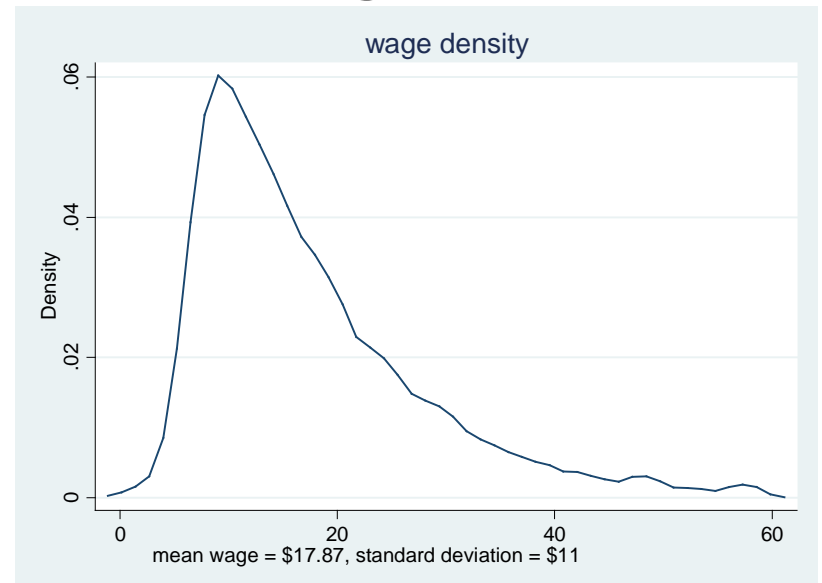


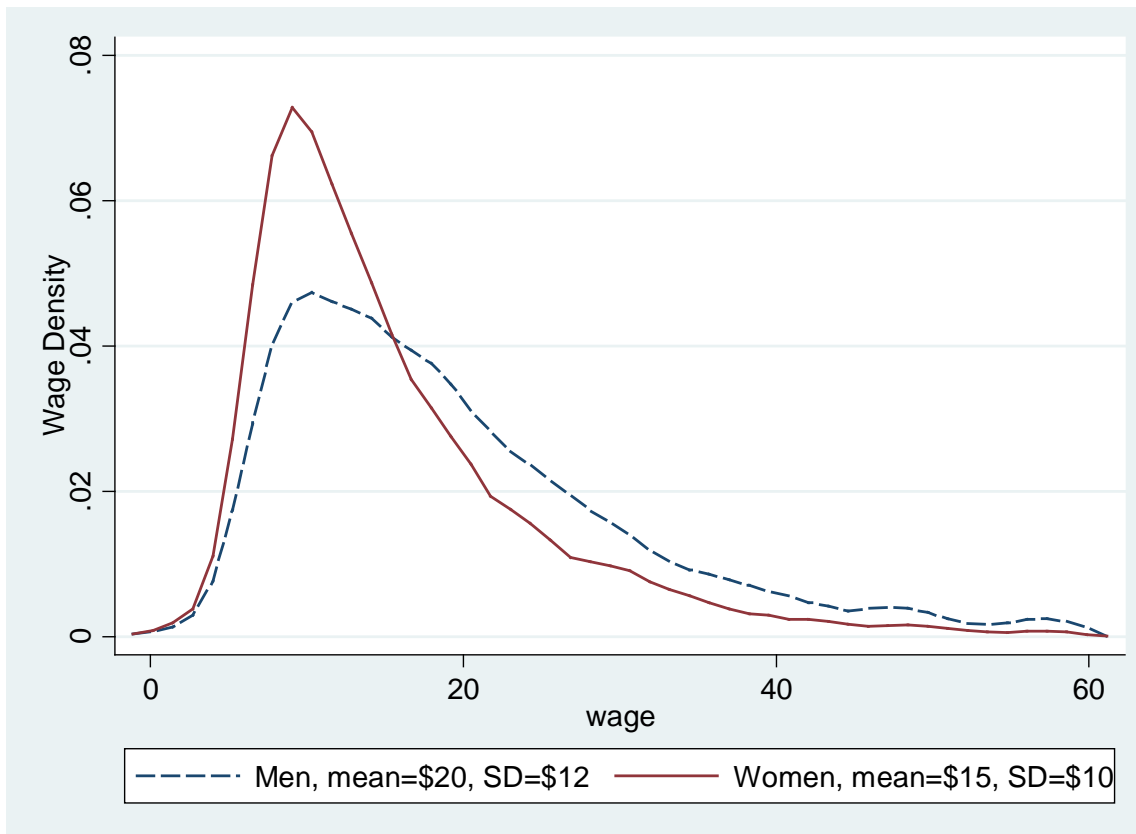
Conditional Forecast

- We had considered forecasting the wage of a random person.
- The distribution is quite diffuse as it includes all wage