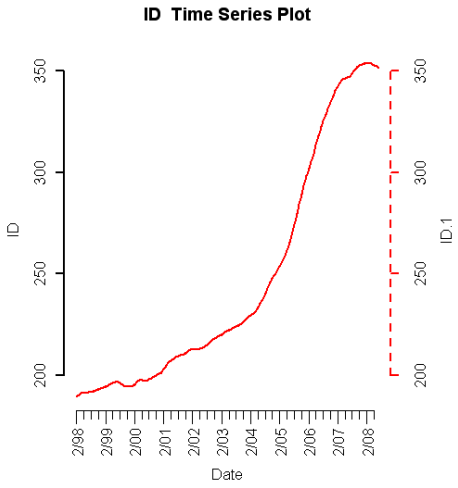
I. PHASE I AND PHASE II RESULTS

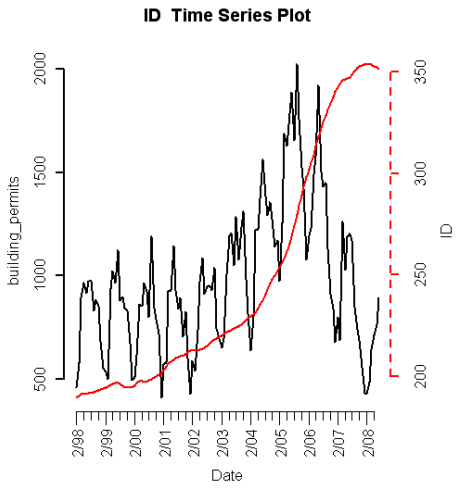
IDAHO (ID)

