+1} - \hat{Y}_{T+1|T})^2]}$$

- There are two sources of error in RMSFE
  1. error because future values of u_t are unknown
  2. error in estimating the coefficients β_0 and β_1
- If the first source of error is much larger than the second (like when sample size T is large), then:

$$\text{RMSFE} \approx \sqrt{var(u_t)} = SER$$

- In practice, we compute RMSFE $\approx \sqrt{var(\hat{u}_t)}$ where $\hat{u}_t$ is the in-sample residuals from the model if T is large

Forecast Intervals

- If we assume u_t is normally distributed, we can compute the standard error of the forecast error of $Y_{T+1|T}$ as:

$$SE(Y_{T+1} - \hat{Y}_{T+1|T}) = \widehat{RMSFE} = \sqrt{var(\hat{u}_t)} = SER$$

- The 95% confidence interval for the forecast $\hat{Y}_{T+1|T}$ can then be computed as:

$$[\hat{Y}_{T+1|T} - 1.96 \times SE(Y_{T+1} - \hat{Y}_{T+1|T}), \hat{Y}_{T+1|T} + 1.96 \times SE(Y_{T+1} - \hat{Y}_{T+1|T})]$$

- Forecast intervals are useful to report because the highlight the range of uncertainty around the model's forecast $\hat{Y}_{T+1|T}$
  - the more narrow a (say 95%) forecast interval, the more confident you are over the range of future events that will potentially occur
  - very wide forecast intervals mean your model is unable to make confident predictions about the future

Forecast Intervals

- Forecast interval with our AR(1) model for $Growth_t$:

$$\widehat{Growth}_t = \underset{(0.680)}{2.672} + \underset{(0.183)}{0.178}\, Growth_{t-1}; \quad \bar{R}^2 = 0.037$$

  - recall the out-sample forecast for $t = 2011$ was 3.03
  - the SER for the AR(1) model is:

$$SER = \sqrt{SSR/n-2} = \sqrt{105.66/(42-2)} = 1.63$$

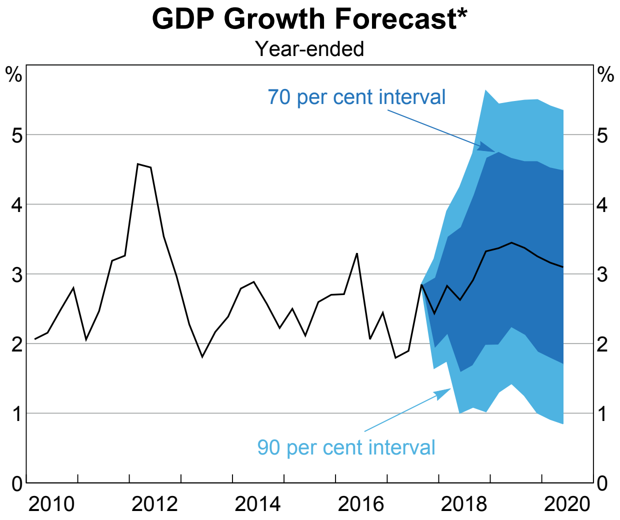
    where there is $n-2$ degrees of freedom as the AR(1) model has $k = 2$ parameters

- 95% forecast interval for $Growth_t$ in 2011 is therefore:

$$[3.03 - 1.96 * 1.63, 3.03 + 1.96 * 1.63] = [-0.165, 6.22]$$

- Interpretation: there is a 95% chance that the growth rate will be between -0.17% and 6.2% in 2011

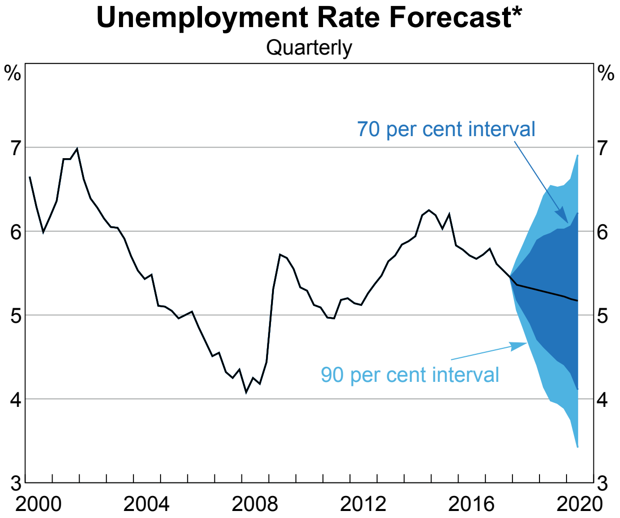
Reserve Bank of Australia Feb 2018 Economic Outlook: Growth Forecast



