Time Series Analysis

Discussion Section 03

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

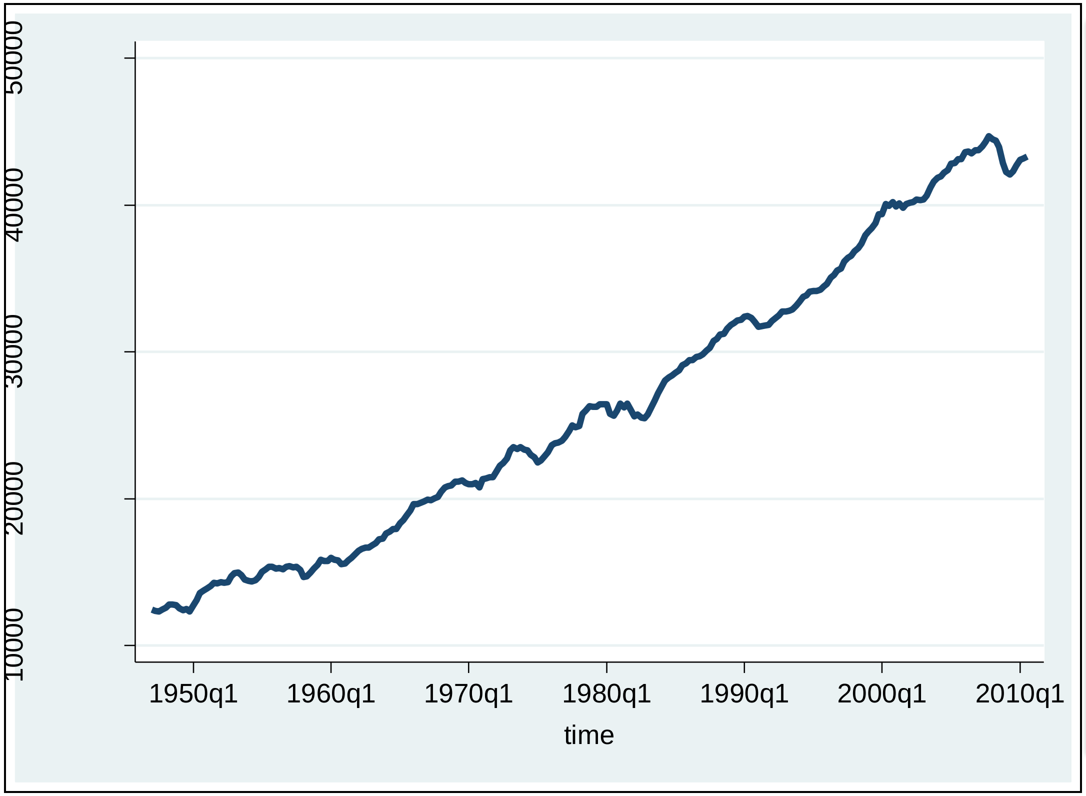
Nonstationary Stochastic Processes

- **Introduction**
- **Nonstationarity and Trends**
- **ARIMA Models**
- Unit Root Tests
- Seasonal ARIMA

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Original Time Series (1947q1 to 2010q3)

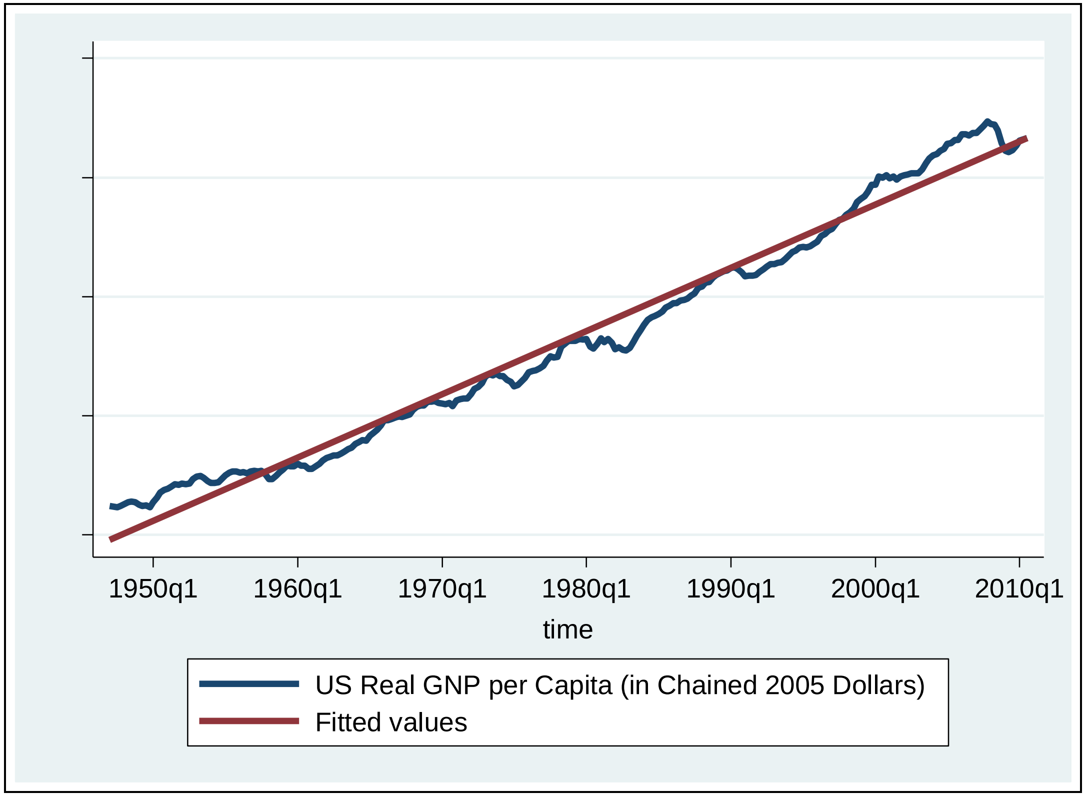
U.S. postwar real GNP per capita (in chained 2005 dollars)



Chained dollars -- A measure used to express real prices. Real prices are those that have been adjusted to remove the effect of changes in the purchasing power of the dollar; they usually reflect buying power relative to a reference year. Prior to 1996, real prices were expressed in constant dollars, a measure based on the weights of goods and services in a single year, usually a recent year. In 1996, the U.S. Department of Commerce introduced the chained-dollar measure. The new measure is based on the average weights of goods and services in successive pairs of years. It is "chained" because the second year in each pair, with its weights, becomes the first year of the next pair. The advantage of using the chained-dollar measure is that it is more closely related to any given period covered and is therefore subject to less distortion over time.

