Supporting Information for "Urban Water Conservation Policies in the United States"

Jonathan M. Gilligan[1,2,3], Christopher A. Wold[3], Scott C. Worland[2], John J. Nay[3,4,5], David J. Hess[3,6], George M. Hornberger[1,2,3]

[1]Department of Earth & Environmental Sciences, Vanderbilt University, Nashville, Tennessee, USA

[2]Department of Civil & Environmental Engineering, Vanderbilt University, Nashville, Tennessee, USA

[3]Vanderbilt Institute for Energy and Environment, Vanderbilt University, Nashville, Tennessee, USA

[4]Information Law Institute, New York University, New York, New York, USA

[5]Berkman Klein Center, Harvard University, Cambridge, Massachusetts, USA

[6]Department of Sociology, Vanderbilt University, Nashville, Tennessee, USA

Contents

Additional Supporting Information (Files uploaded separately)

1. Captions for Datasets S1 to (Replaced: ~~S2~~ replaced with: S4)
2. Data Analysis Scripts S1 (zip file)
3. Captions for Large Tables S1, S2, (Replaced: ~~S5–S7~~ replaced with: S7–S9)

Introduction

This supporting information document presents additional details of the data and analysis.

SI Text

Data

We used VWCI data for 195 cities in 45 states, as shown in Table S1.

Corresponding author: Jonathan M. Gilligan, `jonathan.gilligan@vanderbilt.edu`

At the MSA level (Dataset S1, Table S1), our regression analysis used the following six covariates: ln(population), population growth rate between 2010 and 2014, the Köppen aridity index, the fraction of the municipal water supply coming from surface water (henceforth, surface-water fraction), the Cook Partisan Voting Index (PVI), and the per-capita real personal income (RPI) for 2014 normalized for inflation and regional variations in the cost of living. We used the natural logarithm of the population rather than the raw population because the raw population was (Replaced: ~~sharply peaked~~ replaced with: skewed, with a sharp peak) near 500,000(Added: , and a long tail at higher populations) (Figure S1).

At the state level (Dataset S2, Table S2), our analysis used the following four covariates: PVI, RPI, the Köppen aridity index, and the surface-water fraction.

Analysis

Diagnostics

(Added: Our Monte Carlo analysis sampled four Markov chains, allowing each chain to warm-up and tune sampling parameters for the first 1000 iterations and then sampling each chain for 1000 more iterations, yielding a total of 4000 samples. Each sample is a vector of length 57, with values for each of the parameters α_0, β_j, γ_k, δ_{state}, σ, and ϕ, where j indexes over the six MSA-level covariates, k indexes over the four state-level covariates, and *state* indexes over 44 of the 45 states (leaving one out for identifiability). The samples approximate random draws from the joint posterior probability distribution of the parameters, given the priors and the observed data. Thus, the statistics of the sampled values approximate the statistics of the joint posterior distribution.
) ← Explain in detail what the Monte Carlo samples are.

(Replaced: ~~Pairwise correlation plots of the posterior probability distributions of regressions parameters (Figures S3–S5) are smooth and show little correlation. The Hamiltonian Monte Carlo calculations proceeded without any divergences or excessive tree depths, and the Gelman-Rubin $\hat{R}$ potential scale-reduction factor (Tabs S7–S9) converged to ≥ 0.999 for each parameter.~~ replaced with: Collinearity among the predictor variables is diagnosed by observing correlations in the joint posterior probability distributions of the regression coefficients [*Stan Development Team*, 2016, pp. 288–293]. Inefficient sampling due to varying curvature in the log-probability manifold or poorly chosen priors can be diagnosed by irregularities in joint posterior distributions [*Stan Development Team*, 2016, pp. 316–321]. Pairwise correlation plots of the Monte-Carlo samples for the regression coefficients in our models of VWCI, requirements, and rebates (Figures S2–S4) are smooth with little correlation and give no cause for concern. In addition, the Hamiltonian Monte Carlo calculations proceeded without any divergences or exceessive tree depths, and the Gelman-Rubin $\hat{R}$ potential scale-reduction factor converged to ≥ 0.999 for each parameter [*Stan Development Team*, 2016].)