*   Confidence intervals reflect RBA forecast errors since 1993

Sources: ABS; RBA

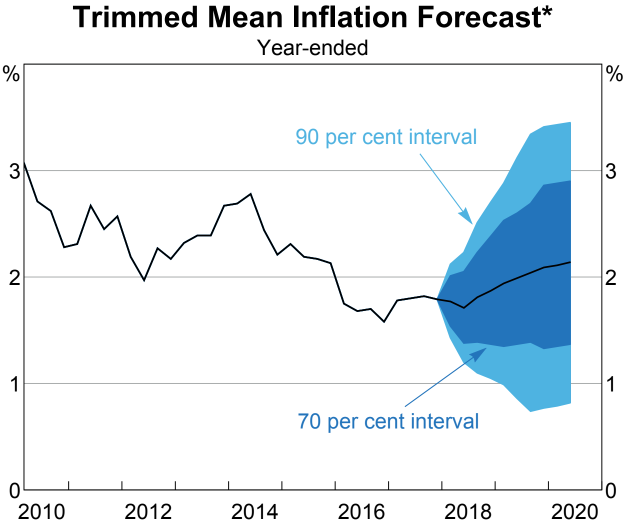
Reserve Bank of Australia Feb 2018 Economic Outlook: Unemployment Forecast



* Confidence intervals reflect RBA forecast errors since 1993

Sources: ABS; RBA

Reserve Bank of Australia Feb 2018 Economic Outlook: Inflation Forecast



* Confidence intervals reflect RBA forecast errors since 1993

Sources: ABS; RBA

p^{th}-Order Autogressive Model

- The AR(1) model uses Y_{t-1} to forecast Y_t but in doing so potentially ignores useful information for earlier periods (e.g., $t-2, t-3, \ldots$) for forecasting
- The p^{th}-order autoregressive model or AR(p) model models Y_t as a linear function of p lags:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2}, \ldots, \beta_p Y_{t-p} + u_t$$

- Idea: adding more lags to the model we can improve:
  - in-sample model fit: $\uparrow R^2$
  - out-of-sample model forecast accuracy: $\uparrow$ RMSFE

AR(2) model example

- Estimated AR(1) model for growth estimated on periods $t = 1969, \ldots, 2010$

$$\widehat{Growth}_t = \underset{(0.680)}{2.672} + \underset{(0.183)}{0.178}\, Growth_{t-1}; \quad \bar{R}^2 = 0.037$$

- Estimated AR(2) model for growth estimated on periods $t = 1969, \ldots, 2010$

$$\widehat{Growth}_t = \underset{(0.623)}{3.194} + \underset{(0.159)}{0.085}\, Growth_{t-1} - \underset{(0.107)}{0.080}\, Growth_{t-2}; \quad \bar{R}^2 = 0.013$$

- Realized growth rates for $Growth_t$
  - in-sample in $t = 2009$ and $t = 2010$ is 1.81 and 2.00
  - out-of-sample in $t = 2011$ realized $Growth_t$ is 2.37
- Forecast errors for the model in 2011 are thus:
  - AR(1): $2.37 - (2.672 + 0.178 \times 2.00) = -0.66$
  - AR(2): $2.37 - (3.194 + 0.085 \times 2.00 - 0.08 \times 1.81) = -0.85$

Autoregressive Distributed Lag Model

- Economic theory often suggests another variable X_t could help forecast our variable of interest Y_t
- Examples:
  - interest rates $\downarrow \rightarrow$ cause GDP $\uparrow$
  - wholesale costs $\downarrow \rightarrow$ cause retail prices $\downarrow$
  - A country's president negatively tweets about a company $\uparrow \rightarrow$ company stock price $\downarrow$
- We can extend our time series model to allow for another time series variable X_t to have a lagged and persistent effect on the time series of a variable of interest Y_t

Presidential Tweets: 31 March and 3 April 2018

While we are on the subject, it is reported that the U.S. Post Office will lose $1.50 on average for each package it delivers for Amazon. That amounts to Billions of Dollars. The Failing N.Y. Times reports that "the size of the company's lobbying staff has ballooned," and that...

10:45 PM - Mar 31, 2018

♡ 82.1K 46K people are talking about this

...does not include the Fake Washington Post, which is used as a "lobbyist" and should so REGISTER. If the P.O. "increased its parcel rates, Amazon's shipping costs would rise by $2.6 Billion." This Post Office scam must stop. Amazon must pay real costs (and taxes) now!

10:52 PM - Mar 31, 2018

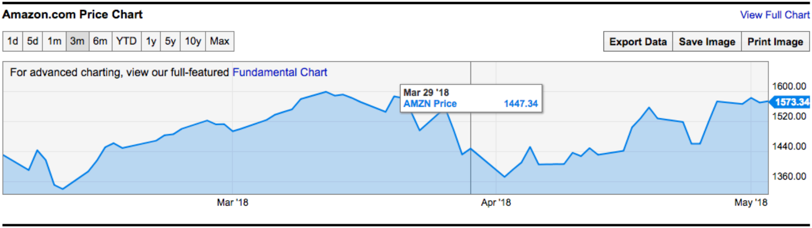
♡ 89.6K 52.9K people are talking about this

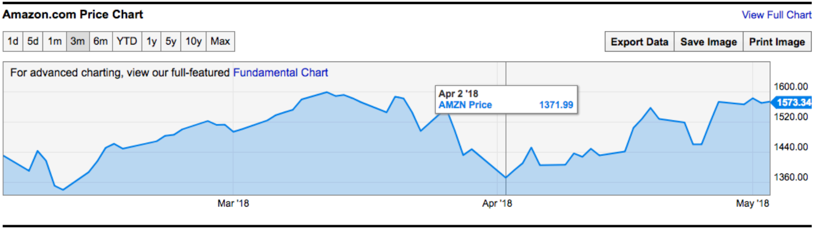
I am right about Amazon costing the United States Post Office massive amounts of money for being their Delivery Boy. Amazon should pay these costs (plus) and not have them bourne by the American Taxpayer. Many billions of dollars. P.O. leaders don't have a clue (or do they?)!