A. Plots of Predictors vs. HPI (non-simple return)

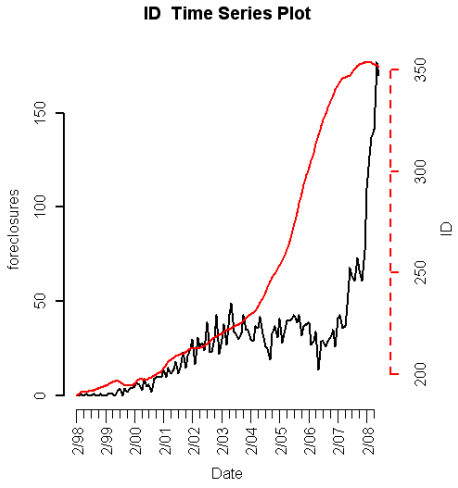
Original HPI Series



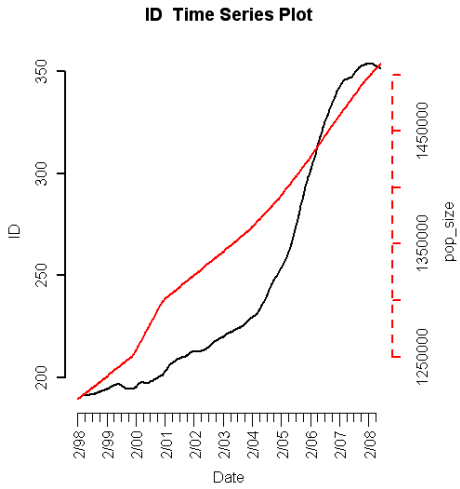
Building Permits



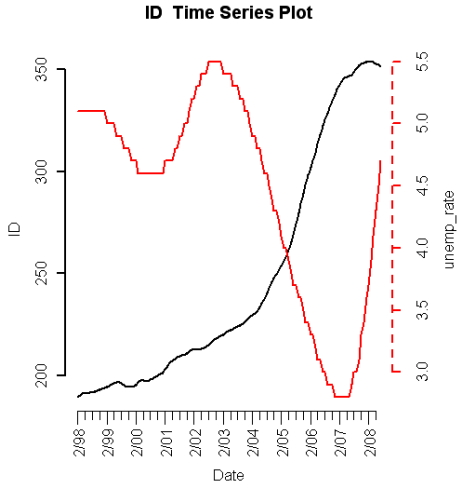
Foreclosures



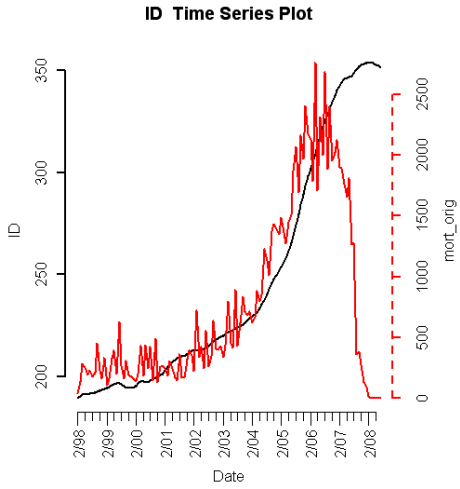
Population Size



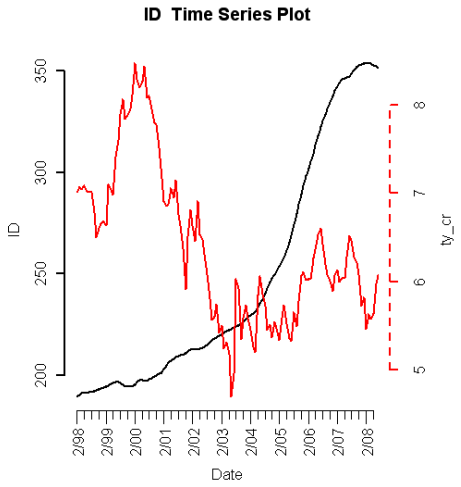
Unemployment Rate



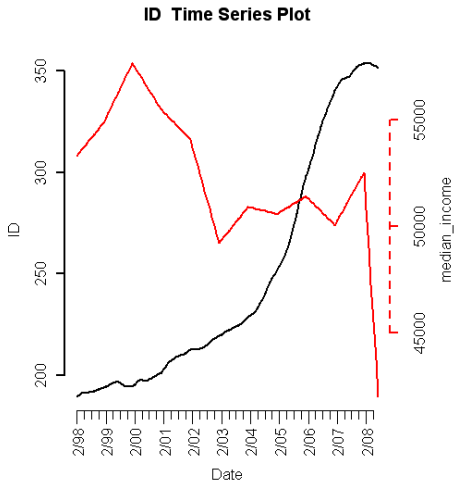
Mortgage Originations



30-Year Commitment Rate

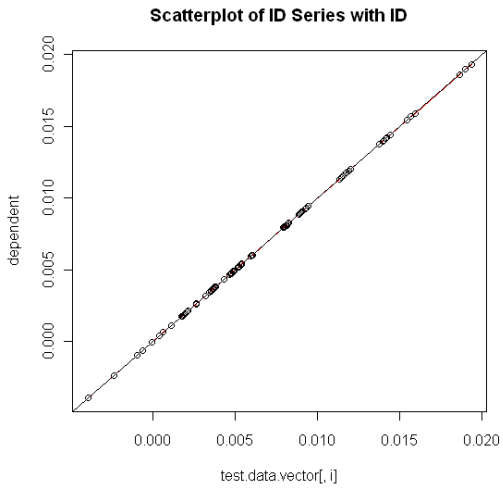


Median Income

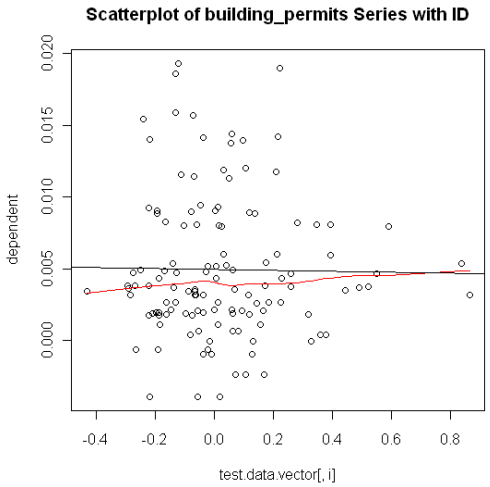


B. Scatterplots (simple return)

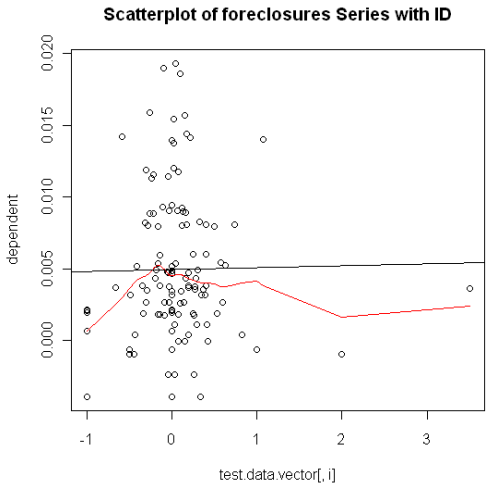
HPI vs. HPI



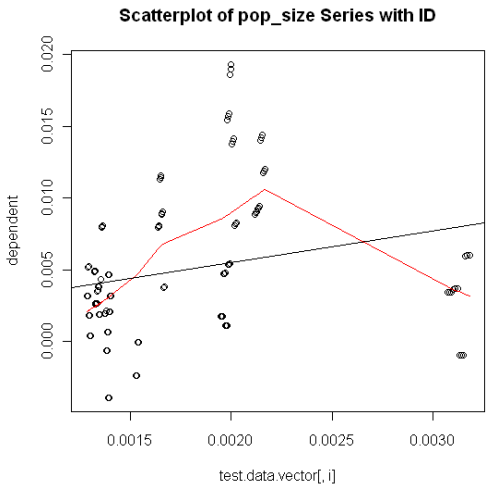
HPI vs. Building Permits



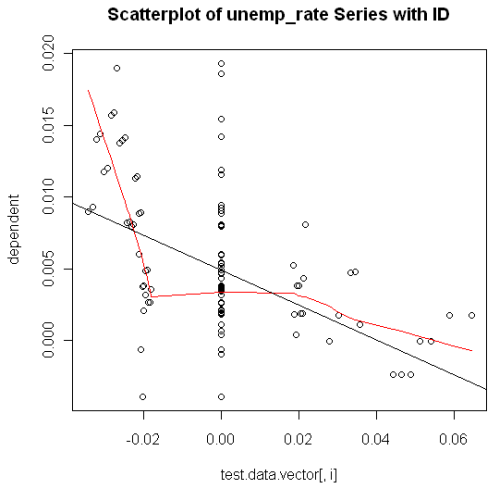
HPI vs. Foreclosures



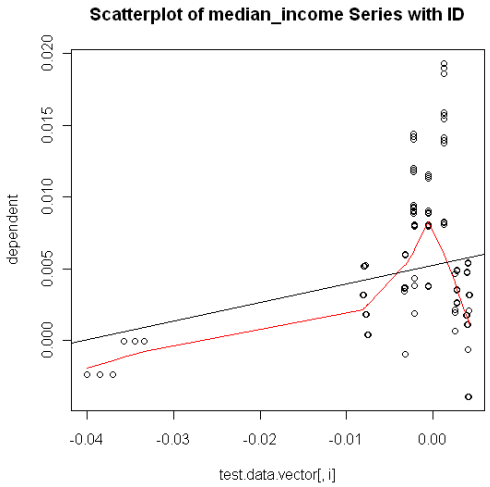
HPI vs. Population Size



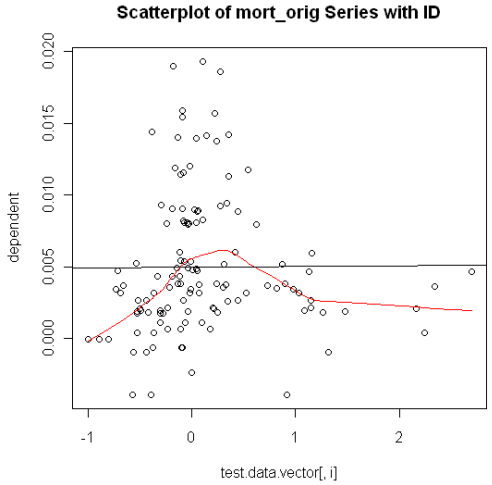
HPI vs. Unemployment Rate



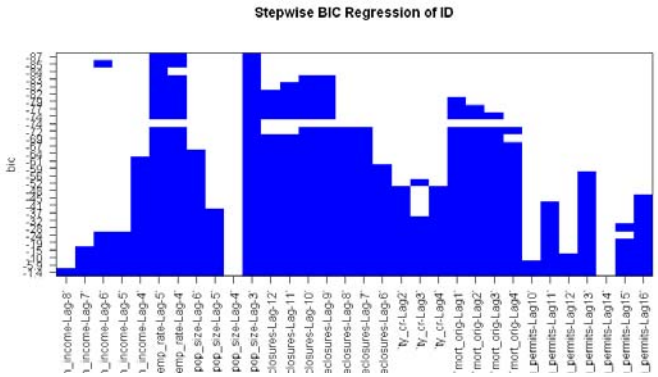
HPI vs. Median Income



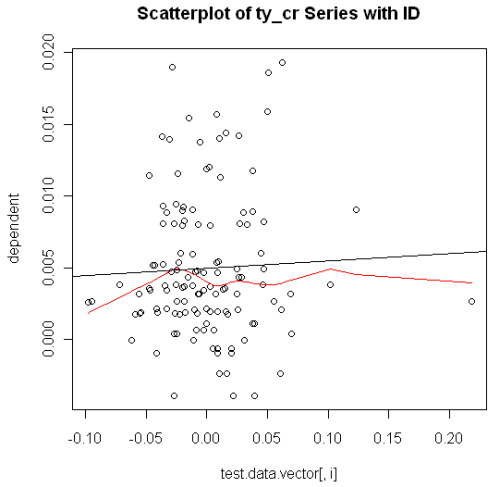
HPI vs. Mortgage Originations



C. Exhaustive Stepwise Regression



HPI vs. 30-Year Commitment Rate



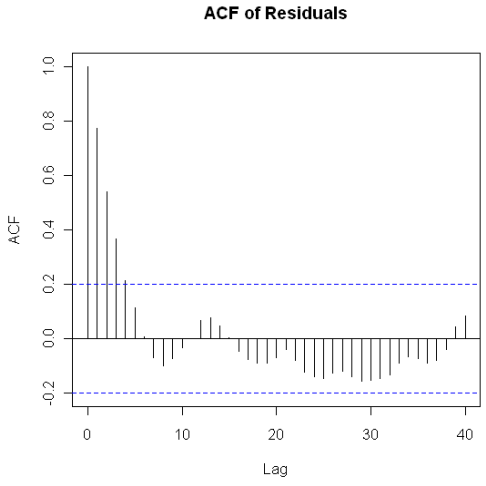
D. Final RLM Model

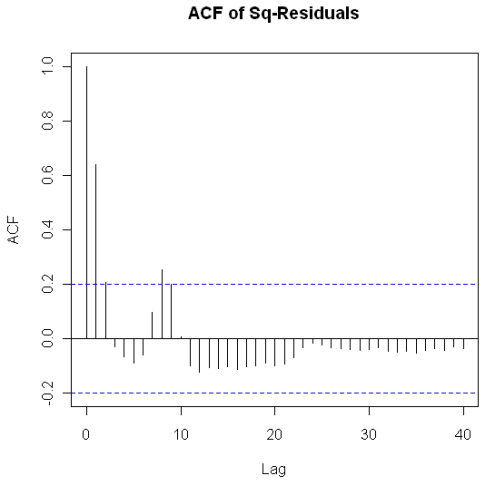
Coefficients:

	Value	Std. Error	t-value	
median_income-Lag-8	0.031	0.128	0.240	
median_income-Lag-7	-0.034	0.165	-0.209	
median_income-Lag-6	0.165	0.320	0.516	
median_income-Lag-5	-0.108	0.417	-0.258	
median_income-Lag-4	-0.326	0.299	-1.093	
unemp_rate-Lag-5	-0.051	0.033	-1.534	
unemp_rate-Lag-4	-0.050	0.035	-1.422	
pop_size-Lag-6	3.880	2.263	1.715	
pop_size-Lag-5	-1.655	3.452	-0.479	
pop_size-Lag-4	-0.712	3.614	-0.197	
pop_size-Lag-3	3.450	2.278	1.514	
foreclosures-Lag-12	-0.002	0.002	-1.173	
foreclosures-Lag-11	-0.002	0.001	-1.928	
foreclosures-Lag-10	-0.005	0.001	-3.567	*
foreclosures-Lag-9	-0.005	0.001	-3.803	*
foreclosures-Lag-8	-0.003	0.001	-2.362	*
foreclosures-Lag-7	-0.003	0.001	-2.004	*
foreclosures-Lag-6	-0.002	0.001	-1.343	
ty_cr-Lag2	-0.007	0.012	-0.549	
ty_cr-Lag3	-0.001	0.011	-0.119	
ty_cr-Lag4	-0.008	0.012	-0.609	
mort_orig-Lag1	-0.003	0.001	-2.650	*
mort_orig-Lag2	-0.004	0.001	-3.009	*
mort_orig-Lag3	-0.004	0.001	-3.128	*
mort_orig-Lag4	-0.003	0.001	-2.467	*
building_permits-Lag10	0.000	0.002	-0.037	
building_permits-Lag11	-0.002	0.003	-0.885	
building_permits-Lag12	0.000	0.003	-0.064	
building_permits-Lag13	-0.001	0.002	-0.344	
building_permits-Lag14	0.002	0.002	0.937	
building_permits-Lag15	0.000	0.002	0.017	
building_permits-Lag16	0.001	0.002	0.546	

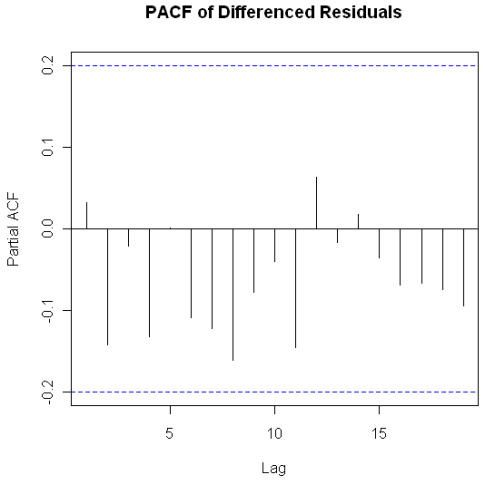
1. Phase I

ACF of Residuals





PACF and ADF Test




```
> adf.test(y$residuals, k = 3)
        Augmented Dickey-Fuller Test
data:  y$residuals
Dickey-Fuller = -3.6903, Lag order = 3, p-value = 0.02902
```

2. Phase II

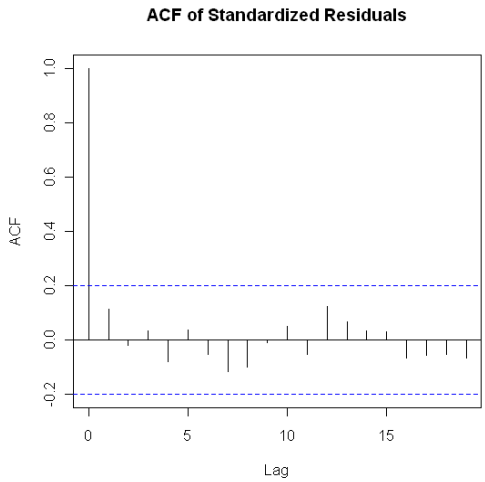
```
ARIMA(1,0,0) with zero mean
Call: arima(x = resid.data, order = c(1, 0, 0), include.mean = FALSE)

Coefficients:
                  ar1
        0.7702
s.e.    0.0629
sigma^2 estimated as 6.112e-06:  log likelihood = 444.17
AIC = -884.33   AICc = -884.2   BIC = -879.18
```

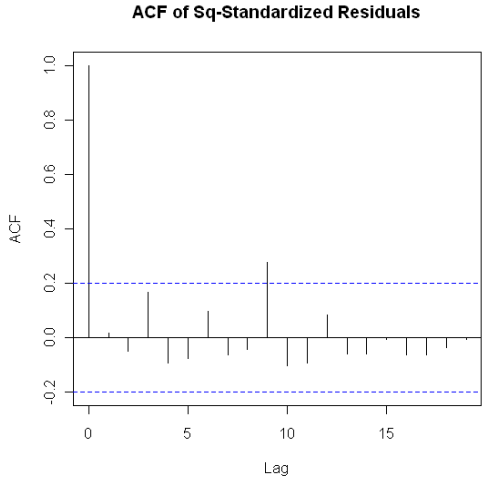
Box-Ljung Test

```
> Box.Ljung.test(arima.fit$resid,lag = 12, adj.DF =  12-params)
   Q(12) P-value df
 7.384673  0.7671 11

> Box.Ljung.test(arima.fit$resid^2,lag = 12, adj.DF =  12)
   Q(12) P-value df
 25.26165  0.0136 12
```

	Estimate	Std. Error	t value	Pr(>|t|)	
ar1	0.7550	0.0716	10.55	<2e-16	***
omega	1.644E-06	0.0000	1.03	0.304	
alpha1	0.0649	0.0652	1.00	0.320	
beta1	0.6636	0.2942	2.26	0.024	*

ACF of Residuals





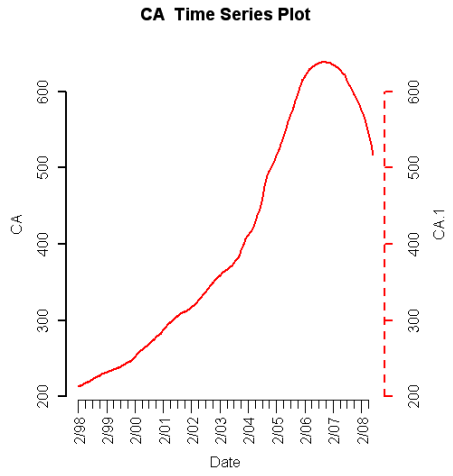
Box-Ljung Test

```
> Box.Ljung.test(sresi, lag = 12, adj.DF =  12-meanParams)
  Q(12) P-value df
 7.3162  0.7729 11
> Box.Ljung.test(sresi^2, lag = 12, adj.DF = 12-volParams )
    Q(12) P-value df
 17.41181  0.0657 10
```

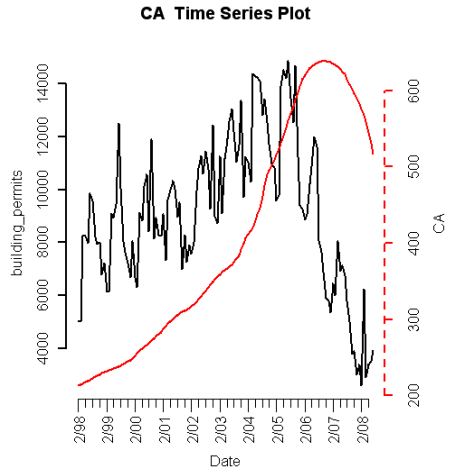
CALIFORNIA (CA)