← Explain using the pair correlation plots to diagnose multicollinearity among predictors.

Model Selection

(Added: We used several model-selection criteria in deciding whether to model the VWCI, requirements, and rebates as binomial or beta-binomial processes. At each joint sample of the model parameters in the Monte-Carlo process, we both computed the log-likelihood of the observed data under the sampled parameters and also generated posterior predictions obtained by drawing simulated observations from binomial or beta-binomial distribution at each joint sample of the model parameters.)

← Detailed discussion of model selection: hierarchical versus single-level and beta-binomial versus ordinary binomial.

(Added: Visual comparisons of distributions of posterior predictions to observed data and comparisons of the posterior predictions of mean, maximum, and minimum VWCI over the cities in our data set showed better agreement for the overdispersed β-binomial process than for a purely binomial one [*Gelman et al.*, 2014].)

(Replaced: ~~Leave-one-out cross-validation (Table S3) and the Widely Applicable Information Criterion (Table S4) were used for model selection (overdispersed beta-binomial versus binomial and hierarchical versus single-level regressions).~~ replaced with: A separate test for overdispersion, which accounts for the danger of overfitting by introducing new free parameters, uses the Leave-One-Out cross-validation Information Criterion (LOO-IC) or the Widely Available Information Criterion (WAIC, also known as the Watanabe-Aikake Information Criterion), obtained by Pareto-smoothed importance sampling [*Vehtari et al.*, 2017]. Both information criteria favored the overdispersed beta-binomial distribution over a pure binomial, and also strongly favored hierarchical over single-level models (Tables S3–S6). Our choice to use very weakly informative priors in our model reduces the accuracy of our estimates of LOO-IC and WAIC [*Vehtari et al.*, 2017], but we do not worry overly about this potential inaccuracy both because the posterior prediction test yields the same results and because a pure binomial model gives very similar results to those presented here.)

← Detailed description of model-selection using information criteria.

(Added: We had initially used the interval 1970–2014 for climatological averaging, but in response to reviewer comments, we investigated other averaging intervals and switched to the 1984–2014 interval because it performed slightly, but insignificantly better according to the information criteria. There are no important differences between between the two analyses (Figure S5).)

← Explain why we replaced the original 1970–2014 averaging period with the shorter 30-year period 1985–2014.

Results

Results of the analysis are summarized in Tables S7–S9.

Robustness Tests

(Added: We chose our explanatory variables based on theoretical considerations, as described in *Hess et al.* [2016]. To test the robustness of our analysis, we compared the results described above to two kinds of alternate regression analyses for the VWCI.)

← Add section with details of robustness checks and alternate predictor variables.

(Added: In the first series, we varied the interval over which we averaged the Köppen aridity index, considering the 45 year period 1970–2014, the 20 year period 1995–2014, and the 10-year period 2005–2014. This tests for sensitivity to recent extreme events versus the longer-term average climate. We observed very little difference in the regression coefficients and very little difference in the information criteria (Figure S5 and Tables S3 and S5). The 1985–2014 interval produced the best results (by a small and insignificant margin) and all of the intervals were statistically consistent with one another.)

(Added: Second, we performed regressions with additional or different explanatory variables (Figure S6 and Tables S4 and S6). Population density has been found to correlate well with voting patterns, and thus might affect water conservation policies [*Rodden*, 2016]. We substituted 2010 population-weighted population density and the rate of change of population density from 2000–2010 [*Wilson et al.*, 2012] for total population. The results were similar to those of our original analysis, and produced slightly and insignificantly inferior information criteria scores.)

(Added: We also considered that the area of an MSA might be important to collecting and distributing water, so we conducted regressions that included an additional explanatory variable representing the total area of the MSA as reported in the 2010 U.S. Census [*Wilson et al.*, 2012], the coefficient for area was consistent with zero and the information criteria scores were slightly and insignificantly inferior to our original analysis.)

