Analysis of Homeownership Rate in US

Yueqiao Faith Zhang

Ashutosh Kumar

Roadmap

- Introduction
- Research Question
- Deciding Series to Work On
- Model 1
- Model 2
- Model 3
- Model 4
- Conclusion

Introduction: Homeownership Rate

Definition:

The proportion of owner-occupied housing units out of all occupied units.

Data:

Non-seasonally adjusted

Quarterly, from Q1, 1965 to Q3, 2018 (215 time periods)

Source: US Census Bureau

Research Questions

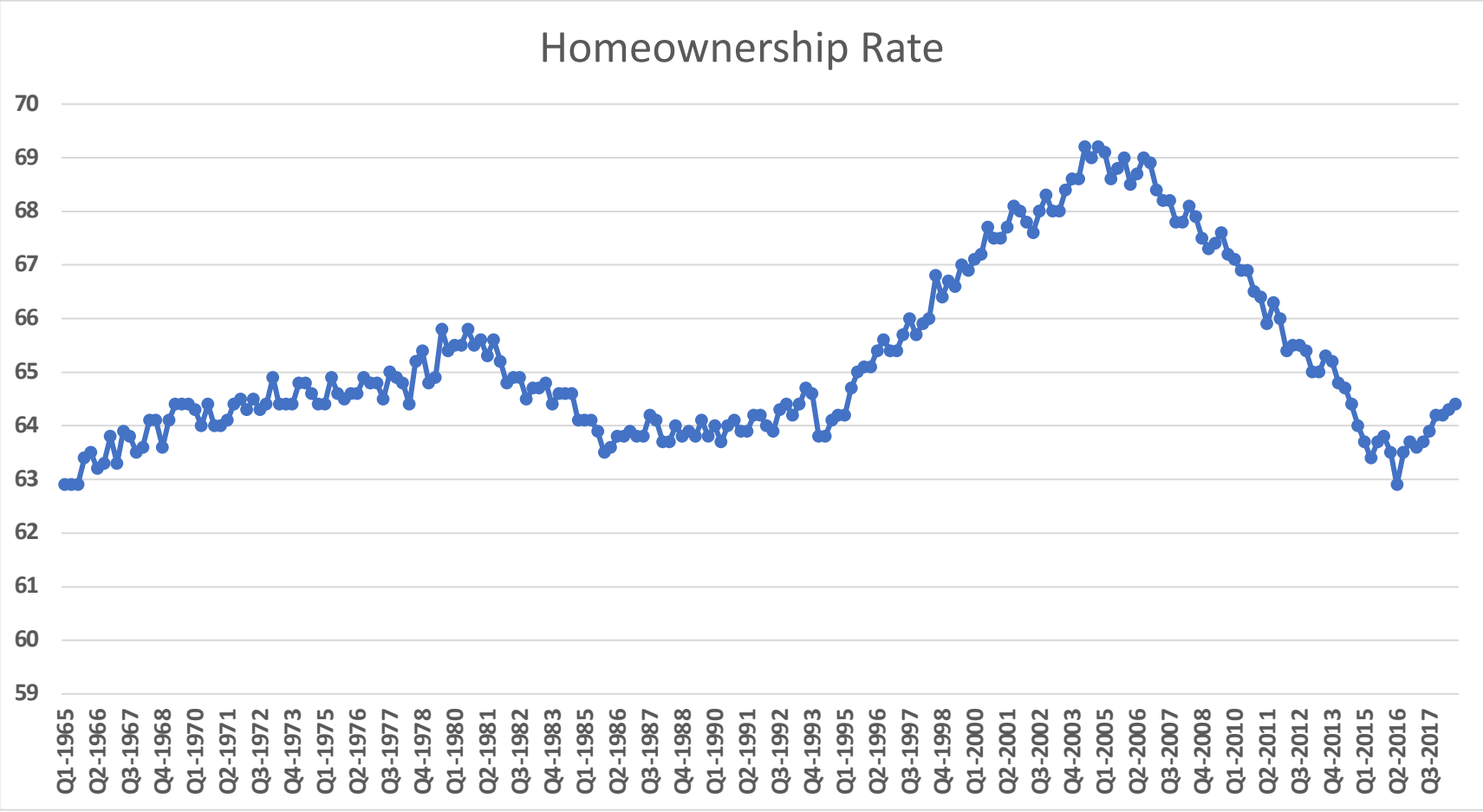
How did the Homeownership Rate change across time in the US?

What was the impact of 2006 Housing Crisis on Homeownership Rate?

What model is the “best fit” for answering above question?

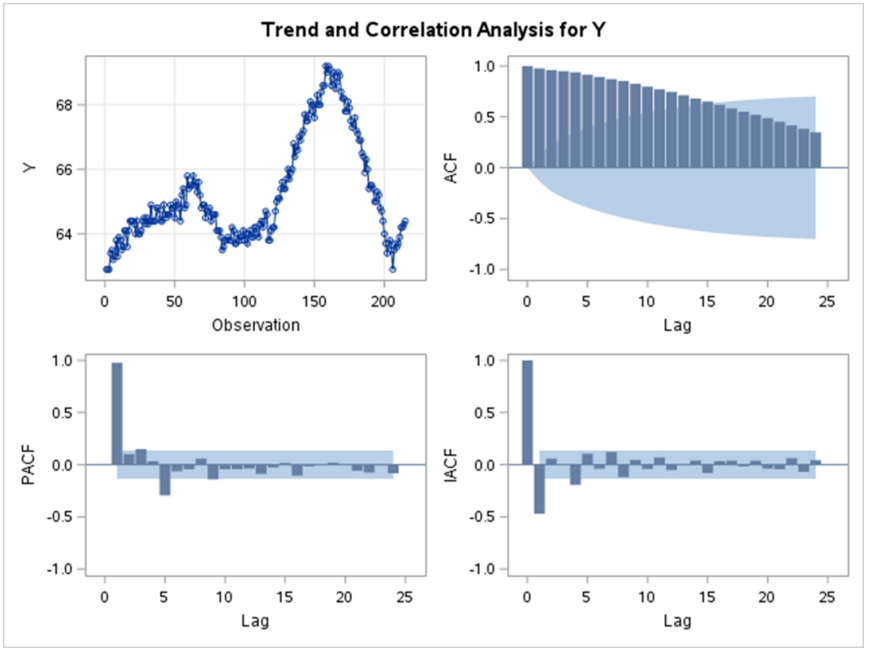
Can we forecast future Homeownership Rate?

ORIGINAL SERIES



Original series Y: Nonstationary

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	0	0.0202	0.6868	0.30	0.7731		
	1	0.0206	0.6869	0.36	0.7882		
	2	0.0207	0.6869	0.46	0.8119		
	3	0.0155	0.6857	0.35	0.7846		
	4	0.0111	0.6847	0.16	0.7331		
	5	0.0146	0.6855	0.19	0.7410		
Single Mean	0	-4.4487	0.4875	-1.68	0.4424	1.46	0.6977
	1	-3.5643	0.5866	-1.54	0.5095	1.27	0.7463
	2	-2.6164	0.7020	-1.41	0.5766	1.12	0.7860
	3	-2.2766	0.7439	-1.24	0.6569	0.84	0.8560
	4	-4.2352	0.5104	-1.50	0.5333	1.14	0.7797
	5	-5.5682	0.3795	-1.70	0.4290	1.47	0.6949
Trend	0	-3.6514	0.9048	-1.20	0.9068	1.55	0.8686
	1	-2.4195	0.9576	-0.92	0.9507	1.58	0.8626
	2	-1.0974	0.9867	-0.53	0.9815	2.04	0.7700
	3	-0.8332	0.9899	-0.41	0.9866	1.70	0.8372
	4	-3.8315	0.8948	-1.19	0.9096	1.14	0.9481
	5	-5.6346	0.7697	-1.47	0.8351	1.44	0.8898



