Supplemental Information

Robert Shaffer

University of Pennsylvania

shafferr@upenn.edu

January 12, 2019

Contents

A Measurement

A.1 Header Regular Expressions

Table 1 gives the set of regular expressions used as inputs to the constitute_tools parser, which I use to parse the American legislative text database I use in this paper. Note that not all of these levels are present in all documents.

Table 1: Regular expressions used to parse American legislative texts

Regular Expression	Sample Plain-Text Match
`(SECTION|SEC\.) [0-9]+[-A-Z]*\.`	SECTION 101; SEC. 446a
`\(([ivx])?([a-hj-uwyz])?\)\s* (?=(?(1)[\s\S]*?\n\s*\([jwy]\)))`	(a)
`\([0-9]+\)\s*`	(32)
`\(([IVX])?([A-HJ-UWYZ])?\)\s* (?=(?(1)[\s\S]*?\n\s*\([JWY]\)))`	(B)
`\([ivx]+\)\s*`	(ii)
`\([IVX]+\)\s*`	(IV)
`` ``\([A-Z0-9a-z]+\) ```	“(A)

A.2 Sample Parsed Text

Table 2: Sample parsed document

Title	Text
SEC 416	FOREIGN STUDENT MONITORING PROGRAM.
(a)	Full «NOTE: 8 USC 1372 note.» Implementation and Expansion of Foreign Student Visa Monitoring Program Required.–The Attorney General, in consultation with the Secretary of State, shall fully implement and expand the program established by section 641(a) of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (8 U.S.C. 1372(a)).
(b)	Integration «NOTE: 8 USC 1372 note.» With Port of Entry Information.–For each alien with respect to whom information is collected under section 641 of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (8 U.S.C. 1372), the Attorney General, in consultation with the Secretary of State, shall include information on the date of entry and port of entry.

USA PATRIOT Act §416(a-b). For original text see the corresponding congress.gov page.

A.3 LSTM Parameter Specification

As described in §3, I used an LSTM to extract named entities from legislative texts. Where not otherwise specified, I left parameter settings at their defaults in the tf_ner library.

For model construction, I used a 300-node layer for word embeddings and a 100-node layer for the LSTM encoder. For the word embedding layer I rely on pre-trained embeddings drawn from Pennington *et al.* (2014)'s GloVe dataset. Like virtually all neural network applications, I trained this model using stochastic gradient descent.[1] I trained the model for up to 25 epochs, with the model set to halt training if no improvement was observed in a held-out development set every 500 batches (with a minimum of 8000 batches run). I used sentences containing 90% of pre-identified named entities for training and 10% as a held-out development set. To avoid overfitting, I use a dropout rate of 0.5 and a batch size of 20 for the gradient descent algorithm used to fit the model.

[1] Specifically, an ADAM optimizer. See Kingma and Ba (2014) for details.

B Descriptive Statistics

Figure 1: Predictor and dependent variable descriptive statistics, for bills with at least one named entity.

Variable	Mean	SD	Source
Average Degree	4.76	14.69	Author
Divided	0.61	-	Casas et al. (2018)
CQ Mention	0.21	-	CQ Almanac
Majority	0.78	-	Casas et al. (2018)
Hitchhiker	0.51	-	Casas et al. (2018)
$\sqrt{\text{Cosponsors}}$	2.81	2.74	Casas et al. (2018)
DW-NOMINATE	0.10	0.44	Casas et al. (2018)
Republican	0.56	-	Casas et al. (2018)

