Replication of Father Founders:

Did Child Gender Affect Voting at the Constitutional Convention?

Prepared for POLI 271: Advanced Statistical Applications

Zayne Sember

Department of Political Science

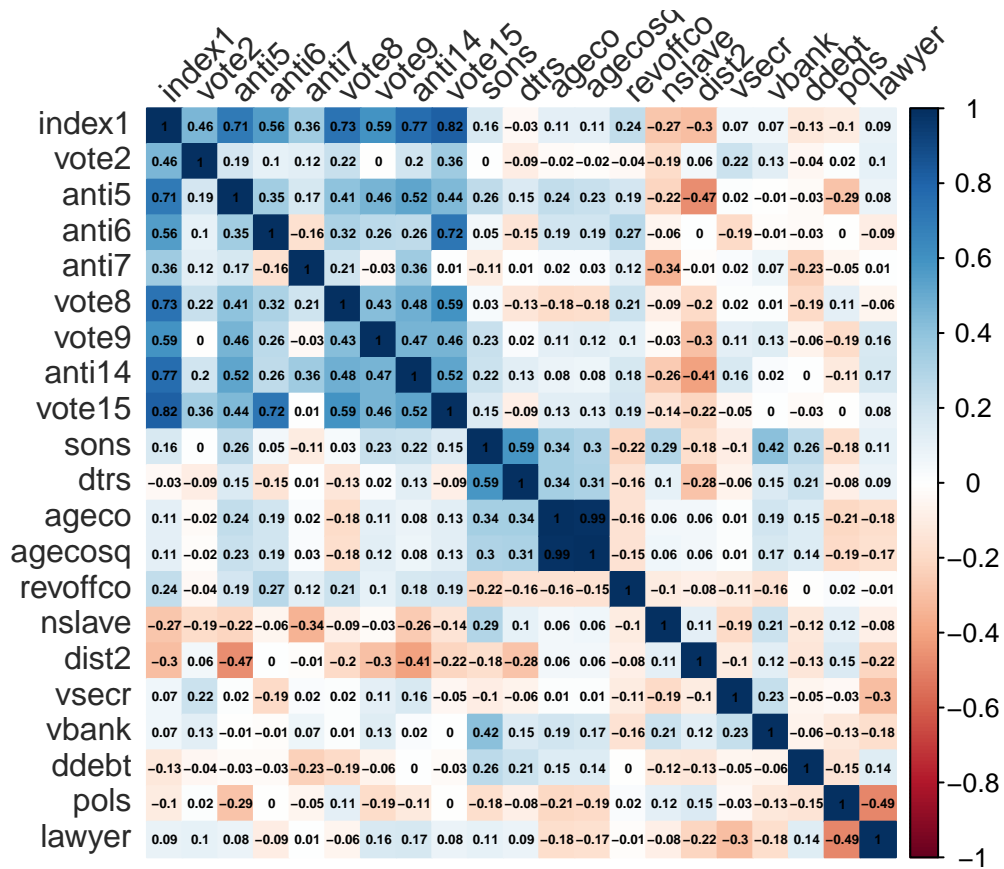University of California, San Diego

Abstract

This article is a replication of Jeremy C. Pope and Soren J. Schmidt's 2021 piece "Father Founders: Did Child Gender Affect Voting at the Constitutional Convention?". In it, they test the hypothesis that the delegates with sons would tend to vote for a stronger national government because they foresaw such a government providing greater opportunities for their sons–for which they find evidence. I begin by replicating their primary model, a Poisson regression on a vote index, and their probit models of individual votes. I then examine the distributions of underlying data to ensure their use of Poisson regression is appropriate. I next assess the differences in standard errors for their model covariates comparing their `STATA` robust standard errors with `R`'s `glm()` standard errors and bootstrapped standard errors. Finally, I use the dataset from which the paper's dataset was derived to assess missingness and compare imputation to the historical supplementation method used by the authors.

Data

The replication data used is provided by Pope and Schmidt (2021). I also employ Dougherty and Heckelman's (2009) dataset which provides delegate vote data without the historical supplementation of votes present in Pope and Schmidt's replication data. All code and data for this project is publicly available at https://github.com/zaynesember/PopeSchmidtReplication.