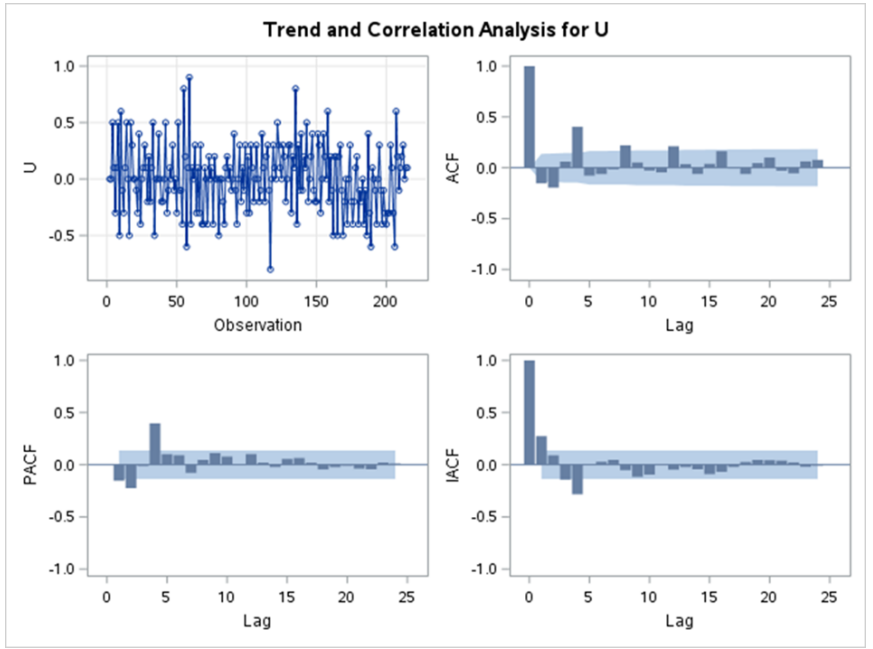
FIRST DIFF (U): (Voila!) STATIONARY

FIRST DIFFERENCE Wednesday,

The ARIMA Procedure

(5% Significance Level)

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-245.611	0.0001	-16.99	<.0001		
	1	-385.894	0.0001	-13.81	<.0001		
	2	-401.464	0.0001	-10.13	<.0001		
	3	-82.8041	<.0001	-5.44	<.0001		
	4	-58.6396	<.0001	-4.47	<.0001		
	5	-44.1120	<.0001	-3.84	0.0002		
Single Mean	0	-245.740	0.0001	-16.96	<.0001	143.76	0.0010
	1	-386.748	0.0001	-13.80	<.0001	95.20	0.0010
	2	-403.421	0.0001	-10.12	<.0001	51.21	0.0010
	3	-83.0361	0.0014	-5.43	<.0001	14.74	0.0010
	4	-58.8490	0.0014	-4.47	0.0004	9.97	0.0010
	5	-44.2628	0.0014	-3.84	0.0031	7.37	0.0010
Trend	0	-247.743	0.0001	-17.08	<.0001	145.80	0.0010
	1	-402.121	0.0001	-14.03	<.0001	98.38	0.0010
	2	-449.241	0.0001	-10.33	<.0001	53.39	0.0010
	3	-88.9924	0.0006	-5.49	<.0001	15.16	0.0010
	4	-63.9229	0.0006	-4.53	0.0018	10.31	0.0010
	5	-48.1323	0.0006	-3.87	0.0150	7.58	0.0182



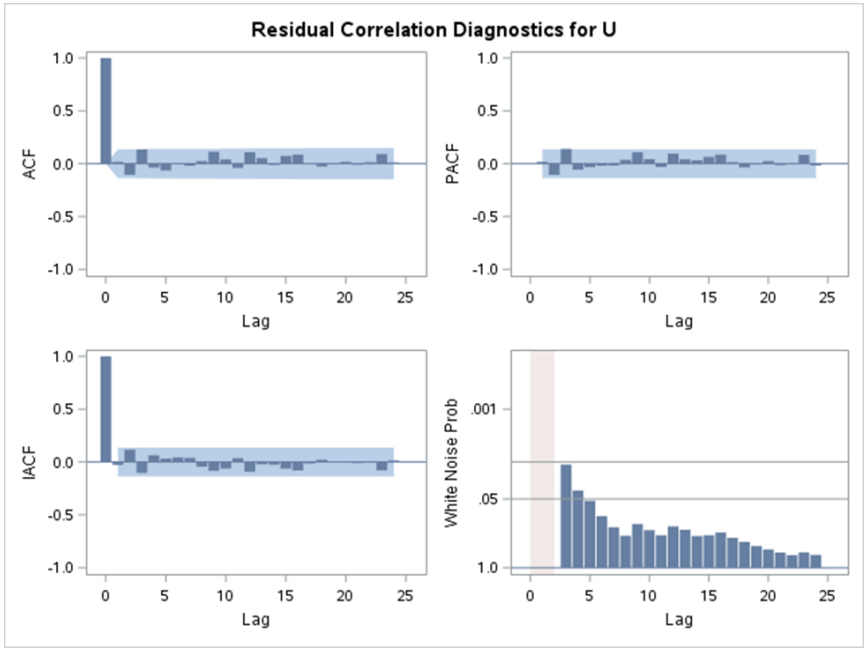
Model 1: Identification and Fitting

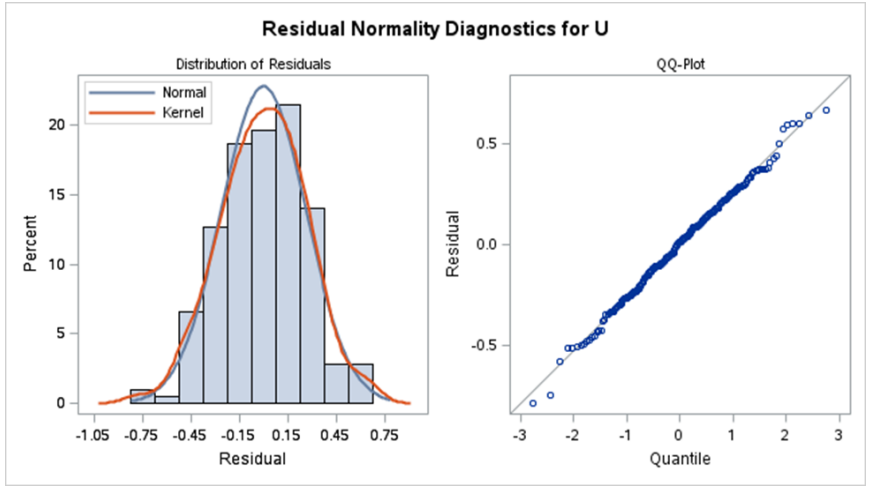
Model for U	AIC	SBC	Variance Estimate	Note
MA(4)	51.83035	68.66023	0.072703	Insig coeff, residuals not white noise
MA((1,2,4))	51.91713	65.38104	0.073067	Residuals not white noise
ARMA(4,4)	41.66577	71.95955	**0.067915**	Insig coeff
ARMA((4),(1,4))	**39.92582**	53.38972	0.068934	Insig coeff
ARMA((4),1)	40.62527	**50.7232**	0.069524	

ARMA(p+d,q) Tentative Order Selection Tests ESACF		
p+d	q	BIC
0	4	-2.59434
3	4	-2.5437
4	4	-2.61296
2	5	-2.5596

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	0.01022	0.02273	0.45	0.6529	0
MA1,1	0.26945	0.06648	4.05	<.0001	1
AR1,1	0.42883	0.06249	6.86	<.0001	4

