# Further Issues in Using OLS with Time Series Data

April 29, 2022

# Stationary and Weakly Dependent Time Series

- **The assumptions used so far seem to be too restricitive**
  - Strict exogeneity, homoscedasticity, and no serial correlation are very demanding requirements, especially in the time series context
  - Statistical inference rests on the validity of the normality assumption
  - Much weaker assumptions are needed if the sample size is large
  - A key requirement for large sample analysis of time series is that the time series in question are stationary and weakly dependent
- **<u>Stationary time series</u>**
  - Loosely speaking, a time series is stationary if its stochastic properties and its temporal dependence structure do not change over time

# Stationary and Weakly Dependent Time Series

- **Stationary stochastic processes**

  A stochastic process $\{x_t : t = 1, 2, \dots\}$ is <u>stationary</u>, if for every collection of indices $1 \leq t_1 \leq t_2 \leq \cdots \leq t_m$ the joint distribution of $(x_{t_1}, x_{t_2}, \dots, x_{t_m})$ is the same as that of $(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_m+h})$ for all integers $h \geq 1$.

- **Covariance stationary processes**

  A stochastic process $\{x_t : t = 1, 2, \dots\}$ is <u>covariance stationary</u>, if its expected value, its variance, and its covariances are constant over time:

  1) $E(x_t) = \mu$, 2) $Var(x_t) = \sigma^2$, and 3) $Cov(x_t, x_{t+h}) = f(h)$.

# Stationary and Weakly Dependent Time Series

- **Weakly dependent time series**

A stochastic process $\{x_t : t = 1, 2, \ldots\}$ is <u>weakly dependent</u>, if $x_t$ is „almost independent" of $x_{t+h}$ if $h$ grows to infinity (for all $t$).

- **Discussion of the weak dependence property**
  - An implication of weak dependence is that the correlation between $x_t$, and $x_{t+h}$ must converge to zero if $h$ grows to infinity
  - For the LLN and the CLT to hold, the individual observations must not be too strongly related to each other; in particular their relation must become weaker (and this fast enough) the farther they are apart
  - Note that a series may be nonstationary but weakly dependent

# Stationary and Weakly Dependent Time Series

- **Examples for weakly dependent time series**

- **Moving average process of order one (MA(1))**

$$x_t = e_t + \alpha_1 e_{t-1}$$ ← The process is a short moving average of an i.i.d. series $e_t$

The process is weakly dependent because observations that are more than one time period apart have nothing in common and are therefore uncorrelated.

- **Autoregressive process of order one (AR(1))**

$$y_t = \rho_1 y_{t-1} + e_t$$ ← The process carries over to a certain extent the value of the previous period (plus random shocks from an i.i.d. series $e_t$)

$$\Rightarrow Corr(y_t, y_{t+h}) = \rho_1^h$$

If the stability condition $|\rho_1| < 1$ holds, the process is weakly dependent because serial correlation converges to zero as the distance between observations grows to infinity.

# Asymptotic Properties of OLS

- **Assumption TS.1' (Linear in parameters)**

  - Same as assumption TS.1 but now the dependent and independent variables are assumed to be <u>stationary</u> and <u>weakly dependent</u>

- **Assumption TS.2' (No perfect collinearity)**

  - Same as assumption TS.2

- **Assumption TS.3' (Zero conditional mean)**

  - Now the explanatory variables are assumed to be only contemporaneously exogenous rather than strictly exogenous, i.e.

$$E(u_t | \mathbf{x}_t) = 0$$ ← The explanatory variables <u>of the same period</u> are uninformative about the mean of the error term

# Asymptotic Properties of OLS

- **Theorem 11.1 (Consistency of OLS)**

$$TS.1' - TS.3' \quad \Rightarrow \quad plim\ \widehat{\beta}_j = \beta_j, \quad j = 0, 1, \ldots, k$$

- **Why is it important to relax the strict exogeneity assumption?**
  - Strict exogeneity is a serious restriction beause it rules out all kinds of dynamic relationships between explanatory variables and the error term
  - In particular, it rules out feedback from the dep. var. on future values of the explanat. variables (which is very common in economic contexts)
  - Strict exogeneity precludes the use of lagged dependent variable as regressors

