Forecasting Project

- It's time to be thinking about your Project
- Project Description is due after Spring Break (Tuesday March 28)
- Describe the variable
  - Source
  - When future observations will be available
  - Where you will find the data
  - Present a time-series of the historical series

Timing is Important

- Your series must have at least one new realization between when you turn in your (second) forecast report and when you turn in your evaluation report.
- For example, if you want to forecast the U.S. GDP for 2017 Q1, the number is announced by the BEA on April 28. You can turn in your forecast up until April 27.

Data

- You select your economic series to forecast
- As this is an economics course, it should be an *economic* series.
- Feel free to ask me about possibilities.
- You can pick a standard economic series, or can be creative.
- I advise against financial series. There is little to forecast, and the project will turn out to be uninteresting.
- I recommend looking for a monthly or quarterly series.
- I recommend looking for a series with at least 20 years of historical data.

Seasonal + Cycle

- Consider a components model with seasonal and AR(1) cycle

$$y_t = S_t + C_t$$

$$C_t = \beta C_{t-1} + e_t$$

- The seasonal S_t is a set of seasonal dummies

$$S_t = \sum_{i=1}^{s} \alpha_i D_{it}$$

Transformation

$$y_t = S_t + C_t$$

$$C_t = \beta C_{t-1} + e_t$$

- Take the first equation and lag it once

$$y_{t-1} = S_{t-1} + C_{t-1}$$

- Multiply it by β

$$\beta y_{t-1} = \beta S_{t-1} + \beta C_{t-1}$$

- Then subtract it from the first equation to find

$$y_t = \beta y_{t-1} + S_t - \beta S_{t-1} + e_t$$

Seasonal Representation

- We find

$$y_t = \beta y_{t-1} + S_t - \beta S_{t-1} + e_t$$

- When the seasonal S_t is a set of seasonal dummies, one for each season, this equation suggests a regression on
  - y_{t-1}
  - Seasonal dummies
  - Lagged Seasonal dummies

Redundant

- But lagged seasonal dummies are redundant with the original seasonal dummies
- The set of lagged dummy variables are collinear with the current dummy variables
- Given that you know this month is February, there is no information in knowing that last month was January.
- The lagged dummies can be (should be) omitted

Seasonal + Cycle

- We have found that the regression model is

$$y_t = \sum_{i=1}^{s} \alpha_i D_{it} + \beta y_{t-1} + e_t$$

or

$$y_t = \alpha_0 + \sum_{i=1}^{s-1} \alpha_i D_{it} + \beta y_{t-1} + e_t$$

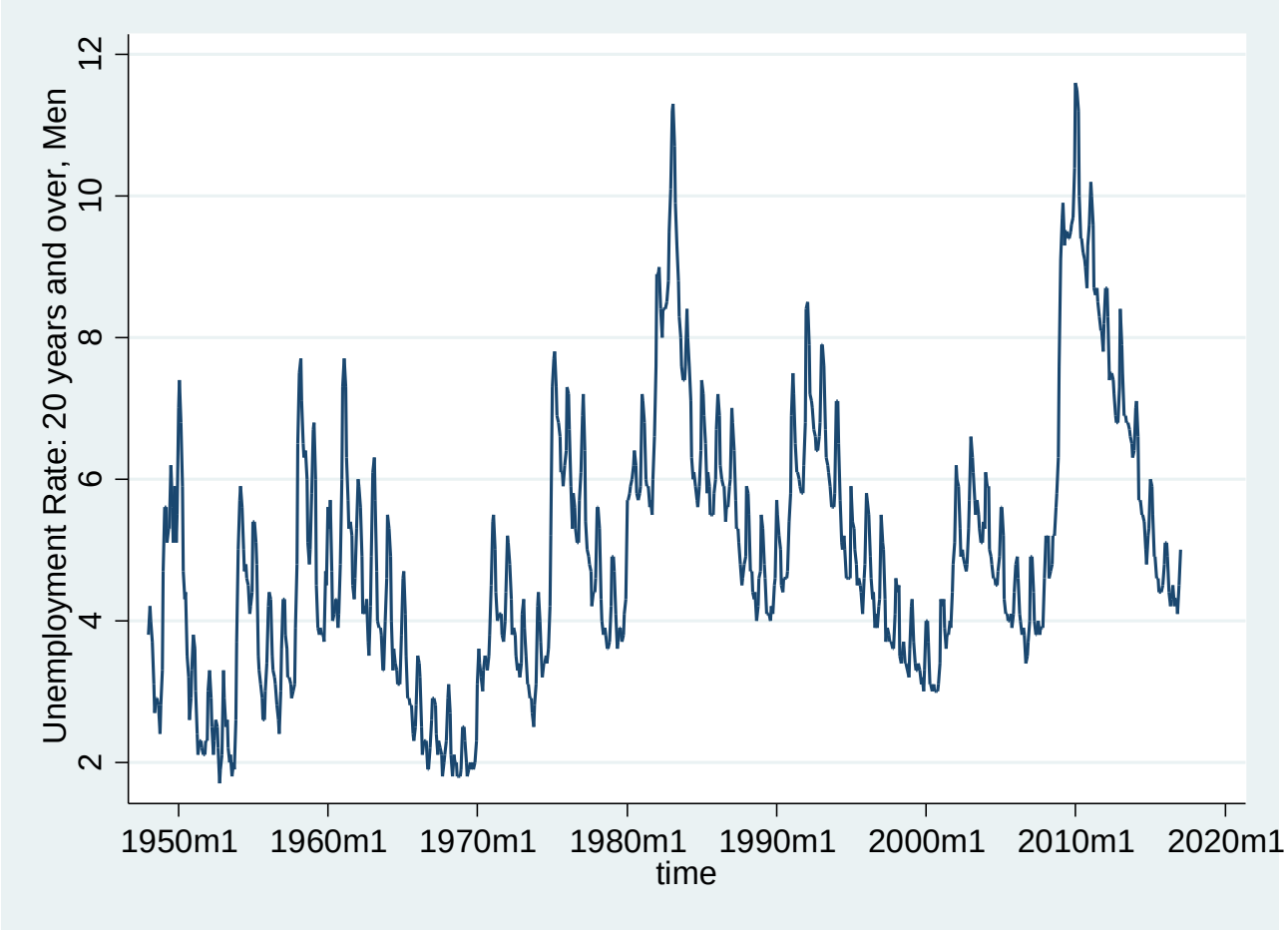
AR(p) Case

- If the cycle is an AR(p) we have

$$y_t = \sum_{i=1}^{s} \alpha_i D_{it} + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + e_t$$

- Estimate by least squares
- Linear Forecasting

Example: Unemployment Rate (Not Seasonally Adjusted - Men)



Regress on Dummies plus AR(12)

