Eco 4306 Economic and Business Forecasting

Lecture 18

Chapter 10: Forecasting the Long Term: Deterministic and Stochastic Trends

Motivation

- we analyzed the behavior and constructed forecasts for time series that contain a deterministic trend
- we will now look at the stochastic trend, and develop tools to be able to distinguish between the two cases

10.2 Stochastic Trends

- a **stochastic trend** is the result of the accumulation over time of random shocks or innovations

$$X_t = \sum_{j=0}^{t-1} \varepsilon_{t-j} = \varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_1$$

  where ε_{t-j} is a white noise
- a process that contains a stochastic trend is called a **unit root process**
- note that for any j the effect of a shock ε_{t-j} on X_t is permanent, it does not diminish over time even as $t \to \infty$

10.2 Stochastic Trends

- an AR(1) process $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ with $\phi_1 = 1$ has a stochastic trend, this process is called a random walk, and we distinguish between
    1. **pure random walk without drift** $Y_t = Y_{t-1} + \varepsilon_t$, so the case with $c = 0$
    2. **random walk with drift** $Y_t = c + Y_{t-1} + \varepsilon_t$, where $c \neq 0$ is a drift
- for a pure random walk we have by repeated backward substitutions

$$\begin{aligned} Y_t &= Y_{t-1} + \varepsilon_t \\ &= Y_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= \dots \\ &= Y_0 + \varepsilon_1 + \dots + \varepsilon_{t-1} + \varepsilon_t \\ &= Y_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j} \end{aligned}$$

10.2 Stochastic Trends

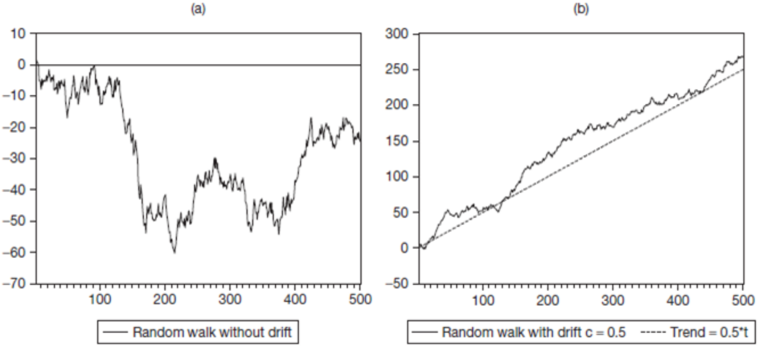
- in case of a random walk with drift, if we substitute backward we obtain

$$
\begin{aligned}
Y_t &= c + Y_{t-1} + \varepsilon_t \\
&= c + (c + Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \dots \\
&= Y_0 + ct + \varepsilon_1 + \ldots + \varepsilon_{t-1} + \varepsilon_t \\
&= Y_0 + ct + \sum_{j=0}^{t-1} \varepsilon_{t-j}
\end{aligned}
$$

- presence of drift term in random walk thus results a deterministic time trend ct

10.2 Stochastic Trends

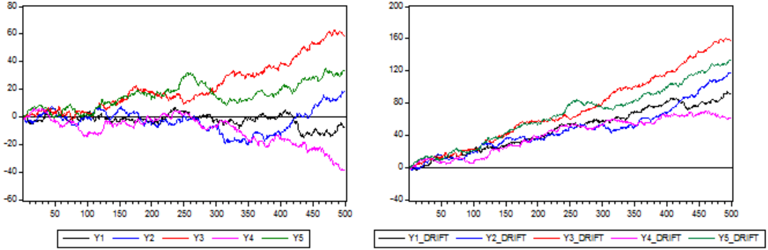
- both pure random walk and the random walk with drift are highly persistent (because effects of shocks are permanent)



10.2 Stochastic Trends

- random walk without drift will either grow, decline or just meander around
- random walk with drift will exhibit a clear upward tendency when $c > 0$ and a downward tendency when $c < 0$

five simulations of random walk without drift vs random walk with drift



10.2.2 Stationarity Properties

for pure random walk without drift $Y_t = Y_{t-1} + \varepsilon_t = Y_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j}$ the first and second moments are

- unconditional mean

$$\mu_t = E(Y_t) = E\left(Y_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j}\right) = Y_0$$

- unconditional variance

$$\gamma_{t,t} = var(Y_t) = var\left(Y_0 + \sum_{j=0}^{t-1} \varepsilon_{t-j}\right) = t\sigma_\varepsilon^2$$

- autocovariance of order k

$$\begin{aligned}\gamma_{t,t-k} = cov(Y_t, Y_{t-k}) &= E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] \\ &= E\left[\left(\sum_{j=0}^{t-1} \varepsilon_{t-j}\right)\left(\sum_{j=0}^{t-k-1} \varepsilon_{t-j}\right)\right] = (t-k)\sigma_\varepsilon^2\end{aligned}$$