Original Time Series (1947q1 to 2010q3)

U.S. postwar real GNP per capita (in chained 2005 dollars)



Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Excursus: Logarithmic Transformation

Before we even remove a deterministic trend in the trend stationary model (TS-model) or difference in the difference stationary model (DS-model), it is often useful to first lake logs of the original series.

This will linearize an exponential trend,

i.e. constant proportional growth.

$$ln(e^{\delta t}) = \delta t$$

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Excursus: Logarithmic Transformation

Moreover, 1st differences of log-series are approximately growth rates (percentage changes) which can be expected to be stationary even if the original series is not.

$$\Delta ln(y_t) = (1-L)ln(y_t) = ln(y_t) - ln(y_{t-1})$$

$$= ln\left(\frac{y_t}{y_{t-1}}\right) = ln\left(\frac{y_{t-1} + y_t - y_{t-1}}{y_{t-1}}\right)$$

$$= ln\left(1 + \frac{y_t - y_{t-1}}{y_{t-1}}\right) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

Recall:

$$ln(1+x) \approx x$$

$$for\ x - small$$

Exercise 3.1:

- Write down the **general formulas** for trend-stationary (TS) and the difference-stationary (DS) models.
- Describe the **difference** between the trend-stationary (TS) and the difference-stationary (DS) model with respect to the persistence of their dynamic response to a random shock to real GNP per capita.

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.2:

Trend-stationary (TS) Model

$$y_t = \sum_{j=0}^{m} \delta_j \cdot t^j + u_t \text{ with } u_t \sim ARMA(p,q)$$

$$\longrightarrow \quad \hat{y}_t = 9.714 + 0.005 \cdot t + \hat{u}_t$$

$$\longrightarrow \quad \hat{u}_t = 1.296 u_{t-1} - 0.212 u_{t-2} - 0.139 u_{t-3}$$

- Show that the series y_t is not stationary if the estimated TS model is the right model.

  Hint: Consider $E[y_t]$.
- Calculate the average percentage annual growth rate of the log GNP per capita.

Exercise 3.3:

Trend-stationary (TS) Model

$$y_t = \sum_{j=0}^{m} \delta_j \cdot t^j + u_t \text{ with } u_t \sim ARMA(p,q)$$

$\longrightarrow$ $$\hat{y}_t = 9.714 + 0.005 \cdot t + \hat{u}_t$$

$\longrightarrow$ $$\hat{u}_t = 1.296u_{t-1} - 0.212u_{t-2} - 0.139u_{t-3}$$

How does a shock today affect the level of y_t one year hence and infinitely far in the future?

Hint: MA representation of y_t

$$y_t = \mu + \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \ldots = \mu + \psi(L)\varepsilon_t$$

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

General Solution: Use “MA representation”

Any stationary ARMA(p,q) process can be written as an infinite MA:

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad \text{with } \psi_0 = 1$$

$$y_{T+l} = \varepsilon_{T+l} + \psi_1 \varepsilon_{T+l-1} + \ldots + \psi_{l-1}\varepsilon_{T+1} + \psi_l \varepsilon_T + \psi_{l+1}\varepsilon_{T-1} + \ldots$$

$$\tilde{y}_{T+l} = \psi_l \varepsilon_T + \psi_{l+1}\varepsilon_{T-1} + \ldots$$

Forecast Error

$$e_{T+l} = y_{T+l} - \tilde{y}_{T+l|T} = \varepsilon_{T+l} + \psi_1 \varepsilon_{T+l-1} + \ldots + \psi_{l-1}\varepsilon_{T+1}$$

$$E\left[e_{T+l}^2\right] = Var\left[e_{T+l}\right] = \left(1 + \psi_1^2 + \ldots + \psi_{l-1}^2\right)\sigma_\varepsilon^2$$

Prediction Interval

$$\left[\tilde{y}_{T+l|T} \pm z_{1-\frac{\alpha}{2}} \left(1 + \psi_1^2 + \ldots + \psi_{l-1}^2\right)^{\frac{1}{2}} \sigma_\varepsilon\right]$$

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

How do do we find $\psi_1, \ldots, \psi_{l-1}$?

$$y_t = \varphi_1 y_{t-1} + \ldots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

$$\left(1 - \varphi_1 L - \varphi_2 L^2 - \ldots - \varphi_p L^p\right) y_t = \left(1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q\right) \varepsilon_t$$

$$a(L) y_t = b(L) \varepsilon_t$$

$$y_t = c(L) \varepsilon_t$$

$$= \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad \text{with} \quad \psi_0 = 1$$

So the ψ_1, ψ_2, ... coefficients in $c(L)$, can be obtained by equating coefficients of L^j, j = 1, 2, ... in $a(L)c(L) = b(L)$.

Exercise 3.4:

Difference-stationary (DS) Model

$$\left(1-0.347L-0.138L^2+0.146L^3\right)\left(1-L\right)y_t = 0.0032+\varepsilon_t$$

How does a shock today affect the level of y_t one year hence and infinitely far in the future?

Hint: MA representation of Δy_t

$$\Delta y_t = \mu + \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-1} + \ldots = \mu + \psi(L)\varepsilon_t$$

Part Availability

“The data for this case are adapted from a series provided by a large U.S. corporation. There are 90 weekly observations showing the percent of the time that parts for an industrial product are available when needed.”

Pankratz (1983) “Forecasting with univariate Box-Jenkins models”

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.5:

- **Identification**: Which model would you chose and why?
- **Estimation**: Estimate your model!
- **Diagnostic checking**: Is the selected model a statistically adequate representation of the available data?

