Online Appendices:
Measuring Constituency Ideology Using Bayesian Universal Kriging

Jeff Gill

Departments of Government and Mathematics & Statistics

American University

jgill@american.edu

A   Online Appendix: Finding Ideological Neighbors for Alaska and Hawaii

Since Alaska and Hawaii are geographic outliers relative to the continental 48 states, it could pose problems to estimate the model and forecast ideology with the states located as they are. Such geographic outliers could exert substantial leverage over both the geographic trend term and the smoothed error structure. Yet, we do want to forecast ideology for these two states and these constituencies. As a result, we did a search for the states' *ideological neighbors*. The idea behind finding ideological neighbors is that we can determine what geographical locales on the Pacific Coast work well as neighbors for the purposes of forecasting.

We searched for ideological neighbors for Alaska and Hawaii as follows: First, we estimated a preliminary point-referenced data model using OLS regression and an error structure on those residuals. In this initial model, we exclude Alaska and Hawaii from the training data. Second, we consider a variety of locations along the Pacific Coast that might serve as ideological neighbors for each state. For each possible location of the state, we compute the sum of squared errors if we forecasted ideology in the CCES data for each state using the model that was trained over the continental 48 states. We chose each state's location based on which locale minimized the sum of squared errors.

For each state, we considered a central block and started out by placing the state so that

its center was as the same northing as the southernmost point on the U.S. Pacific Coast. For Alaska we set the eastings so that the easternmost point in Alaska could never overlap with the westernmost point in the continental states. For Hawaii, we started off by setting the eastings so that the easternmost point in Hawaii would not overlap with the westernmost point of the continental states, but then moved Hawaii 75 kilometers closer: given how the best fits for each state work this adjustment prevents Hawaii from sitting on top of either the continental coast or atop Alaska's new ideological neighbor position.

Figure A.1 illustrates our comparative fit on this process. On each panel the horizontal axis represents the distance from the southernmost point on the continental Pacific Coast, with 0 representing the point at the far south and larger numbers indicating kilometers northward from there. The vertical axis on each panel represents the sum of squared errors (SSE) for the out-of-sample predictions to that state from the continental model. The black solid line represents the SSE for Hawaii at each position, and the dashed blue line represents the SSE for Alaska at each position. The left panel shows all positions along the Pacific Coast, which were considered in 50 km increments. As that panel shows, 1,000 km from the southernmost point the SSE starts rapidly increasing for both states, indicating that the fit becomes remarkably bad if we place either state in the northwest. The right panel therefore focuses on the southern side of the Pacific Coast, ranging from the southernmost point at 0 km to 700 km north of that. For Alaska we see that the smallest SSE emerges 500 km north of the U.S.-Mexico border, and for Hawaii the smallest SSE emerges 450 km north of the border.

Using these placements, Figure 2 in the main text plots the centroids of the census blocks we use for forecasting. The blocks for Alaska and Hawaii are now placed in their new ideological neighbor locations based on the best out-of-sample fit. Hawaii is placed so that its central census block is 450 km north of the U.S.-Mexico border. Substantively, this places the island of Oahu off the coast of the San Francisco Bay, meaning that Honolulu, HI and San Francisco, CA are ideological neighbors. Alaska is placed so that its central census block is 500 km north of the border. Substantively, this means that Juneau, AK is situated just south of San Diego, CA, which also puts the capital city as close as possible to Arizona and Texas for making forecasts of ideology. Anchorage, AK, meanwhile is a bit north of Santa Barbara, CA. Moving forward, when we forecast ideology for Alaska, Hawaii, and each state's respective legislative districts we use these new ideological neighbor locations.