earners. We know nothing about the person being forecast.

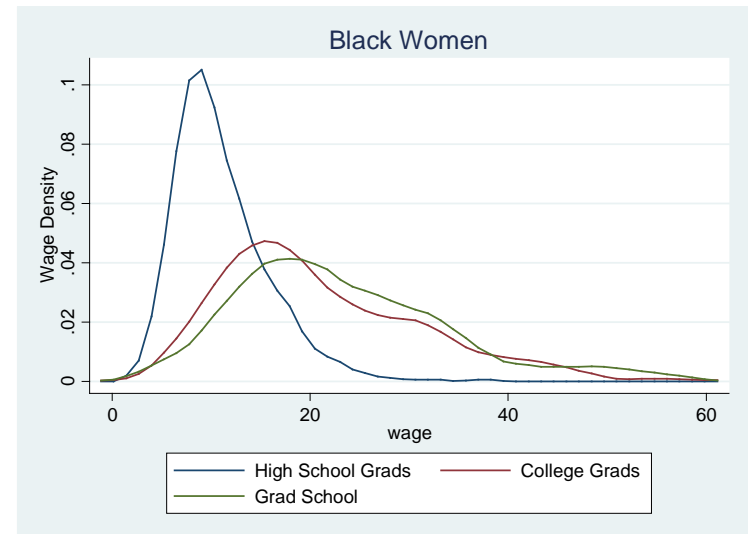
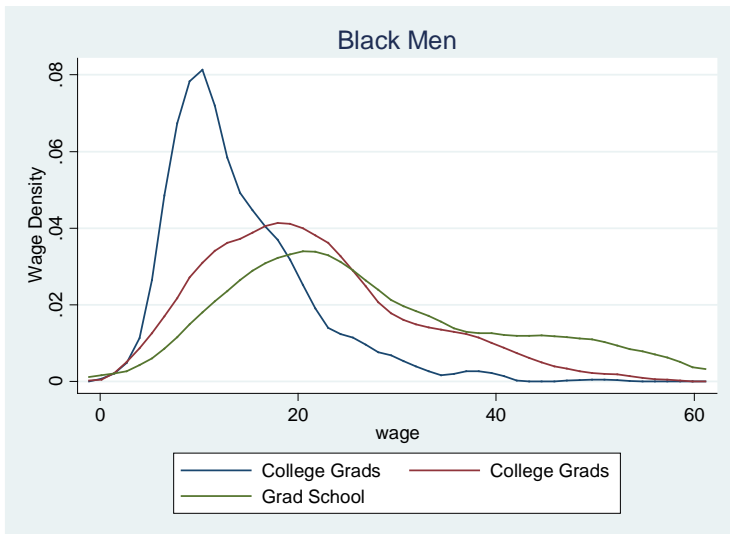
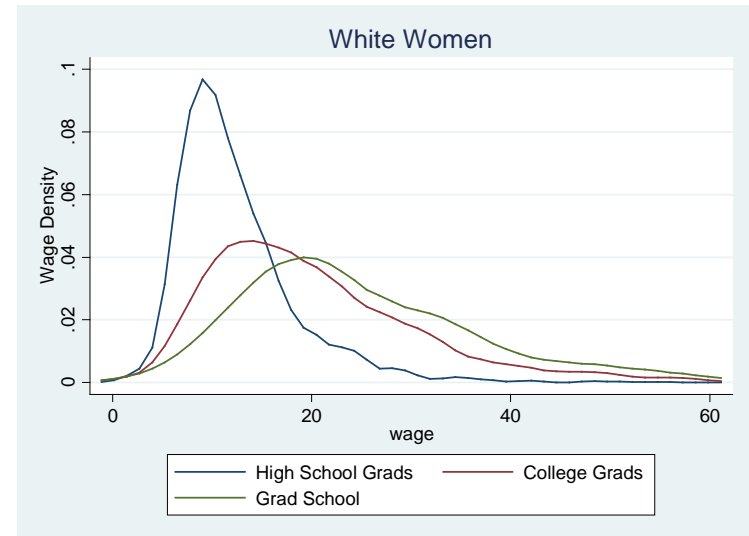
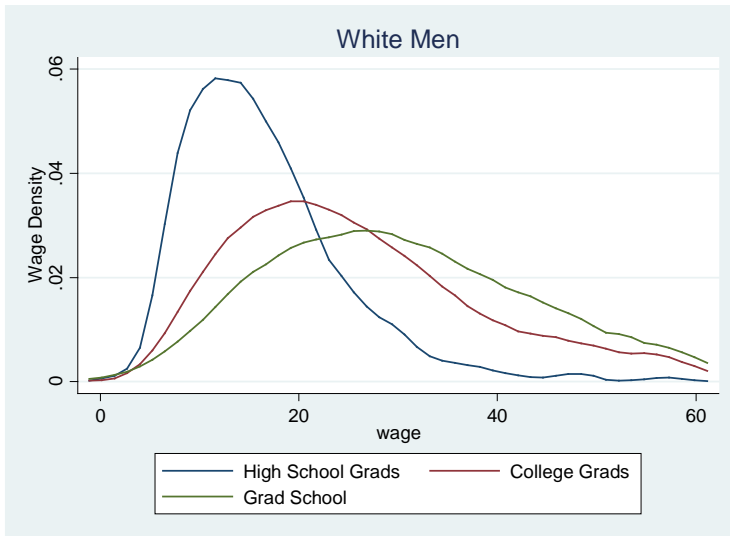


