GOV 52: Replication Project

Liam Hall

2021 May 08

Introduction

The original paper *The Dynamic American Dream*, coauthored by Jennifer Wolak and David A.M. Peterson (2020) and published in the *American Journal of Political Science*[1], analyzes behavior of belief in the American dream over time, and as it relates to economic inequality, social mobility, home ownership, public policy mood (that is, the public demand for liberal policy outcomes), consumer confidence, and the presence of U.S. midterm and presidential elections. The condition of the American dream has long been seen as a barometer of America's perceived condition - a measure of optimism in America's future - and American politicians on both sides of the aisle have historically appealed to the American Dream when discussing their vision of the nation. When times are good, the American Dream is alive as ever. When times are tough, politicians insist that change must happen, that we "need to write a new chapter in American Dream" and "bring it back bigger and stronger and more powerfully than ever before." Wolak and Peterson take a more objective approach in assessing the dynamics of the American Dream. Instead of relying on abstract terms, the authors aggregate responses from quarterly surveys of American citizens from 1973 to 2018 to quantify the percentage of Americans who believe the American Dream is achievable, exploring its dynamics relative to temporal economic, social, and political conditions. Further, the authors explore the responsiveness of belief in the American Dream to these short-term variables in the long run, displaying their findings in both numerical and graphical mediums.

Model Replication and Interpretations

Because the model's variables - belief in the American dream, economic inequality, social mobility, home ownership, public policy mood (that is, the public demand for liberal policy outcomes), consumer confidence, and the presence of U.S. midterm and presidential elections - include both stationary and non-stationary cointegrated time series, the authors employed a generalized error correction model (GECM) that is able to

[1] The replication data was available through the Harvard dataverse. You can access the files here.

effectively combine these different variables (Bannerjee et al. 1993; DeBoef and Keele 2008). Because of the high correlation between social mobility and economic inequality, the authors split the data into two models, one with the gini coefficient variable and one with the social mobility variable, but which are otherwise identical. The GECM model, like the auto-regressive distributed lag model, uses the equation:

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_0 \Delta x_t + \epsilon_t$$

Extending the findings from the GECM models, the original paper also looks at the extent to which the non-stationary explanatory variables have a long run relationship with belief in the American dream, as reflected by long run multipliers and their respective t-values. Ultimately, the authors found statistically significant, positive correlations between belief in the American dream and social mobility, home ownership, and consumer confidence (measured by index of consumer sentiment) and a statistically significant, negative relationship between belief in the American Dream and economic inequality. Intuitively, these findings make sense, because as the economy becomes more lopsided and social mobility wanes, Americans are naturally less optimistic about opportunity in America. Interestingly, the authors also find a positive relationship between election cycles and belief in the American dream, suggesting that non-policy political influences might increase Americans' national optimism, too, however these findings are inconsistently significant.

In this replication, I translate the originally-Latex paper into R and reconstruct the plots and models of the original authors in their entirety. Because variable stationarity affects the GECM model structure, I diverge from the original paper to explore and confirm the stationarity of the variables used in the original GECM model to confirm the model was properly constructed. I also bootstrap the model errors and display the distribution of belief changes in the American dream between time periods (that is, I show how much belief in the American dream changes between time periods to see if the fluctuations might be due to slight variations of human error).

My replication of the models ultimately yielded the same numerical values and, therefore, the same conclusions as the original paper. I explain the different calculations and interpret some of model outputs with some visual aids I constructed from the data.

The model reconstruction was relatively straightforward. I manually created the lag and delta variables, by using the previous time period's values and taking the change in value between consecutive time periods, and input these newly-created variables in the GECM. Because these GECM variables were all calculated manually, I was able to use the lm() function, as opposed to some of the GECM specific functions within the ecm and ARDL packages. Thus, my GECM took the form of a linear regression of the lagged and delta non-stationary variables and the original stationary variables. This yielded the coefficient, standard error, and p-value statistics that are displayed in Table 1. Extending from the original models, I also bootstrapped the variables' standard errors and output them in table 2.