<table>
<tr><td>(a) Full Range of Distance</td><td>(b) Range of 0–700 km</td></tr>
</table>
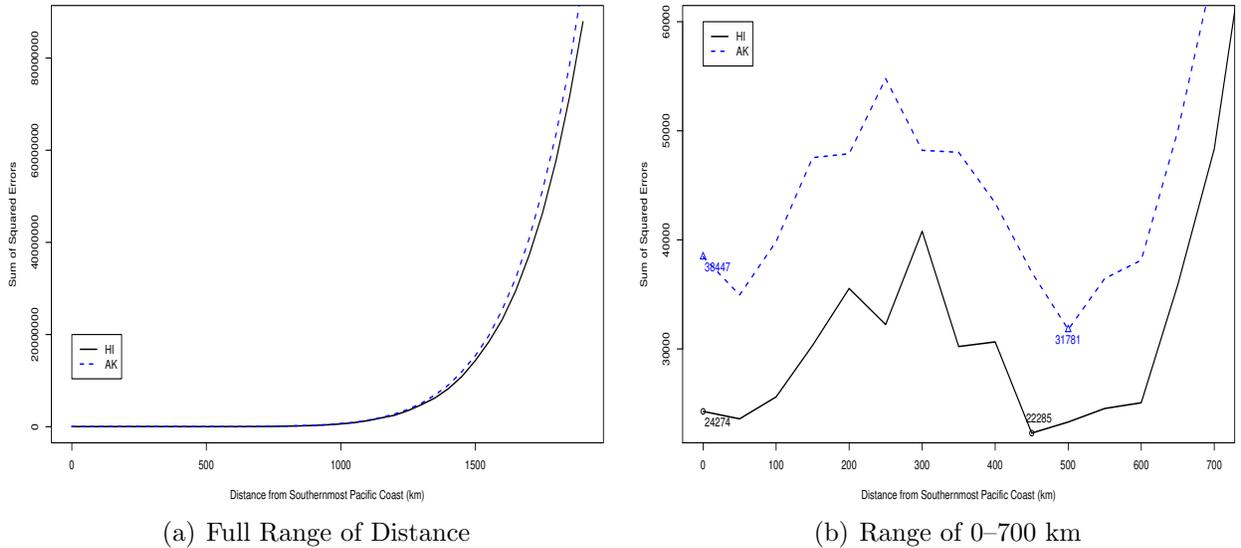
Figure A.1: Sum of squared errors for forecasts of Alaska and Hawaii data for positions along the Pacific Coast.

A.1 Sensitivity of Forecasts to Location

As an additional check of the robustness of our measures for Alaska and Hawaii, we conduct an analysis that compares our forecasts of simulated citizens in these states at various alternative geographic locations to the forecasts we use in our final measures. To do this, we proceeded in the following way: First, we subsetted the 724,814 simulated citizens used in our forecasts to the 1,503 simulated Alaskans and 3,347 simulated Hawaiians. Second, we adjusted the location of these simulated citizens to range from placing each state at 0 km north of the U.S. border with Mexico to 2,000 km north of the border, in 50 km increments. Third, we made kriged predictions of these citizens at each of these locations. For computational feasibility, we made the simplifying assumption that the parameter estimates of the forecasting model is assumed to be the point estimate of each parameter in Table 1 for the main article. Additionally, a kriging model requires observed training data to predict the smoothed error term. Critically for computational time, we only use the first bootstrap subsample from the training data (consisting of 1,088 observations) when making the predictions at each alternative location. Doing this feasibly allowed us to make kriged predictions at 40 different alternative locations and compare each set of predictions to the results we get when we use *all* of our bootstrap samples to forecast ideology in Alaska and Hawaii at their ideological neighbor locations of 500 km and 450 km, respectively.

When we compute the correlations between the kriged values at these various proposed

3

locations against the forecasts that we use in our final measures, the correlations are always large, and they are similar across locations. The correlations for Alaska ranged from 0.81053 to 0.82066, with the points at 500 km (the ideological neighbor location) correlating with the full-bootstrap forecast of 0.81057. (The standard error of these correlations rounds to 0.000228 in all cases.) Meanwhile, the correlations for Hawaii ranged from 0.89039 to 0.89045, with the points at 450 km (the ideological neighbor location) correlating with the full-bootstrap forecast of 0.89045. (The standard error of these correlations rounds to 0.000061 in all cases.) Based on this additional analysis, we conclude that our results for these two states are not particularly sensitive to the ideological-neighbor placement that we use for the states, given the narrow range of these high correlation values. Even with this stability in results, though, we still believe that we have chosen the best ideological neighbor locations based on the analysis of Figure A.1.

B   Online Appendix: Data Sources

- **Survey data of individual ideology in 2008:** The Cooperative Congressional Election Survey, Common Content, 2008. Accessed from `http://hdl.handle.net/1902.1/14003` on April 18, 2013 (Ansolabehere 2011).