Conditional Forecast

- Now suppose we know that the person is a man (or a woman).
- The information improves the forecast.



Conditional on Sex, Race, Education



Conditional Forecasts (Means)

		Men	Women
White	High School	\$17	\$13
	College	\$27	\$20
	Graduate	\$32	\$26
Black	High School	\$14	\$11
	College	\$21	\$21
	Graduate	\$29	\$23

Role of Conditioning

- By conditioning on available information, we can make forecasts more accurate.
- Conditioning reduces the *risk* of the forecast.
- Ignoring estimation, conditioning on more information is always better in the sense of reducing risk.

Forecasting Economic Time Series

Economic forecasts rely on time series data – observations which are recorded sequentially over time.

Time Series Data

- A time series is written as y_t
- The index t denotes the time period.
- A time period may be a year, quarter, month, week, day, transaction, or any other time unit.
- We call this the data frequency.
- In this course, most of the data we will examine will be monthly.
- Some will be quarterly or weekly.

Loading Data into STATA

- See the STATA handout on course webpage
- One method is to copy a text file, or excel file, and paste into the Data Editor
 - This works if the data is well organized as a table, with each row a single observation
- An easy method is to use the FRED interface
 - We will be mostly using this tool

FRED

- St Louis Federal Reserve Bank (FED)
- <https://fred.stlouisfed.org>
- Google “fred”
- A depository of a 424,000 U.S. and international time series from 80 sources
- Plug-in tools to load data directly into Excel and Stata
- To install Stata tool, type
 - `ssc install freduse`
- For example, to load the variables “GDP” and “GDPC1”, type
 - `freduse GDP GDPC1`
 - You need to know FRED label
 - Labels are (mostly) capitalized
 - You must load all variables with one command, with an empty dataset