Figure 1: Correlation Matrix of Variables

	index1	vote2	anti5	anti6	anti7	vote8	vote9	anti14	vote15	sons	dtrs	ageco	agecosq	revoffco	nslave	dist2	vsecr	vbank	ddebt	pols	lawyer
index1	1	0.46	0.71	0.56	0.36	0.73	0.59	0.77	0.82	0.16	−0.03	0.11	0.11	0.24	−0.27	−0.3	0.07	0.07	−0.13	−0.1	0.09
vote2	0.46	1	0.19	0.1	0.12	0.22	0	0.2	0.36	0	−0.09	−0.02	−0.02	−0.04	−0.19	0.06	0.22	0.13	−0.04	0.02	0.1
anti5	0.71	0.19	1	0.35	0.17	0.41	0.46	0.52	0.44	0.26	0.15	0.24	0.23	0.19	−0.22	−0.47	0.02	−0.01	−0.03	−0.29	0.08
anti6	0.56	0.1	0.35	1	−0.16	0.32	0.26	0.26	0.72	0.05	−0.15	0.19	0.19	0.27	−0.06	0	−0.19	−0.01	−0.03	0	−0.09
anti7	0.36	0.12	0.17	−0.16	1	0.21	−0.03	0.36	0.01	−0.11	0.01	0.02	0.03	0.12	−0.34	−0.01	0.02	0.07	−0.23	−0.05	0.01
vote8	0.73	0.22	0.41	0.32	0.21	1	0.43	0.48	0.59	0.03	−0.13	−0.18	−0.18	0.21	−0.09	−0.2	0.02	0.01	−0.19	0.11	−0.06
vote9	0.59	0	0.46	0.26	−0.03	0.43	1	0.47	0.46	0.23	0.02	0.11	0.12	0.1	−0.03	−0.3	0.11	0.13	−0.06	−0.19	0.16
anti14	0.77	0.2	0.52	0.26	0.36	0.48	0.47	1	0.52	0.22	0.13	0.08	0.08	0.18	−0.26	−0.41	0.16	0.02	0	−0.11	0.17
vote15	0.82	0.36	0.44	0.72	0.01	0.59	0.46	0.52	1	0.15	−0.09	0.13	0.13	0.19	−0.14	−0.22	−0.05	0	−0.03	0	0.08
sons	0.16	0	0.26	0.05	−0.11	0.03	0.23	0.22	0.15	1	0.59	0.34	0.3	−0.22	0.29	−0.18	−0.1	0.42	0.26	−0.18	0.11
dtrs	−0.03	−0.09	0.15	−0.15	0.01	−0.13	0.02	0.13	−0.09	0.59	1	0.34	0.31	−0.16	0.1	−0.28	−0.06	0.15	0.21	−0.08	0.09
ageco	0.11	−0.02	0.24	0.19	0.02	−0.18	0.11	0.08	0.13	0.34	0.34	1	0.99	−0.16	0.06	0.06	0.01	0.19	0.15	−0.21	−0.18
agecosq	0.11	−0.02	0.23	0.19	0.03	−0.18	0.12	0.08	0.13	0.3	0.31	0.99	1	−0.15	0.06	0.06	0.01	0.17	0.14	−0.19	−0.17
revoffco	0.24	−0.04	0.19	0.27	0.12	0.21	0.1	0.18	0.19	−0.22	−0.16	−0.16	−0.15	1	−0.1	−0.08	−0.11	−0.16	0	0.02	−0.01
nslave	−0.27	−0.19	−0.22	−0.06	−0.34	−0.09	−0.03	−0.26	−0.14	0.29	0.1	0.06	0.06	−0.1	1	0.11	−0.19	0.21	−0.12	0.12	−0.08
dist2	−0.3	0.06	−0.47	0	−0.01	−0.2	−0.3	−0.41	−0.22	−0.18	−0.28	0.06	0.06	−0.08	0.11	1	−0.1	0.12	−0.13	0.15	−0.22
vsecr	0.07	0.22	0.02	−0.19	0.02	0.02	0.11	0.16	−0.05	−0.1	−0.06	0.01	0.01	−0.11	−0.19	−0.1	1	0.23	−0.05	−0.03	−0.3
vbank	0.07	0.13	−0.01	−0.01	0.07	0.01	0.13	0.02	0	0.42	0.15	0.19	0.17	−0.16	0.21	0.12	0.23	1	−0.06	−0.13	−0.18
ddebt	−0.13	−0.04	−0.03	−0.03	−0.23	−0.19	−0.06	0	−0.03	0.26	0.21	0.15	0.14	0	−0.12	−0.13	−0.05	−0.06	1	−0.15	0.14
pols	−0.1	0.02	−0.29	0	−0.05	0.11	−0.19	−0.11	0	−0.18	−0.08	−0.21	−0.19	0.02	0.12	0.15	−0.03	−0.13	−0.15	1	−0.49
lawyer	0.09	0.1	0.08	−0.09	0.01	−0.06	0.16	0.17	0.08	0.11	0.09	−0.18	−0.17	−0.01	−0.08	−0.22	−0.3	−0.18	0.14	−0.49	1

Results

Model Replication

Pope and Schmidt (2021) present 9 models of interest, all of which share the same covariates (with the exception of a covariate being omitted in two of their probit models due to what the authors deem to be multicollinearity but may be better termed perfect separation). Their primary model is a Poisson regression on a "preferred index" of the eight votes evaluated in each of the additional probit models. The preferred index is calculated by summing the number of "yea" votes for expanding the national government and "nay" votes on limiting the government across the eight key votes examined in their probit models. Appendix A provides the table of all independent and dependent variables. Dependent variables labeled "vote" indicate a vote where a "yea" expands the government and those labeled "anti" denote a vote limiting government. A correlation matrix of all variables is shown in Figure 1. Many of the individual votes (`vote2` through `vote15`) are correlated with one another but this does not threaten any assumptions of the models. Of the independent variables, none are unexpectedly strongly correlated with each other or with the dependent variables.

Table 1 presents the replication of these nine models. All coefficients agree with those presented by Pope and Schmidt, although some standard errors in the probit models differ very slightly. This is due to the original analysis being done in `STATA` with the `robust` command; the similar robust errors reported here were calculated in `R` using the `sandwich` library's `vcovCL` function. The National Exports and State Credit models omit three and twelve observations, respectively. These are omitted to replicate the manner in which `STATA` handles probit models with perfect separation. The debtor and private securities coefficients are omitted from these models due to perfect multicollinearity after the observations are dropped.