The long run multipliers were calculated with the equation $LRM_x = -\beta_x/\alpha_1$, with inputs taken from the GECM model outputs. There are two primary ways to calculate the standard errors of long run multipliers: the delta method and the Bewley transformation, which yield asymptotically accurate values (Nieman & Petersen, 2019). Because the two methods are roughly equal in accuracy, I used the delta method, which used the model covariance and the LRM value to calculate the LRM standard errors. The respective LRM t-values were simply the ratio of the LRM to the LRM's standard error (LRM_x/ϵ_{LRM_x}). These methods were also used by the original authors and yielded the same results as the original paper.

Looking at top half of the model outputs, we can see a lagged home ownership coefficient for the first model of .932 and a delta coefficient of .877. With this, we would expect that a 1% increase in the home ownership rate, all else equal, would increase belief in the American Dream by 2.4 percent. Looking at the LRM values, we can interpret that, for example, an increase of 0.00525 in the gini coefficient (which was the largest between quarters) would lead to a .57 percent decline in belief in the American dream.

Figure 3 plots the estimated lag distributions for belief in the American Dream across four explanatory variables. In other words, the plots show how responsive the American dream variable is to a shift in the each of the four explanatory variables plotted. As we expect, the effect of a shift in a predictor has diminishing effects as time passes, as shown in the lines aymptotically approaching zero. This is encouraging, but non-conclusive, evidence that the GECM model was correctly constructed.

You can access the code for this replication here.

Extensions

As mentioned earlier, treating the respective stationarities of the different variables properly is crucial in yielding accurate results from the GECM model. However, as Webb, Linn, and Lebo (2019, 2020) have explored and documented, the many different tests for stationarity are low-powered and can yield unreliable stationarity assessments. Instead of relying on these tests, Wolak and Peterson judged the stationarity of the variables post-hoc, by using the methods suggested by Webb, Linn, and Lebo to assess their LRM significances: For the given model with k = 2 and n = 150, t-values above the upper bound of |3.56| were deemed statistically significant with t-values below |3.56| deemed insignificant. I applied the same reasoning in my own analysis and reached the same conclusions: that the gini coefficient, social mobility, home ownership rates, and consumer confidence all had significant long term relationships with belief in the American dream. In terms of stationarity, this would suggest that the variables with a statistically significant

Table 1: Explaining Belief in the American Dream

	Dependent variable:	
	Δ in Belief in the American Dream	
	Model 1	Model 2
Belief in the American Dream$_{t-1}$	-0.386^{***}	-0.388^{***}
	(0.057)	(0.058)
Δ Gini coefficient	-315.981	
	(251.617)	
Gini coefficient$_{t-1}$	-41.727^{***}	
	(11.932)	
Δ Social mobility		105.564^{*}
		(61.247)
Social mobility$_{t-1}$		10.237^{***}
		(3.019)
Δ Homeownership	0.877	0.827
	(0.865)	(0.916)
Homeownership$_{t-1}$	0.932^{***}	0.788^{***}
	(0.230)	(0.219)
Δ Policy mood	-0.146	-0.116
	(0.143)	(0.152)
Policy mood$_{t-1}$	0.092	0.121^{*}
	(0.069)	(0.071)
Δ Index of consumer sentiment	-0.025	-0.032
	(0.050)	(0.051)
Index of consumer sentiment$_{t-1}$	0.097^{***}	0.088^{***}
	(0.024)	(0.025)
Midterm election	1.203	1.204
	(1.043)	(1.066)
Presidential campaign	0.535^{**}	0.382^{*}
	(0.221)	(0.221)
Constant	-38.657^{***}	-54.838^{***}
	(12.240)	(15.443)
Long run multipliers		
LRM, Gini coefficient	-108.1521†	
standard error	(26.9571)	
t-value	-4.0120	
LRM, Social mobility		26.3578†
standard error		(6.9230)
t-value		3.8073
LRM, Home ownership	2.4165†	2.0295†
standard error	(.4556)	(.4571)
t-value	5.3041	4.4403
LRM, Policy mood	.2394	.3111
standard error	(.1750)	(.1791)
t-value	1.3681	1.7370
LRM, Index of consumer sentiment	.2527†	.2258†
standard error	(.0578)	(.0560)
t-value	4.3723	3.7635
Observations	175	167
R^2	0.267	0.270

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

† denotes significant LRMs, where t-value exceeds absolute value of 3.560

standard errors wrapped in parentheses

long run relationship with American dream were also non-stationary, since a stationary variable would be too consistent to have a long run multiplier effect. However, extending from the original model, I also measured the stationarity of variables by the Phillips-Perron Test, documented in Table 4, and visualize and compare the stationarity of two different variables, gini coefficient and presidential campaign cycle. More on this in a moment.