- **Population demographic data in 2010:** U.S. Census 2010 block, block group, and tract-level data. Dataset `2010_SF1a` accessed from the National Historical Geographic Information System, `https://www.nhgis.org` on October 13, 2015 (Minnesota Population Center 2011).

- **Census block centroids and area:** U.S. Census TIGER shapefiles for 2010. Accessed from `http://www2.census.gov/geo/tiger/TIGER2010BLKPOPHU/` on December 12, 2015.

- **ZIP code centroids and area:** USA ZIP Code Areas, 2014. TomTom data held by ArcGIS. Accessed from `https://www.arcgis.com/home/item.html?id=8d2012a2016e484dafaac0451f9aea24` on September 2, 2015.

- **Religious affiliation by county:** The 2010 Religious Congregations and Membership Study, provided by the ARDA. Accessed from `http://www.thearda.com/Archive/Files/Descriptions/RCMSCY10.asp` on October 15, 2015 (Grammich et al. 2012).

- **Urban-rural classification continuum by county:** United States Department of Agriculture, 2013 measure. Accessed from `http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx#.UfCK11OE4xc` on October 15, 2015 (United States Department of Agriculture 2013).

- **Ideology of members of Congress:** Subset of Carroll, Lewis, Lo, McCarty, Poole, & Rosenthal's Common Space DW-NOMINATE scores. Restricted to first dimension for 112th Congress. Accessed from `http://voteview.com` on December 20, 2015 (McCarty, Poole and Rosenthal 1997; Poole and Rosenthal 1997).

- **Ideology of state legislators:** Subset of Shor & McCarty's June 2015 update of individual state legislator database, focusing strictly on 2011 scores. Accessed from `http://dx.doi.org/10.7910/DVN/THDBRA` on December 19, 2015 (Shor and McCarty 2011).

C Online Appendix: Crosswalk Between Census State Legislative District Names and Full Names, New Hampshire and Vermont

The 2010 Census files keep track of the state legislative and congressional districts that each census block falls in. Congressional districts follow a numbering scheme or are easily identified as a single at-large district. Most state legislative districts also follow a numbering scheme. All of this allows for merging with information such as legislator ideology scores.

As an important exception, the states of New Hampshire and Vermont name state legislative districts based on the county they are in. For the sake of file size, the names of state legislative districts are abbreviated in the Census files, though files such as the data by Shor and McCarty (2011) use the full name of the district. For the sake of facilitating future merging work, we list a crosswalk that ties the abbreviations to the full district names. Table A.1 lists districts for the New Hampshire House of Representatives, Table A.2 lists districts for the Vermont House of Representatives, and Table A.3 lists districts for the Vermont Senate.

Also of note: Washington has multimember districts. In this case, Shor and McCarty denote the members with their district number, followed by a second index. For example, 021-1 and 021-2 indicate the first and second members from House district 21. To merge with Census data, simply dropping the "-1" and "-2" makes the merge straightforward

Table A.1: Crosswalk between Names of State Legislative Districts and Census Bureau Abbreviations in the New Hampshire House of Representatives