# Asymptotic Properties of OLS

- **Why do lagged dependent variables violate strict exogeneity?**

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

← This is the simplest possible regression model with a lagged dependent variable

Contemporanous exogeneity: $E(u_t | y_{t-1}) = 0$

Strict exogeneity: $E(u_t | y_0, y_1, \ldots, y_{n-1}) = 0$ ←

Strict exogeneity would imply that the error term is uncorrelated with all $y_t$, $t=1, \ldots, n-1$

This leads to a contradiction because:

$$Cov(y_t, u_t) = \beta_1 Cov(y_{t-1}, u_t) + Var(u_t) > 0$$

- **OLS estimation in the presence of lagged dependent variables**

  - Under contemporaneous exogeneity, OLS is consistent but biased

# Asymptotic Properties of OLS

- **Assumption TS.4' (Homoskedasticity)**

$$Var(u_t|\mathbf{x}_t) = Var(u_t) = \sigma^2$$

← The errors are contemporaneously homoskedastic

- **Assumption TS.5' (No serial correlation)**

$$Corr(u_t, u_s|\mathbf{x}_t, \mathbf{x}_s) = 0, \ \ t \neq s$$

← Conditional on the explanatory variables in periods t and s, the errors are uncorrelated

- **<u>Theorem 11.2 (Asymptotic normality of OLS)</u>**

  - Under assumptions TS.1' – TS.5', the OLS estimators are asymptotically normally distributed. Further, the usual OLS standard errors, t-statistics and F-statistics are asymptotically valid.

# Asymptotic Properties of OLS

- **Example: Efficient Markets Hypothesis (EMH)**

The EMH in a strict form states that information observable to the market prior to week $t$ should not help to predict the return during week $t$. A simplification assumes in addition that only past returns are considered as relevant information to predict the return in week $t$. This implies that

$$E(return_t | return_{t-1}, return_{t-2}, \ldots) = E(return_t)$$

A simple way to test the EMH is to specify an AR(1) model. Under the EMH assumption, TS.3' holds so that an OLS regression can be used to test whether this week's returns depend on last week's.

$$\widehat{return}_t = \underset{(.081)}{.180} + \underset{(.038)}{.059} \, return_{t-1}$$

$$n = 689, R^2 = .0035, \bar{R}^2 = .0020$$

There is no evidence against the EMH. Including more lagged returns yields similar results.

# Using Highly Persistent Time Series in Regression Analysis

- **Using trend-stationary series in regression analysis**

  - Time series with deterministic time trends are nonstationary

  - If they are stationary around the trend and in addition weakly dependent, they are called trend-stationary processes

  - Trend-stationary processes also satisfy assumption TS.1'

- **<u>Using highly persistent time series in regression analysis</u>**

  - Unfortunately many economic time series violate weak dependence because they are highly persistent (= strongly dependent)

  - In this case OLS methods are generally invalid (unless the CLM hold)

  - In some cases transformations to weak dependence are possible

# Using Highly Persistent Time Series in Regression Analysis

- **Random walks**

$$y_t = y_{t-1} + e_t$$

The random walk is called random walk because it wanders from the previous position $y_{t-1}$ by an i.i.d. random amount $e_t$

$$\Rightarrow \; y_t = (y_{t-2} + e_{t-1}) + e_t = \ldots = e_t + e_{t-1} + \ldots + e_1 + y_0$$

The value today is the accumulation of all past shocks plus an initial value. This is the reason why the random walk is highly persistent: The effect of a shock will be contained in the series forever.

$$E(y_t) = E(y_0)$$

$$Var(y_t) = \sigma_e^2 t$$

$$Corr(y_t, y_{t+h}) = \sqrt{t/(t+h)}$$

The random walk is <u>not covariance stationary</u> because its variance and its covariance depend on time.