GDP Example

- Example: Load U.S. GDP and real GDP
- Go to <https://fred.stlouisfed.org/>
- Browse data by Category
- National Income & Product Accounts
- GDP/GNP
- “Real Gross Domestic Product”
 - Billions of Chained 2009 Dollars, Quarterly, Seasonally Adjusted Annual Rate
 - Label is “Real Gross Domestic Product (GDPC1)”
 - GDPC1 is the FRED label
- “Gross Domestic Product”
 - Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate
 - FRED label is GDP
- You can change the sample period (generally, use “Max”)
- Use “Edit Graph” to modify graph, say to percentage change
- Use “Download” to obtain an excel file

GDP Example in Stata

- Open Stata
 - `ssc install freduse`
- Memory should be clear, also in case of error you may need to delete a temporary txt file.
 - `freduse GDP GDPC1`
- You now have four variables in memory, “GDP”, “GDPC1”, “date” and “daten”
- If the label names are inconvenient, rename
 - `rename GDPC1 rgdp`

Dates and Times

- STATA requires that one variable be a designated time index
- Time indices created by FRED not in correct format
- To create a time index, easiest method is with tsmktim utility
- First, install on your computer
 - `ssc install tsmktim`
- Use tsmktim to create a time variable, labeled as *time*, *t*, or your choice
- Set as the STATA time index with the tsset command
 - `tsset time`

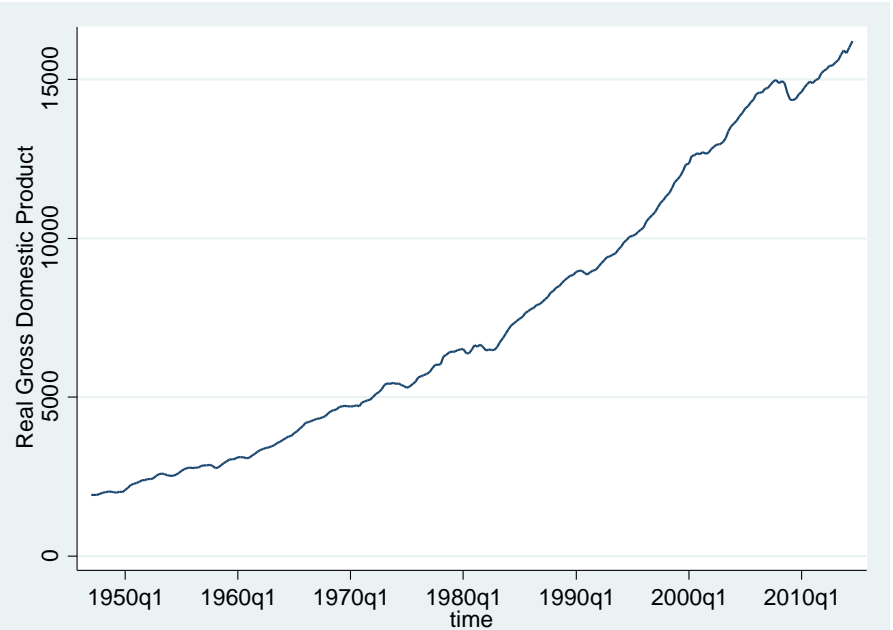
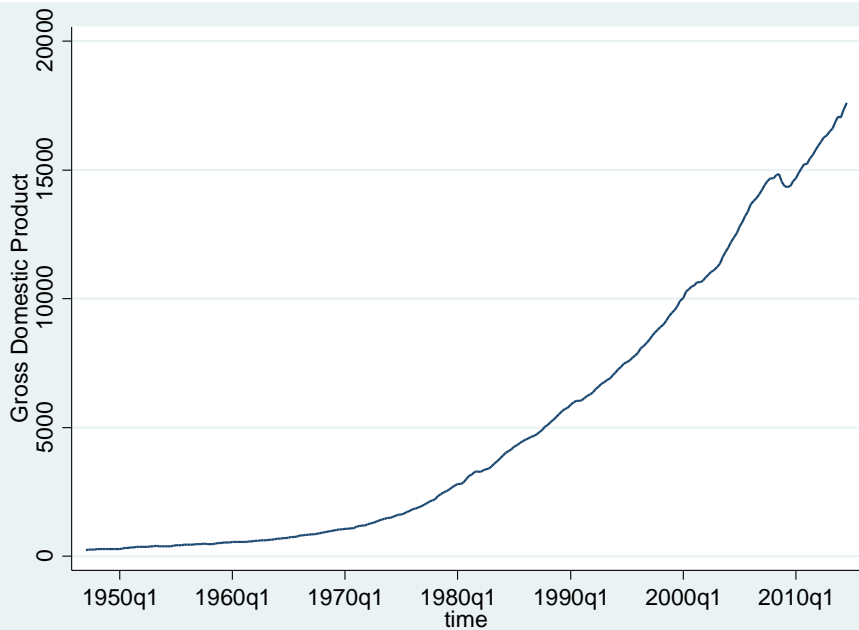
tsmktime