(Added: Finally, we considered that in addition to mean personal income and relative purchasing power, the distribution of income might be important, so we performed regressions that included Gini indices of income inequality at both the MSA and the state levels, taken from the 2014 American Community Survey [*U.S. Census Bureau*, 2017]. The Gini index lies in the range zero (complete equality, with everyone receiving the same income) to one (complete inequality, with one person receiving all of the income and everyone else receiving nothing). In these regressions the coefficient for the state-level Gini index was positive and of comparable magnitude to the state-PVI coefficient, and the coefficient for the MSA-level Gini index was very small and consistent with zero. The information criteria scores were slightly and insignificantly inferior to our original model.)

(Added: All of these different analyses of VWCI consistently found that at the state level, the largest coefficients were for aridity and PVI, and at the MSA-level PVI, population (or population density), and population (or population-density) growth rates were positive and of comparable magnitude. The variations in coefficients across all of the alternate analyses were well within the 95% highest-density intervals of the posterior probability distribution.)

(Added: Similarly, in all of the analyses, the MSA-level PVI, population, and population growth coefficients were positive and distinct from zero. The values and the ranking of these three coefficients changed, but by amounts that were well within the posterior probability distributions. The remaining MSA-level variables were consistent with zero.)

(Added: The posterior distributions were considerably narrower than the prior distributions and lay well within those prior distributions, which indicates that they are not constrained by the priors. We tested this by varying the scales of the priors and by replacing the Cauchy priors on α_0, β, and γ with normal priors. The results were very similar to and consistent with the original analysis.)

(Added: We conclude from this that the results of our analysis are robust against many changes of time-spans, explanatory variables, and assumptions about priors.)

(Added: There are myriad other potential explanatory variables, but our concern that further exploration of alternative models might unintentionally become an exercise in "p-hacking" due to "garden of forking paths" effects [*Gelman and Loken*, 2014] led us to confine this analysis to our original set of variables, which we had previously chosen for theoretical reasons [*Hess et al.*, 2016].)

Figures S1–S6

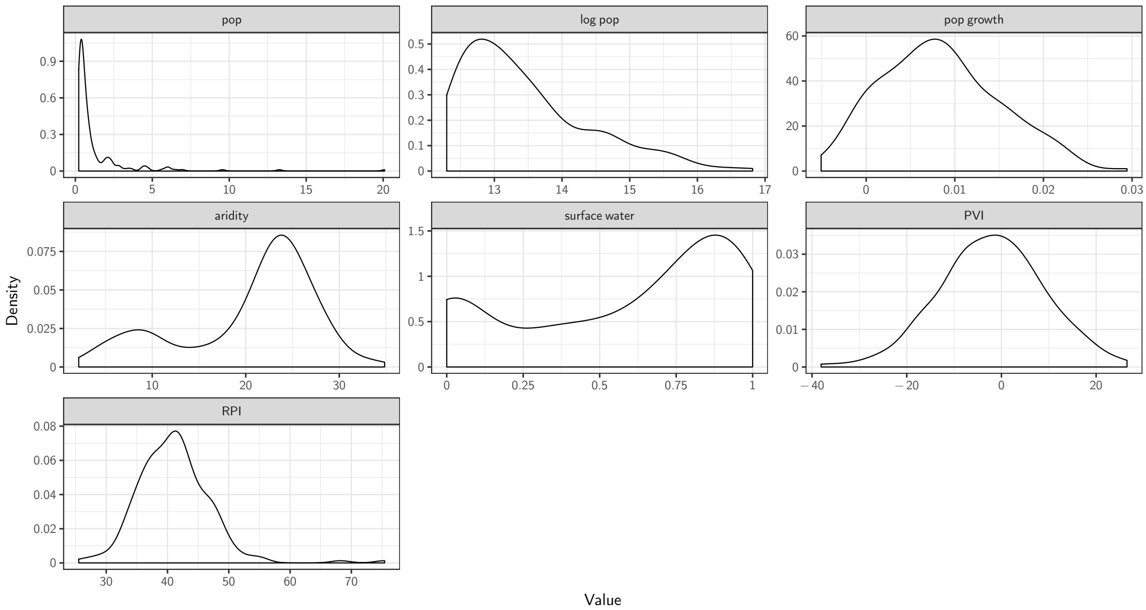

Figure 1. Kernel-density distribution of MSA-level covariates. Population in millions and RPI in thousands of chained 2009 dollars.