It is also <u>not weakly dependent</u> because the correlation between observations vanishes very slowly and this depends on how large $t$ is.
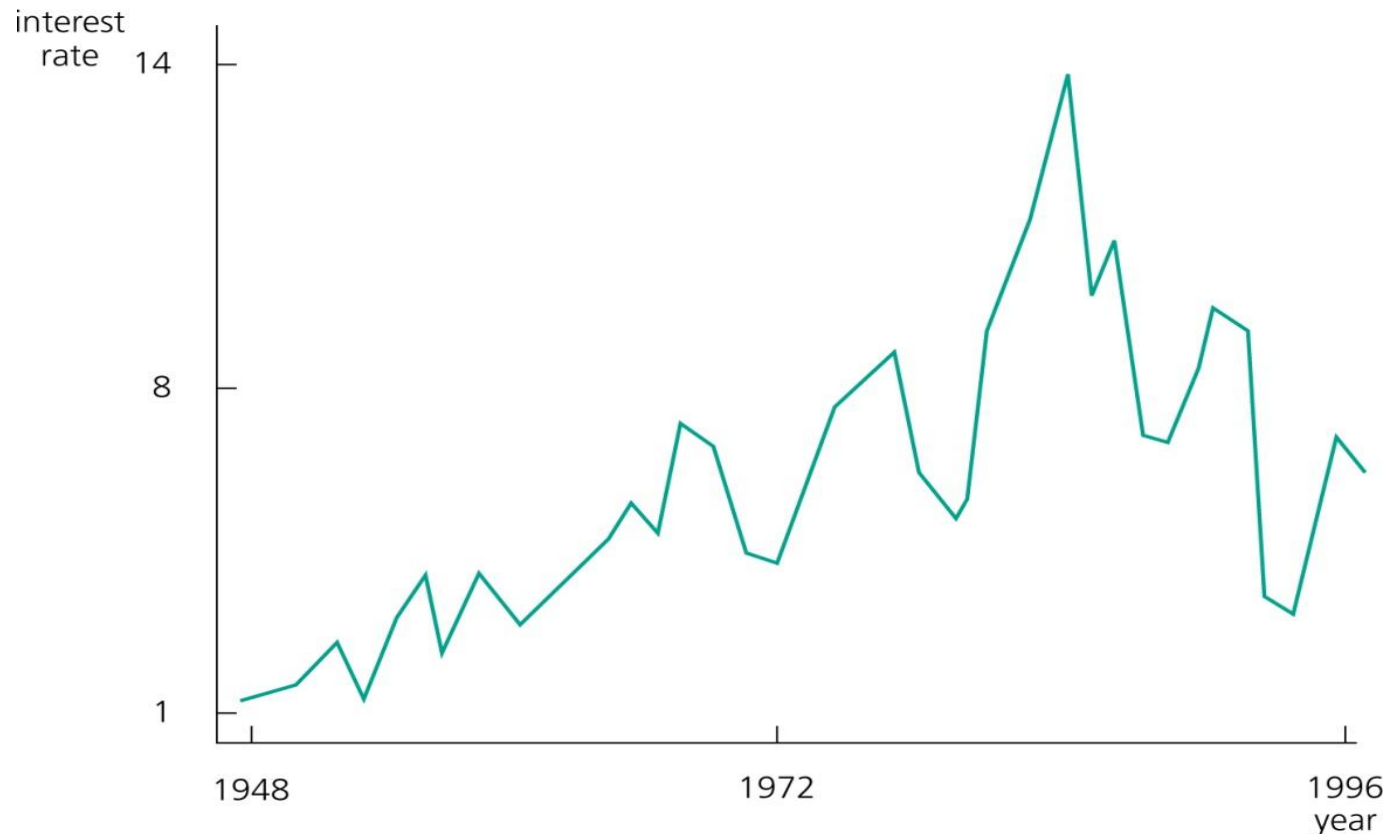
# Using Highly Persistent Time Series in Regression Analysis