Figure 2: Predicted Value Plots
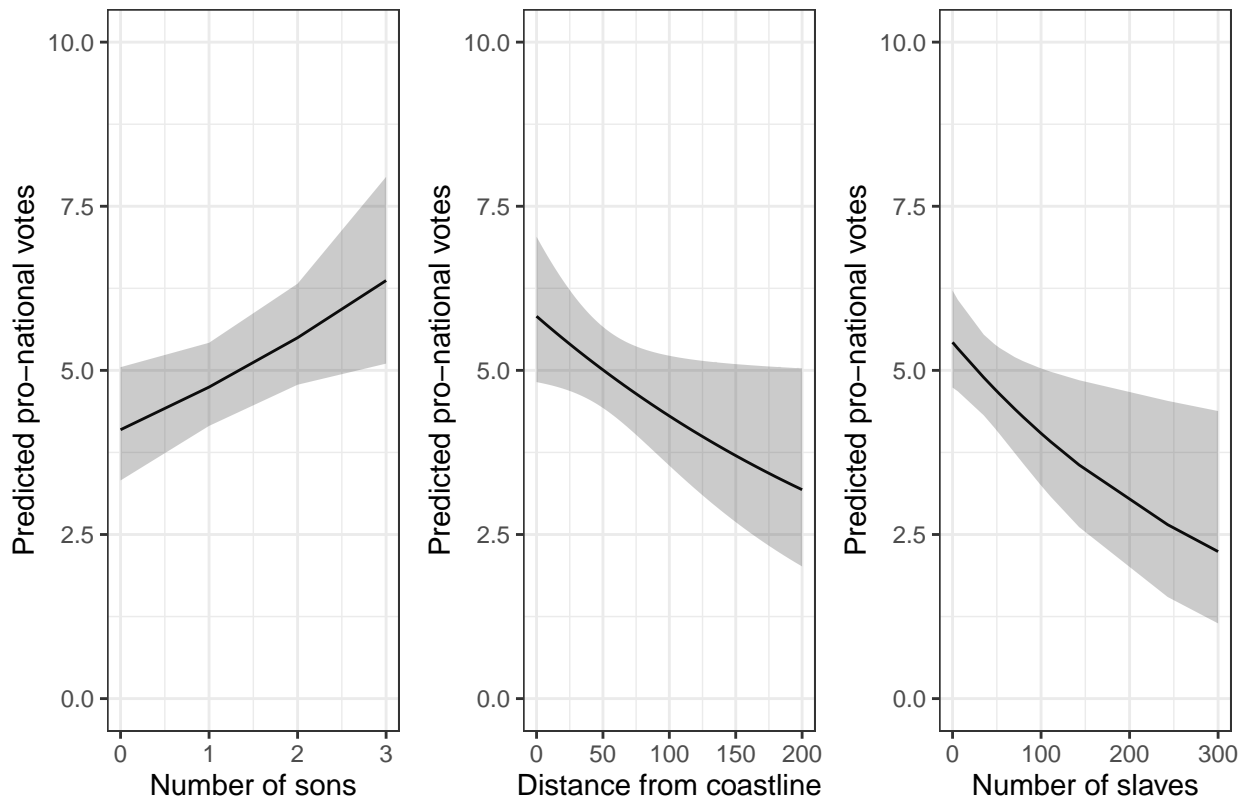
Figure 2 replicates Pope and Schmidt's (2021) Figure 2, providing predicted value plots for the variable of interest (number of sons) and two other influential variables. As they conclude, there is clear evidence that delegates with more sons tended to vote more in favor of expanding the national government at a magnitude comparable to other pertinent variables such as distance from a navigable coastline and number of slaves owned.