- autocorrelation of order k

$$\rho_{t,t-k} = \frac{\gamma_{t,t-k}}{\sqrt{\gamma_{t,t}}\sqrt{\gamma_{t-k,t-k}}} = \frac{(t-k)\sigma_\varepsilon^2}{\sqrt{t\sigma_\varepsilon^2}\sqrt{(t-k)\sigma_\varepsilon^2}} = \sqrt{\frac{t-k}{t}}$$

10.2.2 Stationarity Properties

- unconditional mean is a constant
- unconditional variance and the unconditional autocovariances depend on time - they are increasing functions of time
- pure random walk without drift is thus first but not second order weakly stationary
- since variance is an increasing function of time, potential outcomes of random variable Y_t become more dispersed, increasing the probability of getting very large or very small observations
- autocorrelation are asymptotically 1 regardless of the distance k, sample autocorrelations will be very close to 1, with very slow decay

10.2.2 Stationarity Properties

for the random walk with drift $Y_t = c + Y_{t-1} + \varepsilon_t = Y_0 + ct + \sum_{j=0}^{t-1} \varepsilon_{t-j}$ the first and second moments are

- unconditional mean

$$\mu_t = E(Y_t) = E\left(Y_0 + ct + \sum_{j=0}^{t-1} \varepsilon_{t-j}\right) = Y_0 + ct$$

- unconditional variance

$$\gamma_{t,t} = var(Y_t) = var\left(Y_0 + ct + \sum_{j=0}^{t-1} \varepsilon_{t-j}\right) = t\sigma_\varepsilon^2$$

- autocovariance of order k

$$\begin{aligned}\gamma_{t,t-k} = cov(Y_t, Y_{t-k}) &= E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] \\ &= E\left[\left(\sum_{j=0}^{t-1} \varepsilon_{t-j}\right)\left(\sum_{j=0}^{t-k-1} \varepsilon_{t-j}\right)\right] = (t-k)\sigma_\varepsilon^2\end{aligned}$$

- autocorrelation of order k

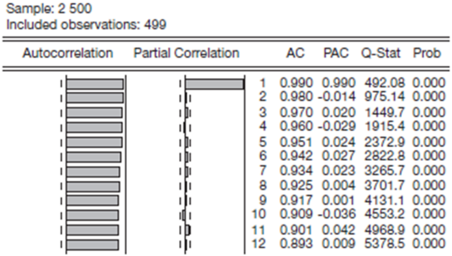
$$\rho_{t,t-k} = \frac{\gamma_{t,t-k}}{\sqrt{\gamma_{t,t}}\sqrt{\gamma_{t-k,t-k}}} = \frac{(t-k)\sigma_\varepsilon^2}{\sqrt{t\sigma_\varepsilon^2}\sqrt{(t-k)\sigma_\varepsilon^2}} = \sqrt{\frac{t-k}{t}}$$

10.2.2 Stationarity Properties

- mean, variance, and autocovariances are all increasing functions of time
- the difference compared to a random walk without drift is that the mean now contains a deterministic linear trend
- random walk with drift is thus neither first nor second order weakly stationary

10.2.2 Stationarity Properties

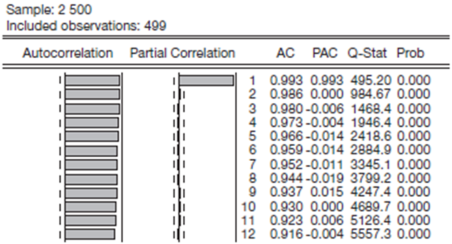
(a) Random Walk Without Drift (Figure 10.8a)



Sample: 2 500
Included observations: 499

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.990	0.990	492.08	0.000
		2	0.980	-0.014	975.14	0.000
		3	0.970	0.020	1449.7	0.000
		4	0.960	-0.029	1915.4	0.000
		5	0.951	0.024	2372.9	0.000
		6	0.942	0.027	2822.8	0.000
		7	0.934	0.023	3265.7	0.000
		8	0.925	0.004	3701.7	0.000
		9	0.917	0.001	4131.1	0.000
		10	0.909	-0.036	4553.2	0.000
		11	0.901	0.042	4968.9	0.000
		12	0.893	0.009	5378.5	0.000

(b) Random Walk with Drift (Figure 10.8b)



Sample: 2 500
Included observations: 499

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.993	0.993	495.20	0.000
		2	0.986	0.000	984.67	0.000
		3	0.980	-0.006	1468.4	0.000
		4	0.973	-0.004	1946.4	0.000
		5	0.966	-0.014	2418.6	0.000
		6	0.959	-0.014	2884.9	0.000
		7	0.952	-0.011	3345.1	0.000
		8	0.944	-0.019	3799.2	0.000
		9	0.937	0.015	4247.4	0.000
		10	0.930	0.000	4689.7	0.000
		11	0.923	0.006	5126.4	0.000
		12	0.916	-0.004	5557.3	0.000

10.2.2 Stationarity Properties

note that

- the first difference of a pure random walk process is

$$\Delta Y_t = Y_t - Y_{t-1} = (Y_{t-1} + \varepsilon_t) - Y_{t-1} = \varepsilon_t$$