- **Examples for random walk realizations**



The random walks wander around with no clear direction
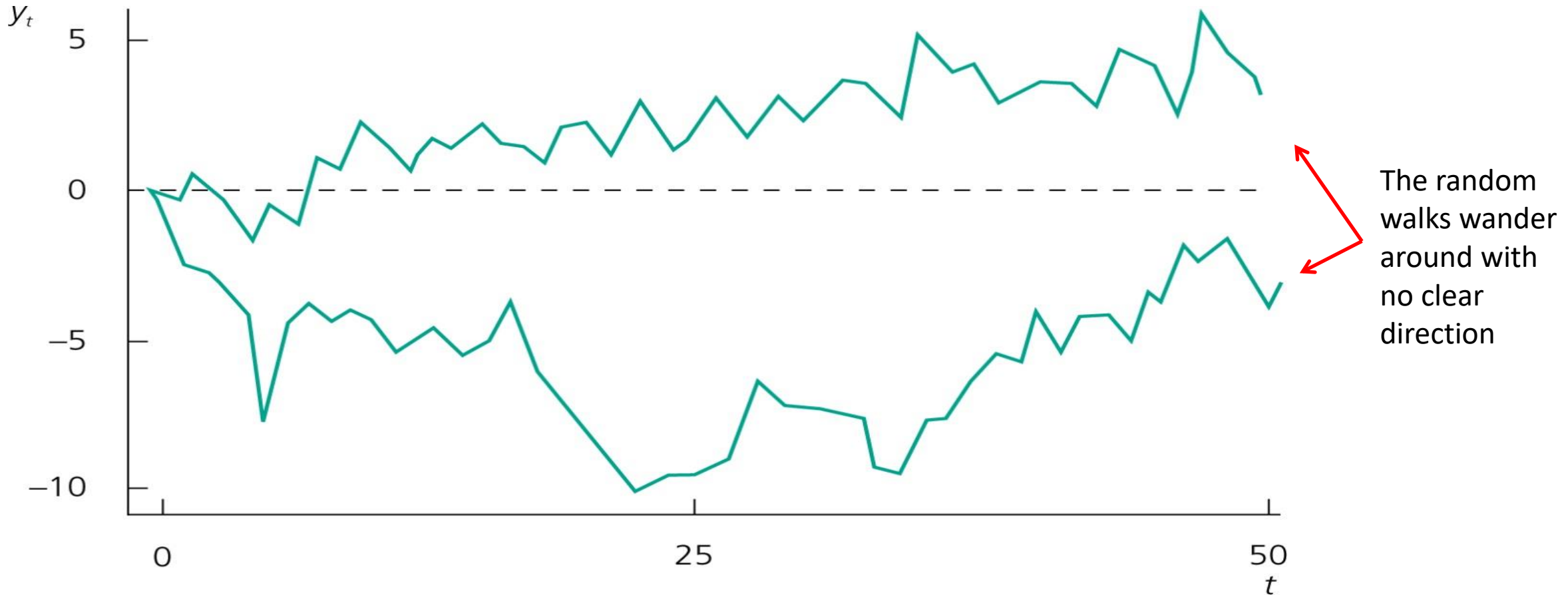
# Using Highly Persistent Time Series in Regression Analysis

- **Three-month T-bill rate as a possible example for a random walk**



A random walk is a special case of a <u>unit root process</u>.

Unit root processes are defined as the random walk but $e_t$ may be an arbitrary weakly dependent process.

From an economic point of view it is important to know whether a time series is highly persistent. <u>In highly persistent time series, shocks or policy changes have lasting/permanent effects</u>, in weakly dependent processes their effects are transitory.

# Using Highly Persistent Time Series in Regression Analysis - OPTIONAL

- **Random walks with drift**

$$y_t = \boxed{\alpha_0} + y_{t-1} + e_t$$

In addition to the usual random walk mechanism, there is a deterministic increase/decrease (= drift) in each period

$$\Rightarrow \quad y_t = \boxed{\alpha_0 t} + e_t + e_{t-1} + \ldots + e_1 + y_0$$

This leads to a linear time trend around which the series follows its random walk behaviour. As there is no clear direction in which the random walk develops, it may also wander away from the trend.