11:55 PM - Apr 3, 2018

♡ 84.1K 71.2K people are talking about this

Amazon Stock Price: 29 March and 2 April 2018





Autoregressive Distributed Lag Model

- We can use the autoregressive distributed lag (ADL) model for modeling the lagged impact of an independent variable X_t on a dependent variable Y_t:

$$\begin{aligned} Y_t = & \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} \\ & + \delta_1 X_{t-1} + \ldots + \delta_q X_{t-q} + u_t \end{aligned}$$

  which includes p lags of Y and q lags of X. We denote this the ADL(p, q) model.
- The AR(p) model is just the ADL(p, q) model with $q = 0$

Autoregressive Distributed Lag Model Example

- P_t = retail petrol price in month t
- C_t = wholesale petrol cost in month t
- ADL(2,0) (or equivalently AR(2)) model estimates:

$$P_t = \underset{(3.040)}{8.501} + \underset{(0.786)}{1.182} P_{t-1} - \underset{(0.776)}{0.249} P_{t-2}; \;\; \bar{R}^2 = 0.911$$

- ADL(2,3) model estimates:

$$\begin{aligned} P_t = & \underset{(3.196)}{4.591} + \underset{(0.214)}{0.329} P_{t-1} + \underset{(0.222)}{0.346} P_{t-2} \\ & + \underset{(0.221)}{1.043} C_{t-1} - \underset{(0.238)}{0.959} C_{t-2} + \underset{(0.086)}{0.222} C_{t-3}; \;\; \bar{R}^2 = 0.924 \end{aligned}$$

Generalised Autoregressive Distributed Lag Model

- The ADL model more generally accommodate k different regressors:

$$
\begin{aligned}
Y_t = & \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} \\
& + \delta_{11} X_{1t-1} + \ldots + \delta_{1q_1} X_{1t-q_1} \\
& + \delta_{21} X_{2t-1} + \ldots + \delta_{2q_2} X_{2t-q_2} \\
& + \cdots + \delta_{k1} X_{kt-1} + \ldots + \delta_{kq_k} X_{kt-q_k} + u_t
\end{aligned}
$$

which is denoted an ADL$(p, q_1, q_2, \ldots, q_k)$ model.

Generalised Autoregressive Distributed Lag Model Example

- P_t = retail petrol price in month t
- C_t = wholesale petrol cost in month t
- E_t = AUD/USD exchange rate in month t
- ADL(2,3,1) model estimates:

$$
\begin{aligned}
P_t =& \underset{(3.253)}{1.170} + \underset{(0.207)}{0.357} P_{t-1} + \underset{(0.215)}{0.303} P_{t-2} \\
&+ \underset{(0.216)}{0.932} C_{t-1} - \underset{(0.231)}{0.893} C_{t-2} + \underset{(0.083)}{0.206} C_{t-3} \\
&+ \underset{(4.467)}{15.031} E_{t-1}; \;\; \bar{R}^2 = 0.929
\end{aligned}
$$

Key Assumptions of the ADL model

- The ADL$(p, q_1, q_2, \ldots, q_k)$ relies on 4 key assumptions

1. Zero conditional mean:

$$E(u_t | Y_{t-1}, \ldots, Y_{t-p}, X_{1t-1}, \ldots, X_{1t-q_1}, \ldots, X_{kt-1}, \ldots, X_{kt-q_k}) = 0$$

2. Large outliers unlikely: $Y_t, X_{1t}, \ldots, X_{kt}$ have nonzero finite fourth moments (kurtosis)
3. No perfect multicollinearity