The first of my extensions was to bootstrap the standard errors and compare them to the standard errors output by the two models. Tables 2 and 3 displays the comparison between the different calculated standard errors. This extension does not help develop greater understanding of the original paper, but was a different way of calculating the standard errors than the original paper used. Ultimately, as you can see in the tables, the errors were very similar between the two methods.

The second and more notable extension was to examine the stationarity of the different variables. Table 3 displays each variables's results for the Phillips-Perron Stationarity test. As mentioned earlier, the Phillips-Perron test as well as the other stationarity tests yield stationarity assessments of varying accuracy. The purpose of including the Phillips-Perron test output in this case was not to arrive at a conclusive stationarity assessment, but to compare the assessment results with the different variables' plots and their residuals, which is graphed in Figure 2, with the hope that pairing the test conclusions with the visuals might help suggest the stationarity of the variable.

Figure 1 is a plot of belief in the American Dream over time, as was included in the original paper but with a linear regression overlay. If we think of stationarity as the degree to which the values are stationary, or fixed and predictable in value, then we would image a variable behaving in a consistent and reliable way. The three types of stationarity are based on how consistent and reliable the variable's behavior is. The three types of stationarity are no drift (i.e. oscillation or wavering) and no trend (i.e. linear regression with slope of 0), with drift and no trend, and drift and trend. Table 4 displays the Phillips-Perron stationarity assessments for each variable, along with the each variables' p-values for these different types of stationarity. According to the test, when the p-value is above .05, we reject the null hypothesis and assume non-stationarity. Rather than going through each individual variable, I chose to look at one variable that is treated like a stationary variable in the model, presidential campaign cycle, and one that's treated like a non-stationary variable, the gini coefficient (see figure 2). In class, we used residual plots to assess model fits. I employed residual plots in this case to help visualize potential stationarity.

The gini coefficient's Phillips-Perron p-values suggest that the variable is non-stationary in all types, because it's p-values all exceed .05. Looking at the gini plots in figure 2, the variable's behavior visually matches what we would expect from the test's assessments: a steady non-zero positive trend and its residuals do not evenly oscillate around 0, suggesting the variable is in fact non stationary. Turning to the presidential campaign variable, the test assessments suggest the variable is stationary in all three types. However, looking at the presidential campaign plots in figure 2, we see that while the linear regression has a slope of around 0, the residuals do not oscillate around 0 evenly. We also know that the presidential campaign cycle occurs every 4 years, and therefore its behavior would be determined from time. In other words, there is seasonality, which would suggest the variable is not stationary by the type 1 standards, but is stationary by the type 2 and type 3 standards. For the purposes of the GECM model, the gini coefficient should be treated like a non-stationary variable and the presidential campaign variable (prezcamp) should be treated as a stationary variable, which the original authors did, albeit due to different reasoning mentioned earlier (i.e. by checking the statistical significance of its LRM according to its t-value, because stationary variables cannot have a long run effect - they don't change enough to cause one).

The third extension I did was to display the distribution of changes in American dream belief between time periods (quarter years). Because very small oscillations around a mean would suggest that belief in the American dream is not truly correlated with the variables as much as the model might suggest (i.e. the variables' behavior might be more determined by human error than statistical significance). I found that the distribution of change in belief in the American dream between quarters was roughly similar to a normal distribution with its standard deviation of 3.660072 and average value of -.0202787, suggesting that the variations in American dream values were likely sizeable enough to have a statistically significant relationship with the explanatory variables.