- For annual data
 - `tsmktime time, start (1947)`
 - `tsset time`
 - This creates an annual variable “time”, starting in 1947
- For quarterly data
 - `tsmktime time, start (1947q1)`
 - `tsset time`
 - This creates a quarterly variable “time”, starting in the first quarter of 1947
- For monthly
 - `tsmktime time, start(1947m3)`
 - `tsset time`
 - Starts in third month of 1947
- For our application, GDP starts in first quarter of 1947
 - `tsmktime time, start (1947q1)`
 - `tsset time`

Graphs

- With time series data, graphs are very useful.
- The place to start is a time-series plot.
 - A plot of the values of the series against time.
- Let's demonstrate by plotting U.S. GDP, nominal and real
 - `tsline GDP`
 - `tsline rgdp`

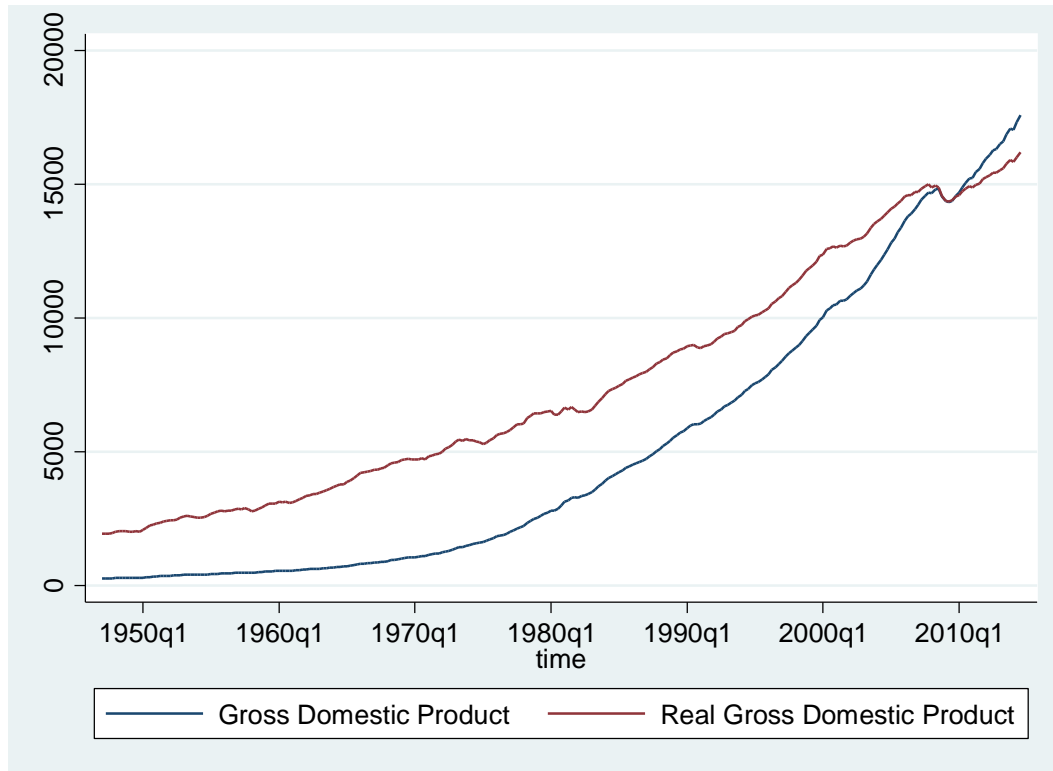
U.S. GDP, Nominal and Real



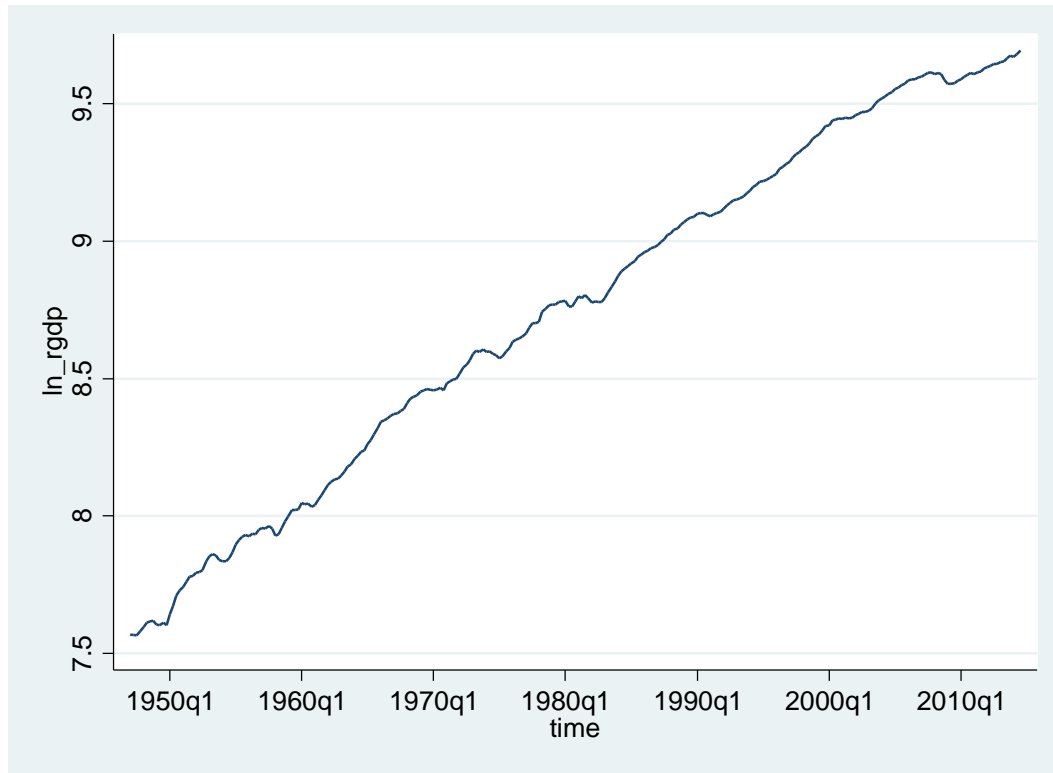
Transformations

- Multiple Plots
 - `tsline GDP rgdp`
- Log transform
 - $\ln(y_t)$
 - `generate ln_rgdpln(rgdp)`
 - `tsline ln_rgdpln_rgdpln`
- Differences
 - $\Delta y_t = y_t - y_{t-1}$
 - `tsline d.rgdpln_rgdpln`
 - `tsline d.ln_rgdpln_rgdpln`
- Percentage Change, annual rate
 - $400(y_t - y_{t-1})/y_{t-1}$
 - Very close to difference of logs