A. Plots of Predictors vs. HPI (non-simple return)

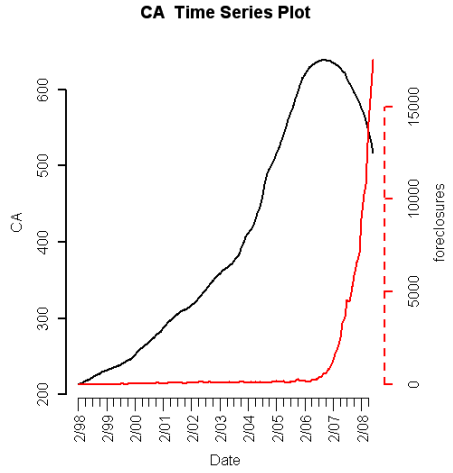
Original HPI Series



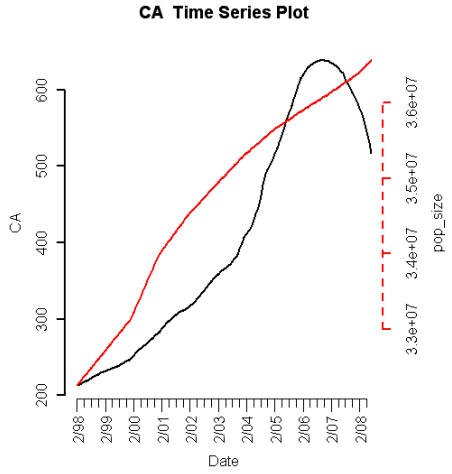
Building Permits



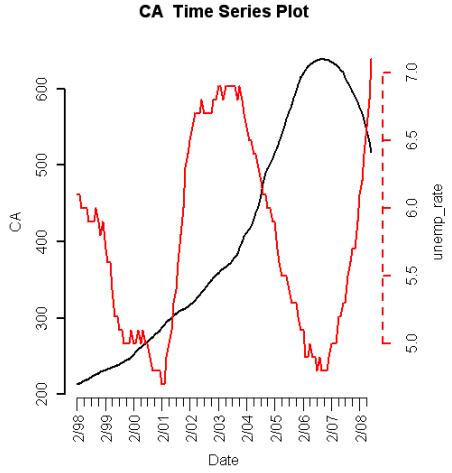
Foreclosures



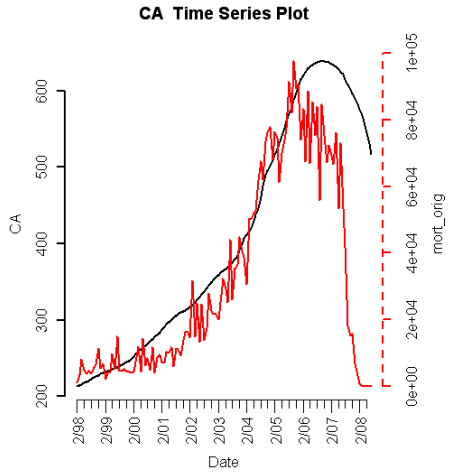
Population Size



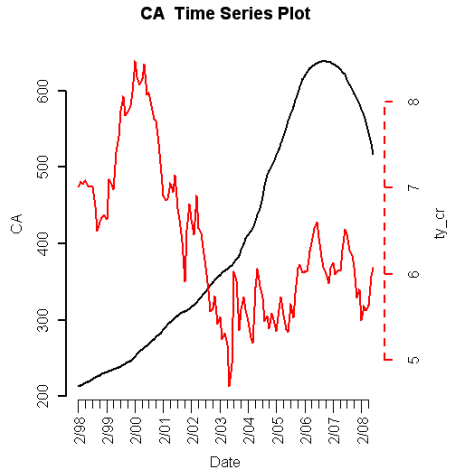
Unemployment Rate



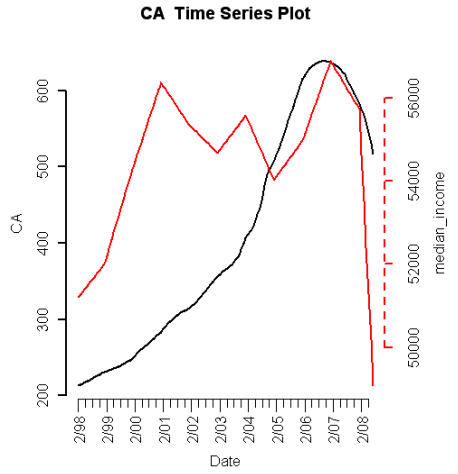
Mortgage Originations



30-Year Commitment Rate

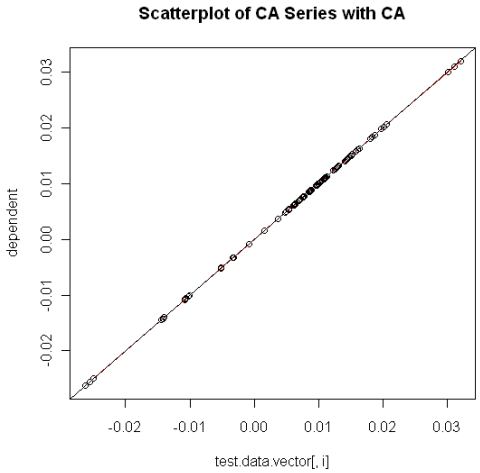


Median Income

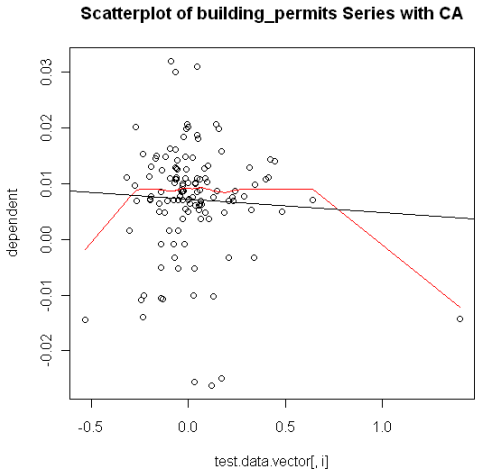


B. Scatterplots (simple return)

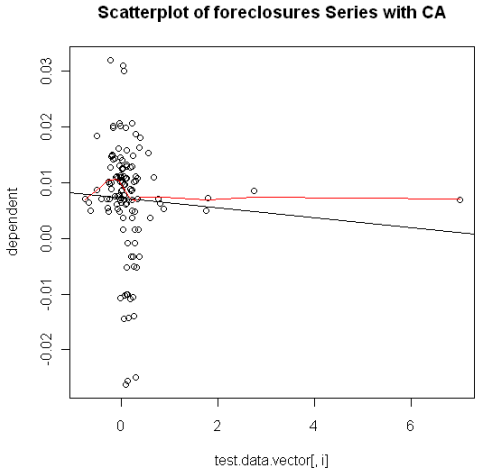
HPI vs. HPI



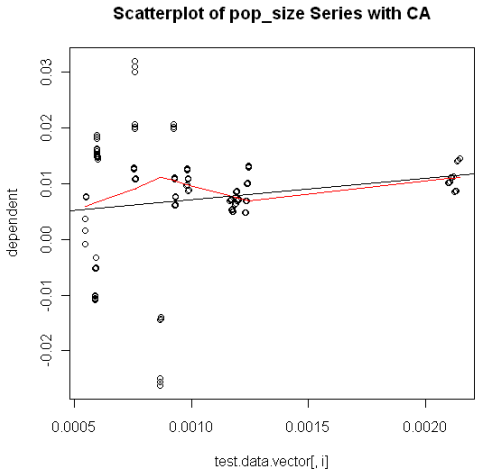
HPI vs. Building Permits



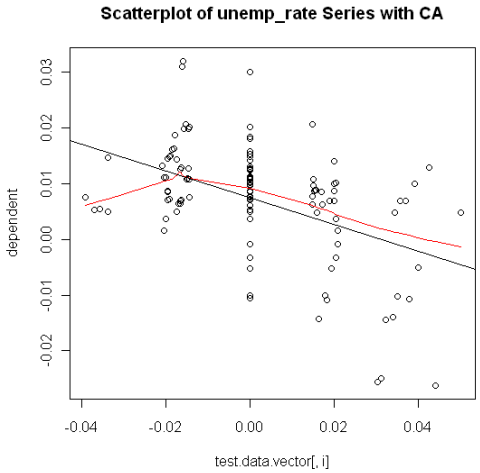
HPI vs. Foreclosures



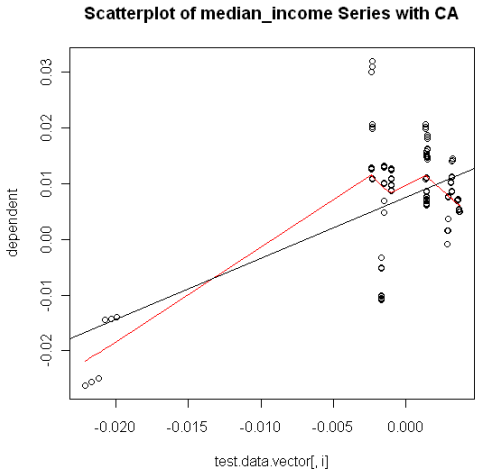
HPI vs. Population Size



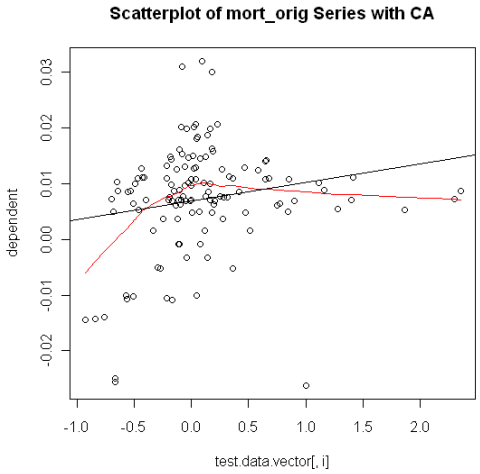
HPI vs. Unemployment Rate



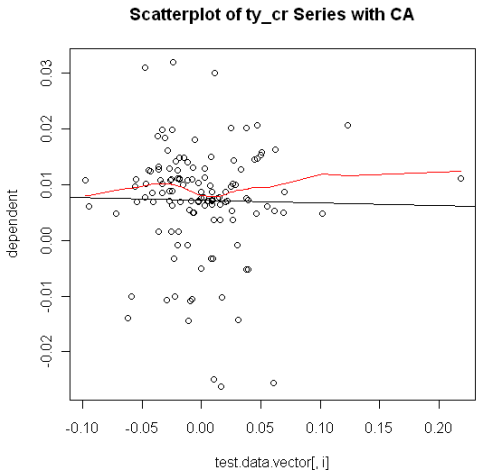
HPI vs. Median Income



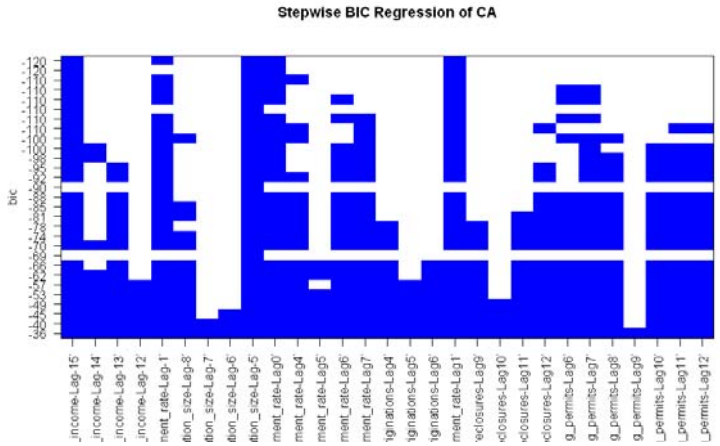
HPI vs. Mortgage Originations



HPI vs. 30-Year Commitment Rate



C. Exhaustive Stepwise Regression

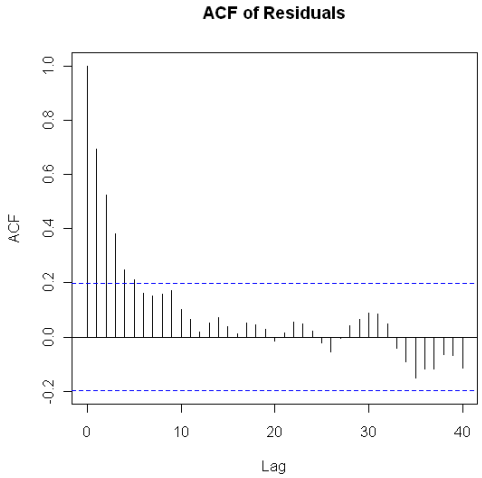


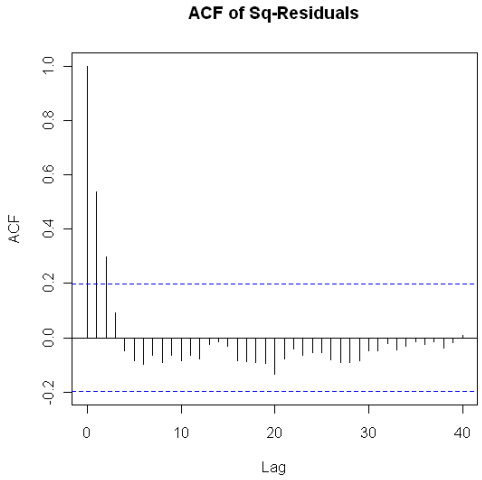
E. Final RLM Model

Coefficients				
	Value	Std. Error	t-value	
median_income-Lag-15	0.4799	0.3456	1.3888	
median_income-Lag-14	-0.0537	0.4965	-0.1083	
median_income-Lag-13	0.105	0.4816	0.2181	
median_income-Lag-12	-0.1179	0.3923	-0.3005	
ty_cr-Lag4	0.0271	0.0174	1.5552	
ty_cr-Lag5	0.003	0.0167	0.1812	
ty_cr-Lag6	-0.0142	0.0168	-0.8429	
