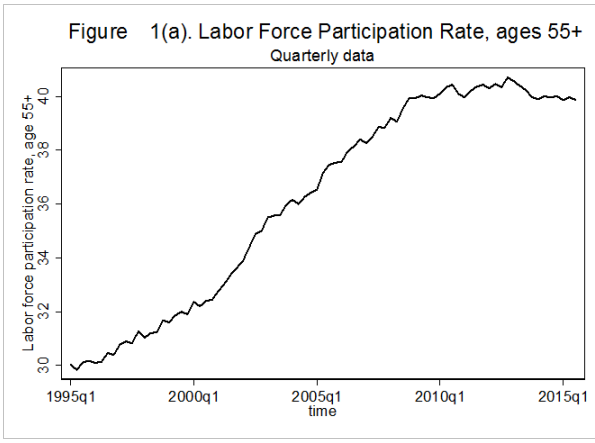
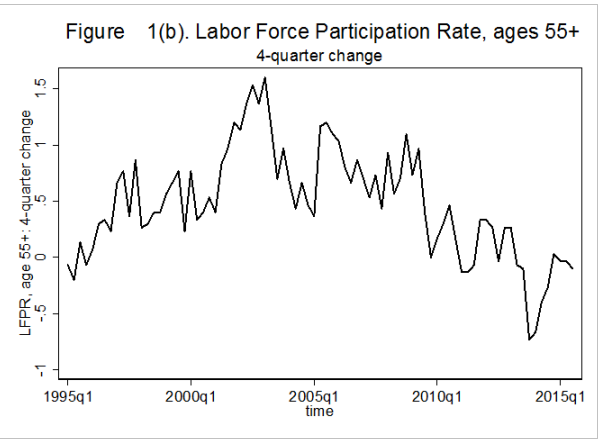
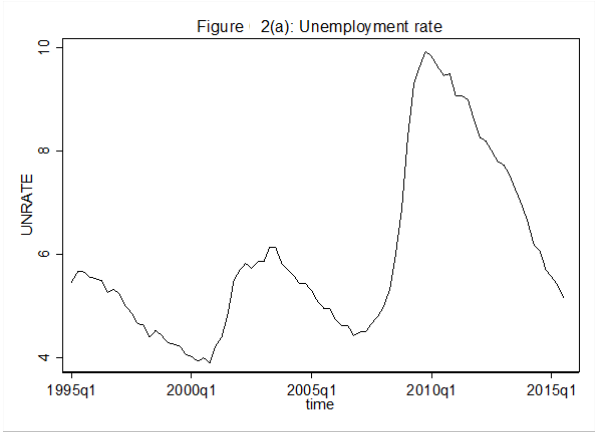
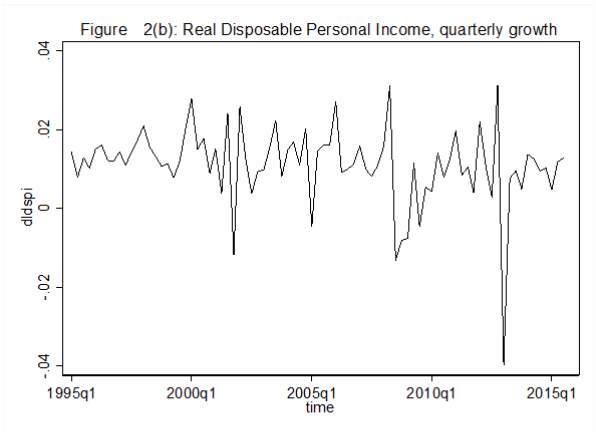
B2. The labour force participation rate (LFPR) is the fraction of workers who are working or are looking for work. From the mid-1990s through approximately 2009, the LFPR of Americans ages 55 and over ("*LFPR55*") increased from roughly 30% to more than 40%. As can be seen in Figure 1(a), since 2010 this fraction has plateaued. Figure 1(b) plots the four-quarter change in the LFPR ages 55+, that is, $y_t - y_{t-4}$, where y_t is LFPR55 quarter t. The plateau post-2009 in Figure 1(a) corresponds to a sharp drop in the post-2009 mean four-quarter change in *LFPR55* in Figure 1(b).

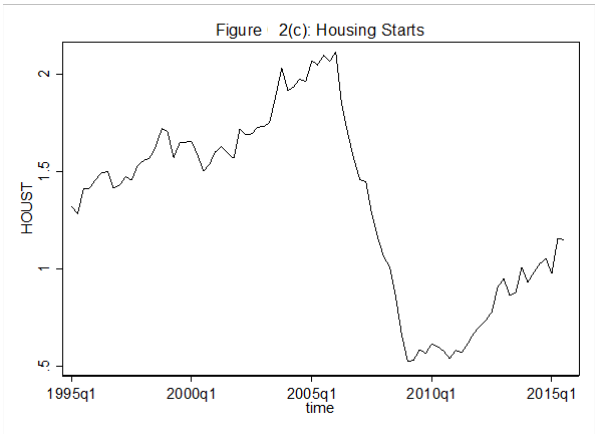
Could this slowdown be the result of the recession of 2008-2009? This question examines this hypothesis using three measures of overall macroeconomic activity plotted in Figure 2: the unemployment rate, real disposable (after-tax) personal income, and housing starts.