Table 2: Model One Standard Error vs. Bootstrapped Standard Errors

	original	bootstrapped	pct_diff
(Intercept)	12.24	11.38	0.07
am_lag	0.06	0.06	0.02
gini_delta	251.62	244.98	0.03
gini_lag	11.93	12.28	-0.03
home_delta	0.87	0.84	0.03
home_lag	0.23	0.20	0.12
mood_delta	0.14	0.15	-0.03
mood_lag	0.07	0.07	0.02
ics_delta	0.05	0.06	-0.23
ics_lag	0.02	0.03	-0.07
midterm	1.04	1.22	-0.17
prezcamp	0.22	0.22	-0.01

Table 3: Model Two Standard Error vs. Bootstrapped Standard Errors

	original	bootstrapped	pct_diff
(Intercept)	15.44	15.65	-0.01
am_lag	0.06	0.06	-0.02
soc_delta	61.25	60.82	0.01
soc_lag	3.02	3.52	-0.16
home_delta	0.92	0.91	0.00
home_lag	0.22	0.22	0.01
mood_delta	0.15	0.16	-0.03
mood_lag	0.07	0.07	0.06
ics_delta	0.05	0.06	-0.18
ics_lag	0.02	0.03	-0.10
midterm	1.07	1.33	-0.25
prezcamp	0.22	0.22	0.01