Table 1: Model Replications

	Preferred Index	National Veto	Debtor Legislators	Cong. Quorum	National Exports	Militia Control	State Credit	Navigation Acts	Military Responsibility
Constant	1.576*	-0.843	5.287	-0.145	1.207	3.558	4.481	-0.538	-0.288
	(0.626)	(2.472)	(5.932)	(2.865)	(2.812)	(4.028)	(5.443)	(2.787)	(2.824)
Number of sons	0.147**	0.134	1.022**	0.277	-0.181	0.460*	0.785**	0.494**	0.554**
	(0.037)	(0.170)	(0.394)	(0.202)	(0.242)	(0.196)	(0.381)	(0.164)	(0.193)
Number of daughters	-0.124**	-0.178	-0.210	-0.501**	0.190	-0.337	-0.395	-0.129	-0.556**
	(0.039)	(0.158)	(0.343)	(0.247)	(0.220)	(0.165)	(0.274)	(0.172)	(0.219)
Age	-0.002	-0.021	-0.184	0.050	-0.051	-0.087	-0.223	-0.001	0.022
	(0.025)	(0.102)	(0.292)	(0.116)	(0.116)	(0.155)	(0.254)	(0.113)	(0.117)
Age squared	0.000	0.000	0.002	0.000	0.001	0.001	0.003	0.000	0.000
	(0.000)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
Revolutionary war officer	0.273**	0.067	1.531**	1.154**	0.337	0.789	0.724	0.888*	0.694
	(0.110)	(0.393)	(0.607)	(0.560)	(0.421)	(0.551)	(0.581)	(0.528)	(0.496)
Logged number of slaves	-0.003**	-0.006	-0.013*	-0.004	-0.015**	-0.006	0.001	-0.009*	-0.007
	(0.001)	(0.004)	(0.005)	(0.003)	(0.006)	(0.003)	(0.006)	(0.004)	(0.004)
Distance to navigable coastline	-0.003**	0.003	-0.018**	-0.004	0.001	-0.008*	-0.009	-0.009*	-0.009*
	(0.001)	(0.005)	(0.008)	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	(0.006)
Public securities (1000s, 1787 dollars)	0.008	0.184*	-0.069	-0.210*	-0.088	-0.049	0.127	0.176	-0.032
	(0.032)	(0.109)	(0.226)	(0.096)	(0.107)	(0.119)	(0.141)	(0.120)	(0.113)
Private securities (1000s, 1787 dollars)	0.000	0.043	-0.067	-0.006	0.054	0.009	NA	0.003	0.006
	(0.004)	(0.032)	(0.049)	(0.022)	(0.056)	(0.026)		(0.035)	(0.040)
Debtor (dummy)	-0.721**	-0.313	-2.513	0.341	NA	-2.293*	-1.077	-1.068	-0.852
	(0.131)	(1.008)	(0.723)	(0.760)		(0.890)	(1.932)	(0.533)	(0.520)
Politician	0.107	0.898	-0.965	-0.227	0.036	0.307	0.064	0.725	0.798
	(0.145)	(0.504)	(0.640)	(0.522)	(0.564)	(0.595)	(0.672)	(0.595)	(0.496)
Lawyer	0.098	1.101*	-0.986	-0.913	0.013	-0.425	0.827	0.855	0.445
	(0.116)	(0.455)	(0.666)	(0.610)	(0.548)	(0.622)	(0.707)	(0.543)	(0.584)
Num.Obs.	53	53	53	53	50	53	41	53	53
AIC	246.7	87.3	57.2	72.8	80.1	76.6	57.8	70.9	72.2
BIC	272.3	112.9	82.8	98.4	103.0	102.2	78.3	96.6	97.8
Log.Lik.	-110.348	-30.661	-15.606	-23.418	-28.052	-25.302	-16.887	-22.474	-23.079

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

NA coefficients omitted due to perfect multicollinearity.

Preferred index model is a Poisson regression, all others are probit. Standard errors are robust.

Figure 3: Poissoness Plot of Preferred Index
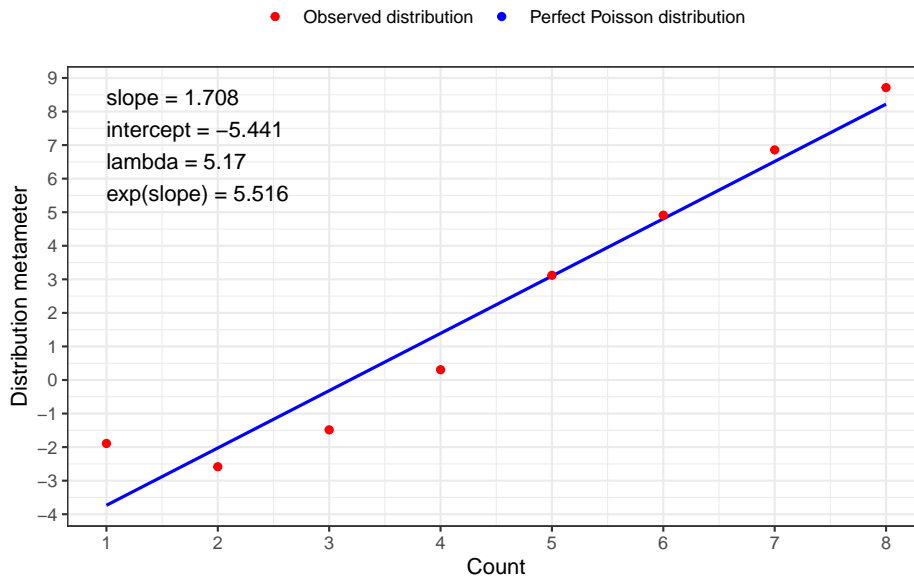
Distribution of Data

I next examine the distribution of the primary dependent variable: the preferred index of individual votes. In order for a Poisson regression on this index to be valid, it needs to follow a Poisson distribution with the key assumption that the count data's mean is equal to its variance. A simple check of this on the preferred index yields a mean of 5.17 and variance of 5.34, which falls 0.07 standard deviations away–a very minor violation of this strict assumption. Figure 3 provides a more visual test of the distribution with a Poissoness plot. If the data were perfectly Poisson then we would expect the intercept of the Poissoness plot to be $-\lambda = -5.17$; in reality it is -5.441. We would also expect the slope to be $\log(\lambda) = 0.742$ whereas it is 1.708 on the plot. Visual inspection of the plot shows the data does generally follow a Poisson distribution with the largest outlier being the number of observations where the preferred index equals one. Figures 4 and 5 compare the fit of the data to the Poisson and negative binomial distributions with rootograms. The data appears to fit both distributions well in this metric, again with an excess of ones being the primary departure from the expected distribution.