Census Code	District Name	Census Code	District Name	Census Code	District Name
101	BELKNAP 1	617	HILLSBOROUGH 17	001	SULLIVAN 1
102	BELKNAP 2	618	HILLSBOROUGH 18	002	SULLIVAN 2
103	BELKNAP 3	619	HILLSBOROUGH 19	003	SULLIVAN 3
104	BELKNAP 4	620	HILLSBOROUGH 20	004	SULLIVAN 4
105	BELKNAP 5	621	HILLSBOROUGH 21	005	SULLIVAN 5
106	BELKNAP 6	622	HILLSBOROUGH 22		
201	CARROLL 1	623	HILLSBOROUGH 23		
202	CARROLL 2	624	HILLSBOROUGH 24		
203	CARROLL 3	625	HILLSBOROUGH 25		
204	CARROLL 4	626	HILLSBOROUGH 26		
205	CARROLL 5	627	HILLSBOROUGH 27		
301	CHESHIRE 1	701	MERRIMACK 1		
302	CHESHIRE 2	702	MERRIMACK 2		
303	CHESHIRE 3	703	MERRIMACK 3		
304	CHESHIRE 4	704	MERRIMACK 4		
305	CHESHIRE 5	705	MERRIMACK 5		
306	CHESHIRE 6	706	MERRIMACK 6		
307	CHESHIRE 7	707	MERRIMACK 7		
401	COOS 1	708	MERRIMACK 8		
402	COOS 2	709	MERRIMACK 9		
403	COOS 3	710	MERRIMACK 10		
404	COOS 4	711	MERRIMACK 11		
501	GRAFTON 1	712	MERRIMACK 12		
502	GRAFTON 2	713	MERRIMACK 13		
503	GRAFTON 3	801	ROCKINGHAM 1		
504	GRAFTON 4	802	ROCKINGHAM 2		
505	GRAFTON 5	803	ROCKINGHAM 3		
506	GRAFTON 6	804	ROCKINGHAM 4		
507	GRAFTON 7	805	ROCKINGHAM 5		
508	GRAFTON 8	806	ROCKINGHAM 6		
509	GRAFTON 9	807	ROCKINGHAM 7		
510	GRAFTON 10	808	ROCKINGHAM 8		
511	GRAFTON 11	809	ROCKINGHAM 9		
501	GRAFTON 1	810	ROCKINGHAM 10		
601	HILLSBOROUGH 1	811	ROCKINGHAM 11		
602	HILLSBOROUGH 2	812	ROCKINGHAM 12		
603	HILLSBOROUGH 3	813	ROCKINGHAM 13		
604	HILLSBOROUGH 4	814	ROCKINGHAM 14		
605	HILLSBOROUGH 5	815	ROCKINGHAM 15		
606	HILLSBOROUGH 6	816	ROCKINGHAM 16		
607	HILLSBOROUGH 7	817	ROCKINGHAM 17		
608	HILLSBOROUGH 8	818	ROCKINGHAM 18		
609	HILLSBOROUGH 9	901	STRAFFORD 1		
610	HILLSBOROUGH 10	902	STRAFFORD 2		
611	HILLSBOROUGH 11	903	STRAFFORD 3		
612	HILLSBOROUGH 12	904	STRAFFORD 4		
613	HILLSBOROUGH 13	905	STRAFFORD 5		
614	HILLSBOROUGH 14	906	STRAFFORD 6		
615	HILLSBOROUGH 15	907	STRAFFORD 7		
616	HILLSBOROUGH 16				

Table A.2: Crosswalk between Names of State Legislative Districts and Census Bureau Abbreviations in the Vermont House of Representatives