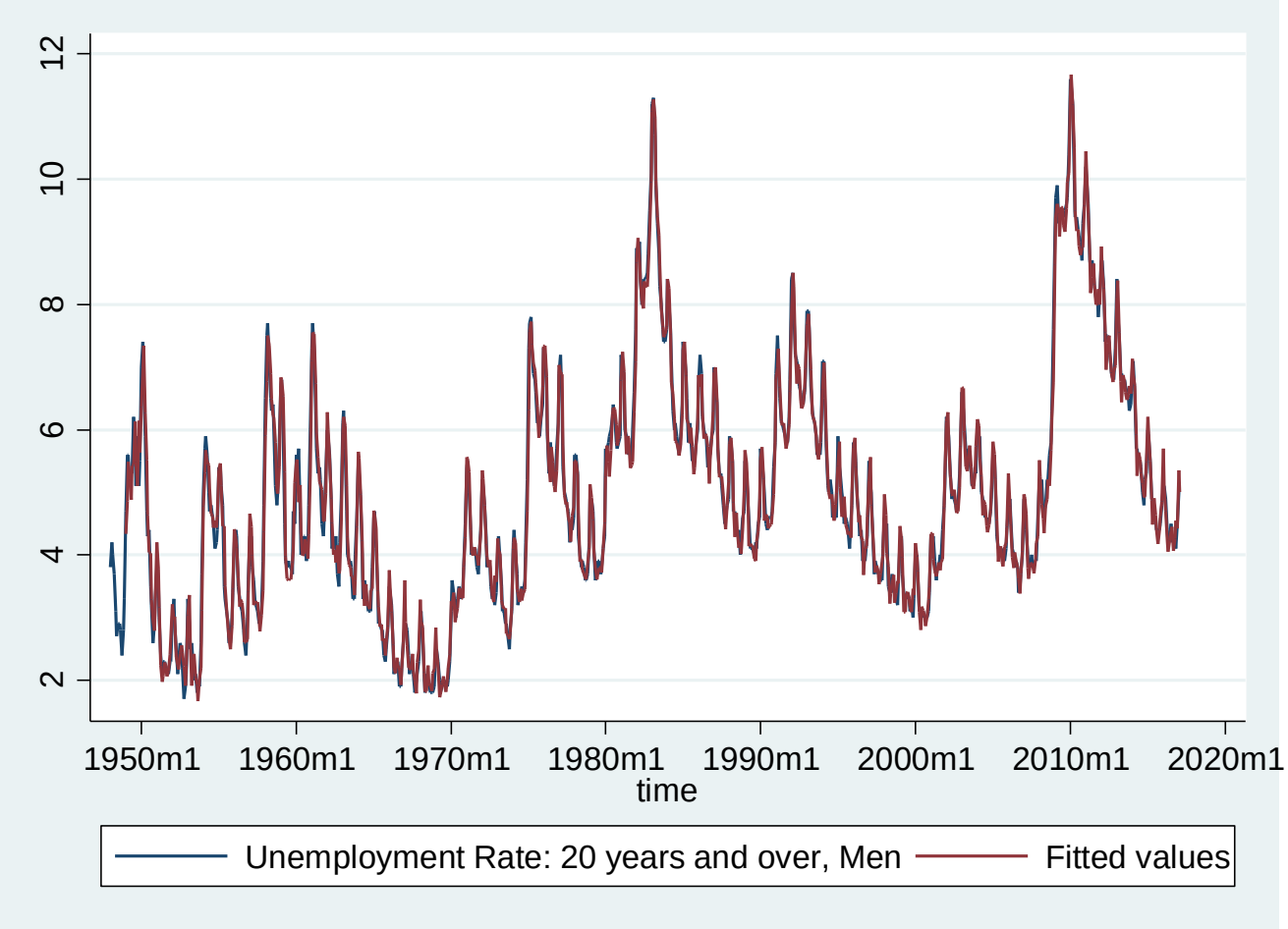
```
. reg ur L(1/12).ur b12.m
```

Source	SS	df	MS		
Model	2786.63037	23	121.157842	Number of obs	= 817
Residual	50.6920731	793	.06392443	F(23, 793)	= 1895.33
				Prob > F	= 0.0000
Total	2837.32245	816	3.47711084	R-squared	= 0.9821
				Adj R-squared	= 0.9816
				Root MSE	= .25283

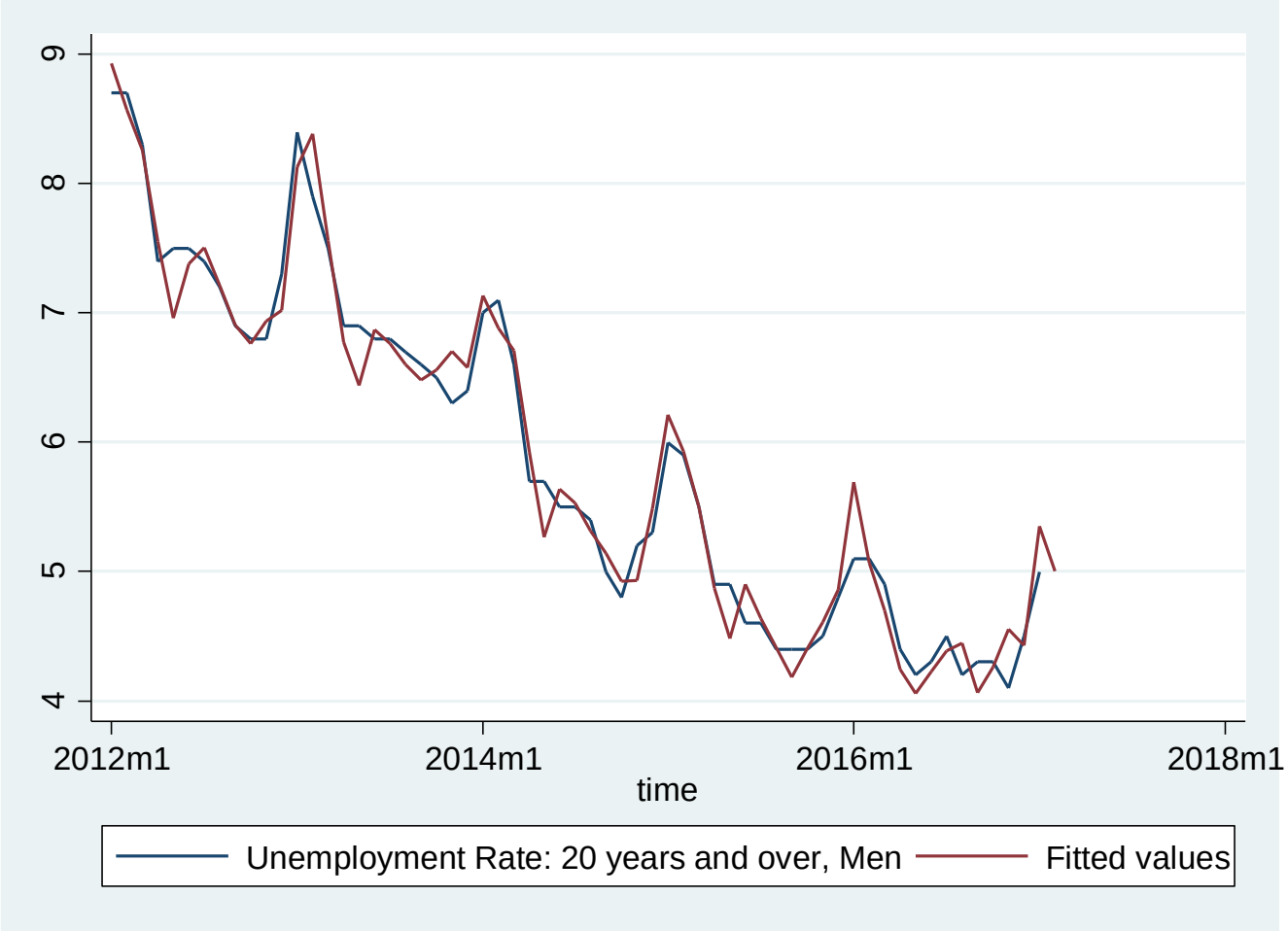
ur	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ur						
L1.	1.077861	.0353814	30.46	0.000	1.008409	1.147313
L2.	.0950265	.0521299	1.82	0.069	-.0073025	.1973555
L3.	-.0938645	.0522103	-1.80	0.073	-.1963512	.0086222
L4.	-.1055048	.0523029	-2.02	0.044	-.2081733	-.0028362
L5.	.0606082	.0524254	1.16	0.248	-.0423008	.1635172
L6.	-.0880589	.0525247	-1.68	0.094	-.1911629	.015045
L7.	.0059334	.0524042	0.11	0.910	-.0969339	.1088007
L8.	.0325417	.0523573	0.62	0.534	-.0702336	.135317
L9.	.0328493	.0521787	0.63	0.529	-.0695755	.135274
L10.	.0085939	.0520795	0.17	0.869	-.093636	.1108238
L11.	.0511347	.0519805	0.98	0.326	-.0509008	.1531703
L12.	-.0982369	.0352268	-2.79	0.005	-.1673856	-.0290881

m						
1	.6099347	.0596986	10.22	0.000	.4927487	.7271207
2	-.3491986	.0791699	-4.41	0.000	-.5046059	-.1937913
3	-.7071303	.0887213	-7.97	0.000	-.8812866	-.532974
4	-.9459756	.0828287	-11.42	0.000	-1.108565	-.7833861
5	-.5267795	.0776436	-6.78	0.000	-.6791909	-.3743681
6	-.1344322	.0750237	-1.79	0.074	-.2817008	.0128363
7	-.1012909	.0810665	-1.25	0.212	-.2604212	.0578394
8	-.3609652	.0892135	-4.05	0.000	-.5360877	-.1858427
9	-.4579354	.0920377	-4.98	0.000	-.6386019	-.277269
10	-.356465	.081047	-4.40	0.000	-.515557	-.197373
11	-.1038898	.0644509	-1.61	0.107	-.2304044	.0226248
_cons	.3914172	.0600225	6.52	0.000	.2735954	.509239