Key Assumptions of the ADL model

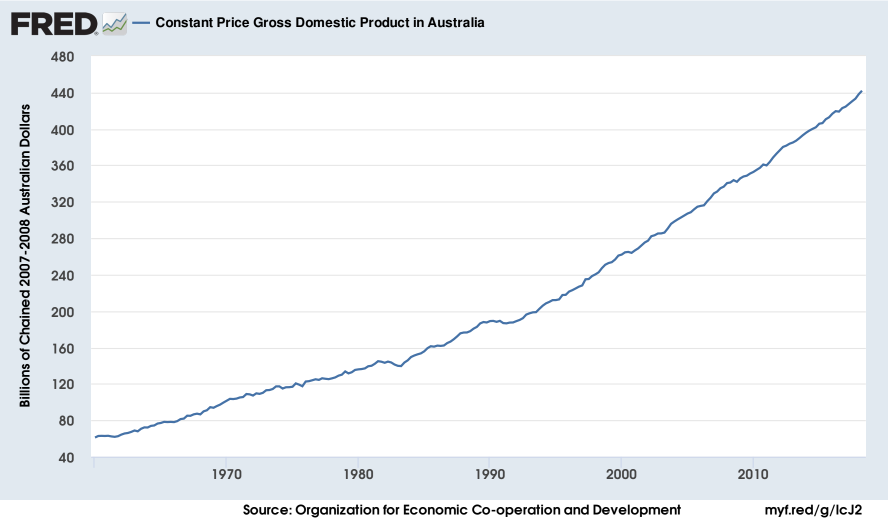
4a. (new) Stationarity: the future tends to be like the past.

- Formally: Y_t is stationary if its probability distribution does not change over time, that is, if the joint distribution of $Y_{s+1}, Y_{s+2}, \ldots, Y_{s+T}$ does not depend on s regardless of the value of T; otherwise Y_t is said to be non-stationary
- Non-stationary processes tend to be characterised by time series that constantly trend over time (either up or down), without reverting to some mean level
- Stationarity processes, in constrast, do not exhibit trends and instead tend to fluctuate around some mean

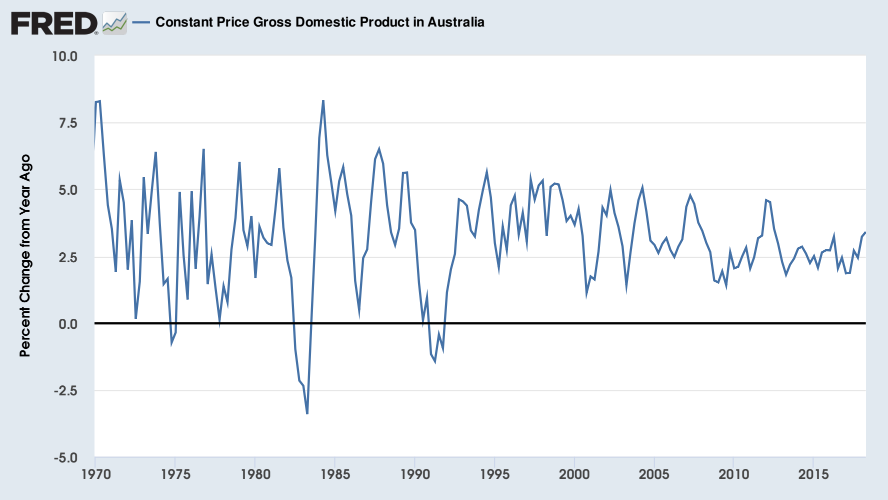
4b. (new) Weak Dependence: $(Y_t, X_{1t}, \ldots, X_{kt})$ and $(Y_{t-j}, X_{1t-j}, \ldots, X_{kt-j})$ become independent as j gets large

- needed so that the LLN and CLT can be applied for hypothesis testing

Australian GDP: Non-Stationary Process



Australian GDP Growth: Stationary Process



Source: Organization for Economic Co-operation and Development myf.red/g/lcJb

Stationary and Non-Stationary Time Series

- Time series models, and the techniques used for estimating and testing them, are different for stationary and non-stationary time series processes
- Future subjects, including ECOM30002: Econometrics 2 and other macroeconometrics subjects develop this rich toolkit for time series analyses for stationary and non-stationary data
- I will discuss future subjects in the last lecture of the semester
- For the rest of this lecture, we <u>assume stationary time series</u> for all Y and X variables in an $\overline{\text{ADL}}(p, q_1, q_2, \ldots, q_k)$ model

Testing Predictive Content: Granger Causality

- Which X_t's are useful, statistically, for predicting Y_t?
- We use Granger Causality Tests, which are based on the F-statistic, to address this question
- Formally, we test the predictive content of X_j for some $j \in [1, k]$ where k is the total number of regressors
  - Recall our general ADL model:

$$\begin{aligned} Y_t = &\beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} \\ &+ \delta_{11} X_{1t-1} + \ldots + \delta_{1q_1} X_{1t-q_1} \\ &+ \delta_{21} X_{2t-1} + \ldots + \delta_{2q_2} X_{2t-q_2} \\ &+ \cdots + \delta_{k1} X_{kt-1} + \ldots + \delta_{kq_k} X_{kt-q_k} + u_t \end{aligned}$$

  - $H_0 : \delta_{j1} = \delta_{j2} = \ldots = \delta_{jq_j} = 0$
    vs.
    H_1 at least one of $\delta_{j1}, \ldots, \delta_{jq_j}$ is not equal to 0
- The Granger Causality Statistic is the F-statistic associated with the joint test described by H_0 vs. H_1 with q_j and $(T - \max\{p, q_1, \ldots, q_k\} - p - \sum_{\ell=1}^{k} q_\ell - 1)$ degs. of freedom