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Percentiles of the chi-squared distribution

Percentiles of the χ^2 Distribution

	Percent									
df	0.005	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99	0.995
1	0.000039	0.000157	0.000982	0.003932	0.015791	2.705544	3.841459	5.023886	6.634897	7.879439
2	0.010025	0.020101	0.050636	0.102587	0.210721	4.605170	5.991465	7.377759	9.210340	10.596635
3	0.071722	0.114832	0.215795	0.351846	0.584374	6.251388	7.814728	9.348404	11.344867	12.838156
4	0.206989	0.297109	0.484419	0.710723	1.063623	7.779440	9.487729	11.143287	13.276704	14.860259
5	0.411742	0.554298	0.831212	1.145476	1.610308	9.236357	11.070498	12.832502	15.086272	16.749602
6	0.675727	0.872090	1.237344	1.635383	2.204131	10.644641	12.591587	14.449375	16.811894	18.547584
7	0.989256	1.239042	1.689869	2.167350	2.833107	12.017037	14.067140	16.012764	18.475307	20.277740
8	1.344413	1.646497	2.179731	2.732637	3.489539	13.361566	15.507313	17.534546	20.090235	21.954955
9	1.734933	2.087901	2.700390	3.325113	4.168159	14.683657	16.918978	19.022768	21.665994	23.589351
10	2.155856	2.558212	3.246973	3.940299	4.865182	15.987179	18.307038	20.483177	23.209251	25.188180
11	2.603222	3.053484	3.815748	4.574813	5.577785	17.275009	19.675138	21.920049	24.724970	26.756849
12	3.073824	3.570569	4.403789	5.226029	6.303796	18.549348	21.026070	23.336664	26.216967	28.299519
13	3.565035	4.106915	5.008751	5.891864	7.041505	19.811929	22.362032	24.735605	27.688250	29.819471
14	4.074675	4.660425	5.628726	6.570631	7.789534	21.064144	23.684791	26.118948	29.141238	31.319350
15	4.600916	5.229349	6.262138	7.260944	8.546756	22.307130	24.995790	27.488393	30.577914	32.801321
16	5.142205	5.812213	6.907664	7.961646	9.312236	23.541829	26.296228	28.845351	31.999927	34.267187
17	5.697217	6.407760	7.564186	8.671760	10.085186	24.769035	27.587112	30.191009	33.408664	35.718466
18	6.264805	7.014911	8.230746	9.390455	10.864936	25.989423	28.869299	31.526378	34.805306	37.156451
19	6.843971	7.632730	8.906517	10.117013	11.650910	27.203571	30.143527	32.852327	36.190869	38.582257
20	7.433844	8.260398	9.590778	10.850812	12.442609	28.411981	31.410433	34.169607	37.566235	39.996846
21	8.033653	8.897198	10.282898	11.591305	13.239598	29.615089	32.670573	35.478876	38.932173	41.401065
22	8.642716	9.542492	10.982321	12.338015	14.041493	30.813282	33.924439	36.780712	40.289360	42.795655
23	9.260425	10.195716	11.688552	13.090514	14.847956	32.006900	35.172462	38.075627	41.638398	44.181275
24	9.886234	10.856362	12.401150	13.848425	15.658684	33.196244	36.415028	39.364077	42.979820	45.558512
25	10.519652	11.523975	13.119720	14.611408	16.473408	34.381587	37.652484	40.646469	44.314105	46.927890
26	11.160237	12.198147	13.843905	15.379157	17.291885	35.563171	38.885139	41.923170	45.641683	48.289882
27	11.807587	12.878504	14.573383	16.151396	18.113896	36.741217	40.113272	43.194511	46.962942	49.644915
28	12.461336	13.564710	15.307861	16.927875	18.939243	37.915923	41.337138	44.460792	48.278236	50.993376
29	13.121149	14.256455	16.047072	17.708366	19.767744	39.087470	42.556968	45.722286	49.587885	52.335618
30	13.786720	14.953457	16.790772	18.492661	20.599235	40.256024	43.772972	46.979242	50.892181	53.671962

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.6:

Forecasting

- Forecast x_t from one to four weeks ahead!
- Forecast y_t from one to four weeks ahead!
- Forecast x_t and y_t from one to four weeks ahead using the information that we know at the end of week 91 that y_{91} = 87 and that we know at the end of week 92 that y_{92} = 86.5.

Forecasting

Optimal forecast:

Information set Ω_T:

Additional assumption:

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Forecasting an ARIMA (p,1,q)

$x_t = y_t - y_{t-1}$ => $\boxed{y_t = y_{t-1} + x_t}$

$$In\ period\ (T+1):\quad y_{T+1} = y_T + x_{T+1} \quad \rightarrow \tilde{y}_{T+1/\Omega_T} = E(y_T + x_{T+1} \,/\, \Omega_T) = y_T + \tilde{x}_{T+1/\Omega_T}$$

$$In\ period\ (T+2):\quad y_{T+2} = y_{T+1} + x_{T+2} = (y_T + x_{T+1}) + x_{T+2}$$

$$\rightarrow \tilde{y}_{T+2/\Omega_T} = E(y_T + x_{T+1} + x_{T+2} \,/\, \Omega_T) = \underbrace{y_T + \tilde{x}_{T+1/\Omega_T}}_{\tilde{y}_{T+1/\Omega_T}} + \tilde{x}_{T+2/\Omega_T}$$

$$In\ period\ (T+3):\quad y_{T+3} = y_{T+2} + x_{T+3} = (y_T + x_{T+1} + x_{T+2}) + x_{T+3}$$

$$\rightarrow \tilde{y}_{T+3/\Omega_T} = E(y_T + x_{T+1} + x_{T+2} + x_{T+3} \,/\, \Omega_T)$$

$$= \underbrace{y_T + \tilde{x}_{T+1/\Omega_T} + \tilde{x}_{T+2/\Omega_T}}_{\tilde{y}_{T+2/\Omega_T}} + \tilde{x}_{T+3/\Omega_T}$$

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

ARMA(*p*,*q*) process at time *T* + *l*:

$$\tilde{x}_{T+l|\Omega_T} = \varphi_1 \tilde{x}_{T+l-1|\Omega_T} + \ldots + \varphi_p \tilde{x}_{T+l-p|\Omega_T} + \tilde{\varepsilon}_{T+l|\Omega_T} - \theta_1 \tilde{\varepsilon}_{T+l-1|\Omega_T} - \ldots - \theta_q \tilde{\varepsilon}_{T+l-q|\Omega_T}$$

Recursive forecasting recipe:

1. replace unknown x_{T+l} by their forecasts for $l > 0$;
2. “forecasts” of x_{T+l} , $l \leq 0$, are simply the known values x_{T+l}
3. since ε_t is white noise, the optimal forecast of ε_{T+l}, $l > 0$, is simply zero
4. “forecasts” of ε_{T+l}, $l \leq 0$, are just the known values ε_{T+l}

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.7:

Forecasting

- Calculate the forecast error and then the MSE for the (true) model

  $x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$ and $y_t = y_{t-1} + x_t = y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$
- Calculate confidence intervals for $\tilde{x}_{91}, \tilde{x}_{92}, \tilde{x}_{93}, \tilde{y}_{91}, \tilde{y}_{92}$, and $\tilde{y}_{93}$.