- the first difference of a random walk process with a drift is

$$\Delta Y_t = Y_t - Y_{t-1} = (c + Y_{t-1} + \varepsilon_t) - Y_{t-1} = c + \varepsilon_t$$

- so in both case by differencing we have obtained a second order weakly stationary process

10.2.2 Stationarity Properties

- random walks without and with drift are thus examples of a **difference stationary process** - a nonstationary process for which differencing yield a stationary process
- a difference stationary process is said to be **integrated of order** d, denoted as $I(d)$ if it need of to be differenced d times to obtain a stationary process
- random walk is thus $I(1)$ since ΔY_t is stationary
- a stationary process is $I(0)$ since Y_t itself is stationary and no differencing is needed
- for economic and business data, we rarely have difference stationary series that need to be differenced more than twice, most need to be differenced only once or not at all

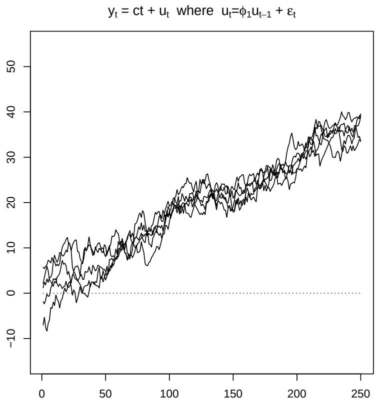
10.2.2 Stationarity Properties

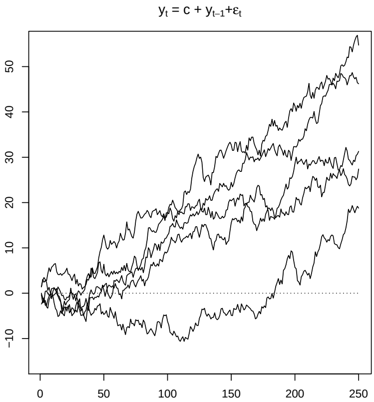
to summarize

- **process with deterministic trend**
  - mean is growing over time, but variance and autocovariances do not depend on time
  - this process is not second order weakly stationary, but can be made second order weakly stationary by detrending
  - often referred to as **trend stationary process**
  - effects of shocks are temporary
- **process with stochastic trend**
  - mean either constant or growing over time, variance and autocovariances depend on time
  - this process is not second order weakly stationary, and can not be made second order weakly stationary by detrending
  - often referred to as **difference stationary process**
  - effects of shocks are permanent

10.2.2 Stationarity Properties

five simulations of trend stationary time series vs random walk with drift

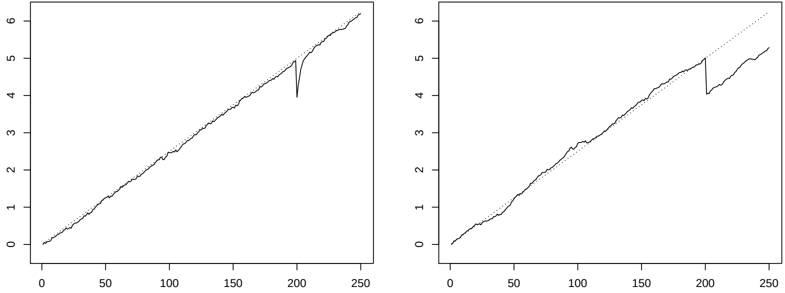




10.2.2 Stationarity Properties

It is important to be able to distinguish between the two cases:

- with trend stationary series shocks have **transitory effects**
- with difference stationary series shocks have **permanent effects**



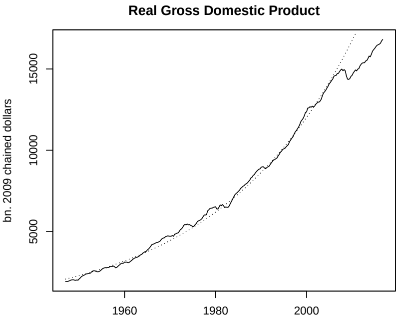
In addition, as we will see later additional issues arise with difference stationary series in the context of multivariate time series analysis

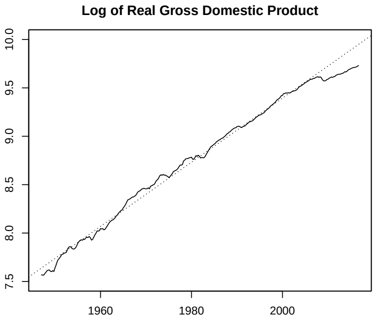
10.2.2 Stationarity Properties

U.S. Real GDP and the effect of 2008-2009 recession

two alternative explanations

- difference stationary process implying permanent effect of a negative shock
- trend stationary process with a structural break (change in intercept and/or trend)





10.2.2.1 Testing for Unit Root

- next goal: develop tools to be able to determine whether a time series has a unit root (contains stochastic trend) or is trend stationary (contains deterministic trend)