Granger Causality Example

- Recall our ADL(2,3,1) model estimates from before from a sample of $T = 157$ observations, which we used to prediced petrol prices in month t:

$$
\begin{aligned}
P_t = & \underset{(3.253)}{1.170} + \underset{(0.207)}{0.357} P_{t-1} + \underset{(0.215)}{0.303} P_{t-2} \\
& + \underset{(0.216)}{0.932} \color{blue}{C_{t-1}} \color{black} - \underset{(0.231)}{0.893} \color{blue}{C_{t-2}} \color{black} + \underset{(0.083)}{0.206} \color{blue}{C_{t-3}} \\
& \color{black} + \underset{(4.467)}{15.031} \color{red}{E_{t-1}}\color{black}; \; \bar{R}^2 = 0.929
\end{aligned}
$$

- Granger Causality Tests:
  - C_t: test of the null that the coefs. on C_{t-1}, C_{t-2}, and C_{t-3} equals 0 yields an $F_{3,147} = 8.80$ with p-value$= 0.000001$
  - E_t: test of the null that the coef. on E_{t-1} equals 0 yields an $F_{1,147} = 11.32$ with p-value$= 0.001$
- Both C_t and E_t "Granger Cause" P_t as per the ADL model.
  - That is, both of these variables contain predictive content for P_t <u>conditional</u> on all the other variables in the ADL model

Model Selection

- How many lags to include in an ADL($p, q_1, q_2, \ldots, q_k$) model?
- In answering this question, there is a similar trade-off to consider that the adjusted R-Squared deals with:
  - having more lags will improve within sample model-fit
  - however, larger models are less precisely estimated and potentially create imperfect multicollinearity
- Consider first an AR(p) model. One method for choosing p is to find the value of p that minimises the Bayes-Schwarz Information Criterion or BIC, which is:

$$\text{BIC}(p) = \ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{\ln(T)}{T}$$

  where $SSR(p)$ is the sum of squared residuals of the estimated AR(p) model and T is sample size.
- The BIC estimator of p, $\hat{p}$, is that which minimises BIC(p).

Determining Model Lag Length

- BIC formula:

$$\text{BIC}(p) = \ln\left[\frac{SSR(p)}{T}\right] + (p+1)\frac{\ln(T)}{T}$$

- Minimising BIC(p) to determine p trade-offs the influence of p on the two components of BIC(p).
- Specifically, as $\uparrow p$ (lag length goes up):
  - $\downarrow \ln\left[\frac{SSR(p)}{T}\right]$, which occurs because SSR falls when a regression includes more regressors (reward for model fit)
  - $\uparrow (p+1)\frac{\ln(T)}{T}$ (penalty for model size/complexity)

Determining Model Lag Length

- For the more general ADL$(p, q_1, q_2, \ldots, q_k)$ model, the BIC is computed as:

$$\text{BIC}(K) = \ln\left[\frac{SSR(K)}{T}\right] + K\frac{\ln(T)}{T}$$

  where K is the number of parameters in the model, which includes all of the regression coefficients and the intercept
- You can choose the ADL$(p, q_1, q_2, \ldots, q_k)$ specification based on the one that minimises BIC(K) among all model combinations, meaning you potentially check many models
- An alternative popular information criterion for model selection, the Akaike information criterion or AIC(K) is similar to the BIC, except it does not penalise model size as severely:

$$\text{AIC}(K) = \ln\left[\frac{SSR(K)}{T}\right] + K\frac{2}{T}$$

Determining Model Lag Length Application

- Consider the AR(p) model for retail petrol prices in Perth P_t:

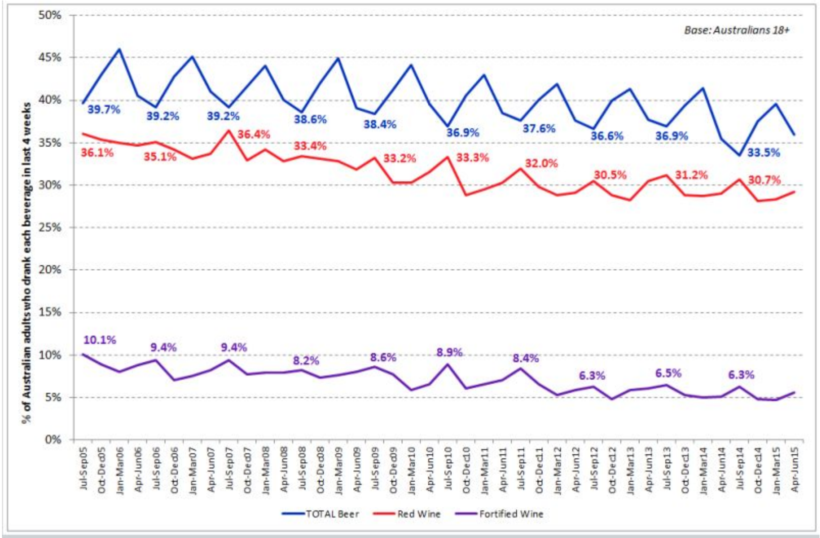
$$P_t + \beta_0 + \beta_1 P_{t-1} + \ldots + \beta_p P_{t-p} + u_t$$