Residual Diagnostics for U

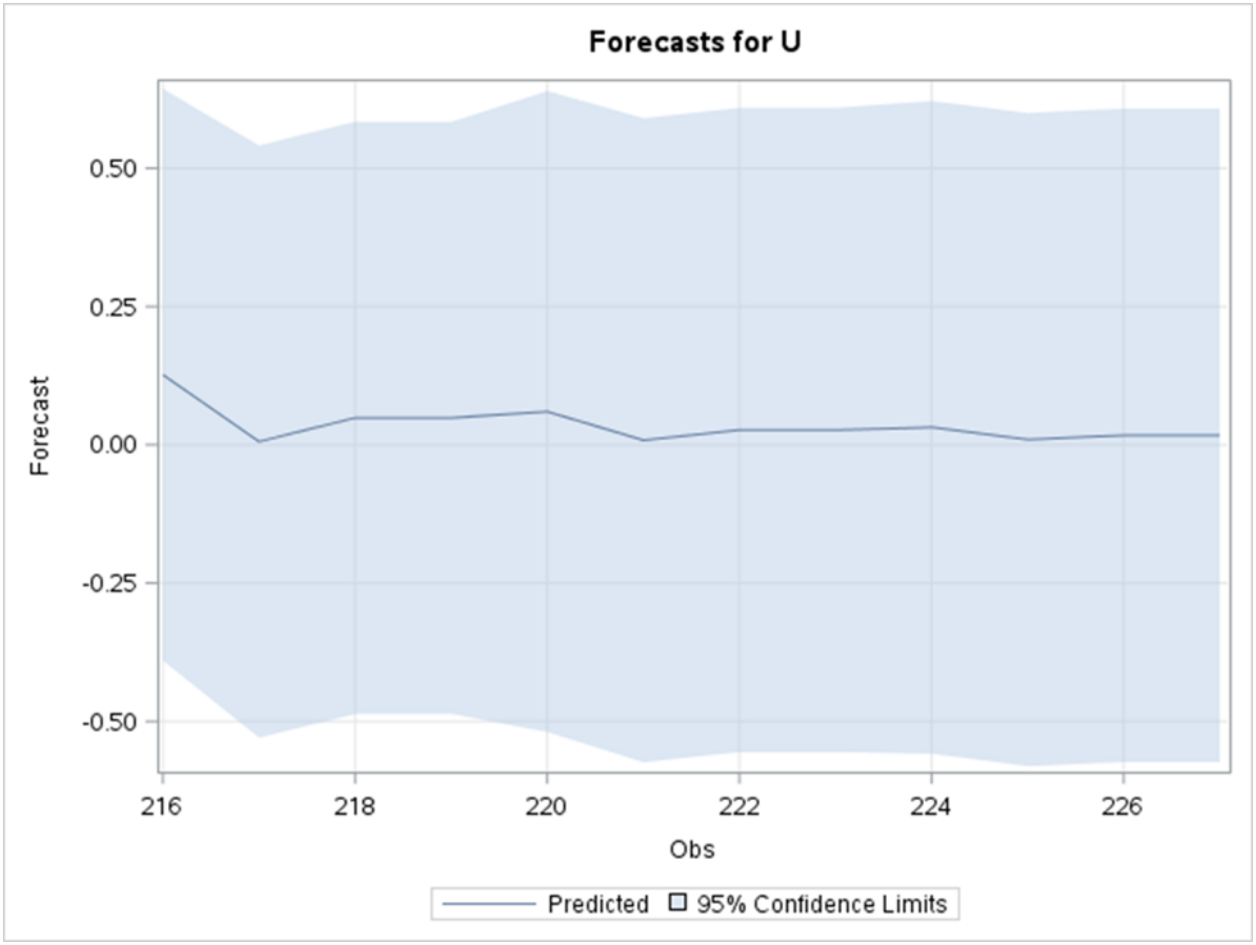




MODEL 1: ARIMA((4),1,1) FORECAST

$$U_t = 0.010223 + \frac{1 - 0.26945\,B}{1 - 0.42883\,B^4} a_t$$

$$Y_t = Y_{t-1} + 0.010223 + \frac{1 - 0.26945\,B}{1 - 0.42883\,B^4} a_t$$



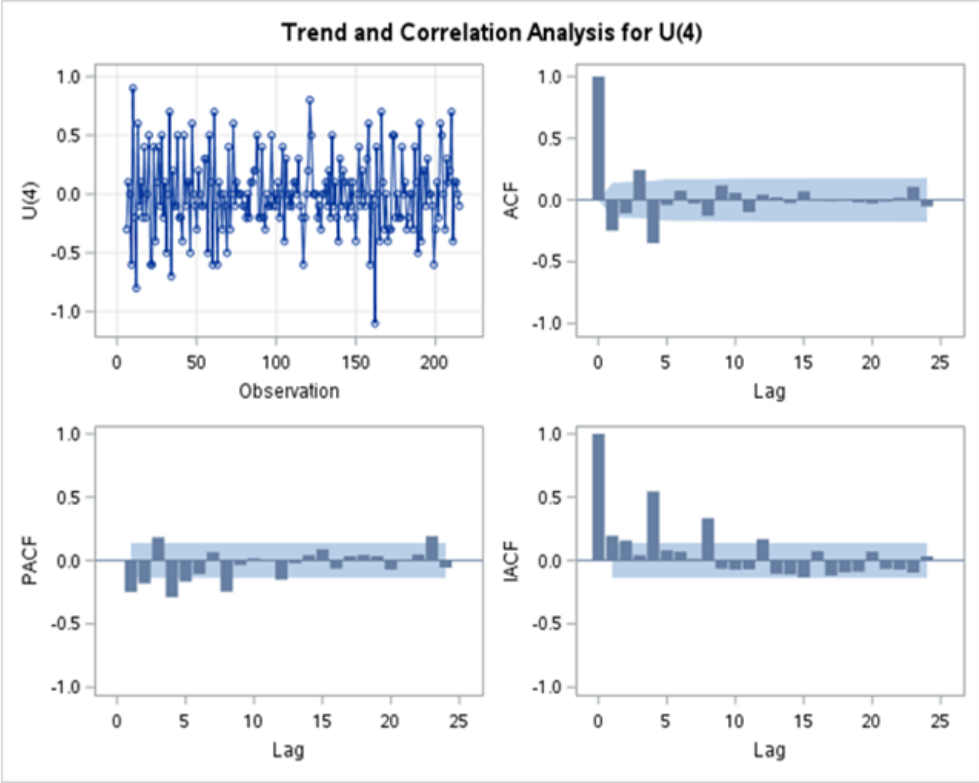
Q4-2018	64.5266
Q1-2019	64.5324
Q2-2019	64.5811
Q3-2019	64.6298
Q4-2019	64.6899
Q1-2020	64.6982
Q2-2020	64.7249
Q3-2020	64.7516
Q4-2020	64.7832
Q1-2021	64.7926
Q2-2021	64.8099
Q1-2019	64.8272

Model 2: Stochastic Seasonal Model

Differencing by 4, i.e. U(4) *Why?*

- Quarterly
- Suggested by CDF of U

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-261.012	0.0001	-18.63	<.0001		
	1	-375.016	0.0001	-13.64	<.0001		
	2	-201.233	0.0001	-8.17	<.0001		
	3	-2069.2	0.0001	-9.67	<.0001		
	4	606.7969	0.9999	-9.33	<.0001		
	5	316.5372	0.9999	-8.56	<.0001		
Single Mean	0	-261.012	0.0001	-18.59	<.0001	172.76	0.001
	1	-375.022	0.0001	-13.6	<.0001	92.51	0.001
	2	-201.242	0.0001	-8.15	<.0001	33.23	0.001
	3	-2070.08	0.0001	-9.65	<.0001	46.52	0.001
	4	606.7453	0.9999	-9.3	<.0001	43.27	0.001
	5	316.5189	0.9999	-8.54	<.0001	36.48	0.001
Trend	0	-261.179	0.0001	-18.55	<.0001	172.13	0.001
	1	-376.223	0.0001	-13.59	<.0001	92.31	0.001
	2	-202.367	0.0001	-8.14	<.0001	33.17	0.001
	3	-2145.33	0.0001	-9.63	<.0001	46.37	0.001
	4	588.5768	0.9999	-9.32	<.0001	43.45	0.001
	5	309.9895	0.9999	-8.57	<.0001	36.7	0.001