a) Regressions (1) and (2) differ only in the way the standard errors were computed. Which method is preferred for this application, or does it not matter, at least in theory? Briefly explain.

b) Suppose you were asked to forecast the change in the *LFPR55+* over the four quarters from 2015Q3 to 2016Q3, and that you are restricted to using one of regressions (2) – (5).
   i. Which specification would you choose, and why?
   ii. Provide the standard error of your forecast (you do not need to provide the value of the forecast).

c) Consider model (2) and the information provided below:

Period	Value of $LFPR55_t$
2014q4	40.00
2014q3	39.93
2014q2	39.97
2014q1	39.93

   i. Provide a forecast for the level of the labour force participation rate for 2015q4.
   ii. Construct a 95% forecast interval for the level of the labour force participation rate

d) Suppose you wanted to test whether there was a structural break in model (2) in 2008q1. Write down a regression model you could use to test the null hypothesis. What statistic would you use to test for a structural break? Be specific about the null hypothesis.

e) Consider the following hypothesis:

> One possible reason for the plateau in the LFPR55+ post-2010 is the recession and slow recovery. For example, absent the recession, older workers might have continued the trend of postponing retirement and would have stayed in the labor force, but given the widespread job loss and difficulty in getting a job during the recession, many older workers simply retired earlier than the 1995-2010 trend would have suggested.

Using Figures 1-3 and Table 2, write a paragraph assessing the evidence on this hypothesis. In your view, does the evidence support or contradict this hypothesis, or is the evidence inadequate to draw a conclusion? Explain.

f) The FRED-MD database is a monthly database consisting of 127 macroeconomic timeseries, with 767 observations for each. Suppose you had this database available to forecast labour force participation.
   i. Describe three methods you could use to incorporate all 127 series into a prediction model.
   ii. Describe how you could compare the prediction performance of models based on each of these methods. Be specific.

Table 1. Variable Definitions
Unit of Observation: U.S., Quarterly, 1995Q1 – 2015Q3 (T = 83)

Variable name	Definition	Mean	Std. Dev.
LFPR55	Labor force participation rate, U.S., age 55 and over (the fraction of Americans age 55+ who are either working or looking for work)	36.2	3.8
Δ*LFPR55*	$LFPR55_t - LFPR55_{t-1}$	0.12	0.22
Δ4 *LFPR55*	$LFPR55_t - LFPR55_{t-4}$ (4-quarter change in *LFPR55*)	0.47	0.48
Unemployment rate	Unemployed as a percentage of the labor force	6.0	1.7
Real Income	After-tax personal income (billions 2010 dollars)	9,989	1,616
Δln(*Real Income*)	$\ln(Real\ Income_t) - \ln(Real\ Income_{t-1})$	0.0069	0.0099
Housing Starts	New home construction started in the quarter (millions of units).	1.334	0.471
Δ*Housing Starts*	$Housing\ Starts_t - Housing\ Starts_{t-1}$	-0.0038	0.084

***Note*:** Δ denotes the first difference, so $\Delta LFPR55 = LFPR55_t - LFPR55_{t-1}$ and $\Delta\ln(Real\ Income) = \ln(Real\ Income_t) - \ln(Real\ Income_{t-1})$.

Figure 3

Four-quarter change in LFPR 55+ (solid) and predicted value (dashed) using regressions (2) – (5). All regressions are estimated 1995q1-2007q4, so dashed values post-2007 are pseudo out-of-sample