Fitted Values



Last few years



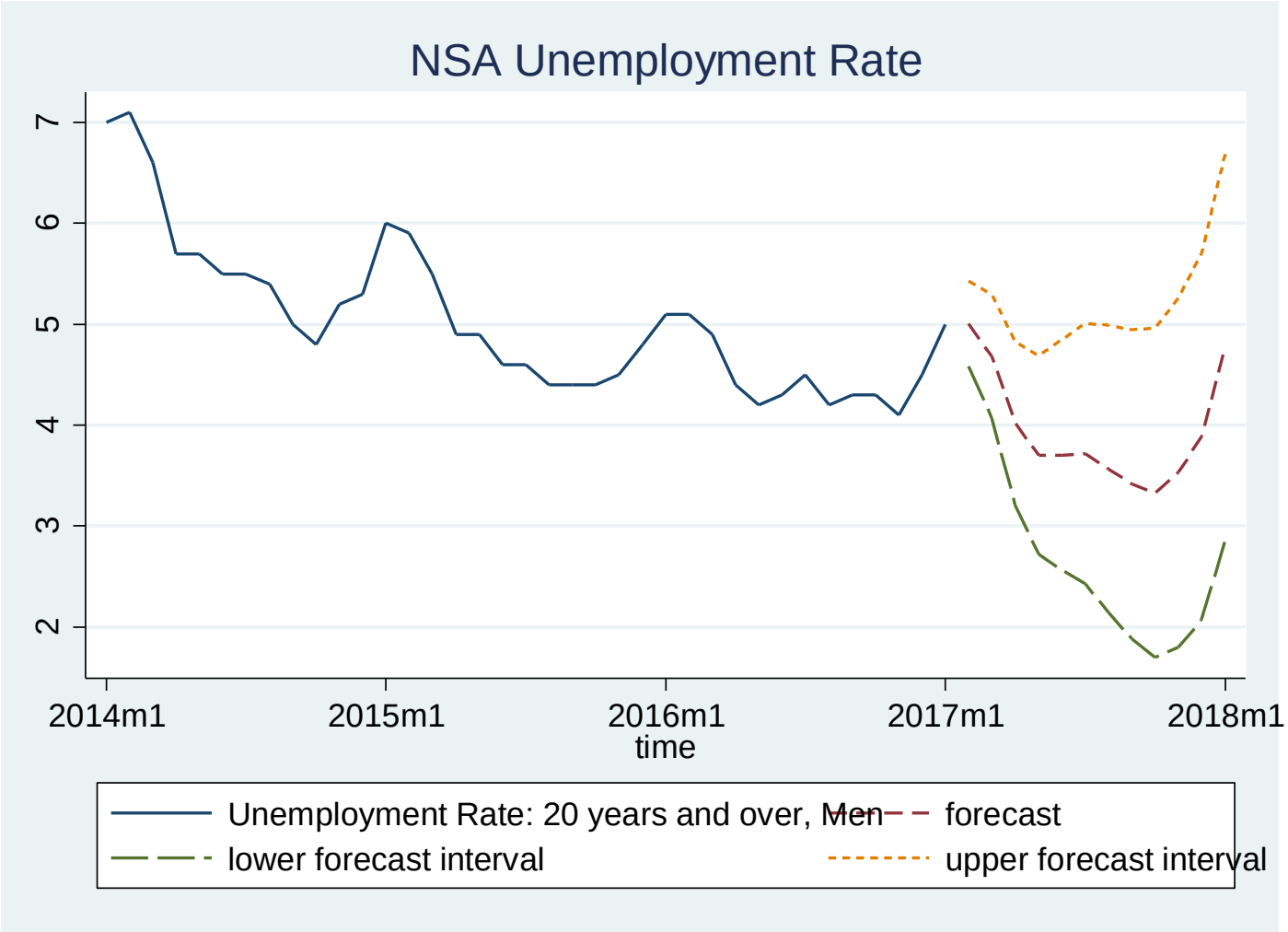
Forecasting with Seasonal Dummy

- To forecast in STATA with seasonal dummies, the dummy variables must be defined for the forecast period
- You must use the **tsappend** command *before* you create the month variable

  .tsappend, add(12)

  .gen m=month(dofm(t))
- If you do the reverse, *m* will have missing values for the forecast period

12-month Direct Forecast



Method

- For each forecast horizon (1 to 12)
  - Estimate regression with full seasonal dummies and 12 autoregressive lags
  - Generate point forecast and standard error of forecast
  - Calculate forecast intervals
- Plot point forecasts and intervals over forecast horizon

Trend+Seasonal+Cycle

- A full model is

$$y_t = T_t + S_t + C_t$$

$$T_t = \mu_1 + \mu_2 t$$

$$S_t = \sum_{i=1}^{s} \alpha_i D_{it}$$

$$C_t = \beta_1 C_{t-1} + \cdots + \beta_p C_{t-p} + e_t$$

Regression Model

- The implied regression model is

$$y_t = \sum_{i=1}^{s} \alpha_i D_{it} + \gamma t + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + e_t$$

- This can be estimated by least-squares
- It is a complete forecasting model

Example: Retail Sales

- U.S. Census Bureau
  - Monthly Retail Sales
  - Not Seasonally Adjusted and Seasonally Adjusted
  - Sales listed by variety of categories
  - 1992-2016
- Census.gov\topics\economy\economic indicators\Advance monthly sales\Monthly Retail Trade Report
  - Time Series

From Census Bureau Spreadsheet

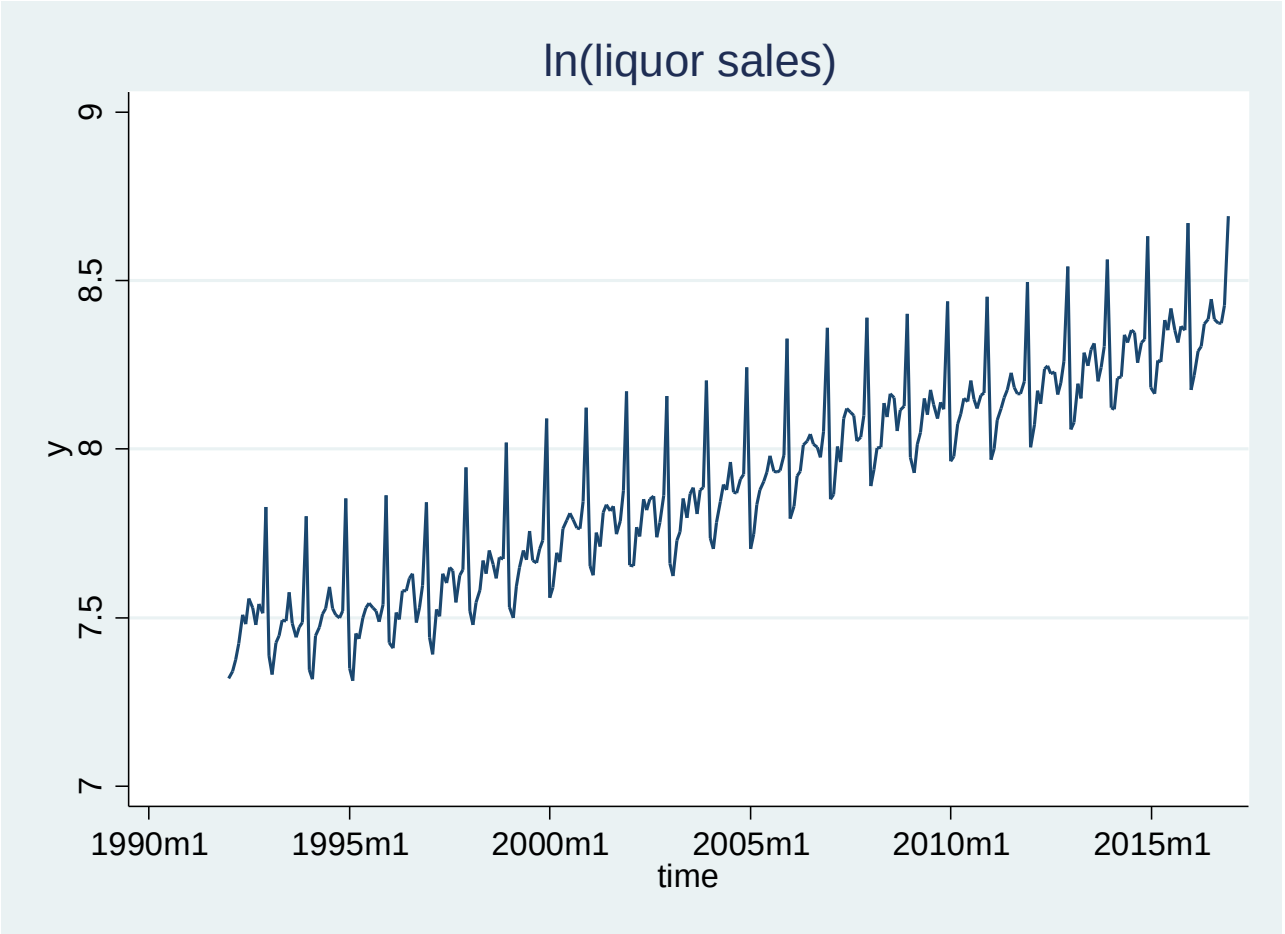
	Jan. 2016	Feb. 2016	Mar. 2016	Apr. 2016	May 2016	Jun. 2016
NOT ADJUSTED						
Retail and food services sales, total	400,928	413,554	460,093	450,730	469,435	464,102
Retail sales and food services excl motor vehicle and parts	321,727	325,292	361,020	355,012	372,473	367,568
Retail sales, total	350,742	361,878	404,012	394,134	412,106	409,191
Retail sales, total (excl. motor vehicle and parts dealers)	271,541	273,616	304,939	298,416	315,144	312,657
GAFO(1)	89,684	94,231	103,670	98,419	103,562	101,448
Motor vehicle and parts dealers	79,201	88,262	99,073	95,718	96,962	96,534
Automobile and other motor vehicle dealers	72,669	81,364	91,366	88,500	89,600	88,822
Automobile dealers	68,984	76,652	84,611	81,446	82,428	81,439
New car dealers	61,792	66,981	75,242	72,883	74,442	73,080
Used car dealers	7,192	9,671	9,369	8,563	7,986	8,359
Automotive parts, acc., and tire stores	6,532	6,898	7,707	7,218	7,362	7,712
Furniture, home furn, electronics, and appliance stores	15,749	16,159	17,159	15,697	16,545	16,835
Furniture and home furnishings stores	8,039	8,203	9,102	8,518	8,852	8,955
Furniture stores	4,464	4,638	5,013	4,611	4,814	4,848
Home furnishings stores	3,575	3,565	4,089	3,907	4,038	4,107
Floor covering stores(2)	1,308	1,342	1,585	1,531	1,552	1,669
All other home furnishings stores	2,140	2,097	2,344	2,207	2,347	2,291
Electronics and appliance stores	7,710	7,956	8,057	7,179	7,693	7,880
Appl.,TV, and other elect. stores	5,388	5,757	5,839	5,317	5,649	5,843
Household appliance stores	1,226	1,321	1,378	1,296	1,415	1,487
Radio, T.V., and other elect. stores	4,162	4,436	4,461	4,021	4,234	4,356
Computer and software stores	(S)	(S)	(S)	(S)	(S)	(S)