Figure 4: Poisson Rootogram
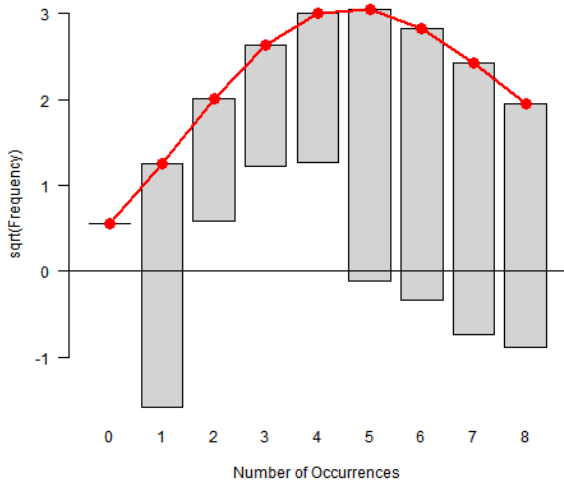
Figure 5: Negative Binomial Rootogram
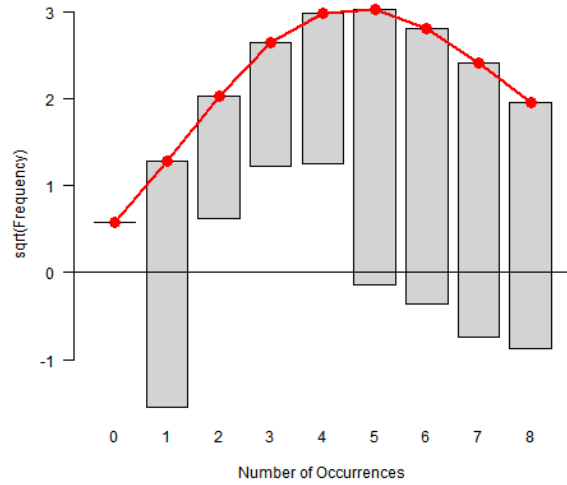
Table 2 shows a comparison of the preferred index model as a Poisson and negative binomial regression. The resulting coefficients and non-robust standard errors are identical. This is due to the fact that a Poisson regression is a special case of the negative binomial in which the variance and mean are equal. Because the mean and variance (as shown earlier) are only 0.07 standard deviations apart, the Poisson and negative binomial regressions are nearly indistinguishable in this case. This is further supported by their equal log likelihoods and nearly equal AIC and BICs and expected given the identical rootograms in Figures 4 and 5 which again highlight the excess of ones in the count data.

Taken in sum, the above evidence strongly indicates that Poisson regression is appropriate for modeling the preferred index. The count data follows an approximately Poisson distribution with nearly equal variance and mean, a lack of overdispersion, and produces identical results when modeled as a negative binomial distribution instead. The data is, of course, not perfectly Poisson, particularly when looking at the number of ones in the preferred index, but overall the assumptions of the Poisson regression are sufficiently met.

Table 2: Comparison of Poisson and Negative Binomial Models of the Preferred Index

	Poisson	Negative Binomial
Constant	1.5761*	1.5761*
	(0.8627)	(0.8627)
Number of sons	0.1471**	0.1471**
	(0.0602)	(0.0602)
Number of daughters	-0.1237**	-0.1237**
	(0.0600)	(0.0600)
Age	-0.0024	-0.0024
	(0.0342)	(0.0342)
Age squared	0.0001	0.0001
	(0.0003)	(0.0003)
Revolutionary war officer	0.2734**	0.2734**
	(0.1339)	(0.1339)
Logged number of slaves	-0.0030**	-0.0030**
	(0.0012)	(0.0012)
Distance to navigable coastline	-0.0030**	-0.0030**
	(0.0015)	(0.0015)
Public securities (1000s, 1787 dollars)	0.0079	0.0079
	(0.0336)	(0.0336)
Private securities (1000s, 1787 dollars)	-0.0003	-0.0003
	(0.0084)	(0.0084)
Debtor (dummy)	-0.7212**	-0.7212**
	(0.3341)	(0.3341)
Politician	0.1066	0.1066
	(0.1828)	(0.1828)
Lawyer	0.0983	0.0983
	(0.1747)	(0.1747)
Num.Obs.	53	53
AIC	246.7	248.7
BIC	272.3	276.3
Log.Lik.	-110.348	-110.349

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Comparison of Preferred Index Model Standard Error Estimates

	R glm	Robust	Bootstrapped
Constant	0.863	0.620	1.250
Number of sons	0.060	0.037	0.057
Number of daughters	0.060	0.039	0.060
Age	0.034	0.025	0.056
Age squared	0.000	0.000	0.001
Revolutionary war officer	0.134	0.109	0.143
Logged number of slaves	0.001	0.001	0.001
Distance to navigable coastline	0.001	0.001	0.002
Public securities (1000s, 1787 dollars)	0.034	0.031	0.046
Private securities (1000s, 1787 dollars)	0.008	0.004	0.052
Debtor (dummy)	0.334	0.129	0.288
Politician	0.183	0.143	0.192
Lawyer	0.175	0.115	0.160