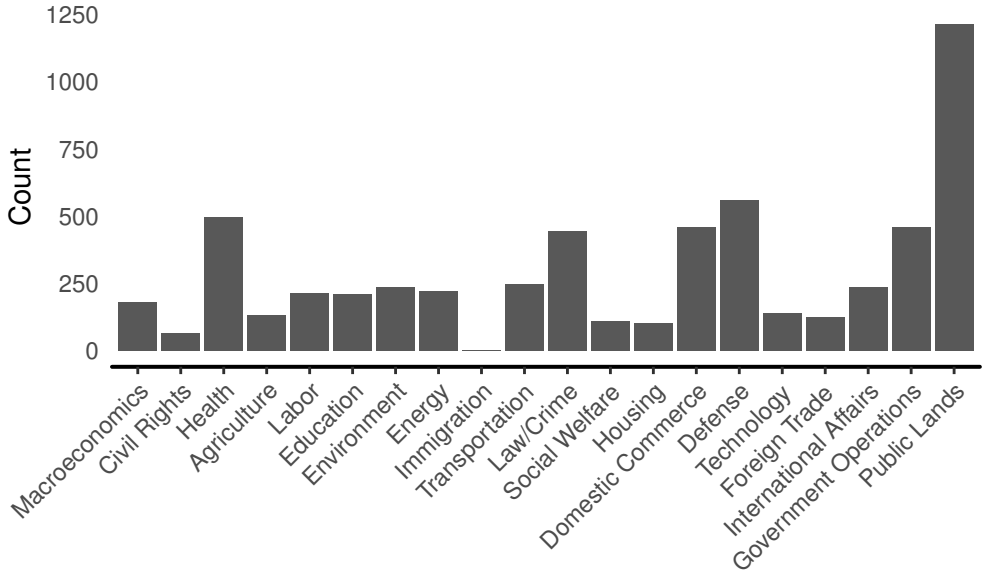
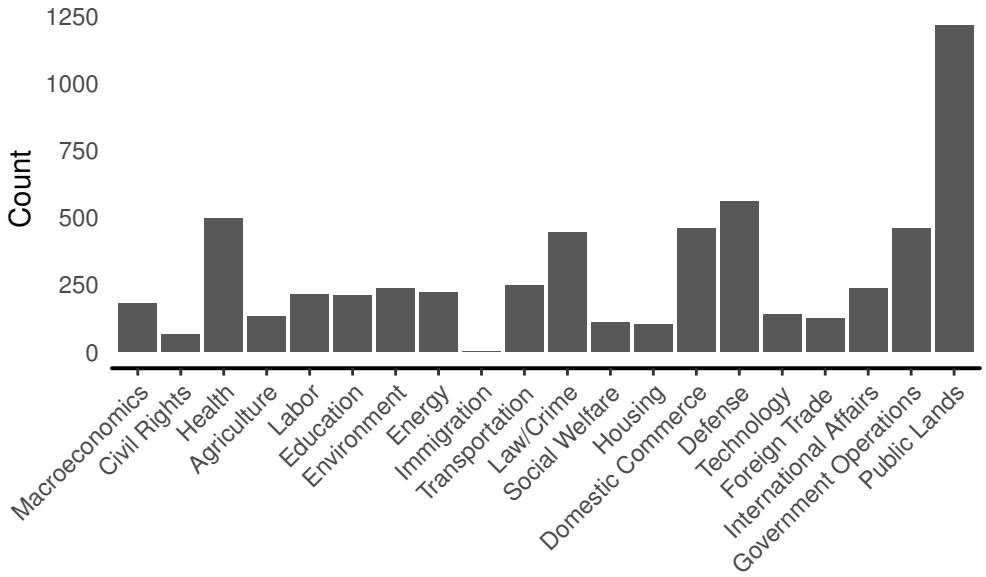


Figure 2: Predictor and dependent variable descriptive statistics, for all bills.

Variable	Mean	SD	Source
Total nodes	5.68	14.83	Author
Divided	0.61	-	Casas et al. (2018)
CQ Mention	0.18	-	CQ Almanac
Majority	0.77	-	Casas et al. (2018)
Hitchhiker	0.52	-	Casas et al. (2018)
$\sqrt{\text{Cosponsors}}$	2.75	2.67	Casas et al. (2018)
DW-NOMINATE	0.11	0.45	Casas et al. (2018)
Republican	0.57	-	Casas et al. (2018)



C Bayesian Model Details

C.1 Specification and Model Fit

To fit all regression models I present in this paper, I use the `brms` library's interface to the Stan programming language (Carpenter *et al.* 2016). For the gamma regression models, I use the likelihood and implementation specified in the `brms` library's `hurdle_gamma()` function. Similarly, for the negative binomial hurdle model I present as part of my robustness checks in Appendix C.2.5, I use the `brms` library's `hurdle_negbinomial()` function.

For priors, I selected weakly informative prior distributions for all variables. These prior values are intended to be uninformative in all cases where a non-trivial quantity of data is present, but restrict parameters from attaining implausible values when data is sparser. Following Ghosh *et al.* (2018), for all regression coefficients I assign a $t(3, 0, 10)$ prior for all intercept coefficients, a $t(3, 0, 2.5)$ prior for all other regression coefficients, and a $Cauchy(0, 5)$ prior on all standard deviation parameters. I additionally placed a $gamma(0.01, 0.01)$ prior on the gamma and negative binomial distribution distribution dispersion parameters.