Liquor Sales

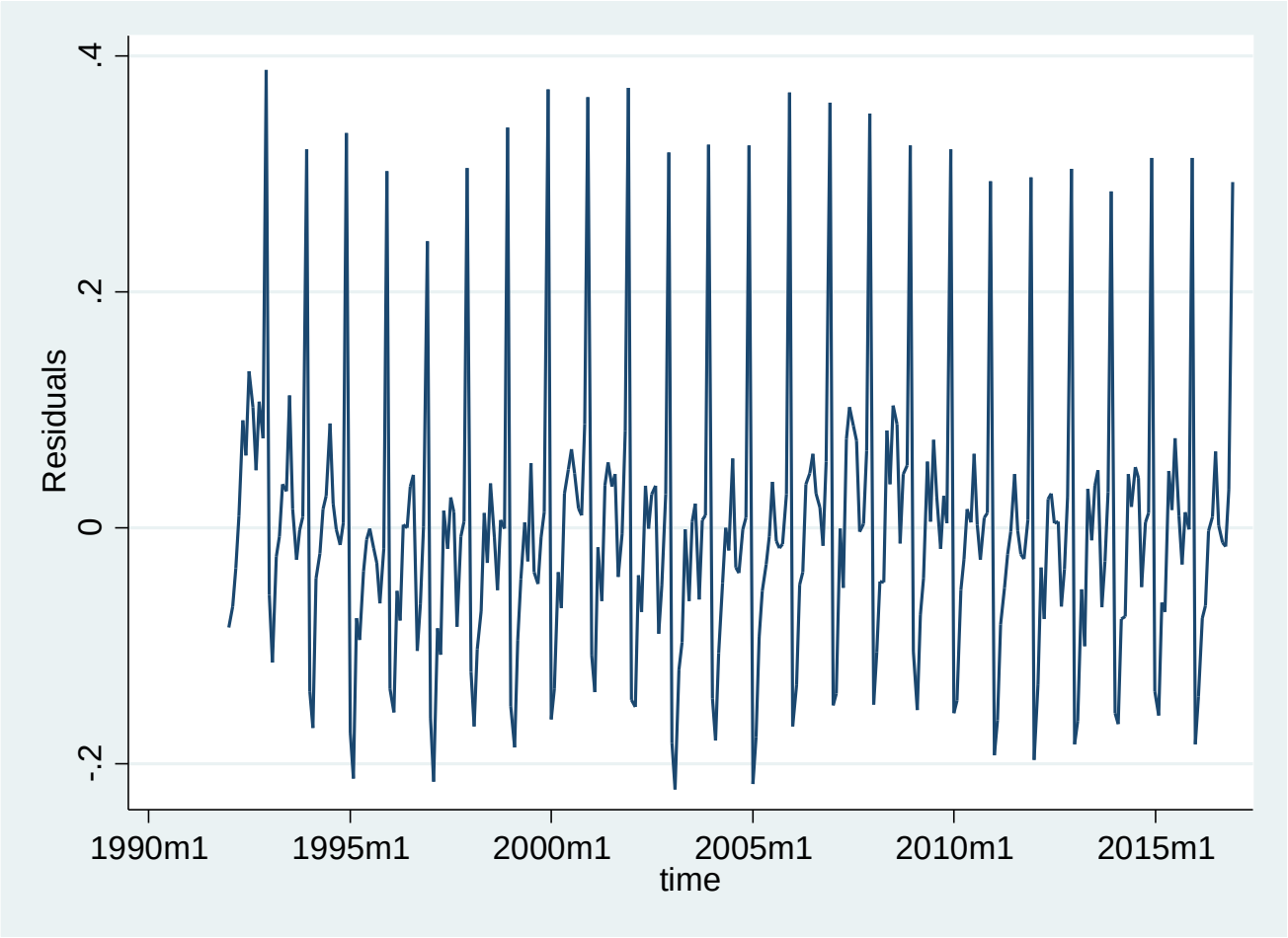
- Beer, Wine and Liquor Stores
- Millions of Dollars
- Sample: 1992-2016
- Not Seasonally Adjusted

Liquor Sales (Millions of $)