$$E(y_t) = \boxed{\alpha_0 t} + E(y_0)$$
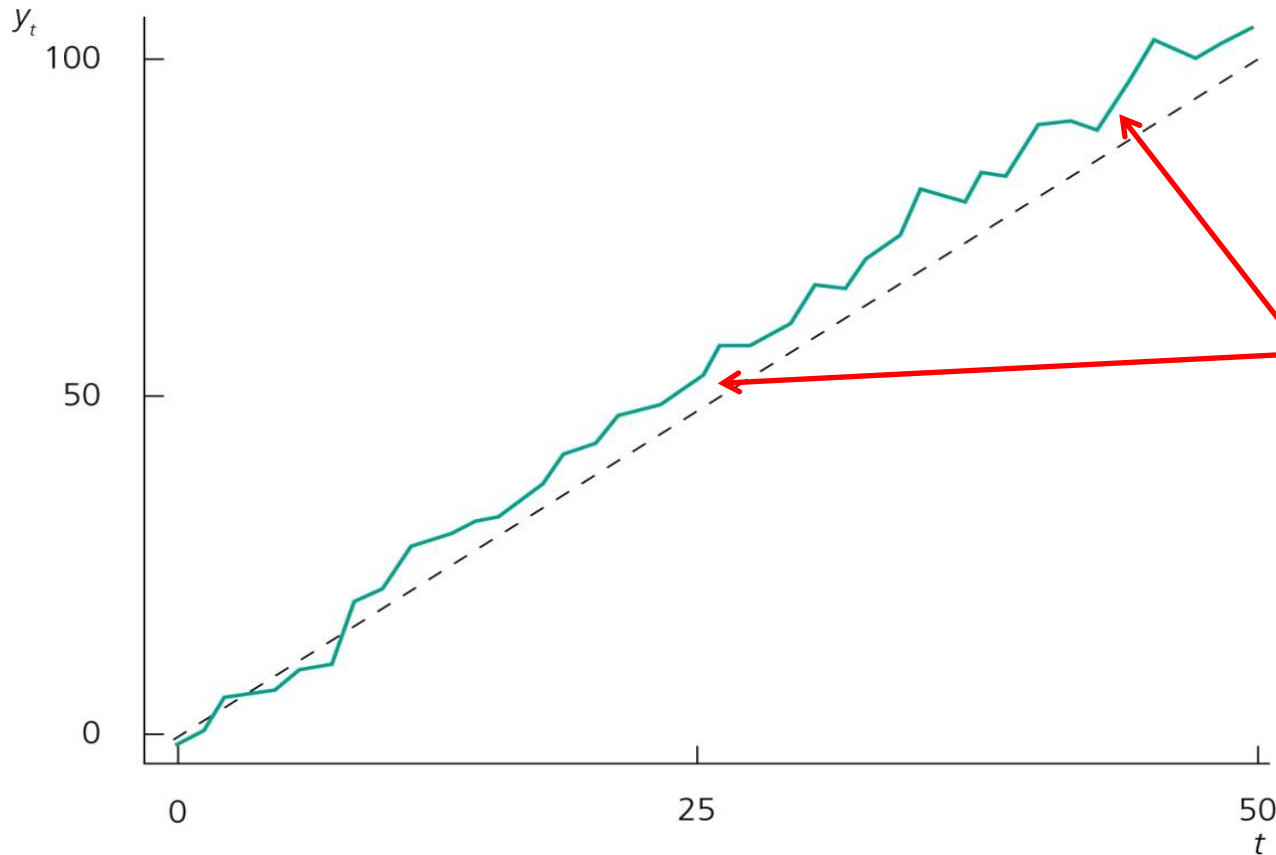
$$Var(y_t) = \sigma_e^2 t$$

$$Corr(y_t, y_{t+h}) = \sqrt{t/(t+h)}$$

Otherwise, the random walk with drift has similar properties as the random walk without drift.

Random walks with drift are <u>not covariance stationary</u> and <u>not weakly dependent</u>.

# Using Highly Persistent Time Series in Regression Analysis - OPTIONAL

- **Sample path of a random walk with drift**



Note that the series does not regularly return to the trend line.

Random walks with drift may be good models for time series that have an obvious trend but are not weakly dependent.

# Using Highly Persistent Time Series in Regression Analysis

- **<u>Transformations on highly persistent time series</u>**

- **Order of integration**

  - Weakly dependent time series are integrated of order zero (= I(0))

  - If a time series has to be <u>differenced</u> one time in order to obtain a weakly dependent series, it is called integrated of order one (= I(1))

- **Examples for I(1) processes**

$$y_t = y_{t-1} + e_t \;\; \Rightarrow \;\; \triangle y_t = y_t - y_{t-1} = e_t$$

After differencing, the resulting series are weakly dependent (because $e_t$ is weakly dependent).

- **Differencing is often a way to achieve weak dependence**

# Using Highly Persistent Time Series in Regression Analysis

- **Deciding whether a time series is I(1)**

    - There are statistical tests for testing whether a time series is I(1) (= unit root tests);

    - Alternatively, look at ***the sample first order autocorrelation***:

$$\widehat{\rho}_1 = \widehat{Corr}(y_t, y_{t-1})$$ ← Measures how strongly adjacent times series observations are related to each other.