Hint:

$$MSE\left(\tilde{y}_{T+s|\Omega_T}\right) = E\left[\left(y_{T+s} - \tilde{y}_{T+s|\Omega_T}\right)^2\right]$$

$$\left[\tilde{y}_{T+s|\Omega_T} \pm 1.96 \cdot \sqrt{MSE\left(\tilde{y}_{T+s|\Omega_T}\right)}\right]$$

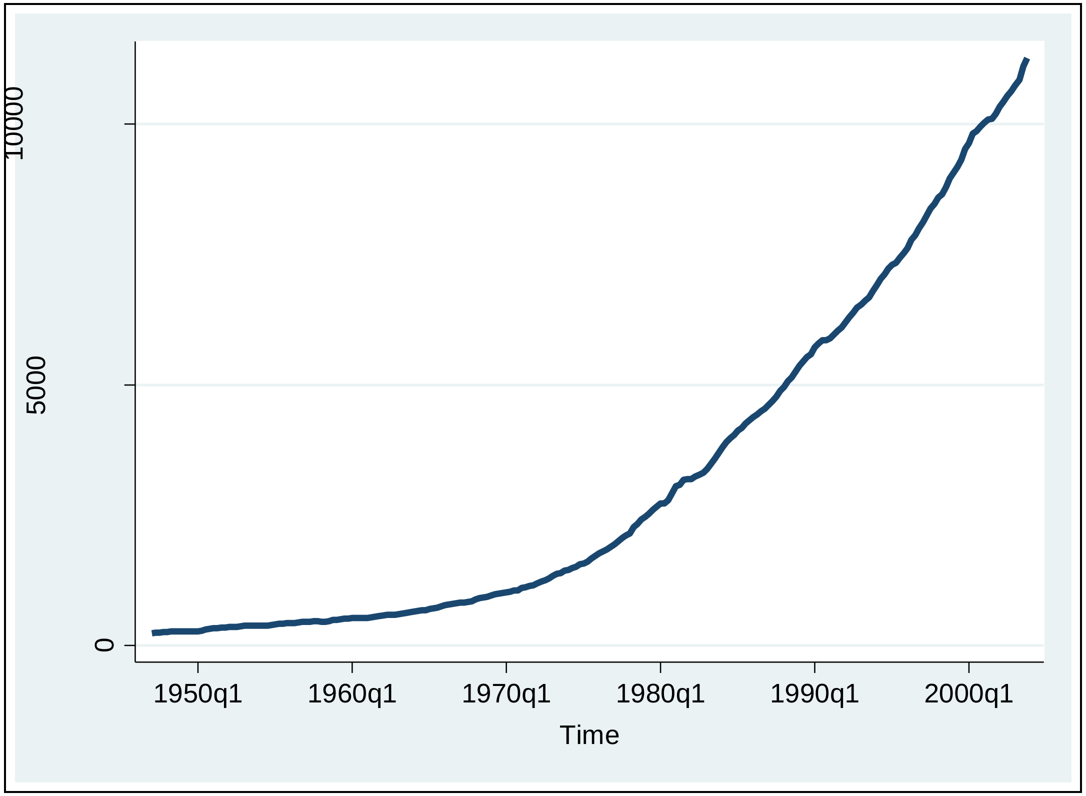
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Nonstationary Stochastic Processes

- Introduction
- Nonstationarity and Trends
- ARIMA Models
- **Unit Root Tests**
- Seasonal ARIMA

Original Time Series

U.S. quarterly GDP (1947q1 – 2003q4)