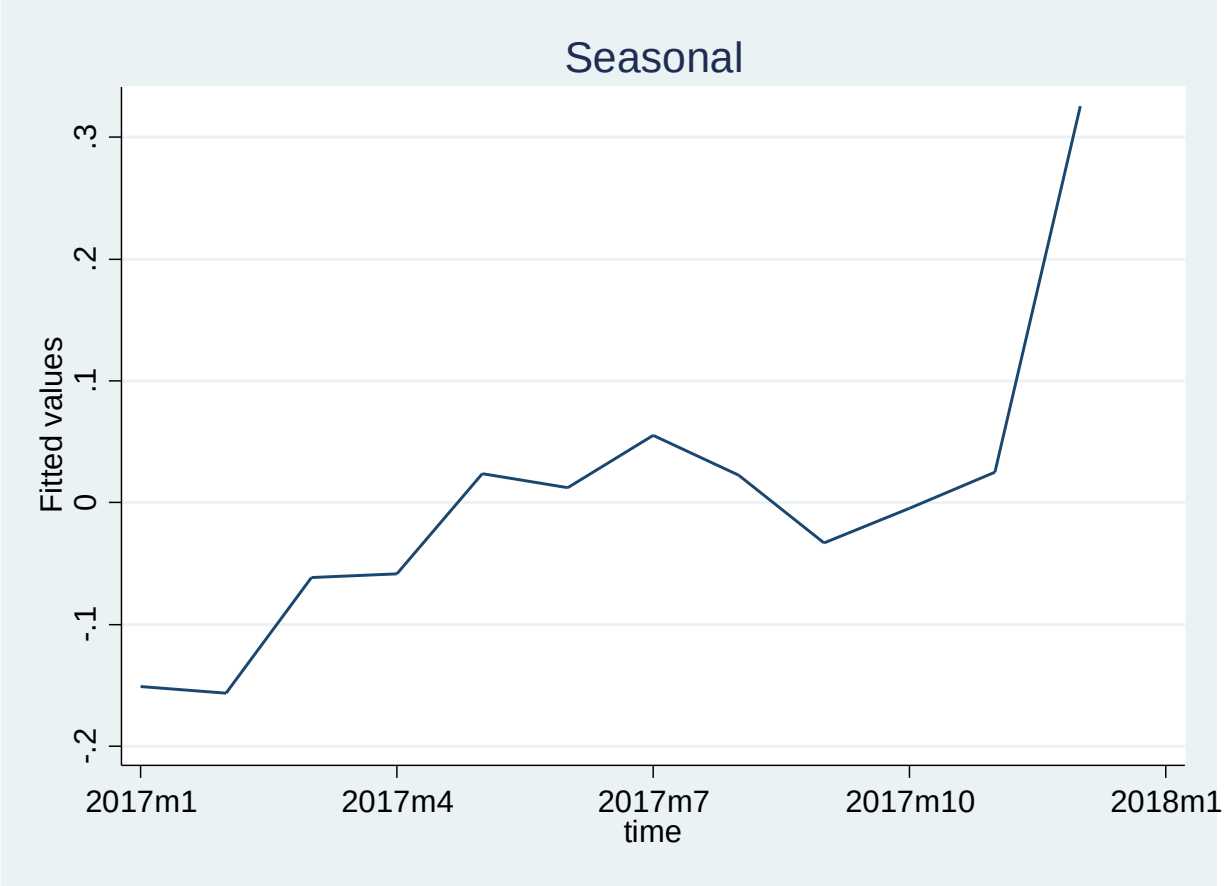
ln(Liquor Sales)



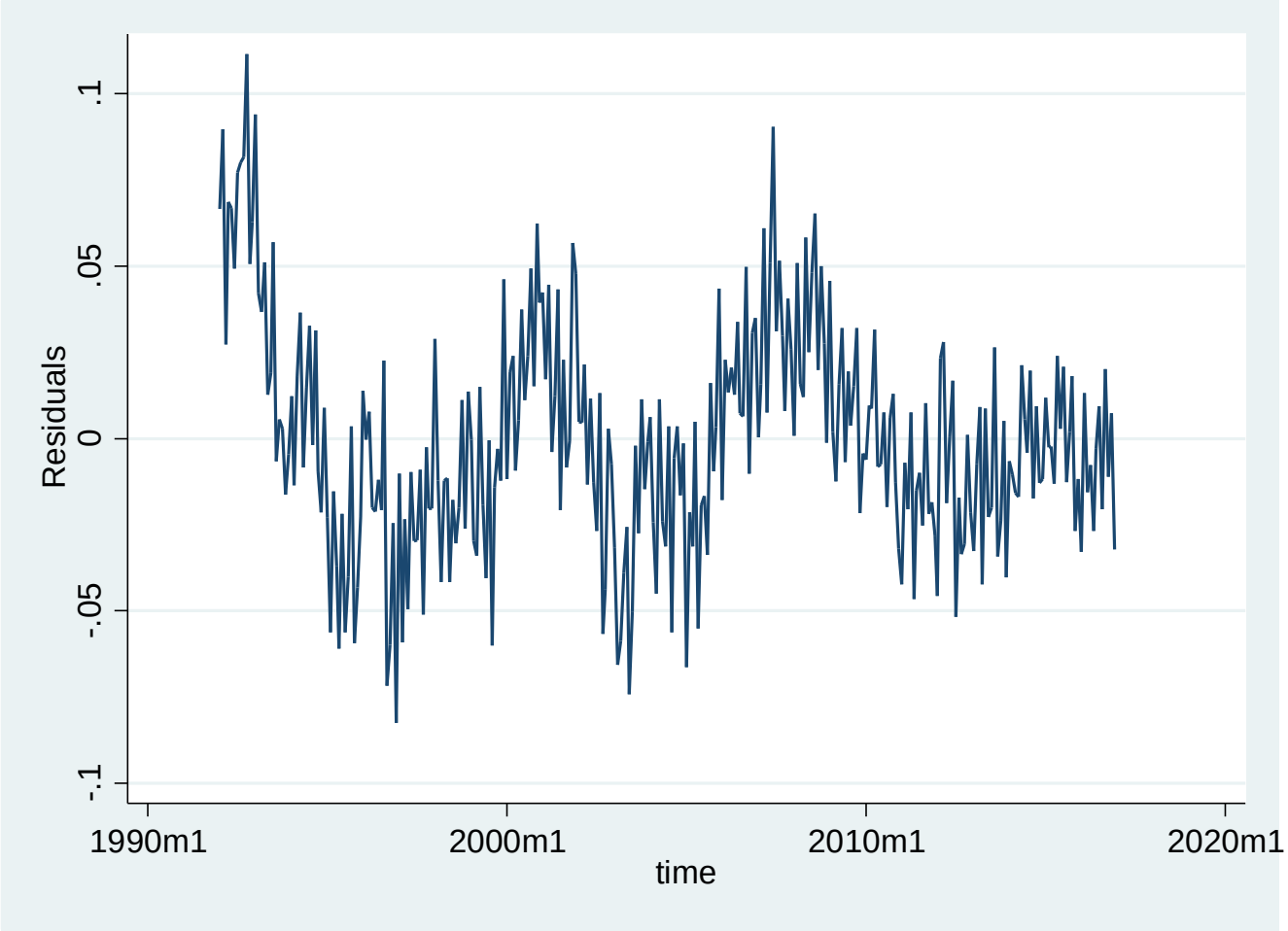
Residual from Linear Trend



Seasonal Dummy



Residuals after Seasonal Dummies



Full Estimation

```
. reg y t b12.m L(12/23).y
```

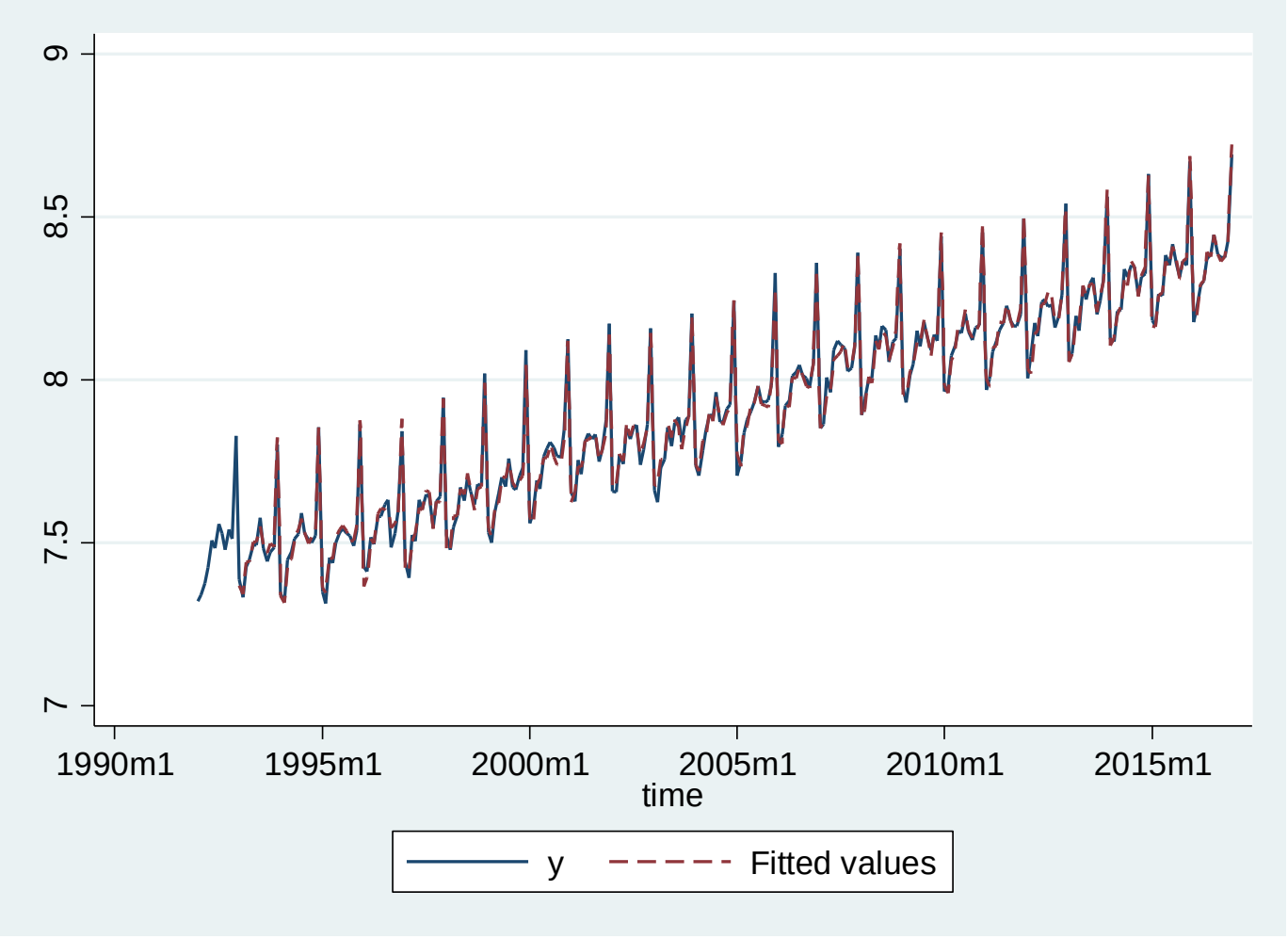
Source	SS	df	MS		
Model	24.4252667	24	1.01771945	Number of obs	= 277
Residual	.169893123	252	.000674179	F(24, 252)	= 1509.57
				Prob > F	= 0.0000
Total	24.5951598	276	.089112898	R-squared	= 0.9931
				Adj R-squared	= 0.9924
				Root MSE	= .02596