Figure 6: Comparison of Preferred Index Model Standard Error Estimates
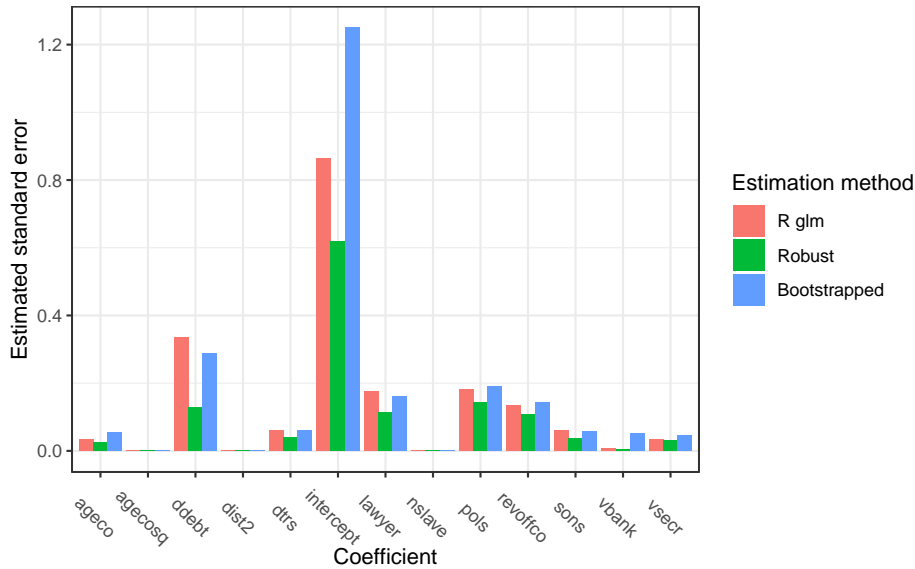
Standard Errors and Underdispersion

Table 3 reports the standard error estimates for coefficients in the preferred index model calculated through three different methods. The first are the errors as reported by R's `glm` function, calculated by inverting the model's Fisher information matrix and taking the square root of its diagonals. Robust standard errors are the same as those reported in Table 1, calculated in an attempt to match the `STATA` robust standard errors reported by Pope and Schmidt (2021). Bootstrapped standard errors are calculated using the `boot` package in R. Evident from Figure 6's bar plot comparing these values, the robust errors are consistently the smallest. The `glm` and bootstrapped standard

errors tend to be of similar magnitude (with the exception of the intercept).

When robust standard errors are smaller than non-robust errors as in this case, it can indicate underdispersion of the count data which inflates non-robust standard errors. Testing the alternative hypothesis that the data is underdispersed (and thus its dispersion parameter is less than one) against the null hypothesis that it is equidispersed (its dispersion parameter is one) as assumed by the Poisson regression yields a statistically significant estimate at the $p < 0.05$ level of the dispersion parameter being 0.71 indicating underdispersion and the possibility that the non-robust errors are inflated.

On average, the non-robust errors in Table 3 are 47% larger (omitting age squared). Changing from robust to `glm` errors (or the nearly equivalent bootstrapped error estimate), the error estimate on the number of sons coefficient grows by 39% and its significance level changes from the $p < 0.01$ reported by Pope and Schmidt (2021) to $p < 0.05$. Because it remains significant and its coefficient estimate of 0.147 remains an order of magnitude larger than its estimated errors, the type of standard error estimates used does not change the substantive interpretation of that particular covariate's impact on delegates' vote choice. For this data, Pope and Schmidt's choice of robust standard errors seems justified as it avoids the inflation of the non-robust errors and does not change the interpretation of the coefficient of interest.

Counterfactual Delegates

Inspection of delegate vote data from Dougherty and Heckelman (2009) reveals two delegates not included by Pope and Schmidt (2021). The first is George Wythe of Virginia's delegation who had no children and left the Convention in June to tend to his dying wife without casting any votes (Holt n.d.). It is not clear whether he was replaced but he was a noted Federalist and thus likely supported a strong national government. The second is William Churchill Houston of New Jersey who left two sons and three daughters when he died one week into the convention, having also cast no votes (Glenn 2013). He was replaced by William Livingston ("The Founding Fathers: New Jersey" 2015).

I use the latter of these delegates as an opportunity to both assess the sensitivity of the preferred index model and to consider a counterfactual Convention in which William Churchill Houston was able to vote. Because the preferred index is calculated by summing eight individual votes with two potential values, creating every possible vote scenario for a counterfactual delegates is computationally trivial[1]. Along with the variable of interest (number of sons), covariates that could be easily found were imputed (age, whether they were a lawyer, etc.), all others were held at the median of the other delegates. Since William Livingston would not have attended had Houston been present, I remove his votes from the counterfactual analysis.

---

[1]Accounting for every combination of two delegates' votes, on the other hand, increases the number of scenarios from $2^8 = 256$ to $2^{16} = 65,536$ scenarios

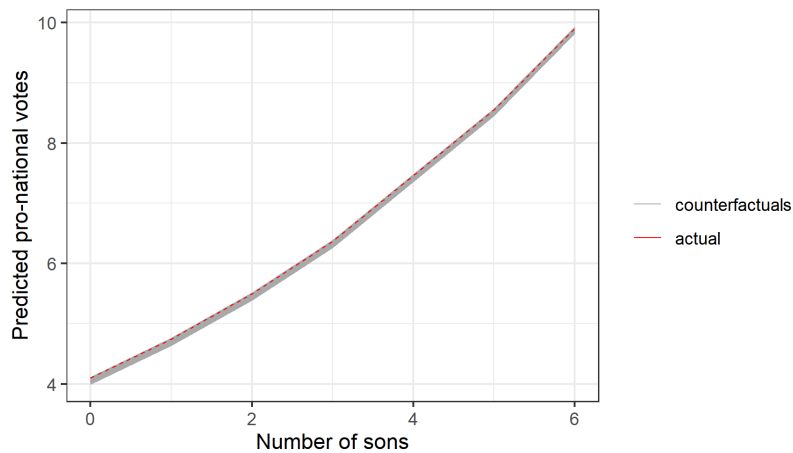Figure 7: Counterfactual Scenario Predicted Value Plot