Figure 1: The Dynamics of Public Belief in the American Dream

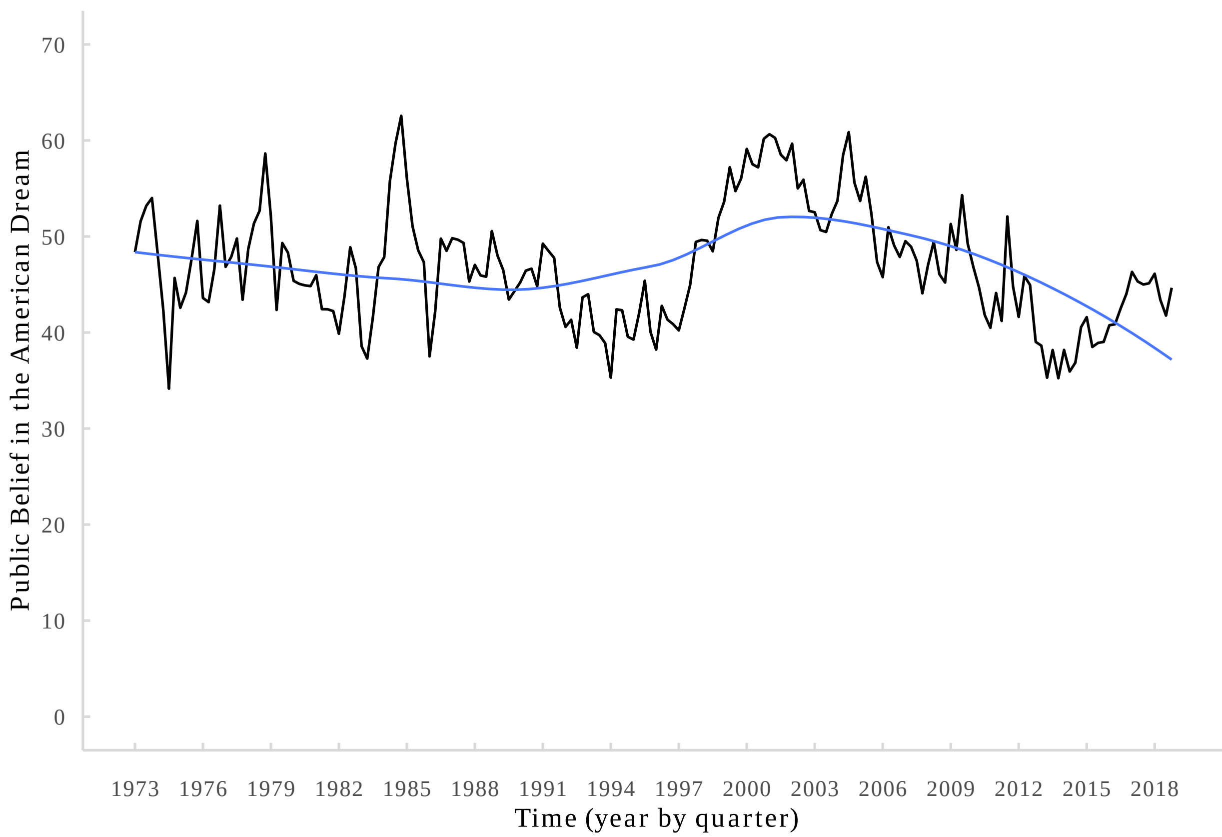


Table 4: Variable Stationarities

Variable	Z_rho	p-value
American Dream		
Type 1: no drift, no trend	-0.4229186	0.5960181
Type 2: with drift, no trend	-29.2144709	0.0100000
Type 3: with drift, and trend	-29.5016390	0.0100000
Gini Coefficient		
Type 1: no drift, no trend	0.1815872	0.7303527
Type 2: with drift, no trend	-0.5720620	0.9183873
Type 3: with drift, and trend	-7.9586331	0.5796473
Social Mobility		
Type 1: no drift, no trend	-0.6523428	0.5449164
Type 2: with drift, no trend	-1.7047092	0.8003862
Type 3: with drift, and trend	-3.7382584	0.9004809
Home Ownership		
Type 1: no drift, no trend	-0.0193342	0.6857035
Type 2: with drift, no trend	-1.7416056	0.7962397
Type 3: with drift, and trend	-0.5692698	0.9900000
Policy Mood		
Type 1: no drift, no trend	-0.0885227	0.6703283
Type 2: with drift, no trend	-19.2059513	0.0135664
Type 3: with drift, and trend	-19.5549765	0.0714619
Consumer Confidence		
Type 1: no drift, no trend	-0.1562859	0.6552698
Type 2: with drift, no trend	-15.7278998	0.0321491
Type 3: with drift, and trend	-16.3027041	0.1642804
Midterm Election		
Type 1: no drift, no trend	-175.0000000	0.0100000
Type 2: with drift, no trend	-169.7220028	0.0100000
Type 3: with drift, and trend	-169.7124886	0.0100000
Presidential Election		
Type 1: no drift, no trend	-53.4972563	0.0100000
Type 2: with drift, no trend	-67.5167961	0.0100000
Type 3: with drift, and trend	-67.8242200	0.0100000