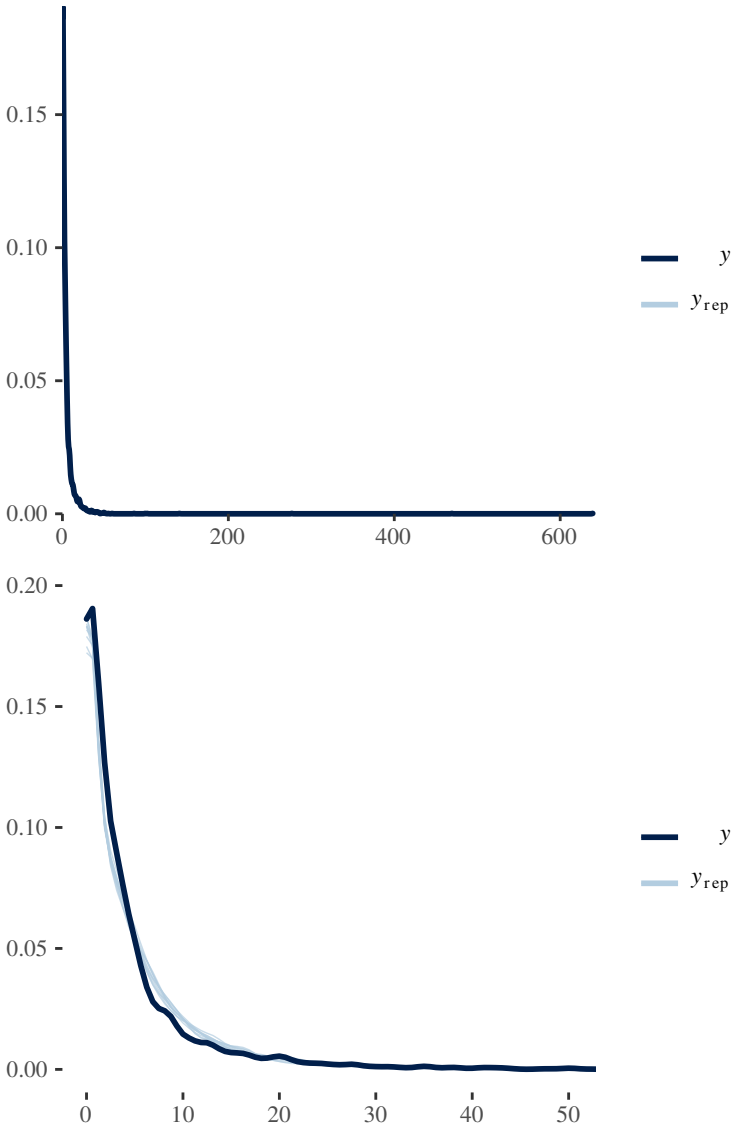
For all models, I ran four chains with 1000 warmup iterations, 1500 post-warmup iterations in each chain, and random initializations for all parameter values. For the negative binomial hurdle model, I additionally set the $adapt\_delta$ value in the sampler to 0.85 to avoid divergent transitions. Visual plots suggested good mixing across chains in all models, with $\hat{R} \leq 1.01$ for all parameters and $n_{eff} \geq 1000$ for all parameters.[2]

Following Gelman *et al.* (2014), in Figure 3 I visually assessed model fit for the "main" model I present in-text using posterior predictive checks. In each plot, I provide the observed density of the node count dependent variable, overlaid on density plots for 10 simulated dependent variable datasets based on randomly-selected post-warmup posterior parameter draws. As shown in the top panel, across the whole dataset the model fit is excellent. Zooming in on smaller values (where most posterior density is located) reveals that the

[2]With $\hat{R}$ a diagnostic quantifying the consistency of an ensemble of Markov chains, and n_{eff} a rough effective sample size calculation (Gelman *et al.* 2014).

model slightly over-fits at small values of the dependent variable ($5 \leq y \leq 10$). Even in this range, however, model fit remains strong.

Figure 3: Posterior predictive plots for the node count dependent variable



C.2 Numerical Coefficient Estimates

For all models in this section, estimates prefixed with (hu) indicate hurdle coefficients. Numerical intercept estimates for all models besides the model presented in-text are suppressed for brevity, but are available in replication materials. For all model variants, key findings regarding the `CQ Mention` and `Divided` coefficients are replicated, at roughly the same scales. Findings regarding all other variables are approximately replicated, though effect sizes are more variable, particularly for the negative binomial and no-hitchhiker model variants.