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
time	.0021533	.0002355	9.14	0.000	.0016896	.002617
m						
1	-.3026131	.0482741	-6.27	0.000	-.3976851	-.2075411
2	-.385364	.0538552	-7.16	0.000	-.4914276	-.2793004
3	-.1593709	.0511461	-3.12	0.002	-.2600992	-.0586426
4	-.2545325	.0412751	-6.17	0.000	-.3358206	-.1732444
5	-.2406469	.0400488	-6.01	0.000	-.3195199	-.1617738
6	-.2230789	.0426811	-5.23	0.000	-.307136	-.1390219
7	-.1311815	.0412751	-3.18	0.002	-.2124697	-.0498934
8	-.2025691	.0463738	-4.37	0.000	-.2938986	-.1112395
9	-.2114963	.0555825	-3.81	0.000	-.3209616	-.102031
10	-.1900004	.0509678	-3.73	0.000	-.2903776	-.0896233
11	-.2378448	.0483877	-4.92	0.000	-.3331406	-.1425489

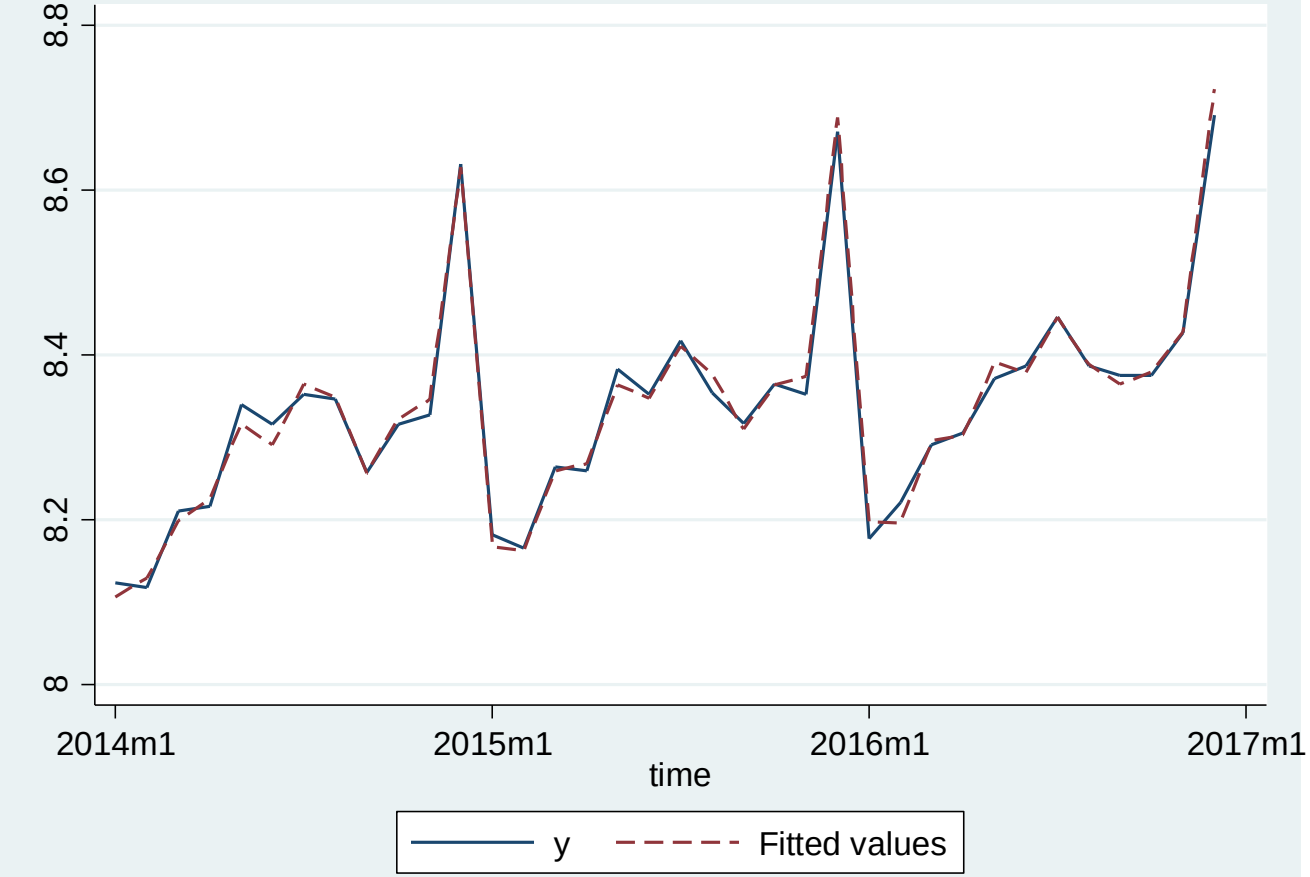
AR Coefficients

y						
L12.	.3288582	.0777133	4.23	0.000	.1758079	.4819085
L13.	-.0718919	.0788974	-0.91	0.363	-.2272743	.0834905
L14.	.212642	.0772127	2.75	0.006	.0605776	.3647064
L15.	-.1866536	.0791559	-2.36	0.019	-.342545	-.0307623
L16.	-.0012333	.078597	-0.02	0.987	-.156024	.1535573
L17.	.0860483	.0806494	1.07	0.287	-.0727844	.244881
L18.	.1064487	.0807929	1.32	0.189	-.0526667	.265564
L19.	-.1125005	.0787637	-1.43	0.154	-.2676195	.0426185
L20.	.0678152	.0796521	0.85	0.395	-.0890533	.2246838
L21.	-.0684206	.0772222	-0.89	0.376	-.2205038	.0836625
L22.	-.0810103	.0774492	-1.05	0.297	-.2335405	.07152
L23.	.0838816	.07556	1.11	0.268	-.0649279	.2326912
_cons	4.098763	.4333599	9.46	0.000	3.245294	4.952232