- consider first an AR(1) model $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$ which has a unit root if $\phi_1 = 1$
- to test whether a unit root is present is same as to test hypothesis $H_0 : \phi_1 = 1$ against on sided alternative $H_1 : \phi_1 < 1$
- in practice, the model is rewritten as $\Delta Y_t = \beta Y_{t-1} + \varepsilon_t$ where $\beta = \phi_1 - 1$ and the unit root is present if $\beta = 0$
- in the rewritten model to test whether a unit root present is same as to test $H_0 : \beta = 0$ against on sided alternative $H_1 : \beta < 0$
- this test can however not be performed using usual t-test - distribution of the test statistic in this case is not Student
- we need to use the so called **Dickey-Fuller test** instead

10.2.2.1 Testing for Unit Root

- there are three variants of the Dickey-Fuller test for the presence of a unit root

- Case I (no constant or trend)

$$\begin{aligned} H_0 &: Y_t = Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(1) \text{ without drift}) \\ H_1 &: Y_t = \phi_1 Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(0) \text{ with zero mean}) \end{aligned}$$

- Case II (constant only)

$$\begin{aligned} H_0 &: Y_t = Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(1) \text{ without drift}) \\ H_1 &: Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(0) \text{ with nonzero mean}) \end{aligned}$$

- Case III (constant and trend)

$$\begin{aligned} H_0 &: Y_t = c + Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(1) \text{ with drift}) \\ H_1 &: Y_t = c + \alpha t + \phi_1 Y_{t-1} + \varepsilon_t && (Y_t \text{ is } I(0) \text{ with deterministic trend}) \end{aligned}$$

10.2.2.1 Testing for Unit Root

- we choose the version of the test that is compatible with the data under the null and under the alternative hypotheses
- trend properties of the data under the alternative hypothesis will determine the form of the test regression used
- in particular
  - case I would be rarely used, it is appropriate only for time series for which there is some strong reason to believe that the unconditional mean is zero
  - case II is appropriate for non-trending economic and financial series like exchange rates, interest rates and spreads, unemployment rate, inflation rate, ...
  - case III is appropriate for trending time series like asset prices or the levels of macroeconomic aggregates like real GDP, industrial production, employment

10.2.2.1 Testing for Unit Root

- for ease of exposition, we have only considered a unit root in AR(1) process $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$
- but in general any ARMA process may contain a unit root
- in this case we need to modify the Dickey-Fuller test to allow for a more complicated dynamics
- this is done **Augmented Dickey Fuller test**, which tests whether $\beta = 0$ in a model

$$\Delta Y_t = \beta Y_{t-1} + \delta_i \sum_{i=1}^{p} \Delta y_{t-i} + \varepsilon_t$$

  where again a constant c and a trend αt may be added if necessary, just like it is done in the Dickey-Fuller test
- the number of lags included in the regression, p, is usually determined by information criteria (AIC or SIC)

10.2.2.1 Testing for Unit Root

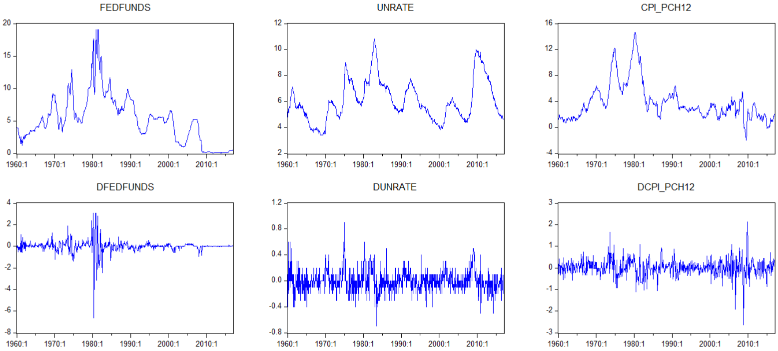
- if we fail to reject the unit root in Y_t, we check for additional unit roots running the Dickey-Fuller test on the successive differences of the series, ΔY_t, $\Delta^2 Y_t$, ... until we reject the unit root in favor of a stationary model
- the general representation of a process with linear dependence: ARIMA(p, d, q)

$$\phi(L)\Delta^d Y_t = \theta(L)\varepsilon_t$$

  where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q$
- when $d = 0$, there isn't a unit root, Y_t is stationary, and we model the time dependence by building an ARMA(p, q) model for Y_t
- when $d = 1$, there is a unit root, ΔY_t is stationary, and we model time dependence by building an ARMA(p, q) model for ΔY_t

Example: Federal Funds Rate, Unemployment Rate, CPI Inflation Rate

- figure below shows the time series plots for Effective Federal Funds Rate `FEDFUNDS`, Unemployment Rate `UNRATE`, CPI inflation (measured as % change from year ago in CPI) `CPI_PCH12`, and their first differences, during the period from 1960M1 to 2016M12