Census Code	District Name	Census Code	District Name
A-1	ADDISON-1	L-2	LAMOILLE-2
A-2	ADDISON-2	L-3	LAMOILLE-3
A-3	ADDISON-3	L-4	LAMOILLE-4
A-4	ADDISON-4	LW1	LAMOILLE-WASHINGTON-1
A-5	ADDISON-5	OG1	ORANGE-1
AR1	ADDISON-RUTLAND-1	OG2	ORANGE-2
B-1	BENNINGTON-1	OA1	ORANGE-ADDISON-1
B21	BENNINGTON-2-1	OGC	ORANGE-CALEDONIA-1
B22	BENNINGTON-2-2	OL1	ORLEANS-1
B-3	BENNINGTON-3	OL2	ORLEANS-2
B-4	BENNINGTON-4	OLC	ORLEANS-CALEDONIA-1
B-5	BENNINGTON-5	OLF	ORLEANS-FRANKLIN-1
BR1	BENNINGTON-RUTLAND-1	R11	RUTLAND-1-1
CA1	CALEDONIA-1	R12	RUTLAND-1-2
CA2	CALEDONIA-2	R-2	RUTLAND-2
CA3	CALEDONIA-3	R-3	RUTLAND-3
CA4	CALEDONIA-4	R-4	RUTLAND-4
CAW	CALEDONIA-WASHINGTON-1	R51	RUTLAND-5-1
C11	CHITTENDEN-1-1	R52	RUTLAND-5-2
C12	CHITTENDEN-1-2	R53	RUTLAND-5-3
C-2	CHITTENDEN-2	R54	RUTLAND-5-4
C31	CHITTENDEN-3-1	R-6	RUTLAND-6
C32	CHITTENDEN-3-2	R-7	RUTLAND-7
C33	CHITTENDEN-3-3	R-8	RUTLAND-8
C34	CHITTENDEN-3-4	RY1	RUTLAND-WINDSOR-1
C35	CHITTENDEN-3-5	W-1	WASHINGTON-1
C36	CHITTENDEN-3-6	W-2	WASHINGTON-2
C37	CHITTENDEN-3-7	W31	WASHINGTON-3-1
C38	CHITTENDEN-3-8	W32	WASHINGTON-3-2
C39	CHITTENDEN-3-9	W33	WASHINGTON-3-3
C35	CHITTENDEN-3-10	W-4	WASHINGTON-4
C-4	CHITTENDEN-4	W-5	WASHINGTON-5
C51	CHITTENDEN-5-1	W-6	WASHINGTON-6
C52	CHITTENDEN-5-1	W-7	WASHINGTON-7
C61	CHITTENDEN-6-1	WC1	WASHINGTON-CHITTENDEN-1
C62	CHITTENDEN-6-2	X-1	WINDHAM-1
C63	CHITTENDEN-6-3	X-2	WINDHAM-2
C71	CHITTENDEN-7-1	X31	WINDHAM-3-1
C72	CHITTENDEN-7-2	X32	WINDHAM-3-2
C-8	CHITTENDEN-8	X33	WINDHAM-3-3
C-9	CHITTENDEN-9	X-4	WINDHAM-4
EC1	ESSEX-CALEDONIA	X-5	WINDHAM-5
EC2	ESSEX-CALEDONIA-ORLEANS	X-6	WINDHAM-6
F-1	FRANKLIN-1	XB1	WINDHAM-BENNINGTON-1
F-2	FRANKLIN-2	XBY	WINDHAM-BENNINGTON-WINDSOR-1
F-3	FRANKLIN-3	Y11	WINDSOR-1-1
F-4	FRANKLIN-4	Y12	WINDSOR-1-2
F-5	FRANKLIN-5	Y-2	WINDSOR-2
F-6	FRANKLIN-6	Y-3	WINDSOR-3
GC1	GRAND ISLE-CHITTENDEN-1-1	Y-4	WINDSOR-4
L-1	LAMOILLE-1	Y-5	WINDSOR-5
		Y61	WINDSOR-6-1
		Y62	WINDSOR-6-2
		YO1	WINDSOR-ORANGE-1
		YO2	WINDSOR-ORANGE-2
		YR1	WINDSOR-RUTLAND-1
		YR2	WINDSOR-RUTLAND-2

Table A.3: Crosswalk between Names of State Legislative Districts and Census Bureau Abbreviations in the Vermont Senate

Census Code	District Name	Census Code	District Name
ADD	ADDISON	LAM	LAMOILLE
BEN	BENNINGTON	ORA	ORANGE
CAL	CALEDONIA	RUT	RUTLAND
CHI	CHITTENDEN	WAS	WASHINGTON
E-O	ESSEX-ORLEANS	WDM	WINDHAM
FRA	FRANKLIN	WSR	WINDSOR
CGI	GRAND ISLE		

References

Ansolabehere, Stephen. 2011. "CCES, Common Content, 2008." *http://hdl.handle.net/1902.1/14003* Ver. 4.

Grammich, Clifford, Kirk Hadaway, Richard Houseal, Dale E. Jones, Alexei Krindatch, Richie Stanley and Richard H. Taylor. 2012. *2010 U.S. Religion Census: Religious Congregations & Membership Study.* Association of Statisticians of American Religious Bodies.

McCarty, Nolan M., Keith T. Poole and Howard Rosenthal. 1997. *Income Redistribution and the Realignment of American Politics.* American Enterprise Institute Studies on Understanding Economic Inequality. Washington: AEI Press.

Minnesota Population Center. 2011. *National Historic Geographic Information System: Version 2.0* [Machine-readable database]. Minneapolis, MN: University of Minnesota.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* New York: Oxford University Press.

Shor, Boris and Nolan McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review* 105(3):530–551.

United States Department of Agriculture. 2013. "2013 Rural-Urban Continuum Codes." *http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx* .