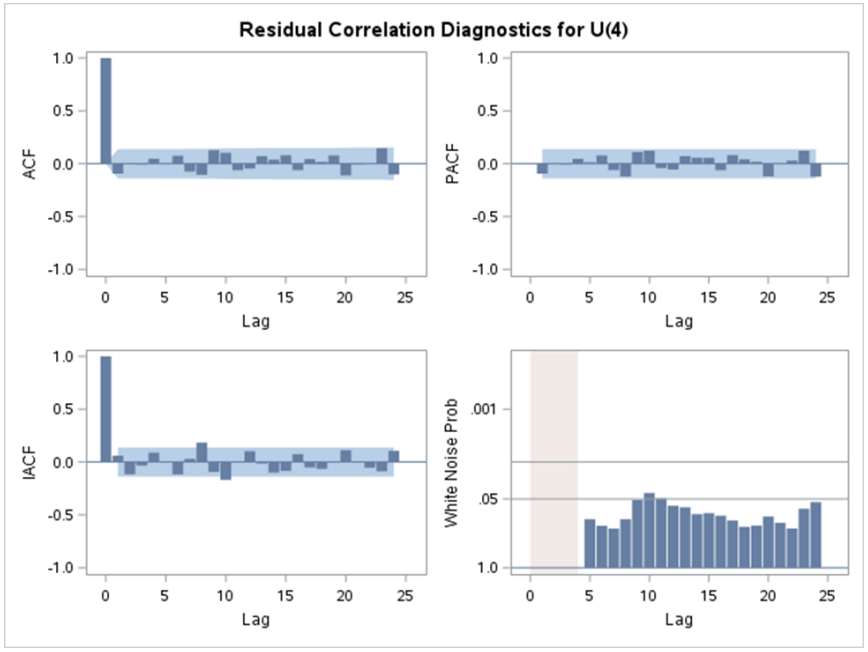


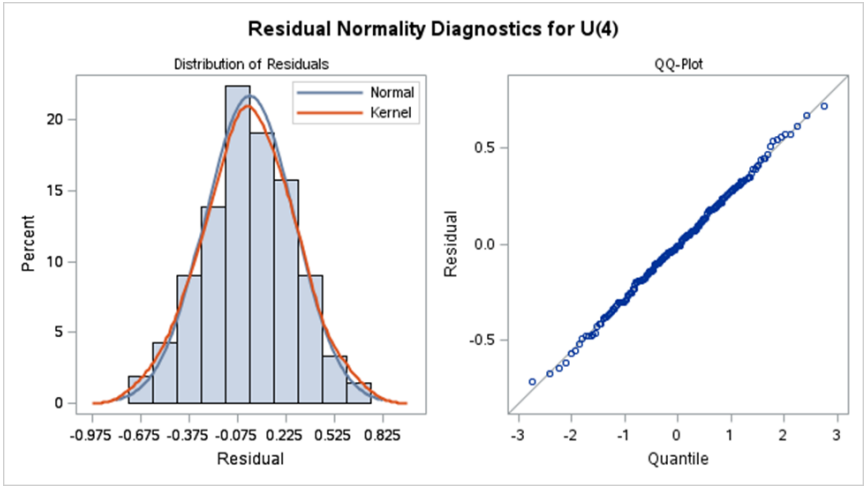
Model 2: Identification and Fitting

ARMA(p+d,q) Tentative Order Selection Tests		
ESACF		
p+d	q	BIC
0	4	-2.56656
3	4	-2.52882
4	4	-2.52207
2	5	-2.51717

Maximum Likelihood Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	-0.0013644	0.0053633	-0.25	0.7992	0
MA1,1	0.1386	0.059	2.35	0.0188	1
MA1,2	0.19274	0.05958	3.24	0.0012	2
MA1,3	-0.15463	0.05955	-2.6	0.0094	3
MA1,4	0.55743	0.06067	9.19	<.0001	4

Model 2: Diagnostics

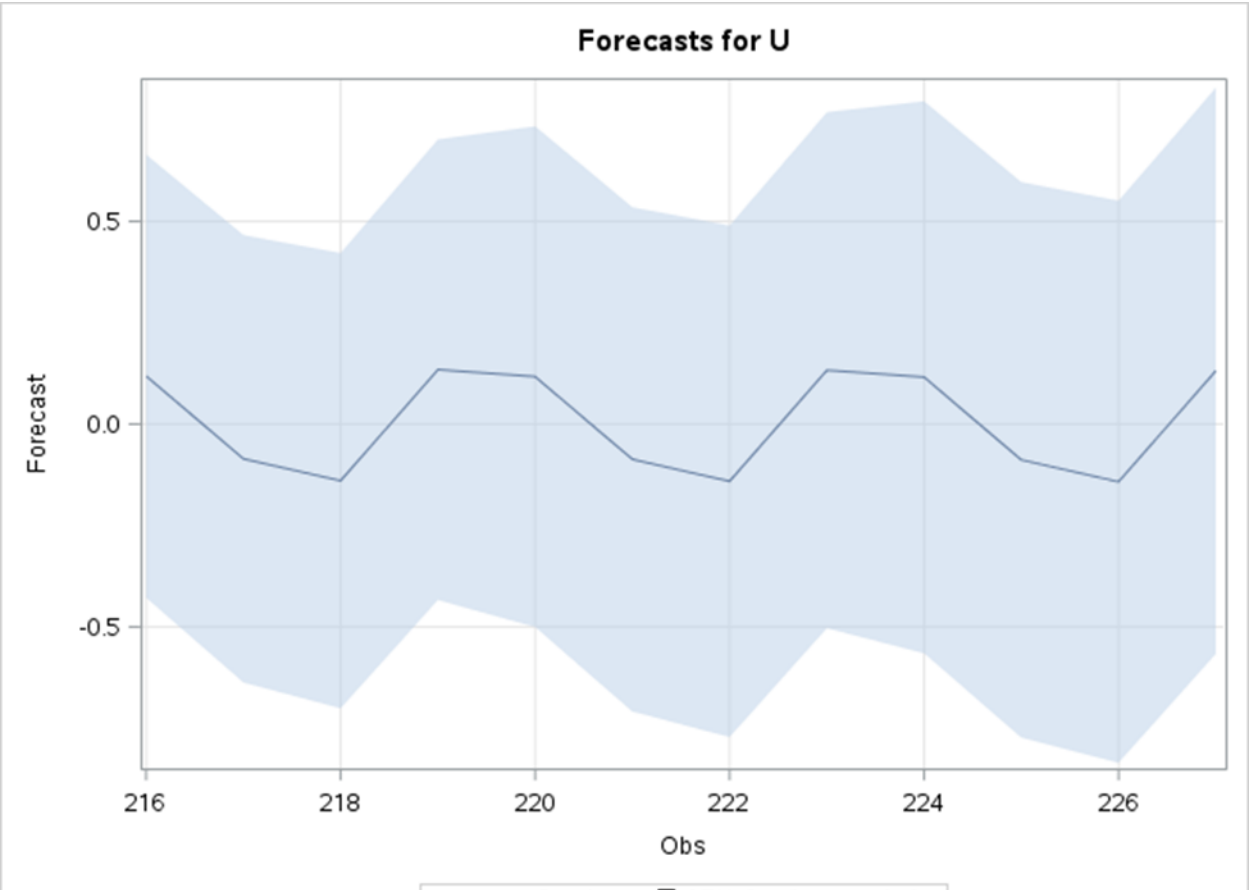




Model 2: Forecast

$$(1 - \mathrm{B}^4)U_t = -0.00136 + (1 - 0.1386\,\mathrm{B} - 0.19274\,\mathrm{B}^2 + 0.15463\mathrm{B}^3 - 0.55743\mathrm{B}^4)a_t$$

$$(1 - \mathrm{B}^4)(Y_t - Y_{t-1}) = -0.00136 + (1 - 0.1386\,\mathrm{B} - 0.19274\,\mathrm{B}^2 + 0.15463\mathrm{B}^3 - 0.55743\mathrm{B}^4)a_t$$