Example: Federal Funds Rate

- to perform Augmented Dickey Fuller test for the presence of unit root in Federal Funds Rate `FEDFUNDS` open the series and choose **View** → **Unit Root Tests**
- since `FEDFUNDS` does not exhibit a clear upward or downward tendency in the option "Include in test equation" we select "Intercept" to perform the Case II variant of the ADF test
- the p-value is 0.1573 so we can not reject the null of a unit root
- in the second step, testing the first difference of `FEDFUNDS` then yields p-value 0.0000, and so we reject the null hypothesis of a unit root in the first difference of `FEDFUNDS`
- we thus conclude that `FEDFUNDS` is integrated of order 1, so $I(1)$

Null Hypothesis: FEDFUNDS has a unit root
Exogenous: Constant
Lag Length: 13 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.348116	0.1573
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(FEDFUNDS) has a unit root
Exogenous: Constant
Lag Length: 12 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-6.504050	0.0000
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

Example: Unemployment Rate

- to perform Augmented Dickey Fuller test for the presence of unit root in Federal Funds Rate UNRATE open the series and choose **View** $\rightarrow$ **Unit Root Tests**
- since UNRATE does not exhibit a clear upward or downward tendency in the option "Include in test equation" we select "Intercept" to perform the Case II variant of the ADF test
- the p-value is 0.0304 so we can reject the null of a unit root at 10% and 5% levels
- we thus conclude that UNRATE is integrated of order 0, so $I(0)$

Null Hypothesis: UNRATE has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.057041	0.0304
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(UNRATE) has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.891697	0.0000
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

Example: CPI Inflation

- to perform Augmented Dickey Fuller test for the presence of unit root in Federal Funds Rate CPI_PCH12 open the series and choose **View** $\rightarrow$ **Unit Root Tests**
- since CPI_PCH12 does not exhibit a clear upward or downward tendency in the option "Include in test equation" we select "Intercept" to perform the Case II variant of the ADF test
- the p-value is 0.0643 so we can not reject the null of a unit root at 1% and 5% levels
- in the second step, testing the first difference of CPI_PCH12 then yields p-value 0.0000, and so we reject the null hypothesis of a unit root in the first difference of CPI_PCH12
- we thus conclude that CPI_PCH12 is integrated of order 1, so $I(1)$

Null Hypothesis: CPI_PCH12 has a unit root
Exogenous: Constant
Lag Length: 15 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.762425	0.0643
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(CPI_PCH12) has a unit root
Exogenous: Constant
Lag Length: 12 (Automatic - based on SIC, maxlag=19)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.389774	0.0000
Test critical values:	1% level	-3.439682	
	5% level	-2.865549	
	10% level	-2.568961	

*MacKinnon (1996) one-sided p-values.

10.2.3 Optimal Forecast

- under quadratic loss function the optimal forecast is a conditional mean $f_{t,h} = \mu_{t+h|t} = E(Y_{t+h}|I_t)$ for $h = 1, 2, \ldots, s$
- properties of a random walk with drift and without drift differ only in the behavior of the mean, their forecasts will also differ in the behavior of the conditional mean

10.2.3 Optimal Forecast

- under pure random walk $Y_t = Y_{t-1} + \varepsilon_t$, for forecasting horizon $h = s$:

1. optimal point forecast

$$f_{t,s} = \mu_{t+s|t} = E(Y_{t+s}|I_t) = E(Y_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}|I_t) = Y_t$$

2. forecast error

$$e_{t,s} = Y_{t+s} - f_{t,s} = Y_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s} - Y_t = \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}$$

3. uncertainty associated with the forecast is summarized by the variance of the forecast error

$$\sigma^2_{t+s|t} = var(e_{t,s}|I_t) = var(\varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}) = s\sigma^2_\varepsilon$$

4. the density forecast is the conditional probability density function $f(Y_{t+s}|I_t)$, assuming ε_{t+s} is normally distributed white noise, we have

$$Y_{t+s}|I_t \sim N(Y_t, s\sigma^2_\varepsilon)$$

10.2.3 Optimal Forecast

- under random walk with drift $Y_t = c + Y_{t-1} + \varepsilon_t$, for forecasting horizon $h = s$:

1. optimal point forecast

$$f_{t,s} = \mu_{t+s|t} = E(Y_{t+s}|I_t) = E(sc + Y_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}|I_t) = sc + Y_t$$

2. forecast error

$$e_{t,s} = Y_{t+s} - f_{t,s} = sc + Y_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s} - sc - Y_t = \varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}$$

3. uncertainty associated with the forecast is summarized by the variance of the forecast error

$$\sigma^2_{t+s|t} = var(e_{t,s}|I_t) = var(\varepsilon_{t+1} + \varepsilon_{t+2} + \ldots + \varepsilon_{t+s}) = s\sigma^2_\varepsilon$$