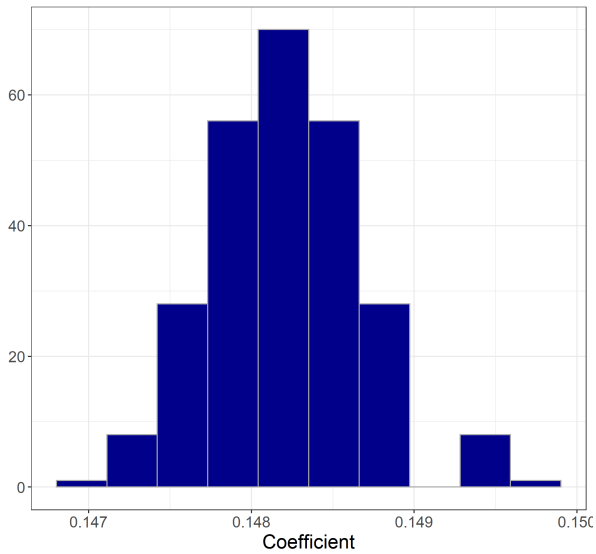Figure 8: Counterfactual Number of Sons Coefficient Distribution

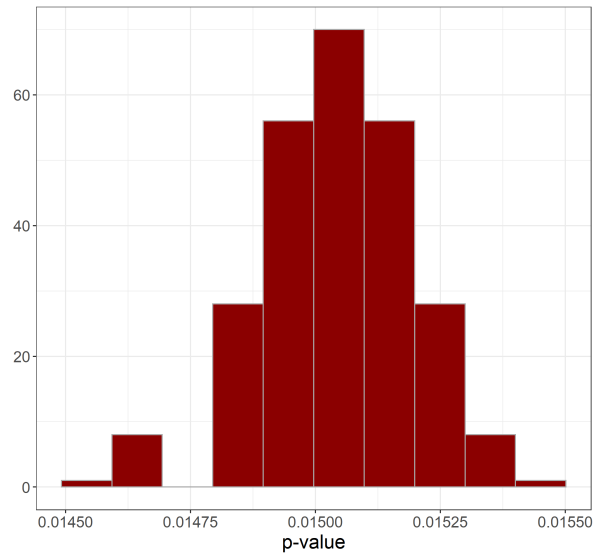Figure 9: Counterfactual Number of Sons p-value Distribution

Figure 7 shows the predicted value plot of the marginal effect of a delegate's number of sons on their predicted pro-national government votes for every counterfactual scenario (gray) and the observed scenario (red). Confidence intervals are omitted for ease of viewing. It is clear from this plot that the model's estimates would be virtually indistinguishable no matter how Houston voted. Figures 8 and 9 show the distribution of number of sons coefficient estimates and p-values for every counterfactual scenario. Since the coefficient distribution is centered around 0.148 and the coefficient reported for the non-counterfactual model is 0.147, most of the counterfactuals resulted in a very slight increase in this particular coefficient. The distribution of p-values reveals that the coefficient remains significant at the $p < 0.05$ level (using non-robust standard errors) for every scenario. These results indicate that, in the alternative history where William Churchill Houston voted at the Convention rather than William Livingston, the number of sons delegates likely would have remained a strong predictor of whether they supported expansion of

10

the national government regardless of how Houston voted.