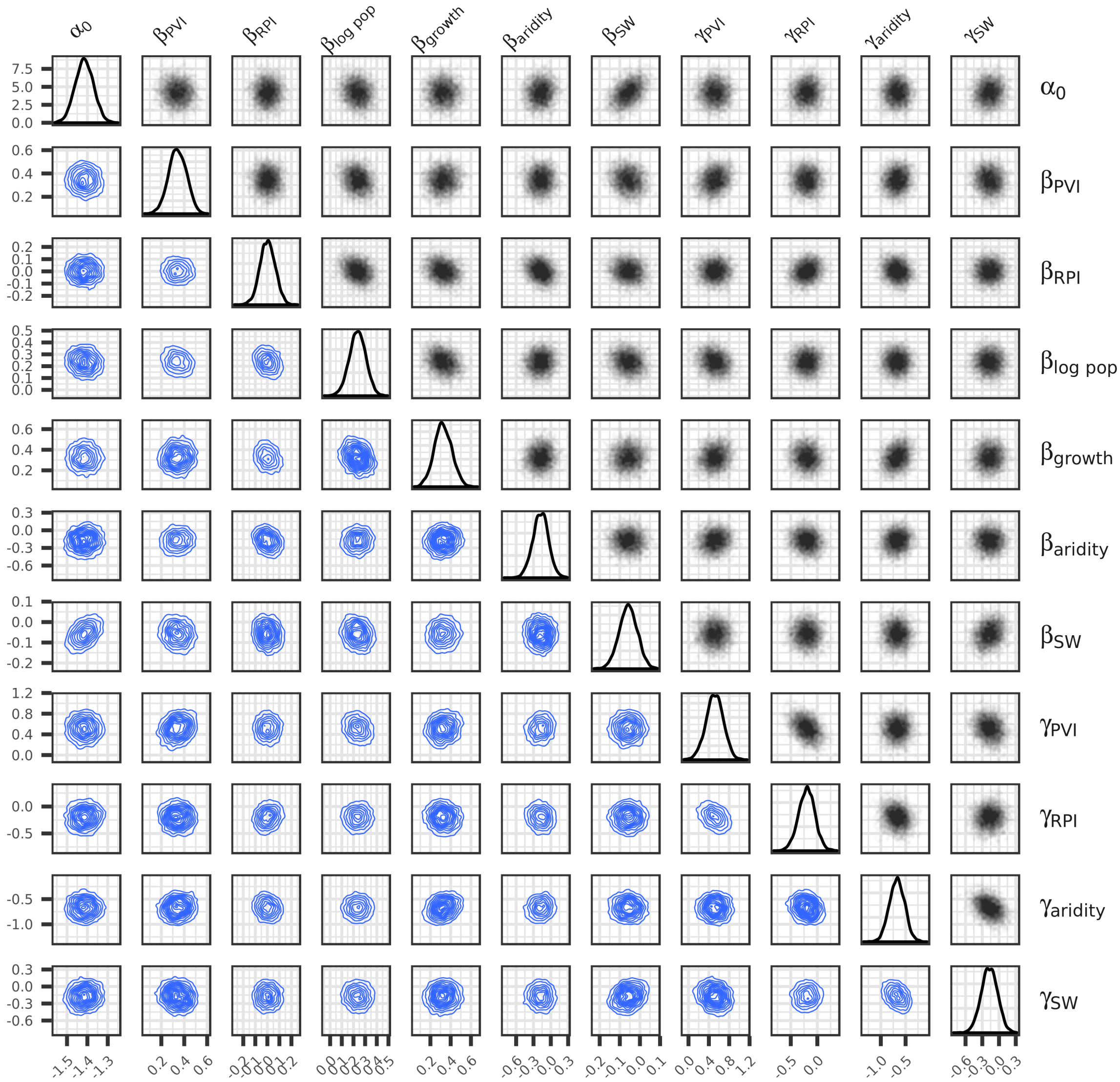

Figure 2. Correlation plot of posterior probability distribution of regression coefficients α, β, and γ for VWCI. The diagonal panels show the probability density for each coefficient, panels in the upper triangle show scatterplots of 4000 HMC samples, and panels in the lower triangle show joint probability density contours corresponding to the scatterplot in the upper triangle. Slight correlations are apparent, as between γ_{aridity} and γ_{SW}, γ_{PVI} and γ_{RPI}, and β_{SW} and α_0, but these are small enough not to pose problems apart from slightly increasing the uncertainty in the parameter estimates.