4. the density forecast is the conditional probability density function $f(Y_{t+s}|I_t)$, assuming ε_{t+s} is normally distributed white noise, we have

$$Y_{t+s}|I_t \sim N(sc + Y_t, s\sigma^2_\varepsilon)$$

10.2.3 Optimal Forecast

to summarize

- for pure random walk $Y_t = Y_{t-1} + \varepsilon_t$ we have

h	$\mu_{t+h\|t}$	$\sigma^2_{t+h\|t}$
1	Y_t	σ^2_ε
2	Y_t	$2\sigma^2_\varepsilon$
$\vdots$		
s	Y_t	$s\sigma^2_\varepsilon$

- for random walk with drift $Y_t = c + Y_{t-1} + \varepsilon_t$ we have

h	$\mu_{t+h\|t}$	$\sigma^2_{t+h\|t}$
1	$c + Y_t$	σ^2_ε
2	$2c + Y_t$	$2\sigma^2_\varepsilon$
$\vdots$		
s	$sc + Y_t$	$s\sigma^2_\varepsilon$

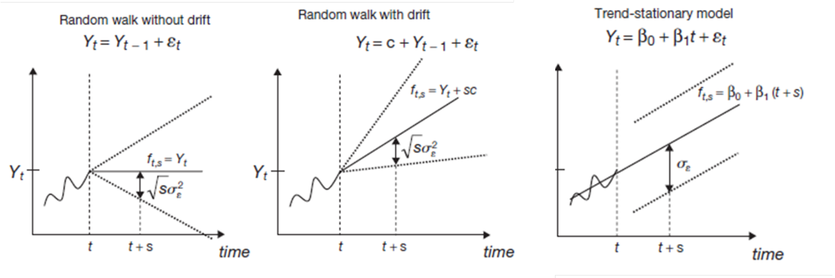
10.2.3 Optimal Forecast

to summarize

- when there is no drift, the point forecast is constant for any forecasting horizon and is equal to the most recent value of the process in the information set; when there is a drift, the point forecast is a line with slope c and intercept Y_t.
- uncertainty of the forecast is the sum of equally weighted future innovations
- variance of the forecast is a linear function of the forecasting horizon with slope σ_ε^2

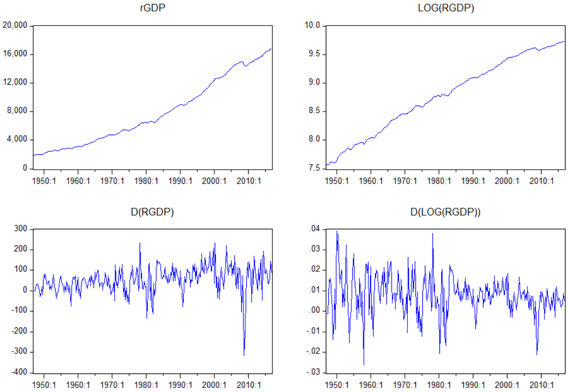
10.2.3 Optimal Forecast

- the uncertainty of the random walk forecasts increases with the forecast horizon while that from a trend-stationary process remains constant for any horizon
- dotted lines represent the uncertainty of the point forecast - the 1 standard deviation interval forecast, $f_{t,s} \pm \sigma_{t+s|t}$



Example: U.S. Real GDP

- figure below shows the time series plots for U.S. Real GDP `rGDP`, log transformed U.S. Real GDP `log rGDP`, and their first differences



Example: U.S. Real GDP

- to perform Augmented Dickey Fuller test for the presence of unit root in log transformed U.S. Real GDP $\log rGDP_t$, generate the log transformed series, then open it and choose **View** $\rightarrow$ **Unit Root Tests**
- since $\log rGDP_t$ exhibits a clear upward tendency in the option "Include in test equation" we select "Trend and intercept" to perform the Case III variant of the ADF test
- the p-value is 0.8805 so we can not reject the null of a unit root
- in the second step, testing the first difference of $\log rGDP_t$, so $\Delta \log rGDP_t$, yields p-value 0.0000, and so we reject the null hypothesis of a unit root in $\Delta \log rGDP_t$
- log of U.S. real GDP, $\log rGDP_t$, is integrated of order 1, so $I(1)$

Null Hypothesis: LRGDP has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.321455	0.8805
Test critical values:	1% level	-3.991412	
	5% level	-3.426073	
	10% level	-3.136231	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(LRGDP) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.51612	0.0000
Test critical values:	1% level	-3.991412	
	5% level	-3.426073	
	10% level	-3.136231	

*MacKinnon (1996) one-sided p-values.