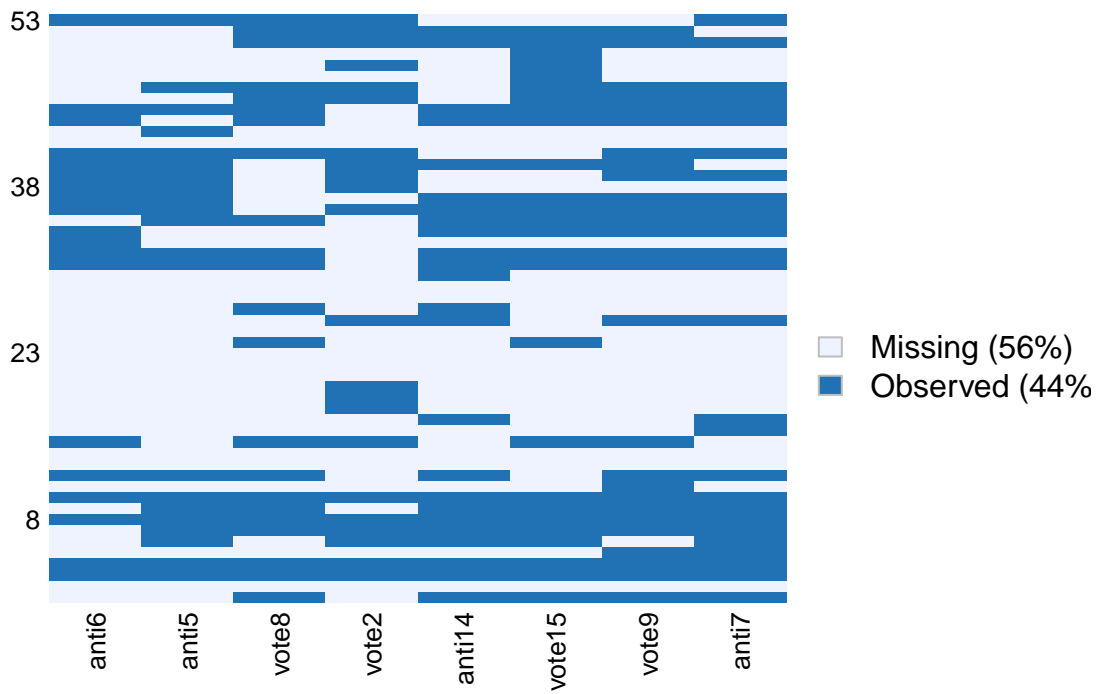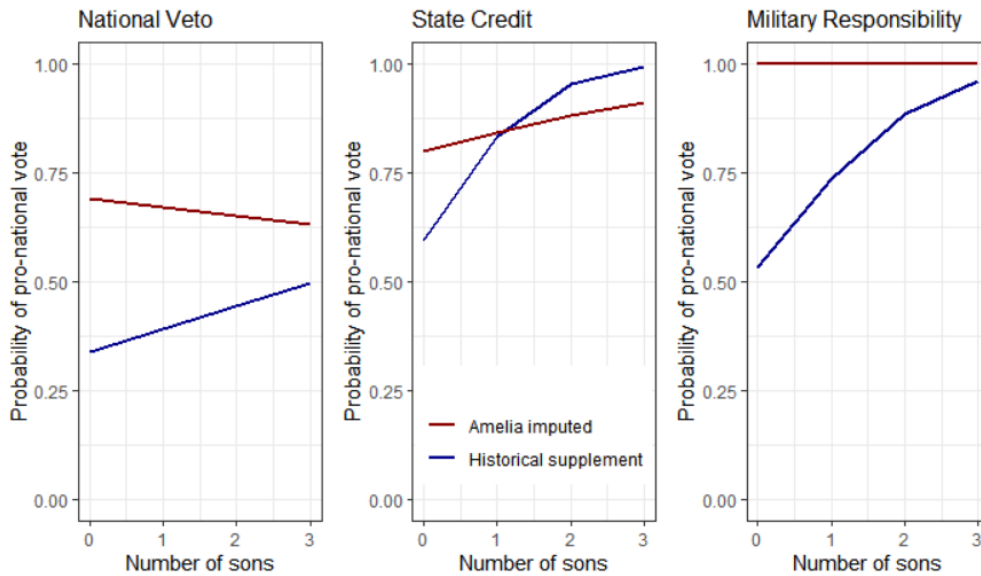## Figure 10: Individual Vote Missingness

Figure 11: Individual Vote Predicted Probability for Imputed Data

Missingness

As Figure 10 shows, the original vote data employed by Pope and Schmidt (2021) is rife with missingness–more than half of all votes are missing. This missingness was originally remedied by McDonald (1958) who used historical sources to supplement the vote data, providing estimates of how delegates *would* have voted. In this section, I compare this historical estimation method to data imputation using R's `Amelia` package. Because such a large percentage of votes are missing from the original data, imputation was possible only on three votes which had enough variance and lacked multicollinearity issues: National Veto, State Credit, and Military Responsibility. The preferred index could thus not be calculated using only the original data; just these three individual votes.

Predicted value plots of this imputation for each of the three votes are compared to the historically supplemented data used by Pope and Schmidt (2021) in Figure 11. While the historically supplemented data consistently predicts that delegates with more sons will have a higher probability of voting in favor of national government expansion, the imputed data shows mixed results. For the National Veto vote the imputed data results in the opposite; more sons lowers the probability of voting to expand the government. For the State Credit vote the results are similar to that of the historical data, albeit less drastic. For Military Responsibility there appears to be no change in vote probability using imputed data. These results indicate, rather unsurprisingly, that addressing missingness through informed historical supplementation is key to seeing an effect for these three votes. Were this supplemental data not available, imputed data may have obfuscated the results of the study.

References

Dougherty, K., & Heckelman, J. C. (2009). Delegate Votes on 28 Motions at the United States Constitutional Convention, 1787: Version 1 [Data set]. ICPSR Data Holdings. Inter-university Consortium for Political and Social Research (ICPSR).

Glenn, T. A. (2013). *William Churchill Houston 1746-1788*. Book on Demand.

Holt, W. W. (n.d.). Wythe, George (1726 or 1727–1806). Retrieved June 4, 2021, from EncyclopediaVirginia.org website: https://encyclopediavirginia.org/entries/wythe-george-1726-or-1727-1806/

McDonald, F. (1958). *We, the People: The Economic Origins of the Constitution*. Chicago, IL: University of Chicago Press.

Pope, J. C., & Schmidt, S. J. (2021). Father Founders: Did Child Gender Affect Voting at the Constitutional Convention? *American Journal of Political Science*, (ajps.12573). doi:10.1111/ajps.12573

The Founding Fathers: New Jersey. (2015, November 6). Retrieved June 4, 2021, from Archives.gov website: https://www.archives.gov/founding-docs/founding-fathers-new-jersey

Appendix A

Table 4: Variable Key

Variable name	Full name
index1	Preferred Index
vote2	National Veto
anti5	Debtor Legislators
anti6	Cong. Quorum.
anti7	National Exports
vote8	Militia Control
vote9	State Credit
anti14	Navigation Acts
vote15	Military Responsibility
sons	Number of sons
dtrs	Number of daughters
ageco	Age
agecosq	Age squared
revoffco	Revolutionary war officer
nslave	Logged number of slaves
dist2	Distance to navigable coastline
vsecr	Public securities (1000s, 1787 dollars)
vbank	Private securities (1000s, 1787 dollars)
ddebt	Debtor (dummy)
pols	Politician
lawyer	Lawyer