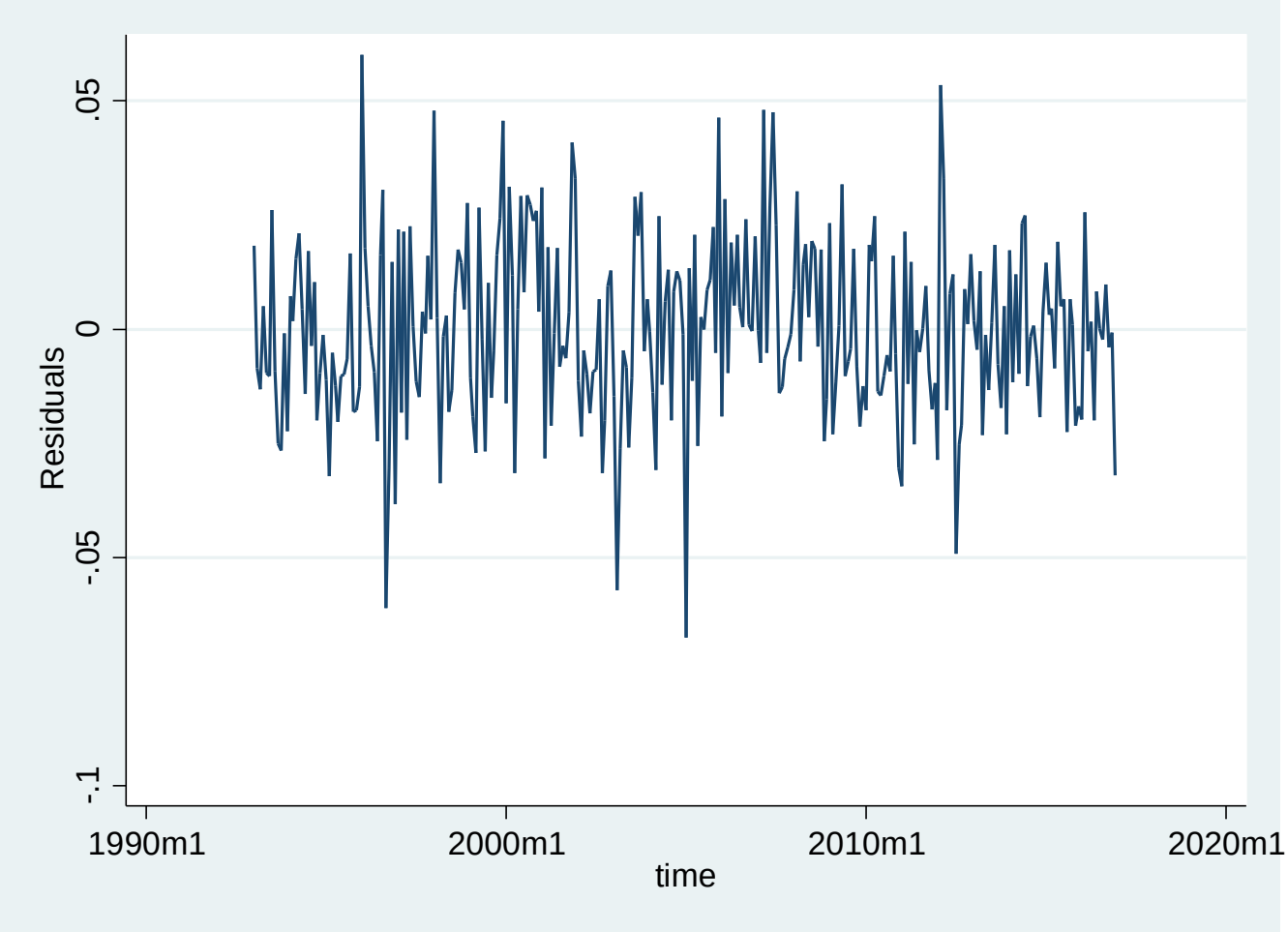
Fitted Values



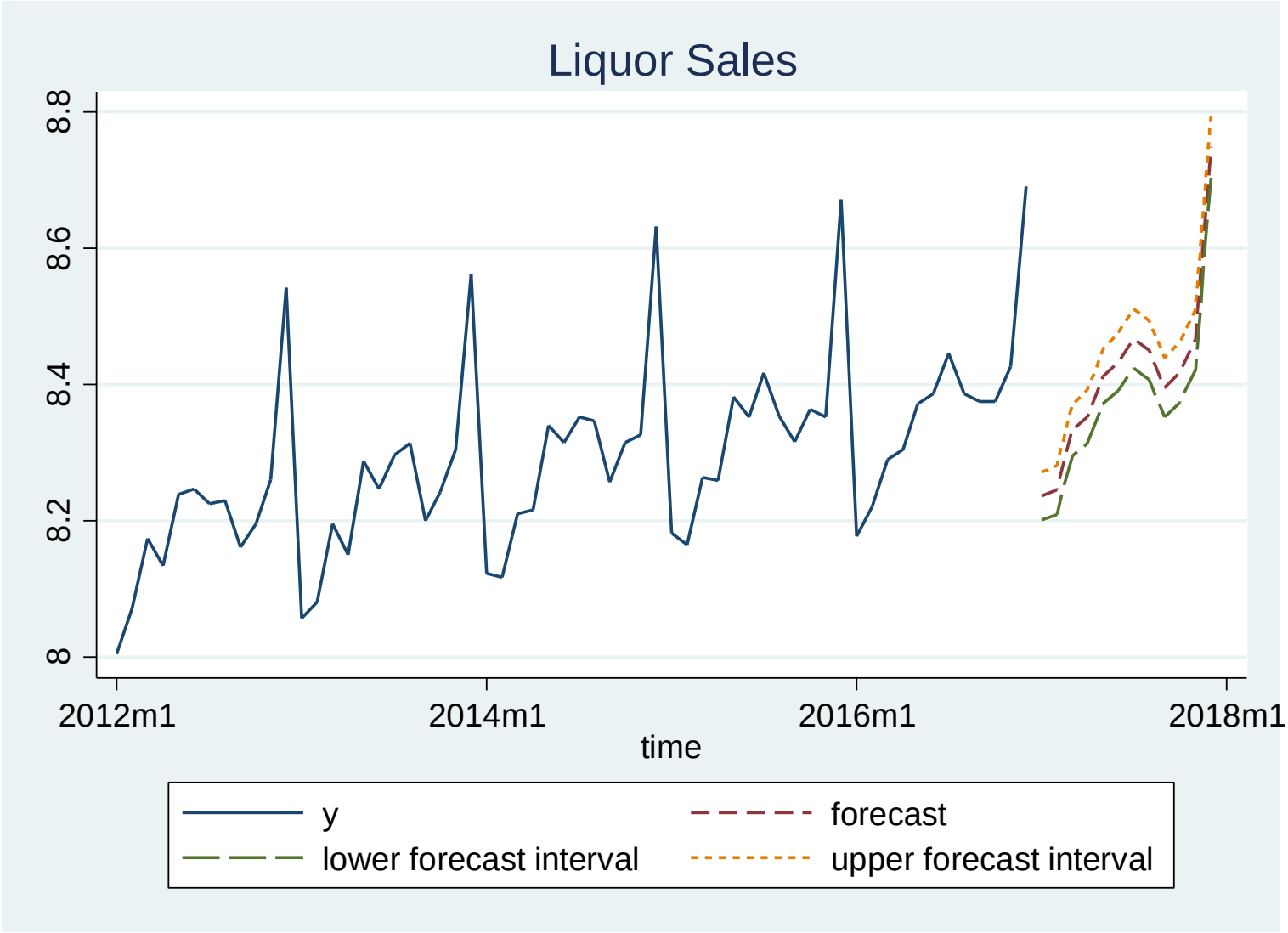
Last 3 years


8.8
8.6
8.4
8.2
8
2014m1
2015m1
2016m1
2017m1
time
y
Fitted values

Residuals



12-Month Forecast



12-month forecast

- The big decrease in the forecast for January 2017 is because of the seasonal dummy effect
- The 12-month forecast is constructed using 12 separate linear regressions, one for each horizon

Application and Evaluation

- In 2015, I was invited to talk at a economics conference in Greece, during the height of a debt restructuring crisis
- My talk described advanced forecast combination methods, using Greek GDP as an illustration
- I made quarterly 2-year extrapolative forecasts
- Greek GDP for 2016 was announced this week, let's see how I did so far

Model

- “Combination Forecasting for Greek GDP Using Multi-Step Cross-Validation”
  - June 12, 2015
- Data for 2000Q1-2015Q1
- Models with AR(1), AR(2), and up to two lags of seven regression predictors