C.2.1 Gamma Hurdle Model (in-text)

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	1.41	0.08	1.25	1.57
$\sqrt{\text{Cosponsors}}$	0.00	0.01	-0.01	0.02
DW-NOMINATE	-0.18	0.04	-0.26	-0.09
DW-NOMINATE^2	0.77	0.11	0.56	0.97
Majority	0.16	0.04	0.08	0.25
CQ Mention	0.36	0.06	0.24	0.47
Divided	-0.06	0.04	-0.13	0.02
Hitchhiker	-0.05	0.03	-0.12	0.02
CQ Mention:Divided	0.53	0.08	0.38	0.68
(hu) Intercept	-0.44	0.13	-0.70	-0.18
(hu) $\sqrt{\text{Cosponsors}}$	-0.05	0.02	-0.08	-0.02
(hu) DW-NOMINATE	0.15	0.10	-0.04	0.34
(hu) DW-NOMINATE^2	0.08	0.24	-0.40	0.55
(hu) Majority	-0.26	0.08	-0.42	-0.09
(hu) CQ Mention	-1.16	0.17	-1.50	-0.83
(hu) Divided	0.06	0.08	-0.09	0.22
(hu) Hitchhiker	-0.16	0.07	-0.30	-0.03
(hu) CQ Mention:Divided	0.14	0.22	-0.28	0.57

	Estimate	Est.Error	Q2.5	Q97.5
Macroeconomics	1.40	0.12	1.16	1.63
Civil Rights	1.32	0.14	1.06	1.59
Health	1.29	0.08	1.14	1.44
Agriculture	1.27	0.11	1.05	1.50
Labor	1.25	0.10	1.05	1.44
Education	1.22	0.11	1.02	1.43
Environment	1.32	0.09	1.14	1.51
Energy	1.33	0.10	1.13	1.53
Immigration	1.41	0.25	0.92	1.90
Transportation	1.48	0.09	1.29	1.66
Law/Crime	1.36	0.08	1.20	1.51
Social Welfare	1.66	0.12	1.43	1.90
Housing	1.45	0.14	1.19	1.73
Domestic Commerce	1.24	0.08	1.08	1.40
Defense	1.63	0.08	1.48	1.78
Technology	1.63	0.10	1.44	1.84
Foreign Trade	1.49	0.11	1.28	1.71
International Affairs	1.32	0.09	1.15	1.49
Government Operations	1.94	0.08	1.78	2.09
Public Lands	1.07	0.06	0.96	1.19
(hu) Macroeconomics	-0.34	0.21	-0.75	0.07
(hu) Civil Rights	-0.40	0.22	-0.83	0.05
(hu) Health	-0.64	0.16	-0.97	-0.32
(hu) Agriculture	-0.55	0.21	-0.96	-0.16
(hu) Labor	-0.46	0.19	-0.83	-0.09
(hu) Education	-0.29	0.19	-0.66	0.09
(hu) Environment	-0.37	0.17	-0.71	-0.03
(hu) Energy	-0.07	0.19	-0.46	0.31
(hu) Immigration	-0.33	0.27	-0.85	0.23
(hu) Transportation	-0.48	0.18	-0.83	-0.13
(hu) Law/Crime	-0.61	0.16	-0.93	-0.29
(hu) Social Welfare	-0.36	0.22	-0.80	0.09
(hu) Housing	-0.30	0.22	-0.73	0.15
(hu) Domestic Commerce	-0.29	0.17	-0.62	0.03
(hu) Defense	-0.32	0.15	-0.62	-0.02
(hu) Technology	-0.66	0.21	-1.09	-0.27
(hu) Foreign Trade	-0.47	0.20	-0.88	-0.06
(hu) International Affairs	-0.66	0.19	-1.03	-0.31
(hu) Government Operations	-0.59	0.16	-0.89	-0.29
(hu) Public Lands	-0.53	0.12	-0.77	-0.30

C.2.2 Gamma Hurdle Model (base DW-NOMINATE coefficient only)

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	1.56	0.08	1.41	1.72
$\sqrt{\text{Cosponsors}}$	0.00	0.01	-0.01	0.02
DW-NOMINATE	-0.02	0.04	-0.10	0.05
Majority	0.17	0.04	0.09	0.25
CQ Mention	0.36	0.06	0.25	0.48
Divided	-0.06	0.04	-0.14	0.02
Hitchhiker	-0.07	0.03	-0.14	0.00
CQ Mention:Hitchhiker	0.52	0.08	0.37	0.67
(hu) Intercept	-0.42	0.12	-0.65	-0.18
(hu)$\sqrt{\text{Cosponsors}}$	-0.05	0.01	-0.08	-0.02
(hu) DW-NOMINATE	0.17	0.08	0.00	0.33
(hu) Majority	-0.26	0.08	-0.42	-0.09
(hu) CQ Mention	-1.15	0.17	-1.48	-0.83
(hu) Divided	0.06	0.08	-0.09	0.22
hu_outcomeHitchhiker	-0.17	0.07	-0.31	-0.02
(hu) CQ Mention:Divided	0.13	0.22	-0.30	0.56