- BIC values for $p = 1, \ldots, 6$
  - $p = 1$: BIC=7.894
  - $p = 2$: BIC=7.829
  - $p = 3$: BIC=7.784
  - $p = 4$: BIC=7.781
  - $p = 5$: BIC=7.758
  - $p = 6$: **BIC=7.738**
- According to the BIC, AR(6) is the preferred model

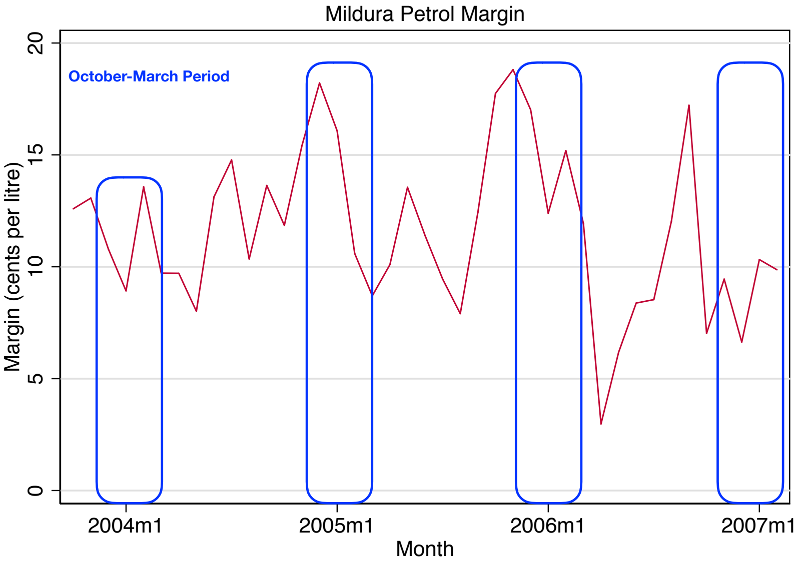
Seasonality

- Dummy variables also play a central role in time series analysis
- Their most common application is to control for seasonality in Y and X in estimating ADL models that relate them
- Seasonality exists when a time series regularly exhibits the same fluctuations at the same points in the calendar year
  - petrol prices always go up in the summer when demand is higher
  - tourism revenues always go up starting around Oct/Nov in Australia as tourists arrive, and start going down in Feb/Mar as tourists leave
  - employment always exhibits strong seasonality year-to-year with regular fluctions in job openings and closings due to temporary demand such as tourism or the holiday season

Seasonality in Australian Alcohol Demand



Seasonality in Midura's Petrol Margins



Seasonality

- We can use dummy variables to account for seasonality in our model for a time series model Y_t
- Consider first the following regression:

$$\begin{aligned} Y_t =& \beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t \\ & + \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sep_t \\ & + \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t + u_t \end{aligned}$$

  where Feb_t is a dummy variable that equals 1 if period t is in February, and similar for the other 10 month dummy variables
- We do not include Jan_t in the regression to avoid the dummy variable trap.
- Given January is a base group, the regression coefficients in the regression are seasonal differences in Y_t relative to January

Seasonality Estimates in Petrol Margins

- Consider a seasonality regression with Y_t is the monthly average city-level petrol, $Margin_t$

$$\begin{aligned} Margin_t = &\beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t \\ &+ \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sep_t \\ &+ \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t \\ &+ \sum_{j=2003}^{2015} \gamma_j dj + u_t \end{aligned}$$

  where dj is a dummy variable for year j
- We estimate this regression for Melbourne, Perth, and Mildura to investigate seasonality in petrol margins in each market

Seasonality Estimates in Petrol Margins

Seasonality in Petrol Margins

	Melbourne	Perth	Mildura
February	-0.635	-0.681	-0.988
	(0.784)	(0.699)	(1.161)
March	0.381	0.612	-1.188
	(0.485)	(0.565)	(1.090)
April	0.382	0.182	-1.010
	(0.585)	(0.544)	(1.246)
May	-0.392	0.293	-0.676
	(0.789)	(0.587)	(1.242)
June	0.423	0.485	-0.200
	(0.609)	(0.600)	(1.421)
July	0.293	0.285	-0.031
	(0.507)	(0.600)	(1.156)
August	1.159	0.321	0.305
	(0.676)	(0.619)	(1.373)
September	0.606	0.990	0.729
	(0.687)	(0.647)	(1.273)
October	1.427^{*}	1.319^{*}	1.166
	(0.588)	(0.573)	(1.156)
November	1.447^{*}	1.217	3.194^{*}
	(0.660)	(0.658)	(1.484)
December	1.575^{*}	1.090	2.105
	(0.724)	(0.695)	(1.487)
Constant	0.218	3.502^{**}	8.787^{**}
	(0.724)	(0.695)	(1.487)
R-Squared	0.671	0.831	0.302
Observations	157	157	157