Q4-2018	64.5184
Q1-2019	64.4329
Q2-2019	64.2934
Q3-2019	64.4277
Q4-2019	64.5448
Q1-2020	64.4579
Q2-2020	64.317
Q3-2020	64.4499
Q4-2020	64.5656
Q1-2021	64.4773
Q2-2021	64.335
Q1-2019	64.4665

- Model identification on Original Series (Y)

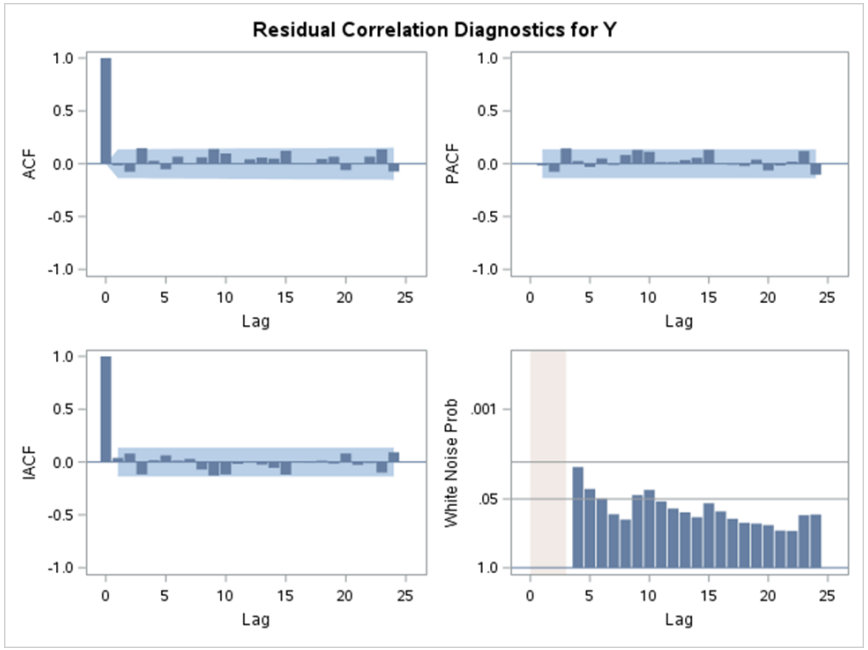
ARMA(p+d,q) Tentative Order Selection Tests		
ESACF		
p+d	q	BIC
1	4	-2.61104
3	4	-2.56748
4	4	-2.54265
5	4	-2.63692

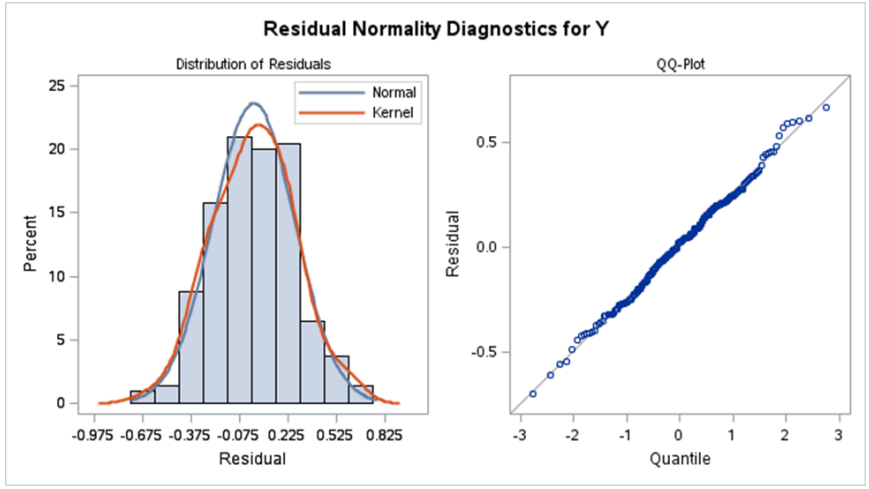
Model for Y	AIC	SBC	Variance Estimate	Note
ARMA(1,4)	55.90339	76.12722	0.072441	Insig coeff, residuals not white noise
ARMA(1,(1,2,4)) INPUT=(T C1 S1 C2)	39.43369	69.76943	0.06631	Insig coeff
ARMA(1,(1,4)) INPUT=(T C1 S1 C2)	38.39846	65.36356	0.066292	Insig coeff of T
ARMA(1,(1,4)) INPUT=(C1 S1 C2)	37.17752	60.77199	0.066144	

Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	64.59498	1.0434	61.91	<.0001	0	Y	0
MA1,1	0.18178	0.06643	2.74	0.0062	1	Y	0
MA1,2	-0.27021	0.06685	-4.04	<.0001	4	Y	0
AR1,1	0.9859	0.01029	95.83	<.0001	1	Y	0
NUM1	0.04422	0.02243	1.97	0.0487	0	C1	0
NUM2	-0.10088	0.0224	-4.5	<.0001	0	S1	0
NUM3	-0.02695	0.01268	-2.12	0.0336	0	C2	0