Example: U.S. Real GDP

- the first difference of real GDP is thus stationary, and so we estimate the model

$$\Delta \log rGDP_t = \beta_0 + \varepsilon_t$$

choose **Object** → **New Object** → **Equation**, in the Equation specification box enter `d(log(rGDP)) c` and in Sample box `1950Q1 2009Q4`

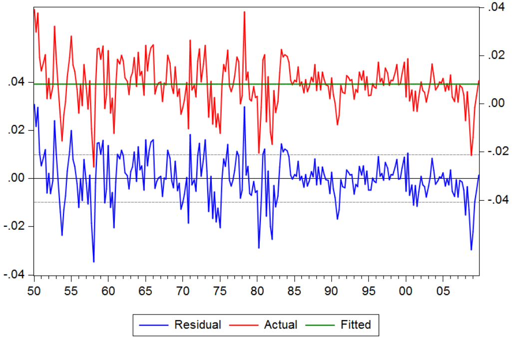
Dependent Variable: D(LOG(RGDP))
Method: Least Squares
Date: 04/02/17 Time: 21:40
Sample: 1950Q1 2009Q4
Included observations: 240

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008257	0.000633	13.03582	0.0000

R-squared	0.000000	Mean dependent var	0.008257
Adjusted R-squared	0.000000	S.D. dependent var	0.009812
S.E. of regression	0.009812	Akaike info criterion	-6.406234
Sum squared resid	0.023010	Schwarz criterion	-6.391731
Log likelihood	769.7480	Hannan-Quinn criter.	-6.400390
Durbin-Watson stat	1.155991		

Example: U.S. Real GDP

residuals shows that the variance is roughly same over time



Example: U.S. Real GDP

- the correlogram for residuals however shows large significant component of PACF at lag 1, and significant components of ACF at lags 1, 2
- the residuals are thus not white noise

Date: 04/02/17 Time: 22:44
Sample: 1950Q1 2009Q4
Included observations: 240

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.401	0.401	39.137	0.000
		2	0.235	0.088	52.558	0.000
		3	0.060	-0.074	53.430	0.000
		4	-0.017	-0.044	53.502	0.000
		5	-0.085	-0.066	55.293	0.000
		6	-0.010	0.070	55.316	0.000
		7	-0.036	-0.035	55.643	0.000
		8	-0.039	-0.036	56.025	0.000
		9	0.056	0.103	56.804	0.000
		10	0.065	0.023	57.862	0.000
		11	0.027	-0.034	58.048	0.000
		12	-0.108	-0.162	61.026	0.000
		13	-0.122	-0.040	64.858	0.000
		14	-0.077	0.061	66.395	0.000
		15	-0.077	-0.045	67.945	0.000
		16	0.048	0.100	68.542	0.000
		17	0.045	-0.009	69.065	0.000
		18	0.042	-0.004	69.518	0.000
		19	0.014	-0.021	69.571	0.000
		20	0.035	-0.000	69.893	0.000
		21	-0.075	-0.078	71.391	0.000
		22	-0.027	0.051	71.586	0.000
		23	-0.053	-0.017	72.343	0.000
		24	-0.001	0.036	72.343	0.000

Example: U.S. Real GDP

- to fix this issue we include the first regular AR lag in the model, so that u_t is now given by an AR(1) specification

$$\Delta \log rGDP_t = \beta_0 + u_t$$
$$u_t = \phi_1 u_{t-1} + \varepsilon_t$$

- to estimate it choose **Object** → **New Object** → **Equation**, in the Equation specification box enter `d(log(rGDP)) c ar(1)` and in Sample box `1950Q1 2009Q4`

Dependent Variable: D(LOG(RGDP))
Method: ARMA Maximum Likelihood (BFGS)
Date: 04/02/17 Time: 21:40
Sample: 1950Q1 2009Q4
Included observations: 240
Convergence achieved after 5 iterations
Coefficient covariance computed using outer product of gradients

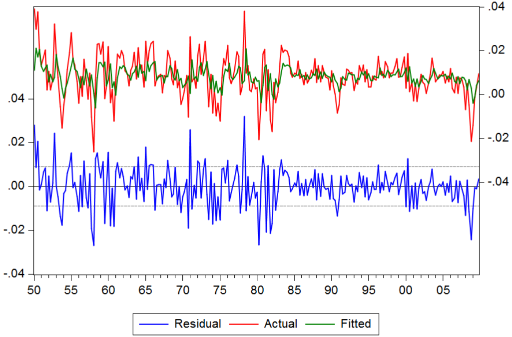
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008352	0.000963	8.671928	0.0000
AR(1)	0.416893	0.051680	8.066826	0.0000
SIGMASQ	7.98E-05	5.78E-06	13.80036	0.0000

R-squared	0.168033	Mean dependent var	0.008257
Adjusted R-squared	0.161012	S.D. dependent var	0.009812
S.E. of regression	0.008988	Akaike info criterion	-6.572734
Sum squared resid	0.019144	Schwarz criterion	-6.529226
Log likelihood	791.7281	Hannan-Quinn criter.	-6.555203
F-statistic	23.93350	Durbin-Watson stat	2.065524
Prob(F-statistic)	0.000000		

Inverted AR Roots	.42

Example: U.S. Real GDP

residuals do not show any systematic pattern



Example: U.S. Real GDP

correlogram also suggests that the residuals are white noise

Date: 04/02/17 Time: 22:44
Sample: 1950Q1 2009Q4
Included observations: 240
Q-statistic probabilities adjusted for 1 ARMA term