Figure 2: Gini and Presidential Campaign Stationarities

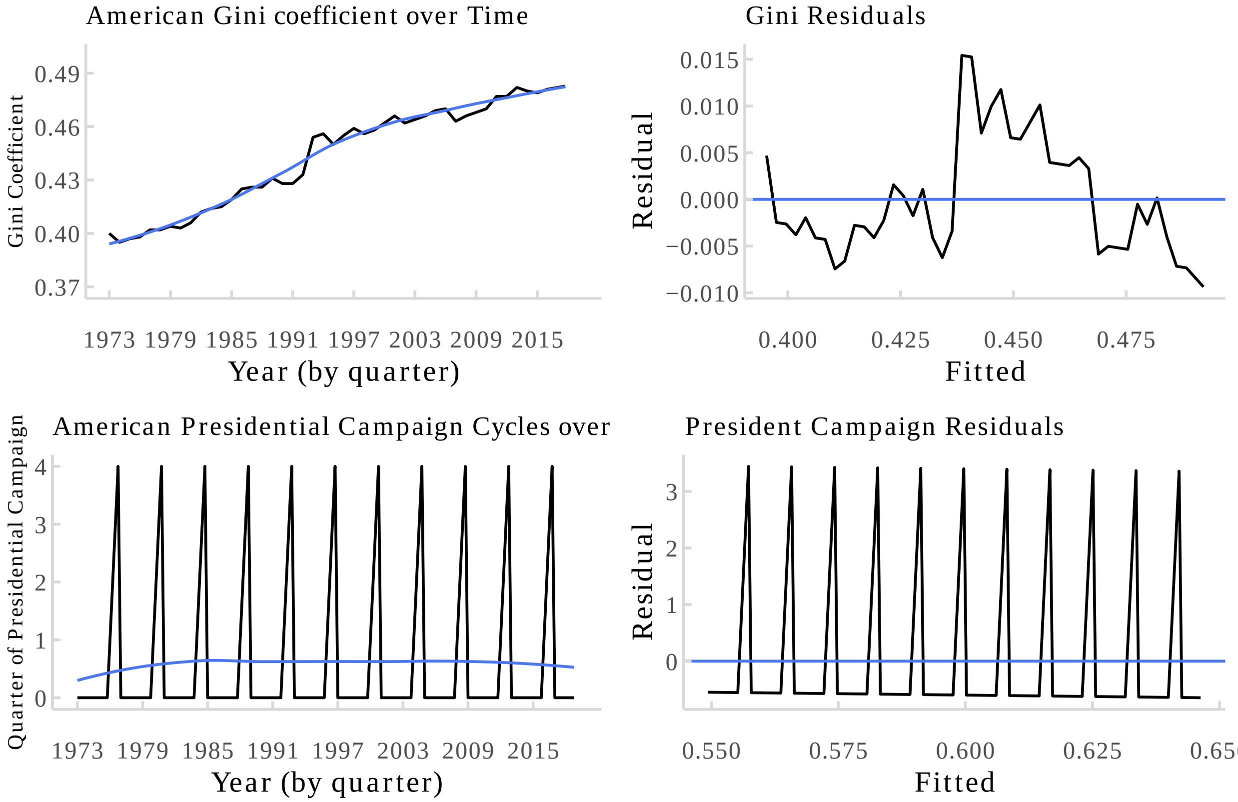


Figure 3: Estimated Lag Distributions for Belief in the American Dream

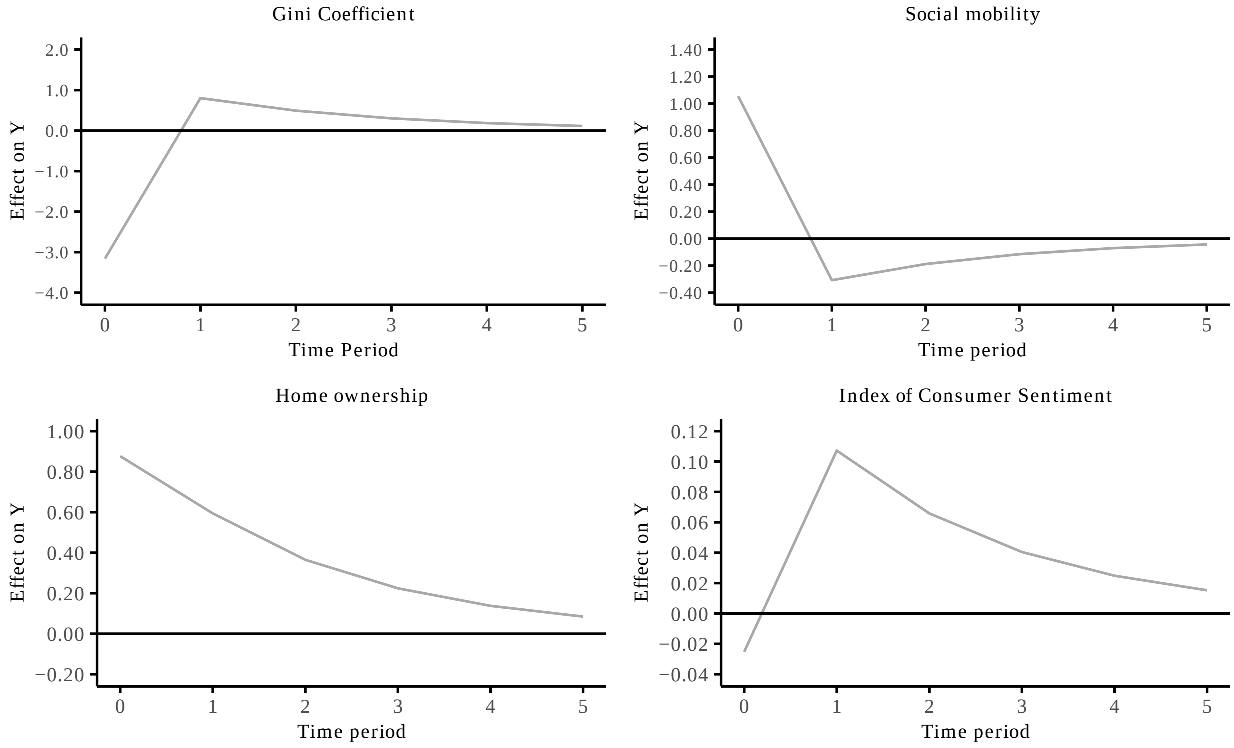
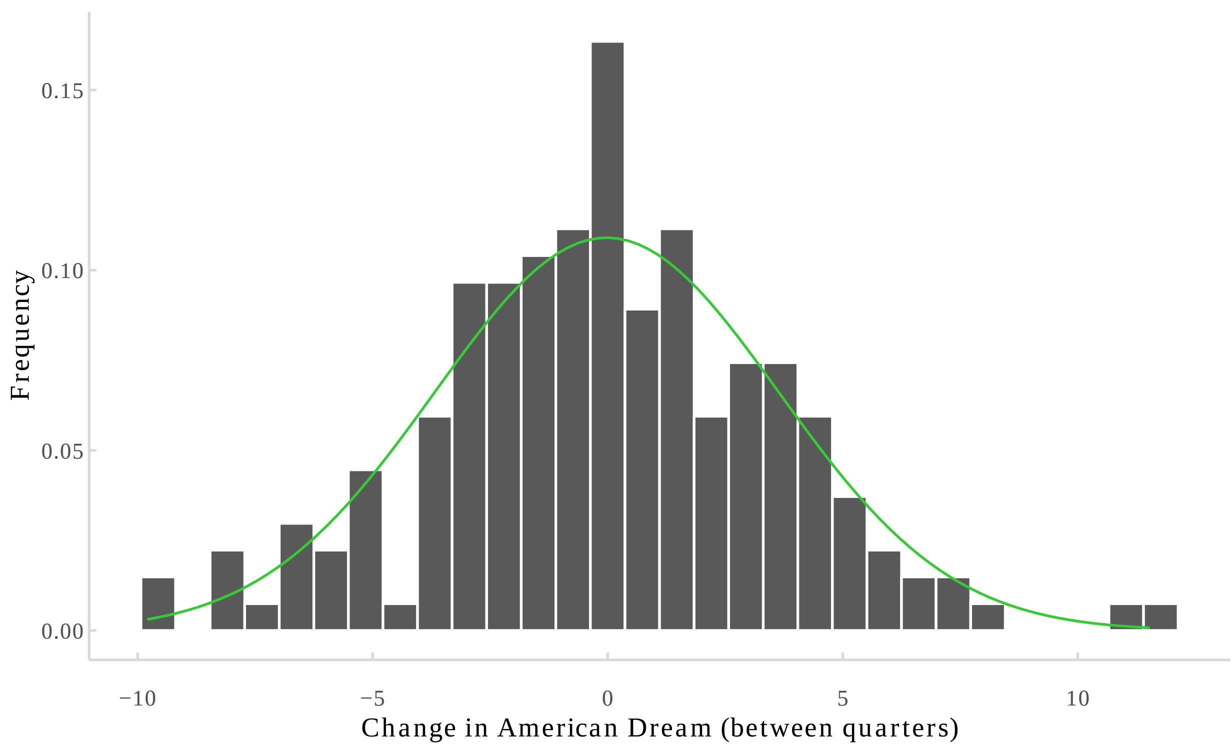


Figure 4: Distribution of Change in Belief in American Dream



Acknowledgements

This paper was helped in large part by the generous and helpful advice of my TF, Le Bao, and my professor, Prof. Jeff Gill.

Bibliography (in order of reference)

Wolak, Jennifer, & Peterson, David A. M. (2020). The Dynamic American Dream. *American Journal of Political Science*, 64(4), 968-981.

Bannerjee, Anindya, Juan Dolado, JohnW. Galbraith, and David F. Hendry. 1993.*Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford: Oxford University Press.

De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52(1): 184–200.

Nieman, Mark David, & Peterson, David A. M. (2019).

Webb, Clayton, Suzanna Linn, and Matthew Lebo. 2020. "Beyond the Unit Root Question: Uncertainty and Inference." *American Journal of Political Science* 64(2): 275–92.

Webb, Clayton, Linn, Suzanna, & Lebo, Matthew. (2019). A Bounds Approach to Inference Using the Long Run Multiplier. *Political Analysis*, 27(3), 281-301.