Model 3: Deterministic Seasonal Model

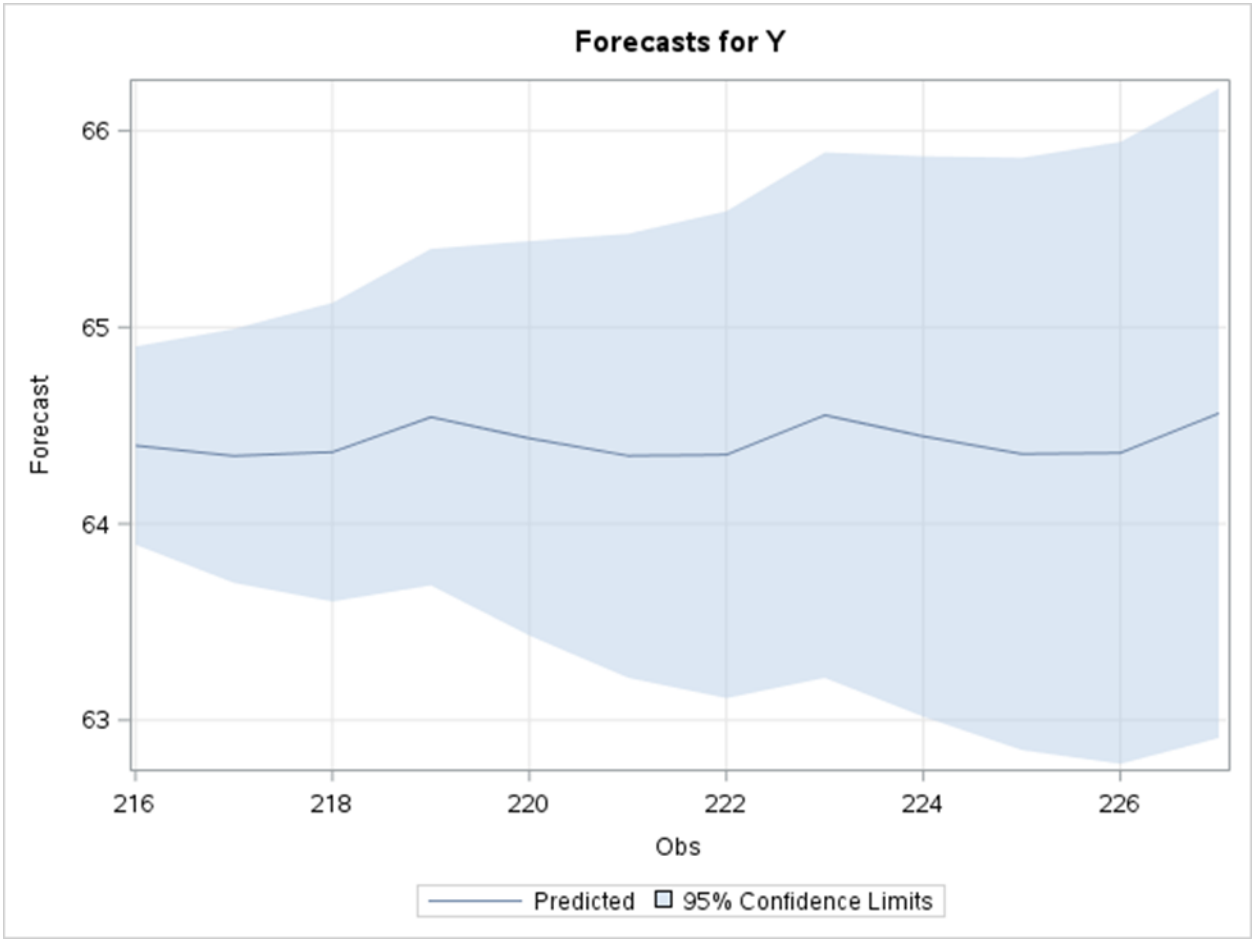
Model 3: Diagnostics





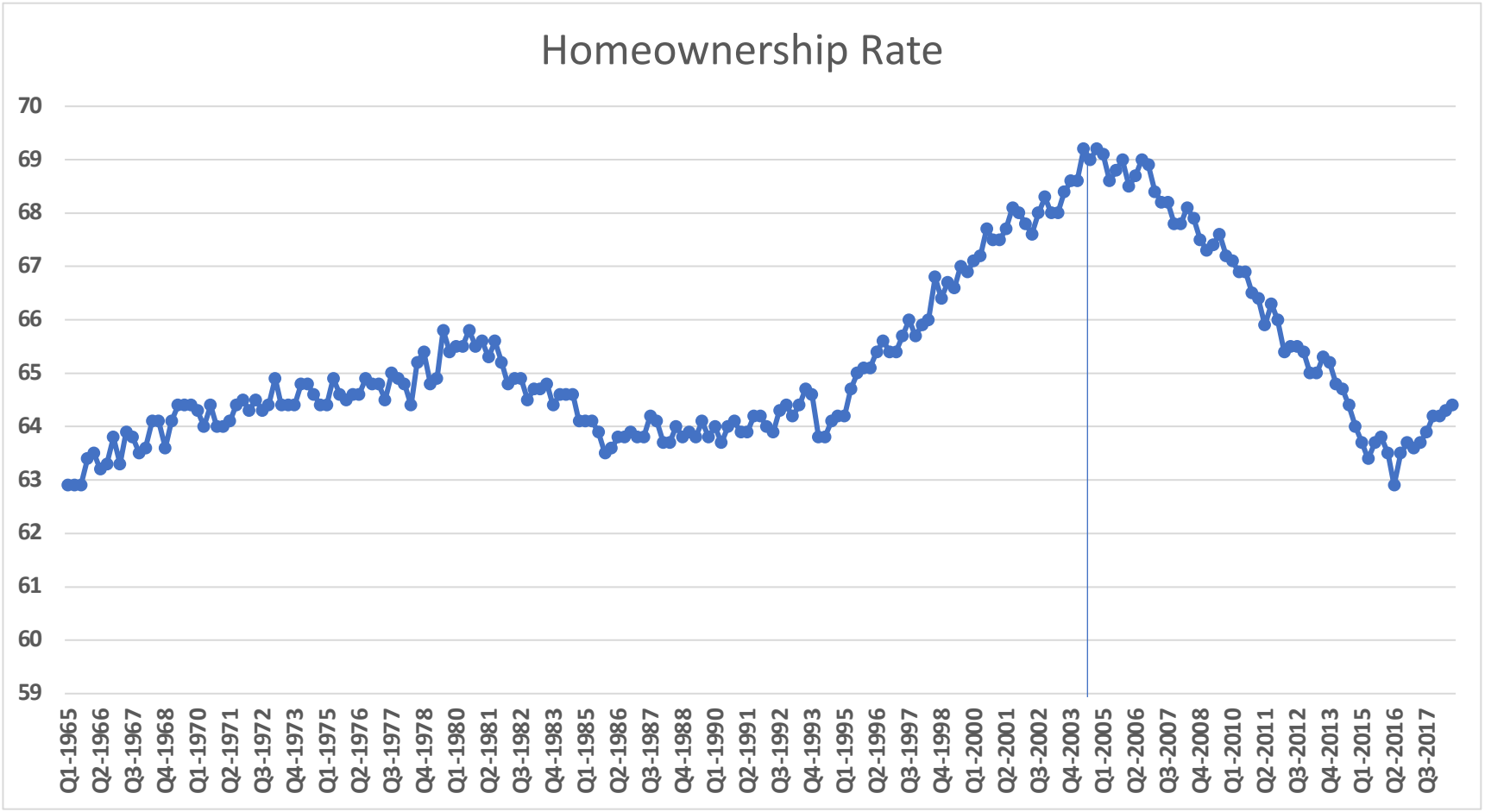
Model 3: Forecast

$$Y_t = 0.910619 + 0.044223\cos\left(\frac{\pi}{2}t\right) - 0.10088\ \sin\left(\frac{\pi}{2}t\right) - 0.02695\cos(\pi t) + \frac{1 - 0.18178\,\mathrm{B} + 0.27021\,B^4}{1 - 0.9859\,\mathrm{B}}a_t$$



Q4-2018	64.398
Q1-2019	64.3455
Q2-2019	64.3646
Q3-2019	64.5429
Q4-2019	64.4348
Q1-2020	64.3462
Q2-2020	64.3514
Q3-2020	64.5528
Q4-2020	64.4446
Q1-2021	64.3558
Q2-2021	64.3609
Q1-2019	64.5622

Model 4: Intervention Analysis



Intervention Analysis

- In 2000 (Before intervention)
  - Fed interest rates dropped to 1 percent to deal with dot com bust
  - Mortgage-based securities became more profitable than traditional bonds and Treasuries.

- In 2004 (Intervention)
  - Securities and Exchange Commission changed the leverage rules for five Wall Street banks, that allowed more risky lending
  - The Office of the Comptroller of the Currency federally overrode anti-predatory State laws

After Intervention

Between 2006 and 2014,

Fed interest rate raised again, unmanageable cost of ownership

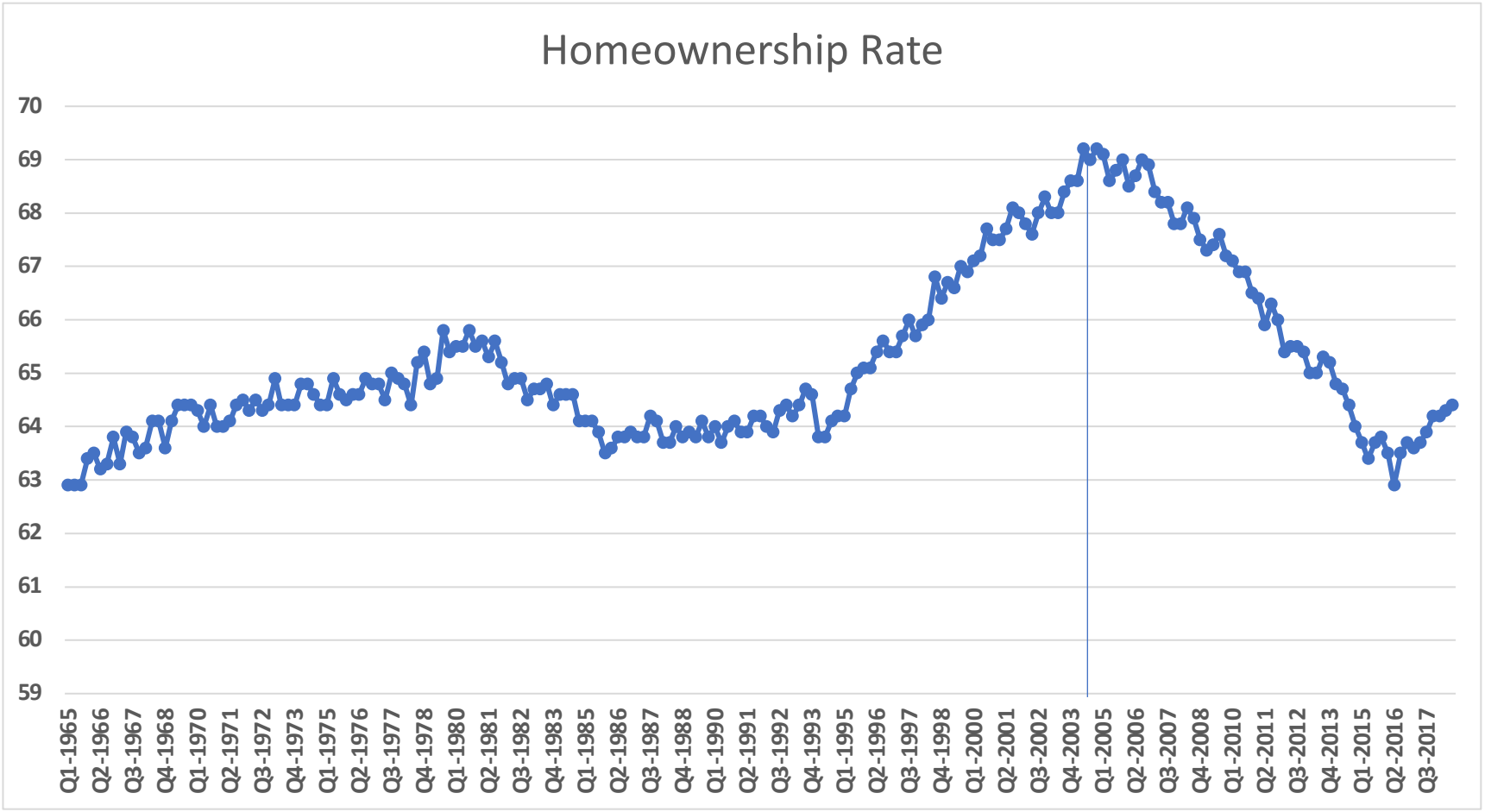
Nearly 10 million homeowners in US experienced