Margins 1.427 cpl higher in October relative to Jan's margin in Melbourne on average

Notes: Dependent variable is petrol margin in month t. Base group is January, meaning the month dummies are seasonal margin differences month-to-month relative to January's margin. All regressions include year dummy variales as controls for each year between 2012-2015. $^{**}p < 0.01$, $^{*}p < 0.05$

Deseasonalising Data

- We can also use seasonal dummy variables to deseasonalise our data
- Consider again the seasonality regression for the monthly average city-level petrol, $Margin_t$

$$\begin{aligned} Margin_t = & \beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t \\ & + \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sep_t \\ & + \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t + u_t \end{aligned}$$

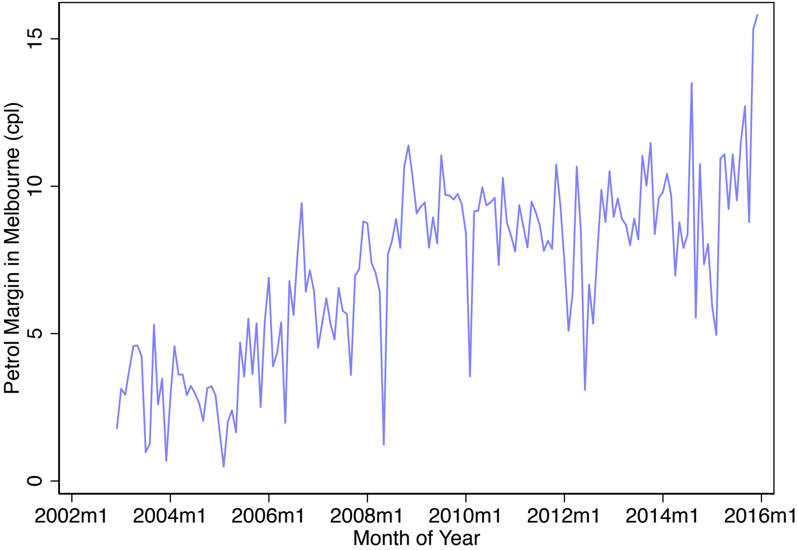
  and note we dropped the year-of-sample dummy variables from before
- After estimating this model, obtained the residuals/in-sample forecast errors $\hat{u}_t$ using the predicted values $\widehat{Margin}_t$:

$$\hat{u}_t = Margin_t - \widehat{Margin}_t$$

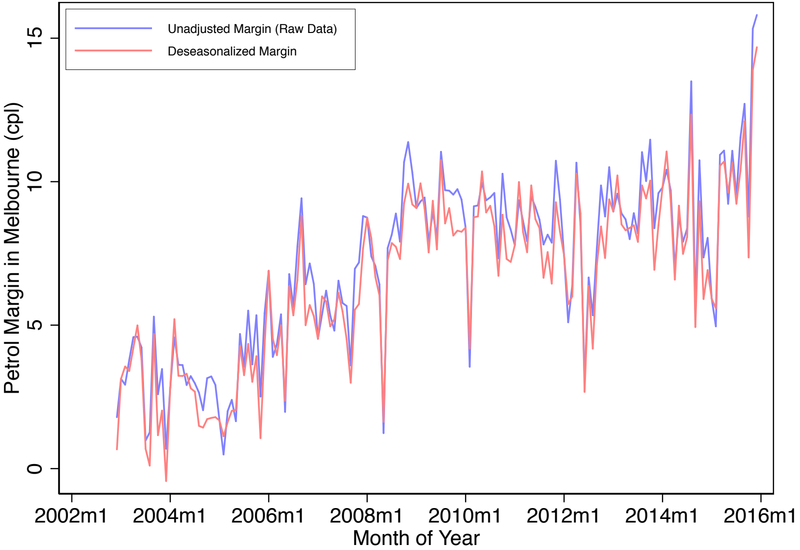
- With this, you can compute the deaseasonalised margin as:

$$\widehat{Margin}_t^{deseas} = \hat{\beta}_0 + \hat{u}_t$$

Deseasonalising Data



Deseasonalising Data



Deseasonalising and Detrending Data

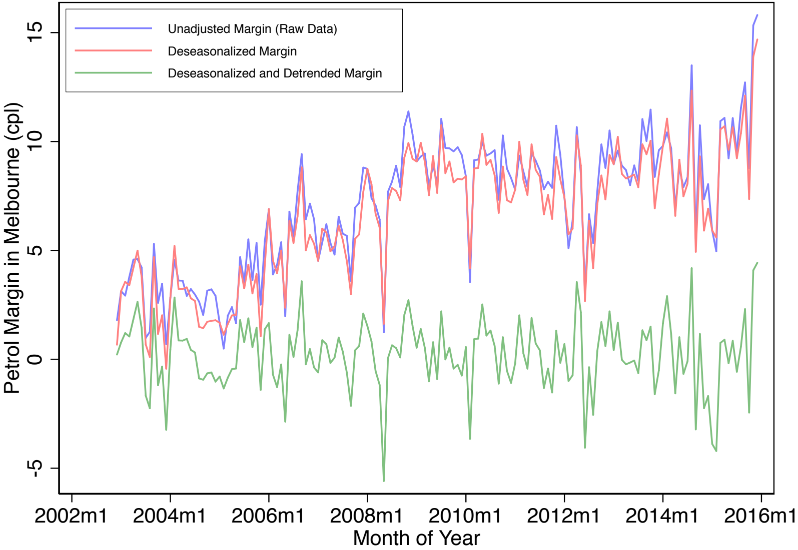
- Now consider the seasonality regression but where we again include the year-of-sample dummy variables:

$$\begin{aligned}
Margin_t = &\beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t \\
&+ \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sep_t \\
&+ \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t \\
&+ \sum_{j=2003}^{2015} \gamma_j dj + u_t
\end{aligned}$$

  where dj is a dummy variable for year j
- Following the same steps as before, after estimating this model, obtained the residuals/in-sample forecast errors $\hat{u}_t$ using the predicted values $\widehat{Margin}_t$
- With this, you can now compute the deaseasonalised <u>and</u> detrended margin as:

$$\widehat{Margin}_t^{deseas,detrend} = \hat{\beta}_0 + \hat{u}_t$$

Deseasonalising and Detrended Data



Seasonality, ADL Model, and Time Frequency

- Finally, returning to our ADL model, we can account for seasonality by simply adding dummies to the model
- Which dummies you can potentially include depend on the frequency of the data
- Quarterly data:
  - quarter-of-year dummies (e.g., summer, fall, winter spring)
- Monthly data:
  - month-of-year dummies (e.g., Jan, Feb, Mar, Apr, ...)
  - alternative: quarter-of-year dummies
- Weekly data:
  - week-of-year dummies (e.g., week 1, week 2, ..., week 52)
  - alternatives: month-of-year or quarter-of-year dummies

Seasonality, ADL Model, and Time Frequency

- Daily data:
  - day-of-week dummies (e.g., Mon, Tue, Wed, ...)
  - alternatives: week-of-year or month-of-year or quarter-of-year dummies
- Hourly data:
  - day-of-week dummies (e.g., 12am, 1am, 2am, ...)
  - alternatives: day-of-week or week-of-year or month-of-year or quarter-of-year dummies
- Familiar trade-off when deciding which set of seasonal dummy variables to include in the model:
  - flexibility: higher frequency dummies (e.g., more dummies) means more flexibility in controlling for flexbility
  - precision: more dummies means larger standard errors and less precise econometric models estimates
- The BIC and AIC are useful for judging this flexibility vs. precision tradeoff in determining which seasonal dummies to include in the model

Application of ADL Model with Seasonality

- ADL(2,3,1) petrol price model estimates without including monthly dummies in the model:

$$
\begin{aligned}
P_t =& \underset{(3.253)}{1.170} + \underset{(0.207)}{0.357} P_{t-1} + \underset{(0.215)}{0.303} P_{t-2} + \underset{(0.216)}{0.932} C_{t-1} - \underset{(0.231)}{0.893} C_{t-2} \\
& + \underset{(0.083)}{0.206} C_{t-2} + \underset{(4.467)}{15.031} E_{t-1}; \;\; \bar{R}^2 = 0.929
\end{aligned}
$$

- ADL(2,3,1) model estimates including month-of-year dummies in the model:

$$
\begin{aligned}
P_t =& \underset{(3.527)}{1.244} + \underset{(0.216)}{0.458} P_{t-1} + \underset{(0.224)}{0.234} P_{t-2} + \underset{(0.228)}{0.785} C_{t-1} - \underset{(0.241)}{0.804} C_{t-2} \\
& + \underset{(0.087)}{0.233} C_{t-3} + \underset{(4.556)}{15.001} E_{t-1} + \underbrace{\sum_{j=2}^{12} \hat{\gamma}_j dj}_{\frac{month}{dummies}}; \;\; \bar{R}^2 = 0.928
\end{aligned}
$$

Summary

- This lecture has introduced another class of econometric models that can be estimated and tested using multiple linear regression: time series models
- We introduced two models:
  - AR(p) autoregression model
  - ADL($p, q_1, \ldots, q_k$) general autoregressive distributed lag model
- Fundamental tools and calculations for time-series analysis using these models include:
  - visualising autocorrelations, which is the "scatter plot" equivalent for time series data
  - forecasting
  - Granger causality for testing predictive content
  - model selection via BIC or AIC
  - using dummies to account for seasonal fluctuations

Summary

- In summary, time series models focus on the inter-temporal relationship between a dependent variable Y_t with itself (e.g., Y_{t-1}), as well as with independent variables X_t across time $t = 1, \ldots, T$
- Contrasts with the multiple linear regression and nonlinear regression models we developed previous which focus on the cross-sectional relationship between a dependent variable Y_i with independent variables X_i across units $i = 1, \ldots, N$
- This is just the tip of the iceberg with time series modelling; there are many variants on the AR(p) and ADL($p, q_1, \ldots, q_k$) models that are explored in future subjects