ty_cr-Lag7	-0.0177	0.0164	-1.0781	
mort_orig-Lag4	-0.0037	0.0018	-2.0486	*
mort_orig-Lag5	-0.007	0.0022	-3.2181	*
mort_orig-Lag6	-0.004	0.0018	-2.2392	*
unemp_rate-Lag1	-0.1261	0.0514	-2.4554	*
unemp_rate-Lag2	-0.0759	0.0502	-1.5111	
unemp_rate-Lag3	-0.0311	0.0504	-0.6156	
unemp_rate-Lag4	-0.0354	0.0482	-0.7356	
unemp_rate-Lag5	-0.0304	0.0478	-0.6358	
unemp_rate-Lag6	0.0301	0.0493	0.61	
pop_size-Lag2	6.0132	5.0598	1.1884	
pop_size-Lag3	-4.2653	7.2838	-0.5856	
pop_size-Lag4	1.6781	7.358	0.2281	
pop_size-Lag5	3.9044	7.558	0.5166	
pop_size-Lag6	1.5299	7.5535	0.2025	
pop_size-Lag7	-6.2109	7.5069	-0.8274	
pop_size-Lag8	9.5897	5.2106	1.8404	
foreclosures-Lag9	-0.0018	0.001	-1.8421	
foreclosures-Lag10	-0.0022	0.0009	-2.3617	*
foreclosures-Lag11	-0.0025	0.0009	-2.7381	*
foreclosures-Lag12	-0.0024	0.0009	-2.5308	*
building_permits-Lag6	-0.0018	0.005	-0.3537	
building_permits-Lag7	0.0009	0.0053	0.1741	
building_permits-Lag8	0.005	0.0052	0.952	
building_permits-Lag9	0.007	0.0051	1.3639	
building_permits-Lag10	0.0005	0.0051	0.0993	
building_permits-Lag11	-0.0038	0.0049	-0.7675	
building_permits-Lag12	-0.0059	0.0051	-1.1608	

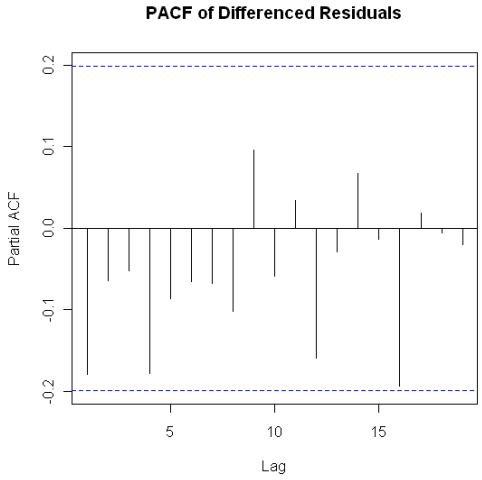
1. Phase 1

ACF of Residuals





PACF and ADF Test




```
> adf.test(y$residuals, k = 2)
        Augmented Dickey-Fuller Test
data:  y$residuals
Dickey-Fuller = -3.5496, Lag order = 2, p-value = 0.04141
```

Box-Ljung Test

```
> Box.test(y$residuals, lag = 12, type = c("Ljung-Box"))
        Box-Ljung test
X-squared = 115.341, df = 12, p-value < 2.2e-16
```