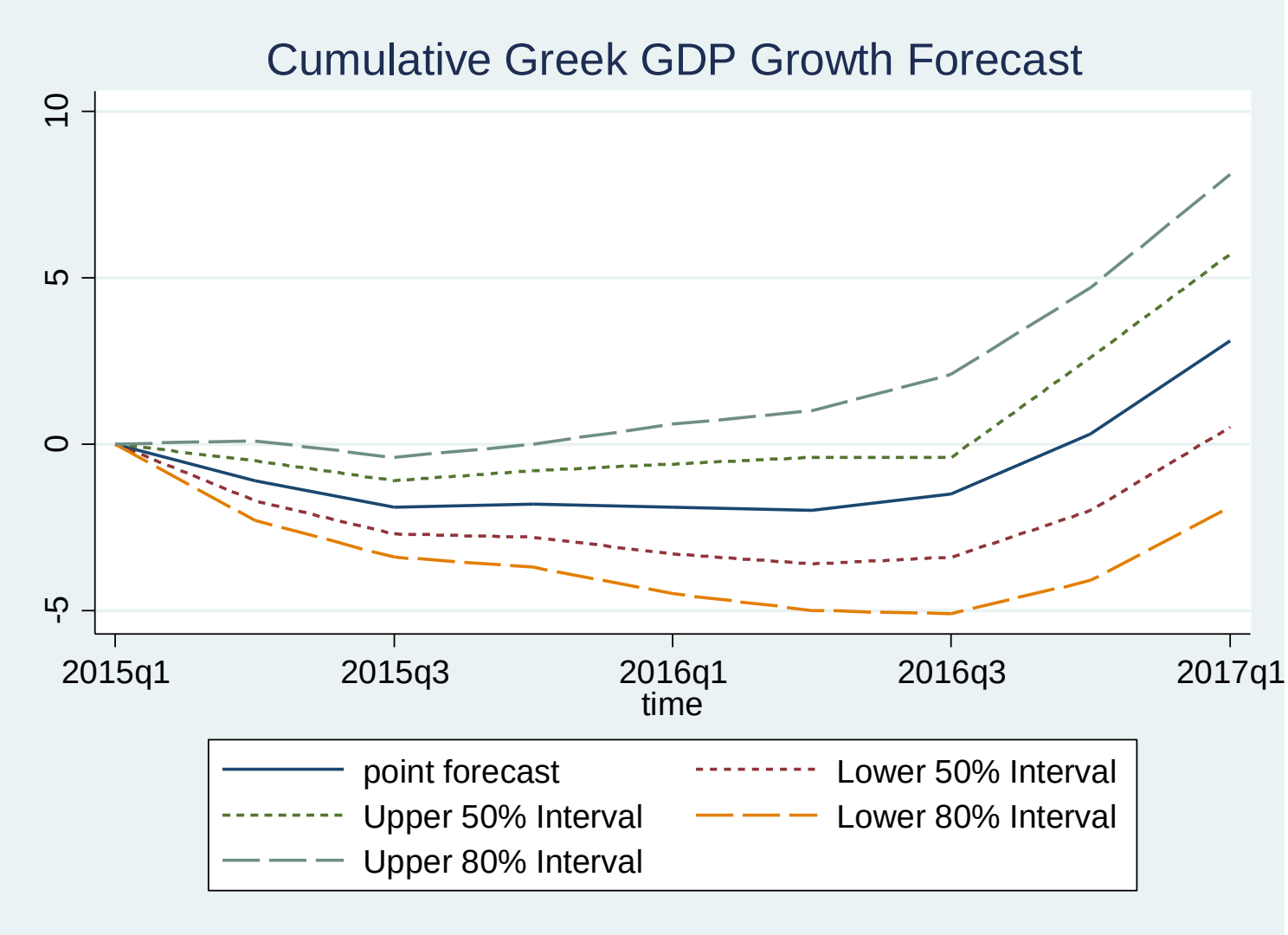
Predictors

- Building Permits
- Unemployment Rate
- Industrial Production Growth Rate
- Retail Sales Growth Rate
- Retail Turnover Growth Rate
- Athens Stock Index Return
- Economic Sentiment Indicator
- Interest Rate Spread of Greek over German Bonds

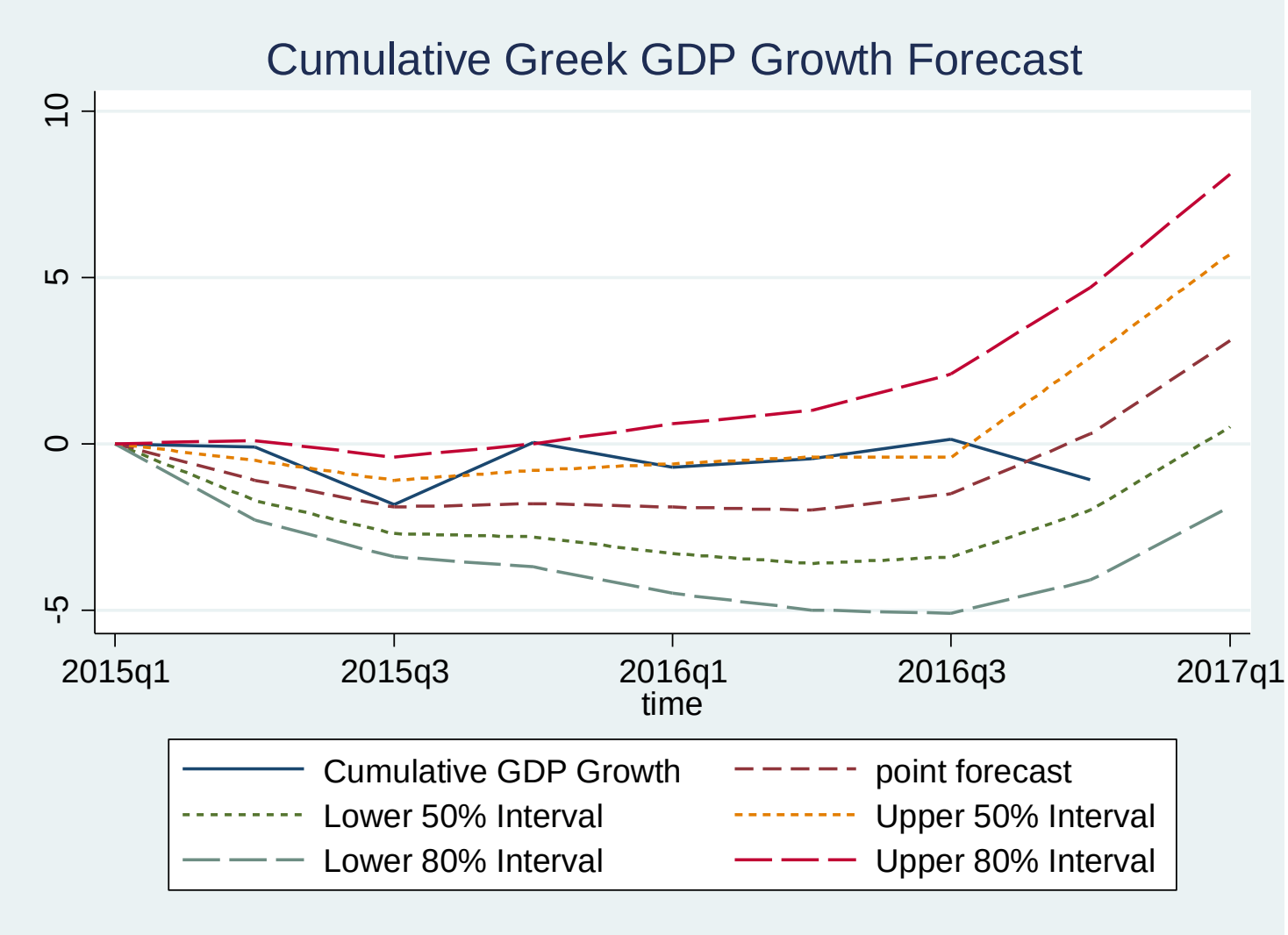
Details

- Averaged over 1701 sub-models
  - Much later in the semester!
  - Forecast Combination
- Direct Forecasting Method
- Forecast horizon h=1,…,8
- Frecasts
  - Point
  - 50% Interval
  - 80% Interval

Point and Interval Forecasts: June 2015



Forecasts and Actual



Comments

- Actual GDP growth performed somewhat better than point forecast at some points
- At end of 7 quarters, cumulative growth is slightly lower than the point forecast
- Out of 7 quarters, 6 realizations are within 80% forecast intervals
- Out of 7 quarters, 4 realizations are within 50% forecast intervals

Assignments

- Read Diebold Chapter 9
- Read Wooldridge Chapter 12.1 and 12.5
  - An electronic copy is in files at Learn@UW
- Problem Set # 7
  - Due Tuesday (3/14)
- Read Chapter 7 from *The Signal and the Noise*
  - Reading Reflection
  - Due Thursday (3/16)
- Forecasting Project
  - Project Description (3/28)