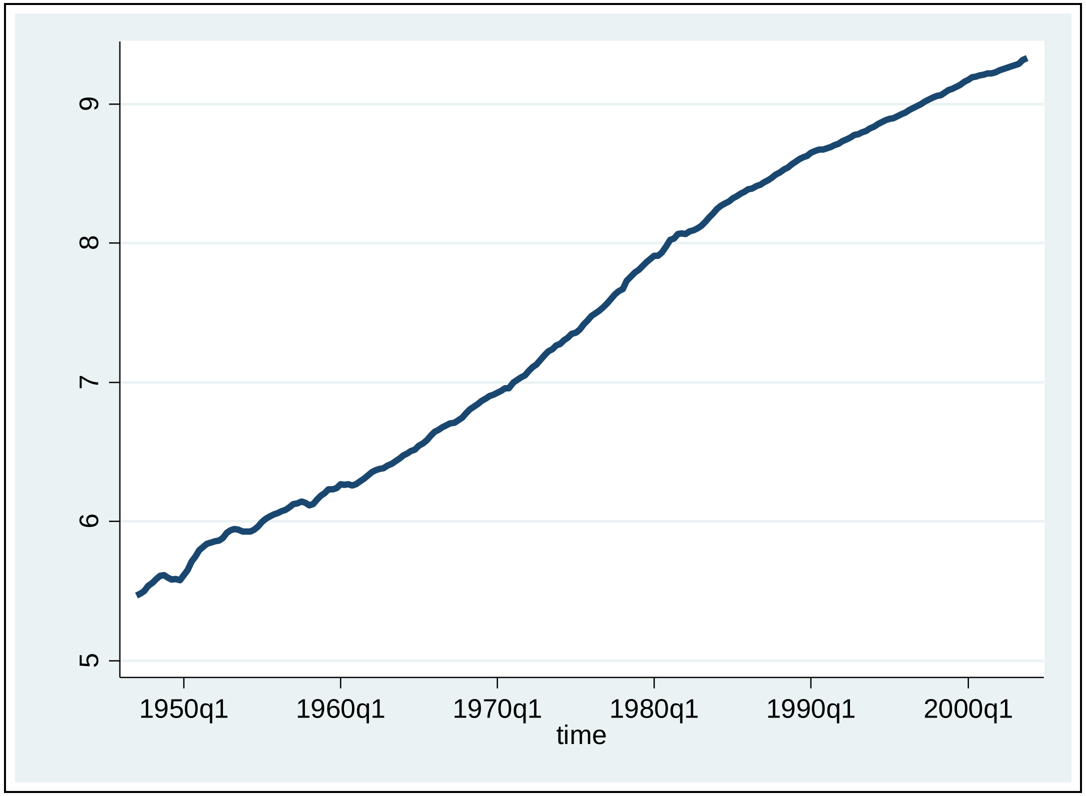


Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Logarithm of GDP

U.S. quarterly GDP (1947q1 – 2003q4), log



Which unit root test is adequate?

„Fit a specification that is a **plausible description of the data** under both the null and the alternative hypothesis.“

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Testing for Unit Roots

no constant, no trend	constant, no trend	constant and trend
$y_t = \varphi_1 y_{t-1} + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \gamma t + \varepsilon_t$
$H_0: \varphi_1 = 1$ $y_t = y_{t-1} + \varepsilon_t$ $H_1: \varphi_1 < 1$ $y_t = \varphi_1 y_{t-1} + \varepsilon_t$	$H_0: \varphi_1 = 1, (\delta = 0)$ $y_t = y_{t-1} + \varepsilon_t$ $H_1: \varphi_1 < 1, (\delta \neq 0)$ $y_t = \varphi_1 y_{t-1} + \delta + \varepsilon_t$	$H_0: \varphi_1 = 1, (\gamma = 0)$ $y_t = y_{t-1} + \delta + \varepsilon_t$ $H_1: \varphi_1 < 1, (\gamma \neq 0)$ $y_t = \varphi_1 y_{t-1} + \delta + \gamma t + \varepsilon_t$
• H_0: pure random walk (no drift) • H_1: stationary AR(1) with mean zero (i.e. strictly speaking $0 \leq \varphi_1 < 1$) • simplest case, mostly educational value • "Testing with zero intercept is extremely restrictive, so much that it is hard to imagine ever using it with economic time series"*	• H_0: pure random walk (no drift) • H_1: stationary AR(1) with arbitrary mean • applies to non-growing series • typical examples: "rates" (interest rates, inflation rates, unemployment rates)	• H_0: random walk with drift • H_1: trend stationary model with AR(1) errors • applies to growing series (but not explosive) • typical examples: GDP, consumption, investment

* Davidson, MacKinnon (1993) "Estimation and inference in econometrics ", p.702

no constant, no trend	constant, no trend	constant and trend
$y_t = \varphi_1 y_{t-1} + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \gamma t + \varepsilon_t$
$H_0 : \varphi_1 = 1$	$H_0 : \varphi_1 = 1, (\delta = 0)$	$H_0 : \varphi_1 = 1, (\gamma = 0)$
$y_t = y_{t-1} + \varepsilon_t$	$y_t = y_{t-1} + \varepsilon_t$	$y_t = y_{t-1} + \delta + \varepsilon_t$
$H_1 : \varphi_1 < 1$	$H_1 : \varphi_1 < 1, (\delta \neq 0)$	$H_1 : \varphi_1 < 1, (\gamma \neq 0)$
$y_t = \varphi_1 y_{t-1} + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \varepsilon_t$	$y_t = \varphi_1 y_{t-1} + \delta + \gamma t + \varepsilon_t$
Estimating equations	Estimating equations	Estimating equations
$y_t = \varphi_1 y_{t-1} + \varepsilon_t$ or	$y_t = \varphi_1 y_{t-1} + \delta + \varepsilon_t$ or	$y_t = \varphi_1 y_{t-1} + \delta + \gamma t + \varepsilon_t$ or
$\Delta y_t = \theta y_{t-1} + \varepsilon_t$	$\Delta y_t = \theta y_{t-1} + \delta + \varepsilon_t$	$\Delta y_t = \theta y_{t-1} + \delta + \gamma t + \varepsilon_t$
$\theta = (\varphi_1 - 1)$	$\theta = (\varphi_1 - 1)$	$\theta = (\varphi_1 - 1)$
Test statistics	Test statistics	Test statistics
$t = \frac{(\hat{\varphi}_1 - 1)}{\sigma_{\hat{\varphi}_1}}$ or $t = \frac{\hat{\theta}}{\sigma_{\hat{\theta}}}$	$t = \frac{(\hat{\varphi}_1 - 1)}{\sigma_{\hat{\varphi}_1}}$ or $t = \frac{\hat{\theta}}{\sigma_{\hat{\theta}}}$	$t = \frac{(\hat{\varphi}_1 - 1)}{\sigma_{\hat{\varphi}_1}}$ or $t = \frac{\hat{\theta}}{\sigma_{\hat{\theta}}}$

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Critical values for Dickey-Fuller tests

U.S. quarterly GDP (1947q1 – 2003q4), log

Sample Size T	No constant, no trend		Constant, no trend		Constant, trend	
	1%	5%	1%	5%	1%	5%
25	-2.66	-1.95	-3.75	-3.00	-4.38	-3.60
50	-2.62	-1.95	-3.58	-2.93	-4.15	-3.50
100	-2.60	-1.95	-3.51	-2.89	-4.04	-3.45
250	-2.58	-1.95	-3.46	-2.88	-3.99	$\boxed{-3.43}$
500	-2.58	-1.95	-3.44	-2.87	-3.98	-3.42
∞	-2.58	-1.95	-3.43	-2.86	-3.96	-3.41

Verbeek (2000) “A Guide to Modern Econometrics ”

Augmented Dickey-Fuller Unit Root Test

Not all time-series processes can be well represented by an AR(1) process. It is possible to use Dickey-Fuller tests in higher-order equations.