- Foreclosure
- Short-sell their property as quickly as possible.

Only 2 out of 5 Wall Street banks survived, that also due to bail out.

Homeownership Rate kept falling between this period

Model 4: Intervention Analysis

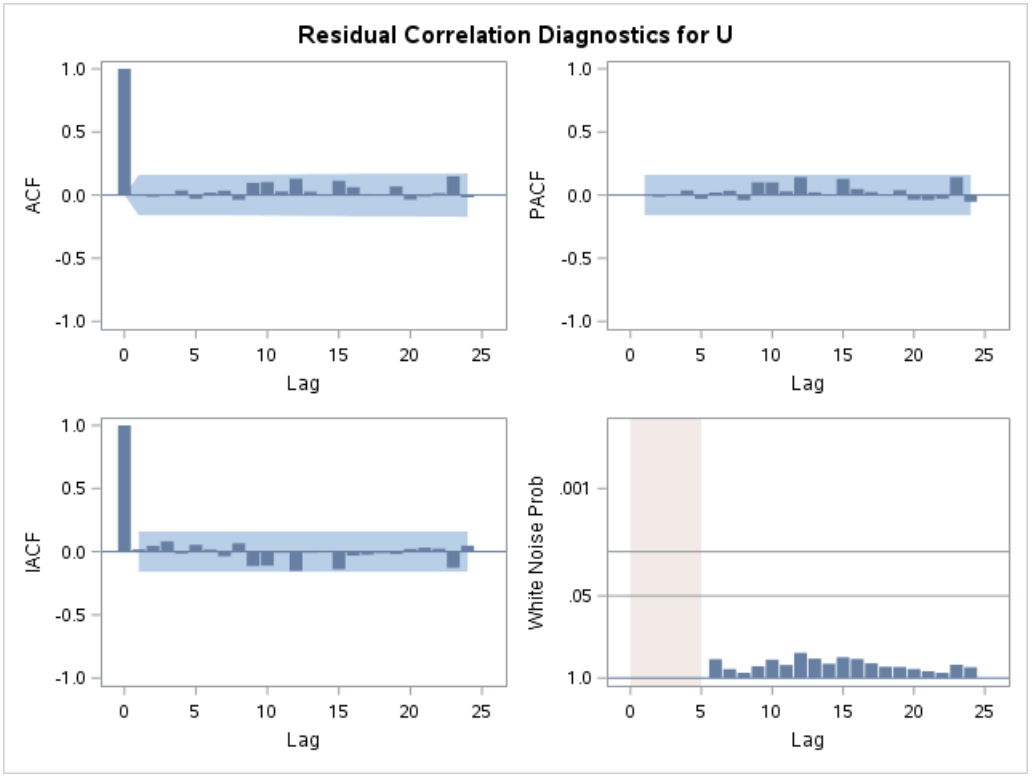


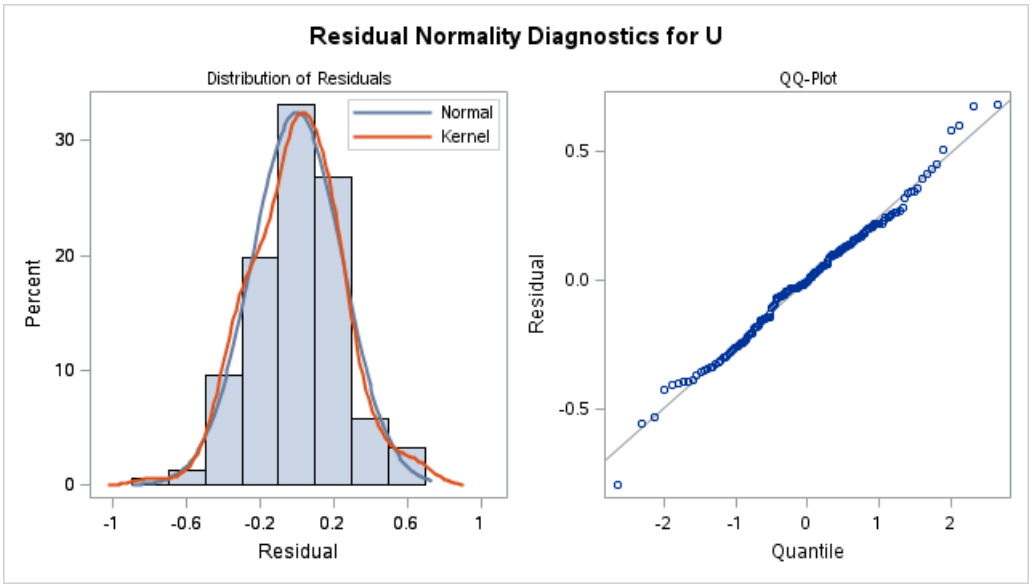
Analysis on Pre-intervention

ARMA(p+d,q) Tentative Order Selection Tests		
ESACF		
p+d	q	BIC
4	3	-2.69814
0	4	-2.68695
3	4	-2.67608
5	3	-2.67919
2	5	-2.6673

Model for U	AIC	SBC	Variance Estimate	Note
ARMA(4,3)	**21.17691**	**45.62688**	**0.063258**	**Insig coeff**
AR ((1,4))	**24.34249**	**33.51123**	**0.066679**	**Residuals not good white noise**
ARMA((1,3,4),(3))	**23.28071**	**38.56194**	**0.065306**	
ARMA((1,4),(3))	**23.88535**	**36.11033**	**0.066034**	**Insig coeff**
ARMA(4,(3))	**17.35414**	**35.69162**	**0.062501**	

Residual Diagnostics for Pre-intervention Model





Analysis on Full Series

Intervention model	AIC	SBC	Variance Estimate	Note
(/(1)S)	34.04859	60.9764	0.065904	Insig coeff
(/(2)S)	34.72146	61.6118	0.066147	Insig coeff
(/(3)S)	35.46509	62.31778	0.066416	Insig coeff
(/(4)S)	31.16093	57.9758	0.0651	
(1$S)	32.45399	56.01582	0.06571	
(1$/(4)S)	33.07999	59.85685	0.06573	
(S)	32.08129	55.64312	0.065595	