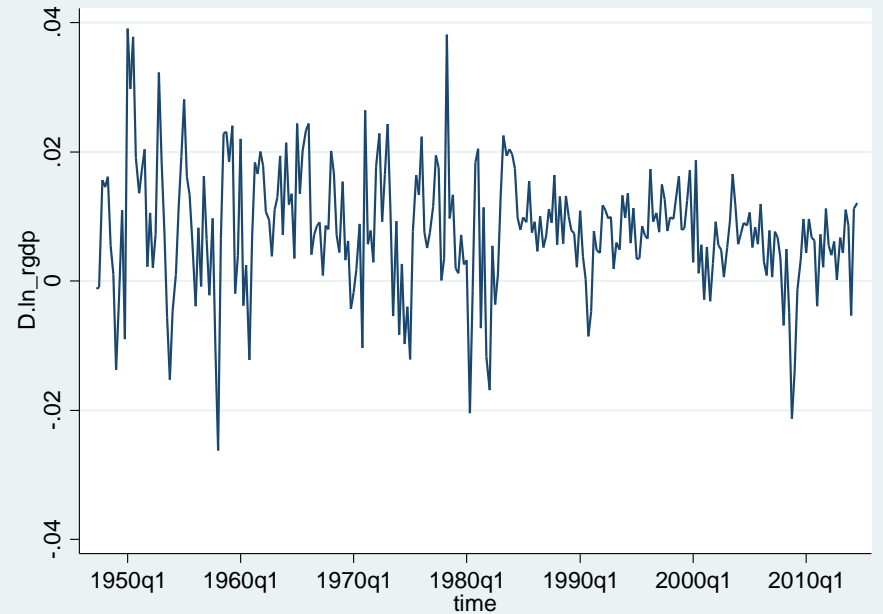
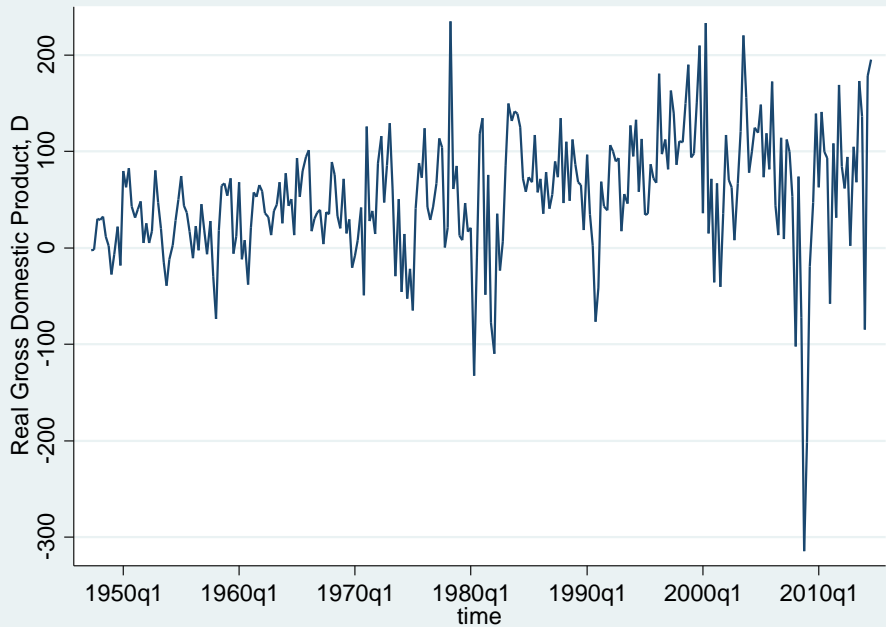
GDP and Real GDP



Ln(Real GDP)



Annual Real GDP Changes (Differenced Real GDP and log rgdp)



Creating a Dataset and Do File

- When done importing data, naming variables, setting the time index, and making transformations, save your file for future use.
- If you want to start over, type
 - `clear all`
- When you create STATA graphs, only the most recent is retained by default. To keep all graphs, you can give them a name
 - `graph rename gdp`
- Collect your commands in a file with the ending `.do`, for example “`problemset1.do`”. Execute the file by typing
 - `do problemset1`
 - This will execute all your commands, create all results, allow you to edit your work later, save your work for future use, etc.

Assignments

- Read Chapters 3-4 from Diebold
- Read Chapter 1 from *The Signal and the Noise*
 - Reading Reflection
 - Due Thursday (1/26)
 - Class Discussion
- Problem Set # 2
 - Due Tuesday (1/31)
- Read Chapter 2 from *The Signal and the Noise*
 - Reading Reflection
 - Due Thursday (2/2)