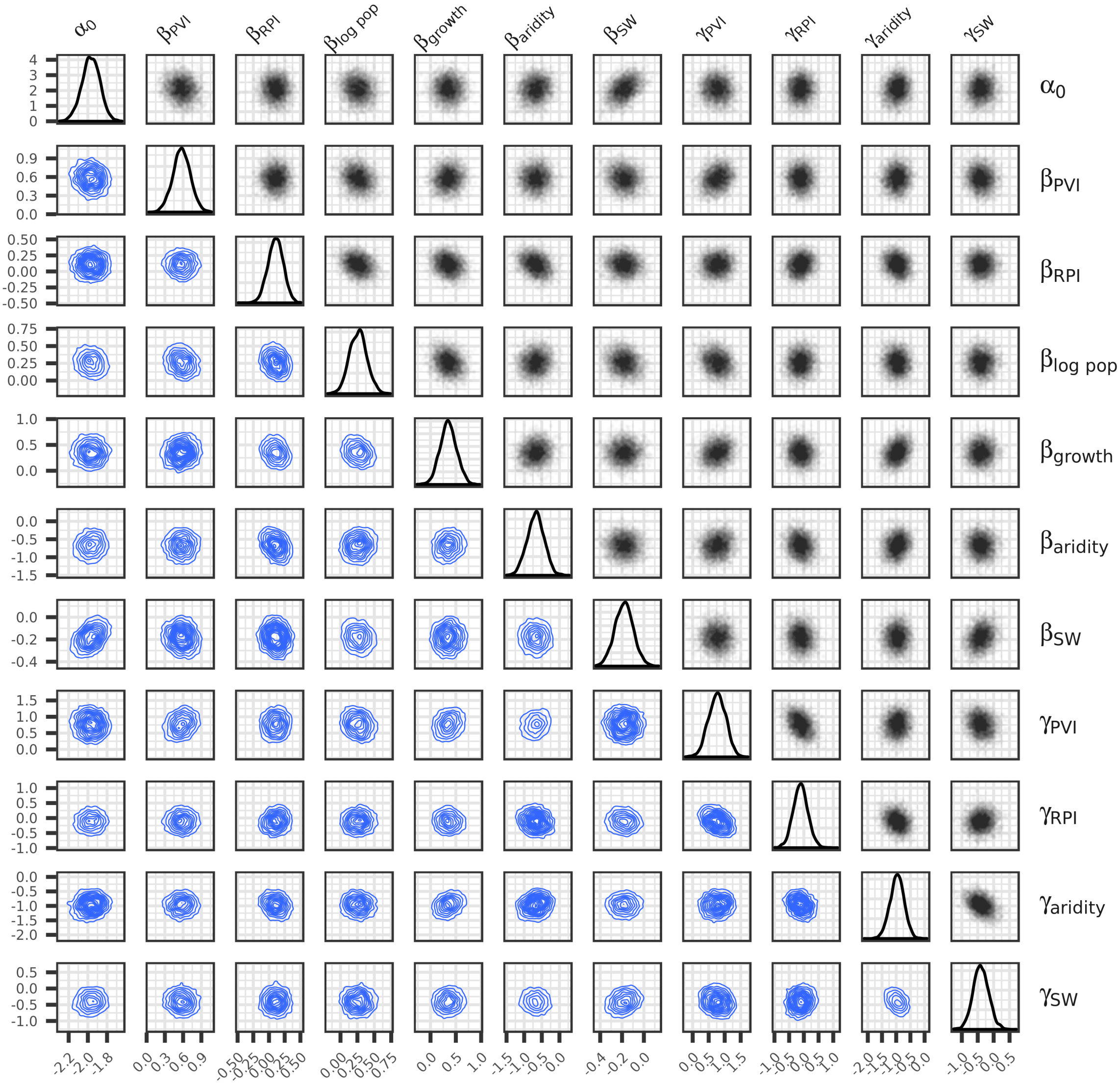

Figure 3. Correlation plot of posterior probability distribution of regression coefficients α, β, and γ for requirements.