Example: AR(2) without constant, no trend

$H_0: \varphi_1 + \varphi_2 = 1$ given $|\varphi_2| < 1$

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t \qquad \Big| + \varphi_2 y_{t-1} - \varphi_2 y_{t-1}$$

$$= (\varphi_1 + \varphi_2) y_{t-1} - \varphi_2 (y_{t-1} - y_{t-2}) + \varepsilon_t \qquad \Big| - y_{t-1}$$

$$\Delta y_t = (\varphi_1 + \varphi_2 - 1) y_{t-1} - \varphi_2 \Delta y_{t-1} + \varepsilon_t \qquad \Big| \text{with } \pi_1 = \varphi_1 + \varphi_2 - 1 \text{ and } \pi_2 = -\varphi_2$$

$$= \pi_1 y_{t-1} + \pi_2 \Delta y_{t-1} + \varepsilon_t$$

$H_0: \pi_1 = \varphi_1 + \varphi_2 - 1 = 0$

In general, for an AR(p): $\Delta y_t = \pi_1 y_{t-1} + \pi_2 \Delta y_{t-1} + \ldots + \pi_p \Delta y_{t-p+1} + \varepsilon_t$

Any **ARMA model** (with an invertible MA polynomial) can be written as an infinite autoregressive process.

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Augmented Dickey-Fuller Unit Root Test

Any unknown **ARIMA(p, d, q)** process can be well approximated by an **ARIMA(p*, d, 0)** of order no more than $\mathbf{T^{1/3}}$. (Said and Dickey(1984), Enders(1995), p.226)

=> The above **augmented regression** can also be used to **test for a unit root** in any **ARMA model**.

$$\Delta y_t = \pi_1 y_{t-1} + \pi_2 \boxed{\Delta y_{t-1}} + \ldots + \pi_p \boxed{\Delta y_{t-p+1}} + \delta + \gamma t + \varepsilon_t$$

"augmentation terms"
Note: order **p** => **(p-1)** augmentation terms

$$\begin{aligned}
\pi_1 &= \varphi_1 + \varphi_2 + \varphi_3 + \ldots + \varphi_p - 1 && \textit{corresponds to } y_{t-1} \\
\pi_2 &= \quad -(\varphi_2 + \varphi_3 + \ldots + \varphi_p) && \textit{corresponds to } \Delta y_{t-1} \\
\pi_3 &= \qquad -(\varphi_3 + \ldots + \varphi_p) && \textit{corresponds to } \Delta y_{t-2} \\
&\ldots \\
\pi_p &= \qquad\qquad -\varphi_p && \textit{corresponds to } \Delta y_{t-p+1}
\end{aligned}$$

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Augmented Dickey-Fuller Unit Root Test

Why is it important to select the appropriate lag length?

Including **too many** lags:

reduces power of the test to reject the null of a unit root:

- because the number of parameters estimated has increased and
- because the number of usable observations has decreased.

Including **too few** lags:

will not appropriately capture the actual error process and φ_1 and its standard error will not be properly estimated.

Augmented Dickey-Fuller Unit Root Test

How to select the appropriate lag length?

- Start with a relatively long lag length (p^*) and **pare down the model by the usual *t*-test**.

  $$\Delta y_t = \pi_1 y_{t-1} + \pi_2 \Delta y_{t-1} + \ldots + \pi_p \Delta y_{t-p+1} + \varepsilon_t$$

  If the null hypothesis $\pi_{p*} = 0$ is accepted, reestimate the regression using a lag length of p^*-1. Repeat the process until the p^*-ℓ is significantly different from zero. If no value of ℓ leads to rejection, the simple Dickey-Fuller test is used.

- Use a model selection criterion to determine the order of the regression, e.g. the **Hannan-Quinn criterion**:

  $$HQ(p) = \log \hat{\sigma}^2(p) + (1+p)\frac{2\ln(\ln(T))}{T} \qquad \hat{\sigma}_\varepsilon^2 = \frac{1}{(T-p)}\sum_{t=1}^{T-p} \hat{\varepsilon}_t^2$$

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Additional Information

For trending series we have to discriminate between deterministic and stochastic trends.

In a Trend-Stationary-Model we have a deterministic trend with *stationary stochastic* fluctuations around this deterministic trend. A Difference-Stationary-Model can have a stochastic trend (e.g. a Random Walk) or a combination of a stochastic and deterministic trend (e.g. a Random Walk with drift). To check if the time series contains a stochastic trend (with a unit root) we can use the (Augmented) Dickey-Fuller-Test. This test is not that powerful, but it gives us a hint, that we can use a Difference-Stationary-Model to deal with the non-stationary time series. If we can reject the hypothesis of a unit root in the data, it means we can still try to get rid of the non-stationarity by using a Trend-Stationary-Model.

For the augmented Dickey-Fuller-Test we consider the general formula (with constant and trend) to test for one unit root (if there is more then one unit root we would have to consider differencing the series more then just once):

$$\Delta y_t = \pi_1 y_{t-1} + \pi_2 \Delta y_{t-1} + \ldots + \pi_p \Delta y_{t-p+1} + \delta + \gamma t + \varepsilon_t$$

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Additional Information

The null hypothesis of the test is:

$$\pi_1 = \sum_k \varphi_k - 1 = 0$$

We estimate the equation:

$$\Delta y_t = \pi_1 y_{t-1} + \pi_2 \Delta y_{t-1} + \ldots + \pi_p \Delta y_{t-p+1} + \delta + \gamma t + \varepsilon_t$$

By ordinary least squares and focus on the estimate of π_1.

For the coefficients π_j with j>1 of the OLS-Regression we can use the standard t-statistic, but for π_1, we have to consider the Dickey-Fuller-Distribution to get correct critical values.

You can either test for various p or you use information criteria, such as AIC, BIC and the Hannan-Quinn-Criteria (HQIC), to get the correct number of lags. All three criteria tend to come to the same conclusion. AIC sometimes overestimates the lag length, because it is the least strict one to penalize an high order of lags. BIC is the strictest in penalizing loss of degree of freedom by having more parameters in the fitted model. The HQIC holds the middle ranking in penalizing and is therefore often used.

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Additional Information

After we have found evidence for a unit root and a DS-Model we have to identify the correct ARMA-Model for the differenced series. For this purpose we use our well-known Box-Jenkins-Approach and apply it to the differenced series.