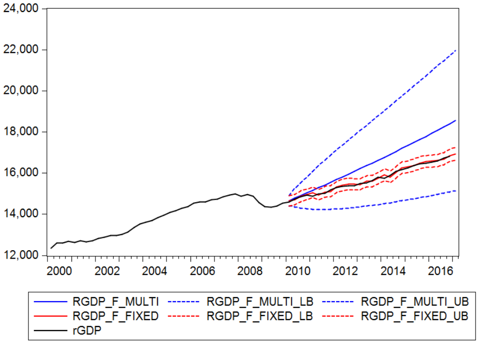
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.054	-0.054	0.6984	
		2	0.097	0.094	2.9845	0.084
		3	-0.025	-0.015	3.1329	0.209
		4	-0.012	-0.023	3.1672	0.367
		5	-0.108	-0.108	6.0694	0.194
		6	0.045	0.039	6.5810	0.254
		7	-0.025	-0.001	6.7360	0.346
		8	-0.066	-0.081	7.8238	0.348
		9	0.067	0.061	8.9382	0.348
		10	0.050	0.062	9.5602	0.387
		11	0.056	0.057	10.366	0.409
		12	-0.105	-0.122	13.177	0.282
		13	-0.079	-0.116	14.773	0.254
		14	-0.007	0.034	14.786	0.321
		15	-0.092	-0.072	16.980	0.257
		16	0.085	0.074	18.840	0.221
		17	0.019	0.023	18.930	0.272
		18	0.028	0.013	19.139	0.321
		19	-0.020	-0.017	19.242	0.377
		20	0.078	0.024	20.834	0.346
		21	-0.110	-0.088	24.037	0.241
		22	0.026	0.023	24.214	0.283
		23	-0.060	-0.024	25.189	0.288
		24	0.012	0.023	25.228	0.339

Example: U.S. Real GDP

- to create h-quarter ahead forecasts for $h = 1, 2, \ldots, 25$, so 2010Q1-2017Q1: choose **Forecast** and set "Series to forecast" to "rGDP"","Method" to "Dynamic forecast" and "Forecast sample" to "2010Q1 2017Q1"
- to create a sequence of 1-quarter ahead forecasts, from 2010Q1-2017Q1: choose **Forecast** and set "Series to forecast" to "rGDP"","Method" to "Static forecast" and "Forecast sample" to "2010Q1 2017Q1"

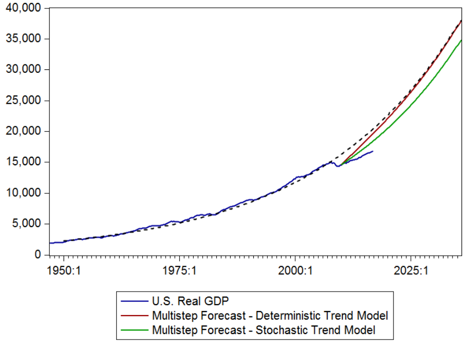
Example: U.S. Real GDP

- sequence of 1-step ahead forecasts is more precise than the multistep forecast - RMSE is 77.3231 for the former and 905.1898 for the latter
- confidence interval is narrower in the case of the 1-step ahead forecasts
- note that in the case of the for the multistep forecast the confidence intervals are getting larger with increasing h, just like in the above stylized diagram for the forecast from a random walk with drift



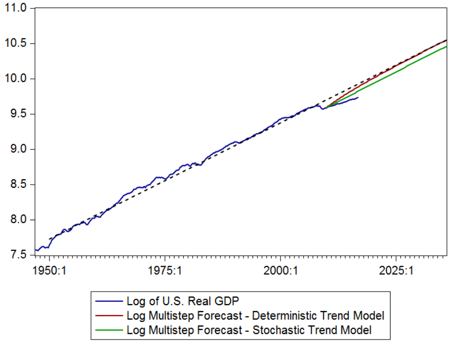
Example: U.S. Real GDP

- comparing the forecast from the deterministic model from HW06.pdf shows that in the deterministic trend model the effects of the negative shock of the 2008-2009 disappear over time, but in the stochastic trend model the effects are permanent



Example: U.S. Real GDP

- comparing the forecast from the deterministic model from HW06.pdf shows that in the deterministic trend model the effects of the negative shock of the 2008-2009 disappear over time, but in the stochastic trend model the effects are permanent



Example: U.S. Real GDP

- comparing the forecast from the deterministic model from HW06.pdf shows that in the deterministic trend model the effects of the negative shock of the 2008-2009 disappear over time, but in the stochastic trend model the effects are permanent
- in case of the deterministic trend model the sequence of 1-step ahead forecasts has RMSE=103.459 and the multistep forecast has RMSE=1649.069
- in case of the stochastic trend model the sequence of 1-step ahead forecasts has RMSE=77.3231 and the multistep forecast has RMSE=905.1898
- for U.S. real GDP, the stochastic trend model thus yields a more precise forecast than the stochastic trend model
- the difference in RMSE appears large, but formally we still should perform the equal predictive power test to compare the two forecasts - see lec13slides.pdf