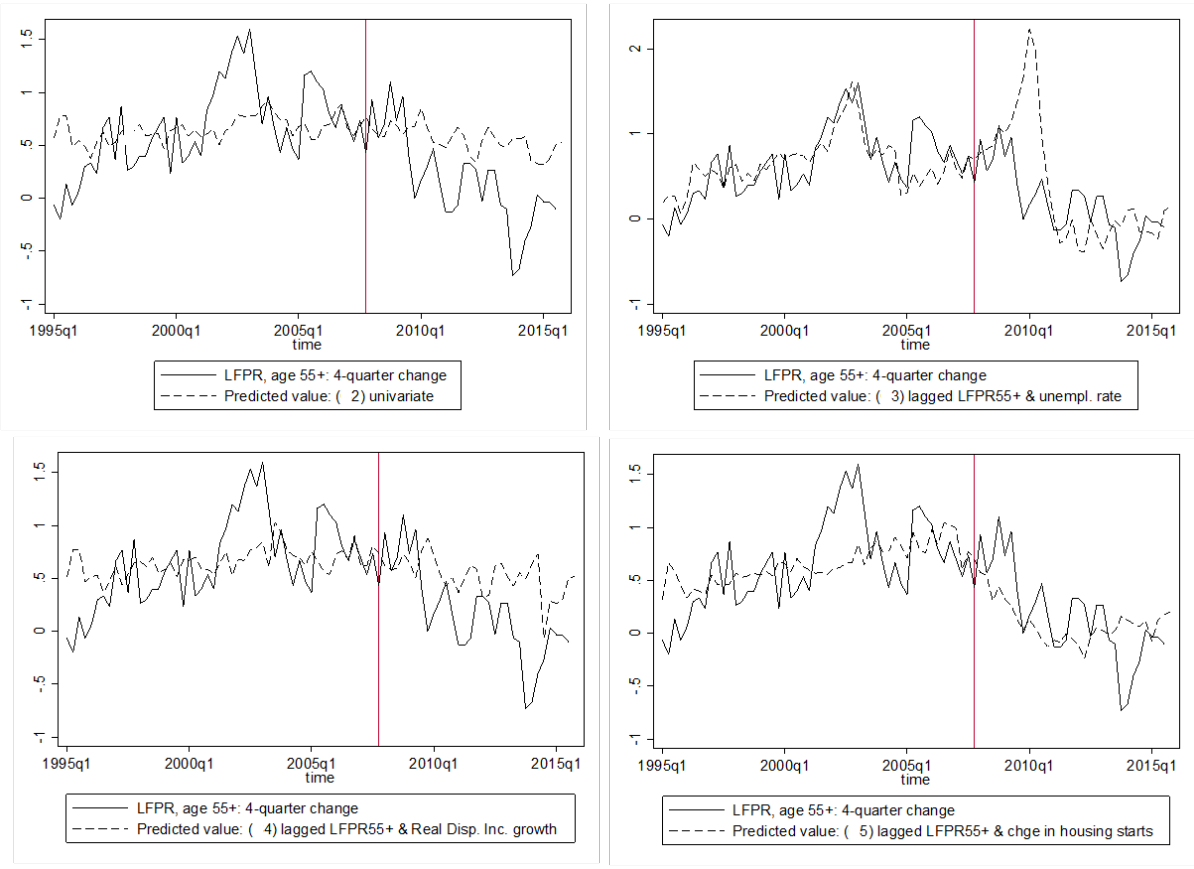


Table 2. OLS Forecasting regressions for the four-quarter change in the LFPR ages 55+

Dependent variable: $\Delta 4\_LFPR55_t = LFPR55_t - LFPR55_{t-4}$

All regressions are estimated using quarterly data, 1995q1-2007q4

	(1)	(2)	(3)	(4)	(5)
Standard errors & tests	HR	HAC	HAC	HAC	HAC
Regressors:					
$\Delta LFPR55_{t-4}$	0.138 (0.245)	0.138 (0.211)	-0.059 (0.149)	0.169 (0.244)	0.024 (0.313)
$\Delta LFPR55_{t-5}$	0.380 (0.289)	0.380 (0.234)	0.216 (0.148)	0.409 (0.245)	0.212 (0.367)
$\Delta LFPR55_{t-6}$	0.430^{+} (0.255)	0.430* (0.166)	0.269* (0.135)	0.458** (0.154)	0.270 (0.224)
$Unemployment\ rate_{t-4}$			0.756** (0.217)		
$Unemployment\ rate_{t-5}$			-0.236 (0.332)		
$Unemployment\ rate_{t-6}$			-0.675** (0.215)		
$\Delta\ln(Real\ Income_{t-4})$				-3.08 (3.16)	
$\Delta\ln(Real\ Income_{t-5})$				-0.205 (6.79)	
$\Delta\ln(Real\ Income_{t-6})$				7.65 (9.64)	
$\Delta Housing\ Starts_{t-4}$					1.01** (0.30)
$\Delta Housing\ Starts_{t-5}$					-0.89 (0.61)
$\Delta Housing\ Starts_{t-6}$					0.51 (0.38)
Constant	0.497** (0.088)	0.497** (0.112)	1.447** (0.412)	0.444 (0.291)	-0.460 (0.712)
BIC	-1.641	-1.641	-2.039	-1.433	-1.537
Adjusted R^2	0.021	0.021	0.441	0.002	0.077
RMSFE, 2008q1-2015q3	0.558	0.558	0.685	0.554	0.381
Total no. observations	52	52	52	52	52
***F*-test (*p*-value) testing coefficients on:**					
$\Delta LFPR55_{t-4}$, …, $\Delta LFPR55_{t-6}$	1.51 (0.224)	4.55 (0.007)	3.33 (0.028)	6.40 (0.001)	2.33 (0.087)
$Unemployment\ rate_{t-4}$,…, $Unemployment\ rate_{t-6}$			11.76 (0.000)		
$\Delta\ln(Real\ Income_{t-4})$, …, $\Delta\ln(Real\ Income_{t-6})$				1.20 (0.320)	
$\Delta Housing\ Starts_{t-4}$, …, $\Delta Housing\ Starts_{t-6}$					4.21 (0.010)

Notes: Standard errors are (in parentheses below coefficients) are heteroskedasticity-robust (HR) or heteroskedasticity- and autocorrelation-consistent (HAC). The RMSFE is the root mean squared forecast error, computed over the period 2008q1-2015q3 (note that the units of the RMSFE are the same as for the dependent variable). Coefficients are significant at the $^{+}$10%, *5%, **1% significance level.