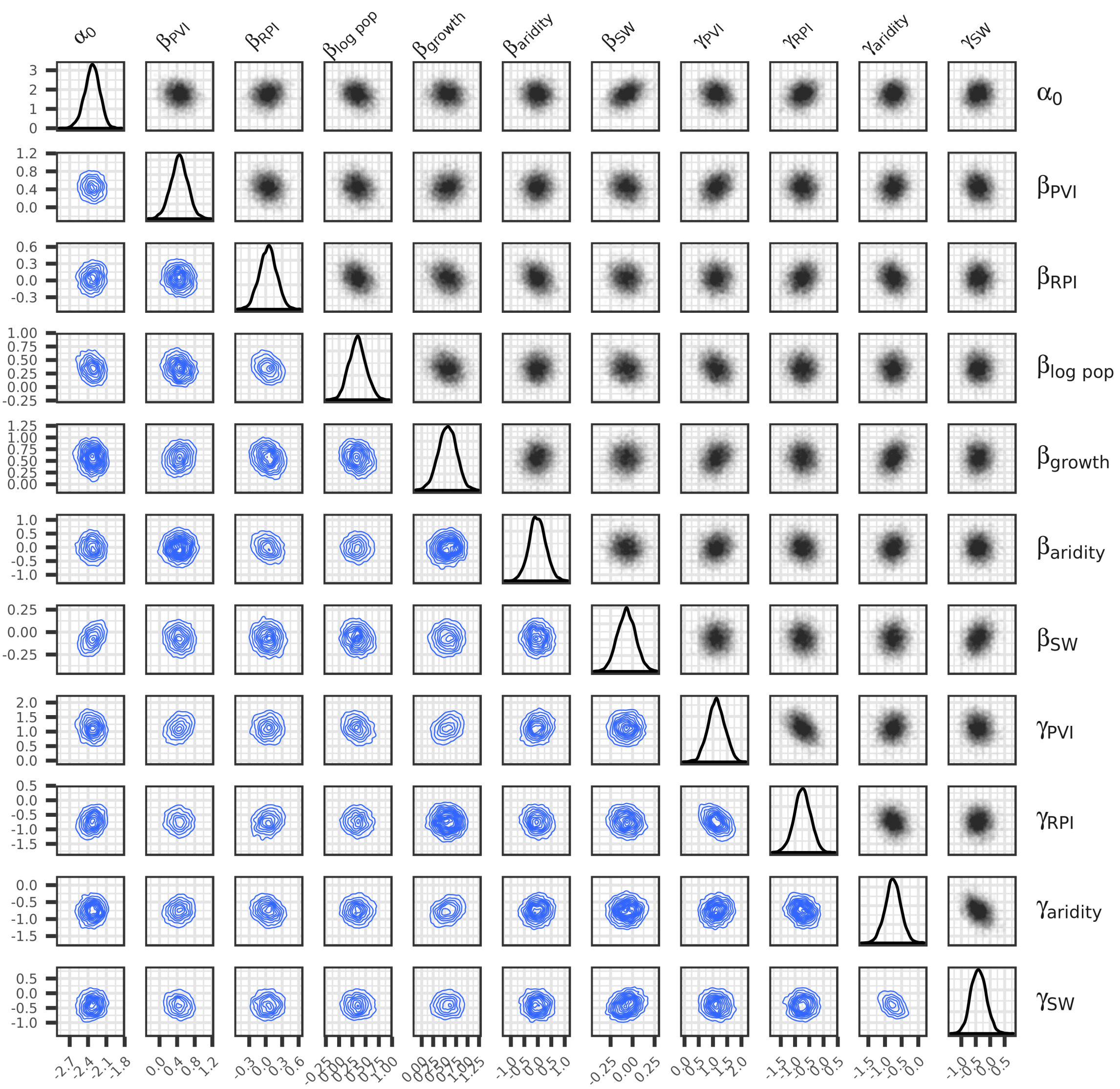

Figure 4. Correlation plot of posterior probability distribution of regression coefficients α, β, and γ for rebates.