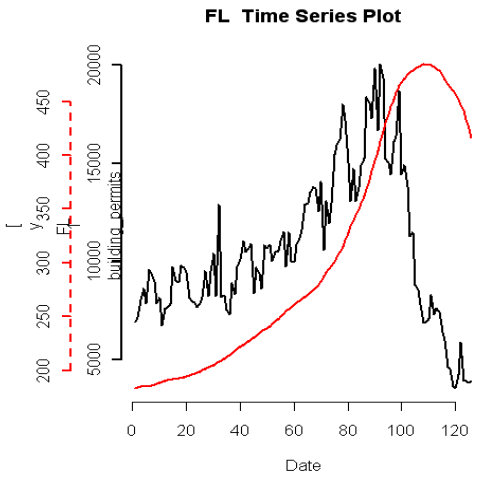
2. Phase II

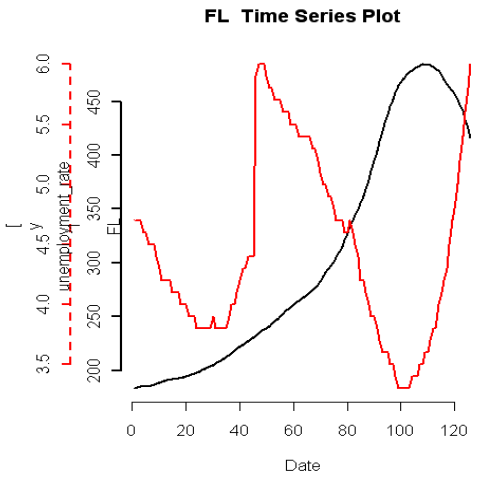
```
ARIMA(1,0,0) with zero mean
Call: auto.arima(x = resid.data, stationary = TRUE)

Coefficients:
         ar1
      0.7238
s.e.  0.0710
sigma^2 estimated as 1.539e-05:  log likelihood = 403.57
AIC = -803.14   AICc = -803.02   BIC = -797.97
```

Box-Ljung Test

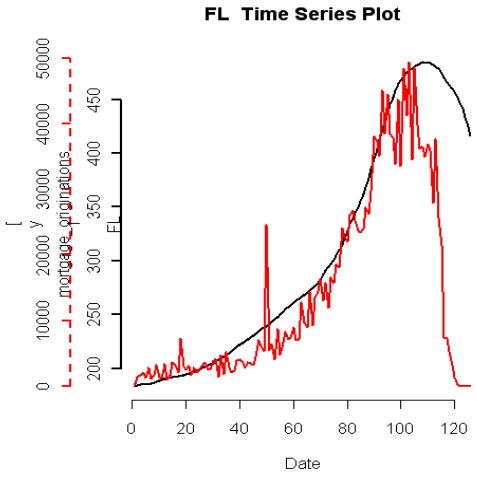
```
> Box.Ljung.test(arima.fit$resid,lag = 12, adj.DF =  12-params)
        Box-Ljung test
X-squared = 6.2959, df = 11, p-value = 0.8529

> Box.Ljung.test(arima.fit$resid^2,lag = 12, adj.DF =  12)
        Box-Ljung test
X-squared = 17.9114, df = 12, p-value = 0.1184
```

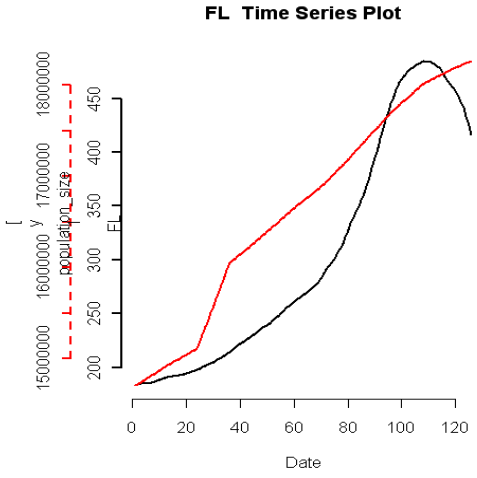
FLORIDA (CA)