Difference Stationary Model

- **Identification**

  Which model would you chose and why?
- **Estimation**

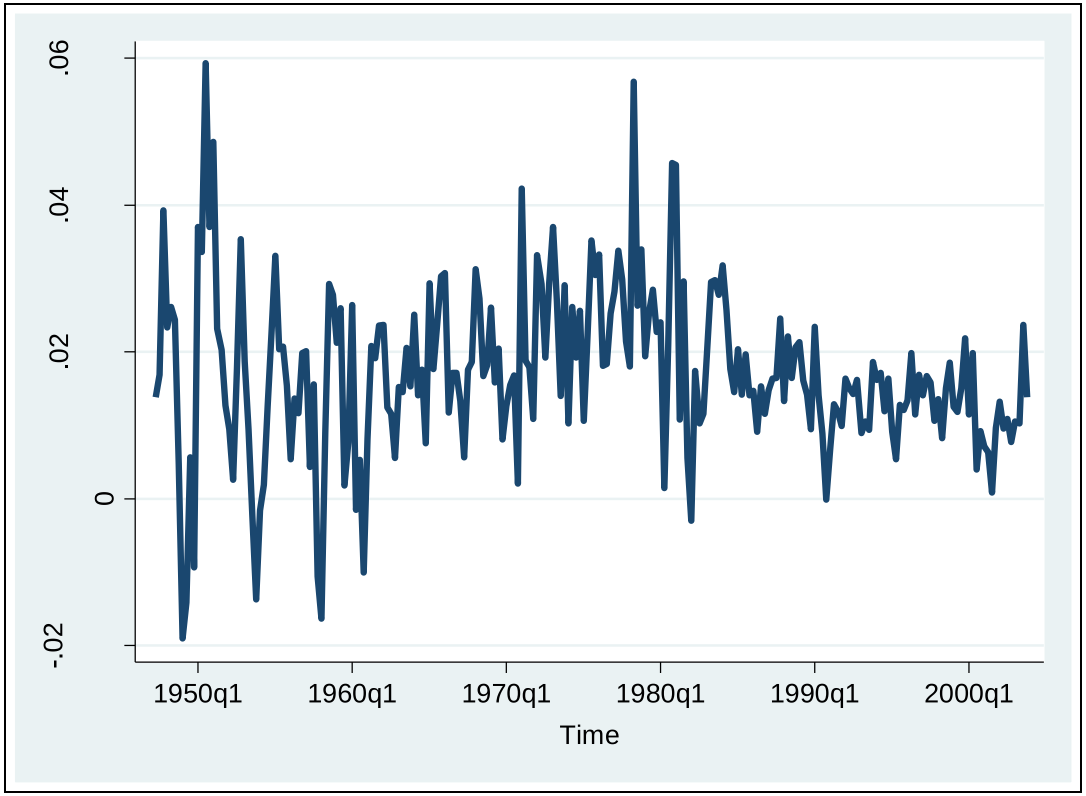
  Estimate your model.
- **Diagnostic checking**

  Is the selected model a statistically adequate representation of the available data?

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Identification

First difference of logarithm of U.S. quarterly GDP (1947q1 – 2003q4)



Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.8:

Forecasting (without using Stata's forecast commands)

- Forecast x_t from one to four quarters ahead.
- Forecast y_t from one to four quarters ahead.
- Forecast x_t and y_t from one to four quarters ahead using the information that we know at the end of 2004q1 that y_{04q1} = 9.36 and that we know at the end of 2004q2 that y_{04q2} = 9.38.

$$\tilde{x}_t = 0.0169737 + 0.3981594\,\varepsilon_{t-1} + 0.2694164\,\varepsilon_{t-2}$$

```
. list time lnGDP res_MA2 in 226/228

     +--------------------------------+
     |   time       lnGDP     res_MA2 |
     |--------------------------------|
226. | 2003q2   9.2916164   -.0037882 |
227. | 2003q3   9.3153305    .0091118 |
228. | 2003q4   9.3291893   -.0057223 |
     +--------------------------------+
```

Nonstationary Stochastic Processes

- Introduction
- Nonstationarity and Trends
- ARIMA Models
- Unit Root Tests
- **Seasonal ARIMA**

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

ARIMA(p,d,q)

$$a(L)\underbrace{\underbrace{(1-L)^d y_t}_{\Delta^d y_t}}_{x_t} = \delta + b(L)\varepsilon_t$$

$$\underbrace{x_t - \varphi_1 x_{t-1} - \ldots - \varphi_p x_{t-p}}_{a(L)x_t} = \delta + \underbrace{\varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}}_{b(L)\varepsilon_t}$$

with y_t ~ **ARIMA(p,d,q)** and x_t ~ **ARMA(p,q)**

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Seasonality

When observations are available on a monthly (s = 12) or a quarterly (s = 4) basis, some allowance must be made for seasonal effects. Two approaches:

- Work with seasonally adjusted data
- Incorporate seasonality into time series models

Multiplicative Seasonal ARIMA

  - allows for stationary seasonal pattern that tends to disappear with increasing lag or lead time
  - allows also for nonstationary seasonal pattern ("seasonal trend", slowly changing seasonal pattern)

Harvey (1981) "Time Series Models", p. 171-185

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Why incorporate seasonality in the model?

Seasonal adjustment procedures tend to result in over-adjustment, so that there is a tendency for seasonally adjusted series to exhibit negative autocorrelations at the seasonal lags.

It is **generally preferable to work with the unadjusted series**. Seasonal adjustment can introduce considerable distortion into a series, and, at the same time, there is no guarantee that the adjusted series will be free from seasonal effects.