    - If the sample first order autocorrelation is close to one, this suggests that the time series may be highly persistent (= contains a unit root)

    - Alternatively, the series may have a deterministic trend

    - Both unit root and trend may be eliminated by differencing

# Using Highly Persistent Time Series in Regression Analysis

- **Example: Fertility equation**

$$gfr_t = \alpha_0 + \delta_0 pe_t + \delta_1 pe_{t-1} + \delta_2 pe_{t-2} + u_t$$

This equation could be estimated by OLS if the CLM assumptions hold. These may be questionable, so that one would have to resort to large sample analysis. For large sample analysis, the fertility series and the series of the personal tax exemption have to be stationary and weakly dependent. This is questionable because the two series are highly persistent:

$$\widehat{\rho}_{gfr} = .977, \quad \widehat{\rho}_{pe} = .964$$

It is therefore better to estimate the equation in first differences. This makes sense because if the equation holds in levels, it also has to hold in first differences:

$$\widehat{\triangle gfr} = -\ .964\ -\ .036\ \triangle pe\ -\ .014\ \triangle pe_{-1} +\ .110\ \triangle pe_{-2}$$
$$\qquad\qquad (.468)\quad (.027)\qquad\qquad (.028)\qquad\qquad (.027)$$

$$n = 69, R^2 = .233, \bar{R}^2 = .197$$

Estimate of $\delta_2$

# Testing for Unit Roots

- For the validity of regression analysis, it is crucial to know whether or not dependent or independent variables are highly persistent

- **Dickey-Fuller test**

The test is based on an AR(1) regression

$$y_t = \alpha + \rho y_{t-1} + e_t$$

$$H_0 : \rho = 1, \ H_1 : \rho < 1$$

Under the null hypothesis, the process has a unit root. Under the alternative, it is a stable AR(1) process

- One can use the t-statistic to test the hypothesis, <u>but under the null, it does not have the t-distribution</u> but the Dickey-Fuller distribution
- The Dickey-Fuller distribution has to be looked up in tables

# Testing for Unit Roots

- **Alternative Formulation of the Dickey-Fuller test**

$$\triangle y_t = \alpha + \theta y_{t-1} + e_t$$

← The alternative representation is obtained by subtracting $y_{t-1}$ from both sides

$$H_0 : \theta = 0, \quad H_1 : \theta < 0$$

The critical value is much more negative than it would be in a t-distribution

- **Critical values for Dickey-Fuller test**

| TABLE 18.2 Asymptotic Critical Values for Unit Root *t* Test: No Time Trend | | | | |
|---|---|---|---|---|
| Significance level | 1% | 2.5% | 5% | 10% |
| Critical value | −3.43 | −3.12 | −2.86 | −2.57 |

© Cengage Learning, 2013

# Testing for Unit Roots

- **Example: Unit root test for three-month T-Bill rates**

$$\Delta \widehat{r3}_i = \underset{(.261)}{.625} \; - \; \underset{(.037)}{.091} \; r3_{t-1}$$

$$n = 123, \; R^2 = .048$$

The t-statistic is -2.46. As consequence, the null hypothesis of a unit root <u>cannot be rejected</u>

Include lagged differences of dependent variable.

- **Augmented Dickey-Fuller test**

$$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + \cdots + \gamma_p \Delta y_{t-p} + e_t$$

  - The augmented Dickey-Fuller test allows for more serial correlation
  - The critical values and the rejection rule are the same as before

# Testing for Unit Roots

- **Dickey-Fuller test for time series that have a time trend**

$$\triangle y_t = \alpha + \delta t + \theta y_{t-1} + \gamma_1 \triangle y_{t-1} + \ldots + \gamma_p \triangle y_{t-p} + e_t$$

Under the alternative hypothesis of no unit root, the process is trend-stationary

- **Critical values for Dickey-Fuller test with time trend**

**TABLE 18.3  Asymptotic Critical Values for Unit Root $t$ Test: Linear Time Trend**

| Significance level | 1% | 2.5% | 5% | 10% |
|---|---|---|---|---|
| Critical value | −3.96 | −3.66 | −3.41 | −3.12 |

© Cengage Learning, 2013

Even more negative

- **There are many other unit root tests …**

# Next Class – Quiz 2

# **Panel Data Models**