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.04255	0.01909	2.23	0.0258	0	U	0
MA1,1	-0.30172	0.142	-2.12	0.0336	3	U	0
AR1,1	-0.2613	0.067	-3.9	<.0001	1	U	0
AR1,2	-0.1945	0.06667	-2.92	0.0035	2	U	0
AR1,3	-0.23945	0.13271	-1.8	0.0712	3	U	0
AR1,4	0.30919	0.07517	4.11	<.0001	4	U	0
NUM1	-0.12817	0.03664	-3.5	0.0005	0	S	0

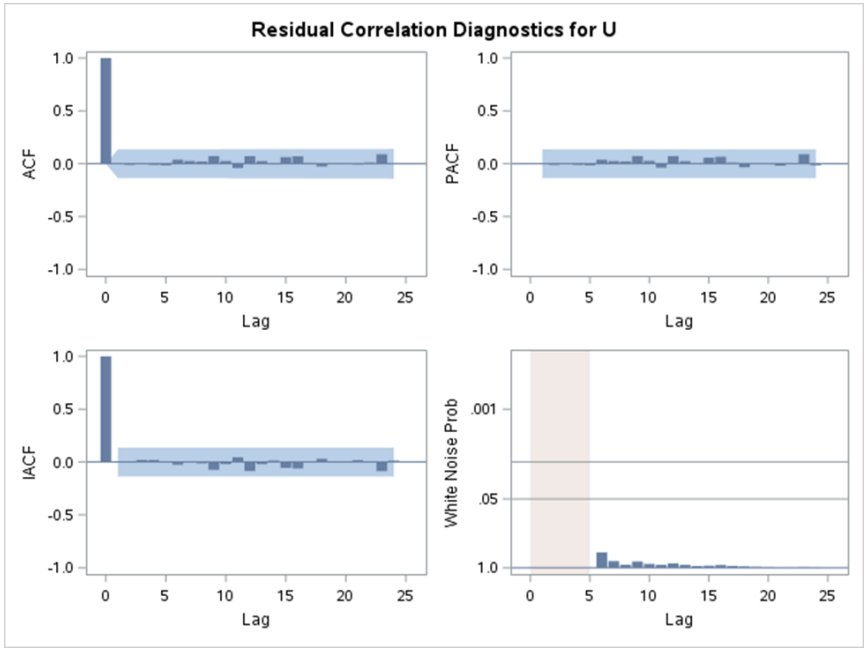
Model 4: Intervention Analysis

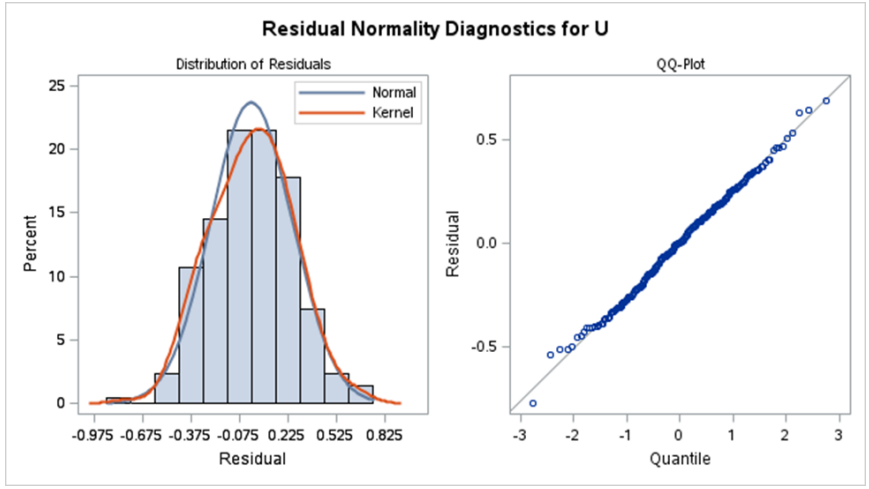
$$U_t = 0.042546 - 0.12817 S_t^{(216)} + \frac{1 + 0.30172\ \mathrm{B}^3}{1 + 0.2613\ \mathrm{B} + 0.1945\ \mathrm{B}^2 + 0.23945\ \mathrm{B}^3 - 0.30919\ \mathrm{B}^4} a_t$$

$$Y_t = Y_{t-1} + 0.042546 - 0.12817 S_t^{(216)} + \frac{1 + 0.30172\ \mathrm{B}^3}{1 + 0.2613\ \mathrm{B} + 0.1945\ \mathrm{B}^2 + 0.23945\ \mathrm{B}^3 - 0.30919\ \mathrm{B}^4} a_t$$

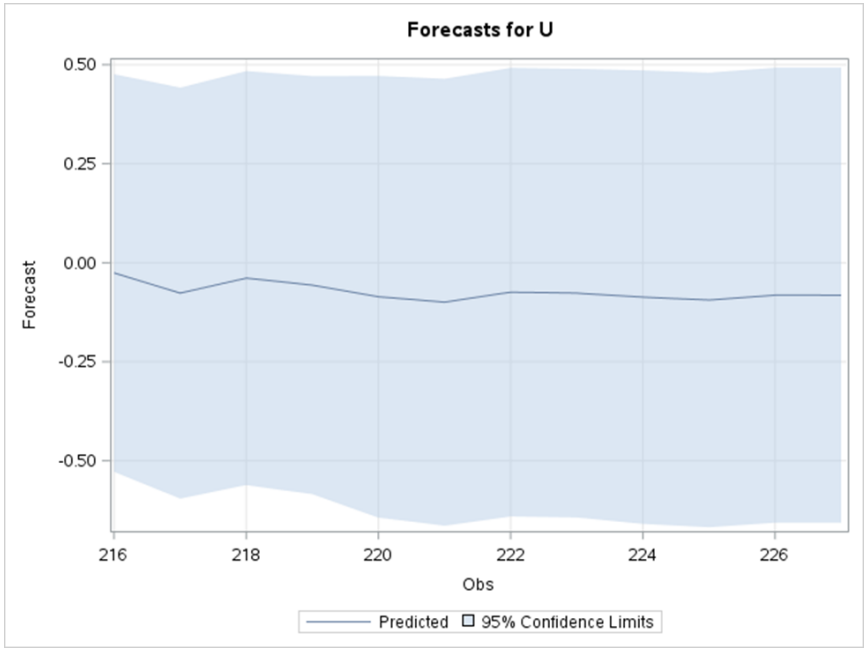
- It tells us that there was significant impact of intervention that led to risky subprime lending, and homeownership fell.
- Rented accommodation or other housing options looked more promising.

Diagnostics for Full series





Q4-2018	64.3742
Q1-2019	64.2975
Q2-2019	64.2588
Q3-2019	64.2022
Q4-2019	64.1162
Q1-2020	64.0165
Q2-2020	63.9421
Q3-2020	63.8653
Q4-2020	63.7784
Q1-2021	63.6843
Q2-2021	63.6025
Q1-2019	63.5206



Model 4: Forecast

Conclusion

Comparison between models

Let the "distance" between models i and j be

$$D_{ij} = \sum_{t=216}^{227} |Y_i(t) - Y_j(t)| \qquad (i = 1,2,3,4)$$

D_{23} =1.1173 and D_{14} =8.6888

Model 2 and 3 are the "closest"

Model 1 and 4 are the "farthest"

However, they are all close enough, good models.

Model 1	Model 2	Model 3	Model 4
64.5266	64.5184	64.398	64.3742
64.5324	64.4329	64.3455	64.2975
64.5811	64.2934	64.3646	64.2588
64.6298	64.4277	64.5429	64.2022
64.6899	64.5448	64.4348	64.1162
64.6982	64.4579	64.3462	64.0165
64.7249	64.317	64.3514	63.9421
64.7516	64.4499	64.5528	63.8653
64.7832	64.5656	64.4446	63.7784
64.7926	64.4773	64.3558	63.6843
64.8099	64.335	64.3609	63.6025

Thank You and Questions