Harvey (1981) “Time Series Models”, p. 171-185

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Seasonal ARMA Processes

General Formulation to allow for both AR and MA terms at specified seasonal lags

$$\left(1-\varphi_1^s L^s - \ldots - \varphi_P^s L^{P_s}\right) y_t = \left(1-\theta_1^s L^s - \ldots - \theta_Q^s L^{Q_s}\right) \zeta_t$$

$$\varphi^s\left(L^s\right) y_t = \theta^s\left(L^s\right) \zeta_t$$

with s denoting the number of seasons in the year and with ζ_t denoting a white noise disturbance term

→ **pure seasonal ARMA** process of order $(P,Q)_s$

→ **ACF will contain ‘gaps’** at non-seasonal lags

Harvey (1981) “Time Series Models”, p. 171-185

Franziska Plitzko

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Primary distinguishing characteristics of theoretical ACF's and PACF's for purely seasonal stationary processes

Process	ACF	PACF
S-AR	Tails off toward zero at lags $k \times s$, k = 1, 2, …	Cuts off after lag P_S
S-MA	Cuts off after lag Q_S	Tails off at lags $k \times s$, k = 1, 2, …
S-ARMA	Tails off at lags $k \times s$	Tails off at lags $k \times s$

Shumway, Stoffer (2000) "Time Series analysis and Its Applications", p. 158

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Seasonal ARMA Processes

A **pure seasonal ARMA** is not appropriate unless seasonal movements are the only predictable feature of the series

→ **Multiplicative** seasonal ARMA

Harvey (1981) “Time Series Models”, p. 171-185

Seasonal ARMA Processes

Multiplicative Seasonal ARMA

Replace the white noise disturbance ζ_t term by a non-seasonal ARMA(p,q) process, u_t:

$$\varphi(L)u_t = \theta(L)\varepsilon_t$$

Plugged in $\varphi^s(L^s)y_t = \theta^s(L^s)\zeta_t$ yields

$$\varphi^s(L^s)\varphi(L)y_t = \theta^s(L^s)\theta(L)\varepsilon_t$$

$\rightarrow$ ARMA process of order $(p,q)\times(P,Q)_s$

Harvey (1981) “Time Series Models”, p. 171-185

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

A General Class of Models

Box and Jenkins propose that the conventional and seasonal differencing operators be applied until the series is stationary and that this stationary series be modeled by a multiplicative seasonal ARMA

$$\varphi^{s}\left(L^{s}\right)\varphi(L)\Delta^{d}\Delta_{s}^{D}y_{t}=\theta^{s}\left(L^{s}\right)\theta(L)\varepsilon_{t}$$

where *D* and *d* are integers denoting the number of times the seasonal and first difference operators are applied respectively.

Multiplicative Seasonal ARIMA of order $(p,d,q)\times(P,D,Q)_{s}$

Harvey (1981) “Time Series Models”, p. 171-185

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Air-carrier freight

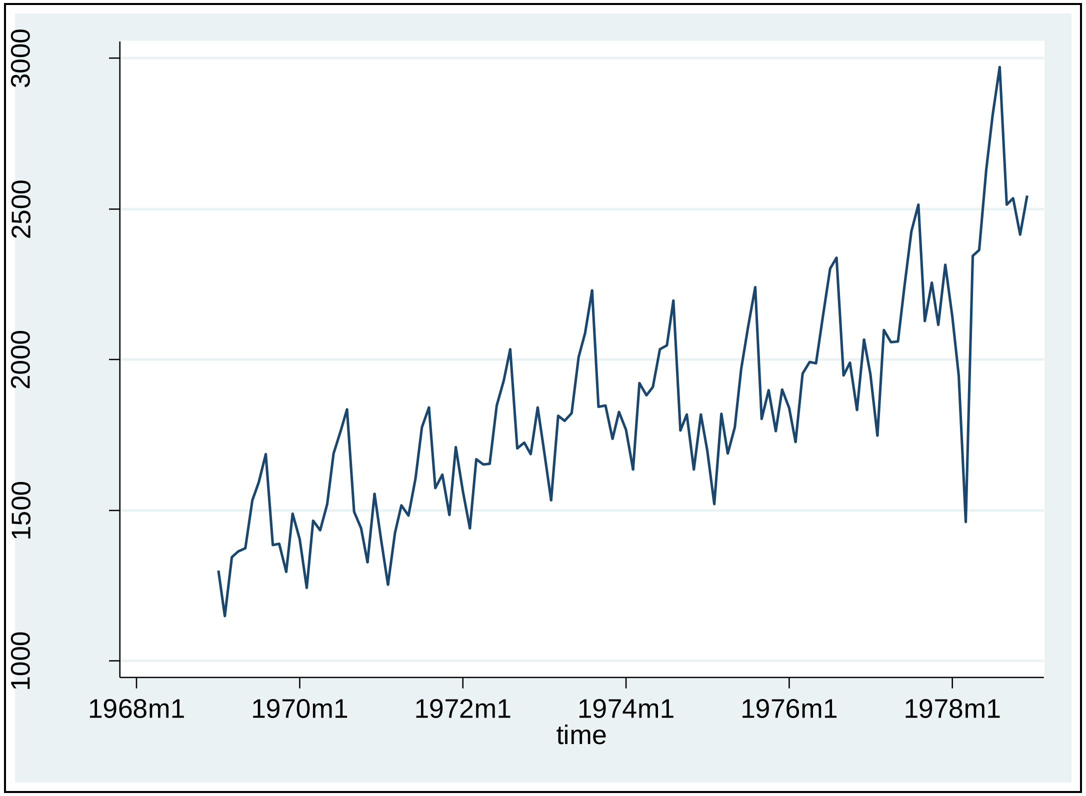
“Volume of freight, measured in ton-mile, hauled by air carriers in the United States. There are 120 monthly observations covering the years 1969-1978.”

Pankratz (1983) “Forecasting with univariate Box-Jenkins models”

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Air-carrier freight

Original series



Pankratz (1983) “Forecasting with univariate Box-Jenkins models”

Franziska Plitzko
Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de

Exercise 3.9:

Write down the following SARIMA-Models with this formula:

$$\varphi^{s}\left(L^{s}\right)\varphi(L)\Delta^{d}\Delta_{s}^{D}y_{t}=\theta^{s}\left(L^{s}\right)\theta(L)\varepsilon_{t}$$

and in our usual notation ($y_t = \ldots$).

- SARIMA $(0,1,1)(0,1,1)_{12}$
- SARIMA $(0,1,0)(1,1,0)_{12}$
- (here the SARIMA-Formula only): SARIMA $(2,2,2)(2,2,2)_{12}$

Raum: H 5103 D, **Tel.:** 314-78734, **E-Mail:** franziska.plitzko@tu-berlin.de