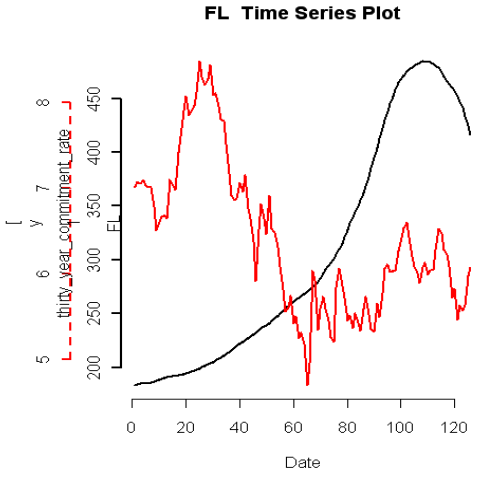

FL Time Series Plot
building_permits
Date

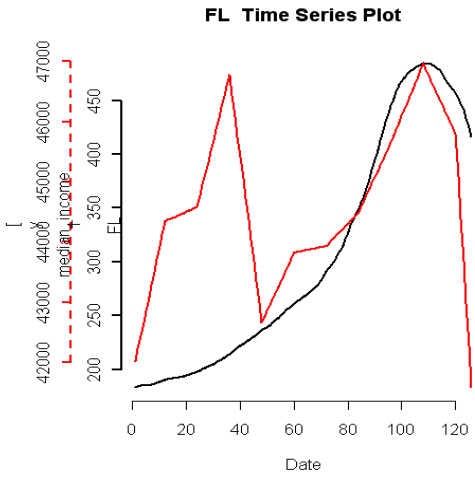

FL Time Series Plot
unemployment_rate
Date

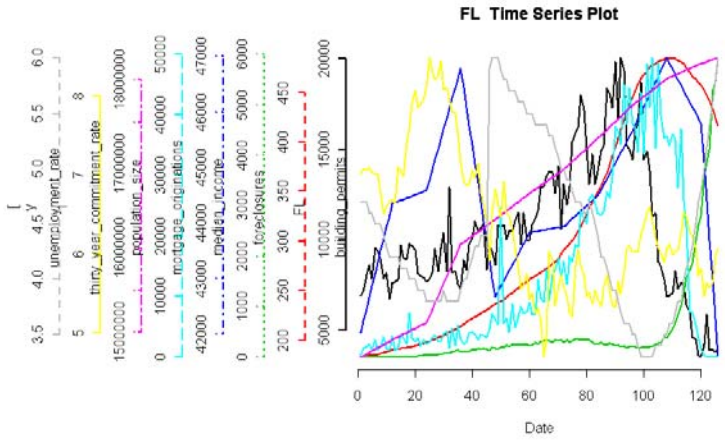

FL Time Series Plot
foreclosures
Date


FL Time Series Plot
mortgage_originations
Date


FL Time Series Plot
population_size
Date


FL Time Series Plot
thirty_year_commitment_rate
Date

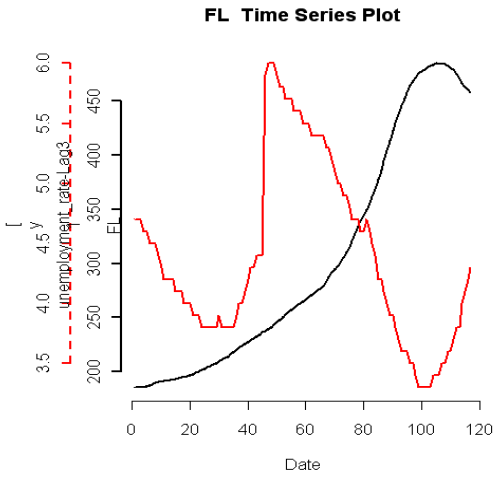


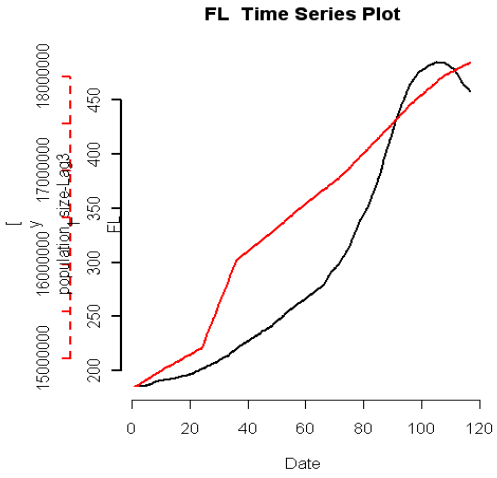


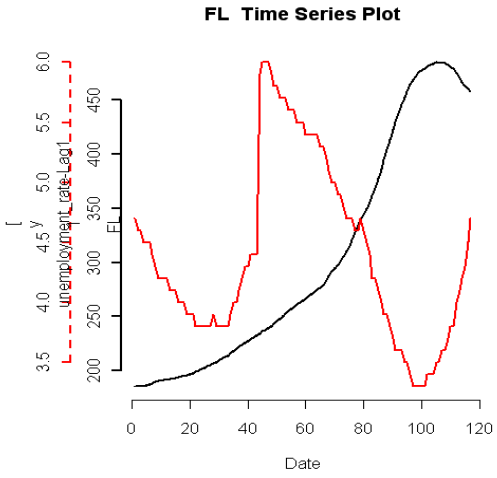
Based on these time series of the data, we picked the following lags to test:

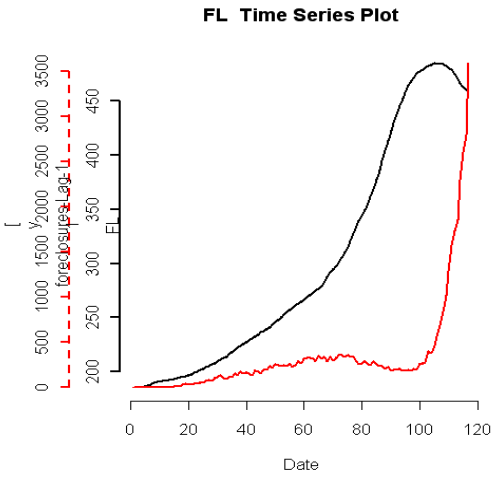
Predictor Variable	Lag(s)
Median income	0
30yr commitment rate	0
Mortgage originations	0
Unemployment rate	-4, -3, 1, 3
Population size	-6, 3
Foreclosures	-3,-1
Building permits	0

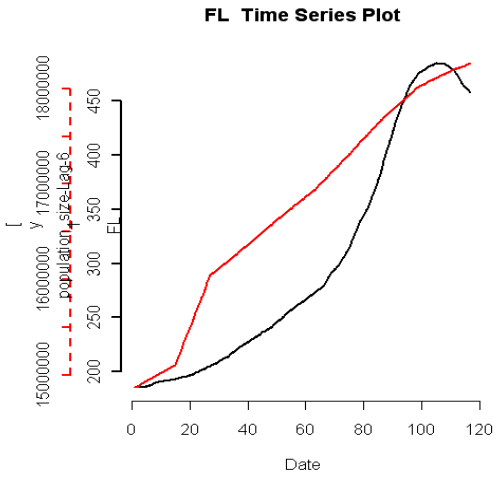
Plots of lagged predictor variables and FL HPI:

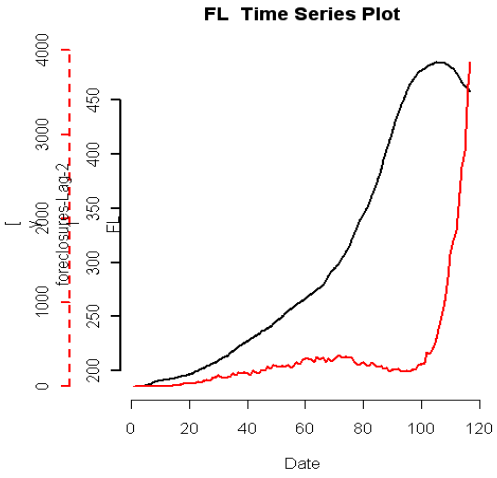


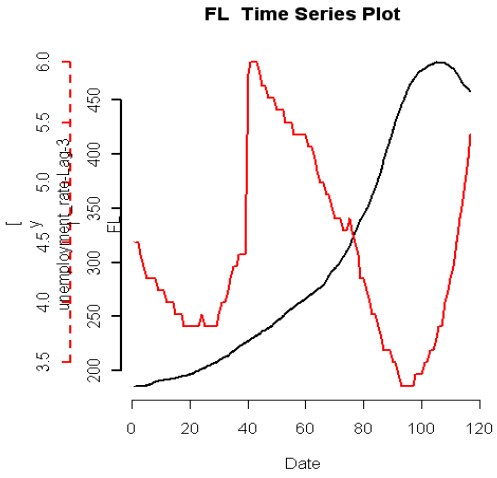


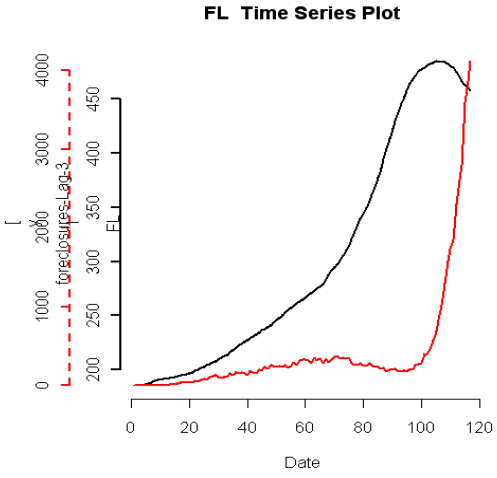


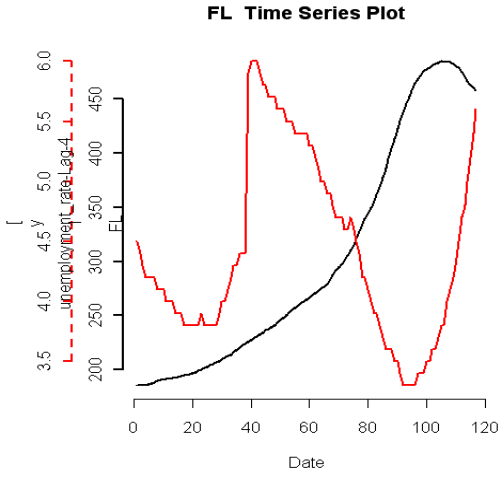

FL Time Series Plot
foreclosures Lag 1
FL
Date


FL Time Series Plot
population size Lag 6
FL
Date

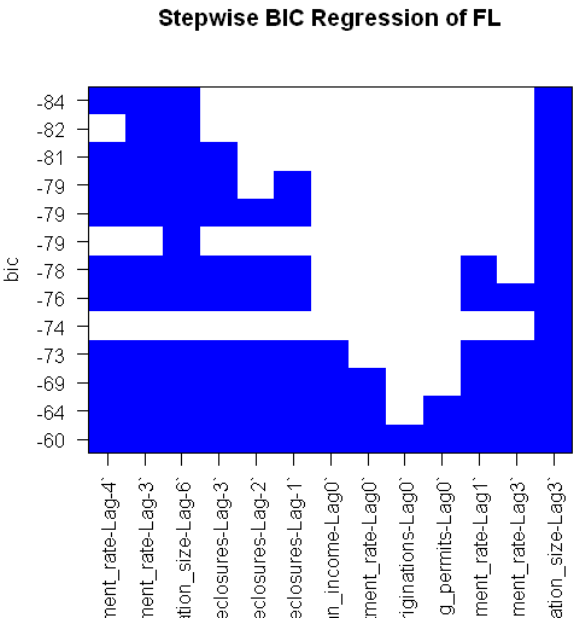

FL Time Series Plot
foreclosures Lag 2
FL
Date


FL Time Series Plot
unemployment rate Lag 3
FL
Date


FL Time Series Plot
foreclosures Lag 3
FL
Date


FL Time Series Plot
unemployment rate Lag 4
FL
Date

Stepwise BIC values for Regression of Lagged Predictor Variables with Florida

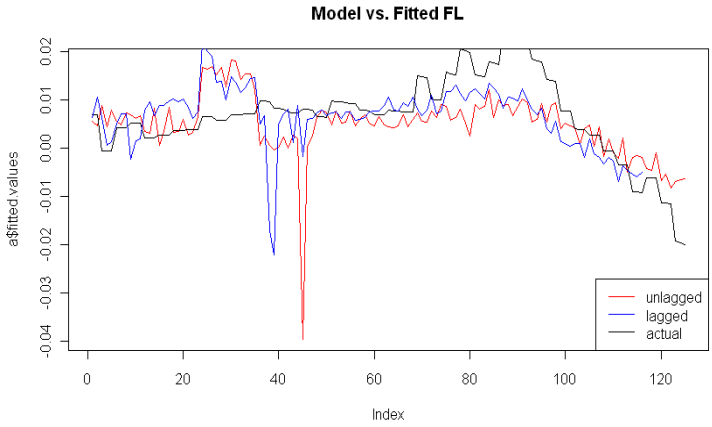


Fitted Robust Linear Model with all lagged variables

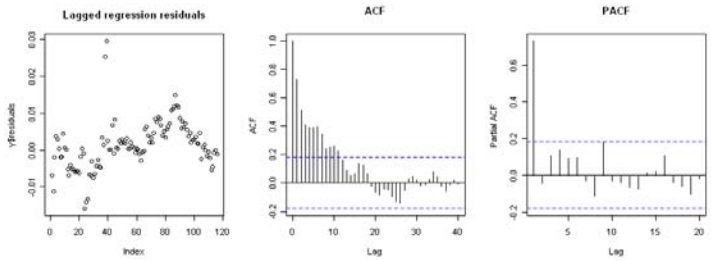
```
Call: rlm(x = shifted.multivar.test.data.vector[, 1:
(n - 1)], y = shifted.multivar.test.data.vector[,
    n])
Residuals:
       Min         1Q     Median         3Q        Max
-0.0159690 -0.0020288  0.0008338  0.0045595  0.0295564

Coefficients:
                                 Value   Std. Error t value
unemployment_rate-Lag-4          -0.0712  0.0162    -4.3983
unemployment_rate-Lag-3          -0.0836  0.0165    -5.0723
population_size-Lag-6             1.3061  0.4229     3.0884
foreclosures-Lag-3               -0.0026  0.0016    -1.6460
foreclosures-Lag-2               -0.0026  0.0015    -1.6699
foreclosures-Lag-1               -0.0026  0.0015    -1.7549
median_income-Lag0               -0.2167  0.2059    -1.0522
thirty_year_commitment_rate-Lag0  0.0009  0.0024     0.3597
mortgage_originations-Lag0       -0.0002  0.0010    -0.2033
building_permits-Lag0             0.0017  0.0045     0.3742
unemployment_rate-Lag1           -0.0285  0.0176    -1.6213
unemployment_rate-Lag3           -0.0237  0.0160    -1.4781
population_size-Lag3              2.9363  0.4187     7.0134

Residual standard error: 0.005344 on 103 degrees of freedom
```

Fitted Robust Linear Model with unlagged variables

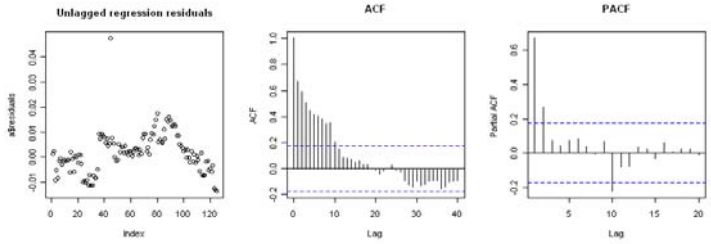
```
Call: rlm(x = test.data.vector[, 1:(length(test.data.vector)
- 1)],
    y = test.data.vector[, length(test.data.vector)])
Residuals:
      Min        1Q    Median        3Q       Max
-0.013701 -0.002252  0.001575  0.006133  0.047483

Coefficients:
                            Value   Std. Error t value
median_income                0.2237  0.1666     1.3423
thirty_year_commitment_rate  0.0002  0.0030     0.0686
mortgage_originations        0.0002  0.0011     0.2014
unemployment_rate           -0.1259  0.0219    -5.7532
population_size              2.9903  0.3543     8.4397
foreclosures                -0.0009  0.0017    -0.5132
building_permits             0.0040  0.0051     0.7731

Residual standard error: 0.007476 on 118 degrees of freedom
```



Lagged Regression Residuals Analysis:



Unlagged Regression Residuals Analysis:



The PACF shows a lag of 2 to use in the ADF test for stationarity.

```
        Augmented Dickey-Fuller Test

data:  y$residuals
Dickey-Fuller = -3.5887, Lag order = 2,
p-value = 0.03735
alternative hypothesis: stationary
```

Residuals are stationary so phase 1 complete. Continue to Phase 2.

First, check Box-Ljung.

```
        Box-Ljung test

data:  y$residuals
X-squared = 220.422, df = 12, p-value < 2.2e-16
```

Box-Ljung fails. Implies residuals are autocorrelated. So we will now model the residuals as time series.

EACF shows that we should try to fit ARMA(1,0).

Fit ARMA:

```
Series: y$residuals
ARIMA(1,0,0) with non-zero mean

Call: arima(x = y$residuals, order = c(1, 0, 0))

Coefficients:
         ar1  intercept
      0.7356     0.0011
s.e.  0.0624     0.0015

sigma^2 estimated as 1.921e-05:  log likelihood =
464.91
AIC = -923.82   AICc = -923.61   BIC = -915.56
```

Check adequacy of ARMA

```
> Box.Ljung.test(arima.fit$residuals,lag = 12,
adj.DF =  12-params)

        Box-Ljung test

data:  data
X-squared = 12.6459, df = 11, p-value = 0.3171
```

The mean equation passes.

```
> Box.Ljung.test(arima.fit$residuals^2,lag = 12,
adj.DF =  12)

        Box-Ljung test

data:  data
X-squared = 24.6408, df = 12, p-value = 0.01662
```

Residuals^2 fails. GARCH modeling needed.

FIT GARCH: Garch estimates

```
         mu          ar1        omega
alpha1        beta1
4.111261e-04 3.241833e-01 3.089253e-06
1.000000e+00 2.568742e-01
```

Check adequacy of GARCH model (using standardized residuals)

```
> sresi=garch.fit@residuals/garch.fit@sigma.t
> Box.Ljung.test(sresi, lag = 12, adj.DF =  12-
meanParams)

        Box-Ljung test

data:  data
X-squared = 10.337, df = 11, p-value = 0.5004

> Box.Ljung.test(sresi^2, lag = 12, adj.DF = 12-
volParams )

        Box-Ljung test

data:  data
X-squared = 5.7986, df = 10, p-value = 0.8319
```

Both Box-Ljung tests for the standardized residuals and squared standardized residuals of the ARMA(1,0) + GARCH(1,1) model pass.

The final fitted models are as follows.

Phase 1 Regression: Drift Modeling

$y_t=\Sigma\beta_i x_{t,i} + res_t$,
where res_t are the residuals

Variable	$\beta_{i,}$
unemployment_rate-Lag-4	-0.0712
unemployment_rate-Lag-3	-0.0836
population_size-Lag-6	1.3061
foreclosures-Lag-3	-0.0026
foreclosures-Lag-2	-0.0026
foreclosures-Lag-1	-0.0026
median_income-Lag0	-0.2167
thirty_year_commitment_rate-Lag0	0.0009
mortgage_originations-Lag0	-0.0002
building_permits-Lag0	0.0017
unemployment_rate-Lag1	-0.0285
unemployment_rate-Lag3	-0.0237
population_size-Lag3	2.9363

Phase 2 Time Series: Volatility Modeling

$(1-\phi_1 B)res_t = \phi_0 + a_t$

$\sigma_t^2 = \alpha a_{t-1}^2 + \beta\sigma_{t-1}^2$
$a_t = \varepsilon_t\sigma_t$
$\varepsilon_t \sim N(0,1)$

Coefficient	Estimate
ϕ_0	0.0011
ϕ_1	0.7356
α	1.00
β	0.26