C.2.3 Gamma Hurdle Model (party specification)

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	1.59	0.08	1.44	1.75
$\sqrt{\text{Cosponsors}}$	0.00	0.01	-0.01	0.02
Republican	-0.08	0.03	-0.15	-0.02
Majority	0.18	0.04	0.10	0.26
CQ Mention	0.35	0.06	0.24	0.47
Divided	-0.05	0.04	-0.13	0.02
Hitchhiker	-0.07	0.03	-0.14	-0.01
CQ Mention:Divided	0.53	0.08	0.38	0.68
(hu) Intercept	-0.47	0.13	-0.72	-0.22
(hu) $\sqrt{\text{Cosponsors}}$	-0.05	0.02	-0.08	-0.02
(hu) Republican	0.10	0.08	-0.04	0.25
(hu) Majority	-0.25	0.09	-0.42	-0.08
(hu) CQ Mention	-1.16	0.17	-1.50	-0.83
(hu) Divided	0.07	0.08	-0.09	0.22
(hu) Hitchhiker	-0.17	0.07	-0.31	-0.03
(hu) CQ Mention:Divided	0.14	0.22	-0.29	0.57

C.2.4 Gamma Hurdle Model (no hitchhikers)

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	1.37	0.11	1.15	1.59
$\sqrt{\text{Cosponsors}}$	0.00	0.01	-0.02	0.01
DW-NOMINATE	-0.19	0.07	-0.32	-0.05
DW-NOMINATE^2	1.15	0.15	0.86	1.46
Majority	0.10	0.07	-0.04	0.24
CQ Mention	0.35	0.08	0.20	0.50
Divided	-0.08	0.06	-0.20	0.04
CQ Mention:Divided	0.63	0.10	0.43	0.84
(hu) Intercept	-0.31	0.19	-0.69	0.07
(hu) $\sqrt{\text{Cosponsors}}$	-0.10	0.02	-0.15	-0.05
(hu) DW-NOMINATE	0.11	0.14	-0.17	0.40
(hu) DW-NOMINATE^2	0.16	0.34	-0.52	0.82
(hu) Majority	-0.24	0.14	-0.51	0.04
(hu) CQ Mention	-1.69	0.25	-2.19	-1.21
(hu) Divided	0.10	0.11	-0.13	0.32
(hu) CQ Mention:Divided	0.74	0.30	0.15	1.36

C.2.5 Negative Binomial Hurdle Model (total node dependent variable)

	Estimate	Est.Error	l-95% CI	u-95% CI
Intercept	0.58	0.12	0.35	0.81
$\sqrt{\text{Cosponsors}}$	0.03	0.01	0.01	0.05
DW-NOMINATE	-0.20	0.06	-0.33	-0.08
DW-NOMINATE^2	0.35	0.16	0.05	0.66
Majority	0.37	0.06	0.25	0.49
CQ Mention	1.18	0.09	0.99	1.36
Divided	-0.05	0.06	-0.16	0.06
Hitchhiker	0.04	0.05	-0.06	0.14
CQ Mention:Divided	0.40	0.12	0.16	0.65
(hu) Intercept	-0.61	0.16	-0.91	-0.30
(hu) $\sqrt{\text{Cosponsors}}$	-0.03	0.01	-0.06	-0.01
(hu) DW-NOMINATE	0.15	0.08	-0.01	0.32
(hu) DW-NOMINATE^2	0.53	0.20	0.14	0.92
(hu) Majority	-0.23	0.08	-0.37	-0.08
(hu) CQ Mention	-1.58	0.18	-1.95	-1.24
(hu) Divided	0.05	0.07	-0.08	0.19
(hu) Hitchhiker	-0.08	0.06	-0.20	0.05
(hu) CQ Mention:Divided	0.65	0.22	0.23	1.09

References

Bob Carpenter, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *J Stat Softw*, 2016.

Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. Taylor & Francis, Boca Raton, FL, 3 edition, 2014.

Joyee Ghosh, Yingbo Li, Robin Mitra, et al. On the use of cauchy prior distributions for bayesian logistic regression. *Bayesian Analysis*, 13(2):359–383, 2018.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.