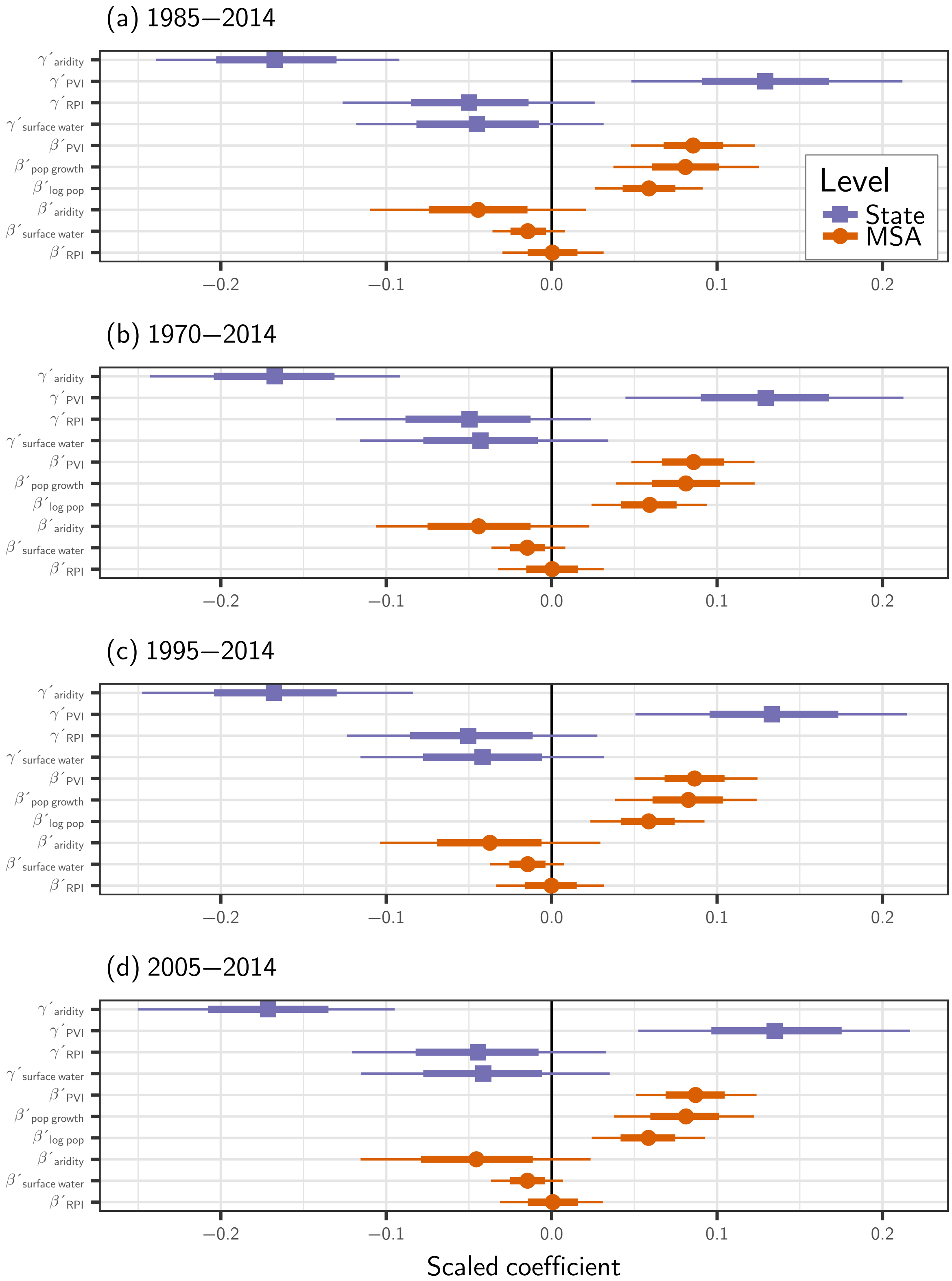

Figure 5. Regression coefficients for VWCI, averaging state and MSA aridity over different intervals.

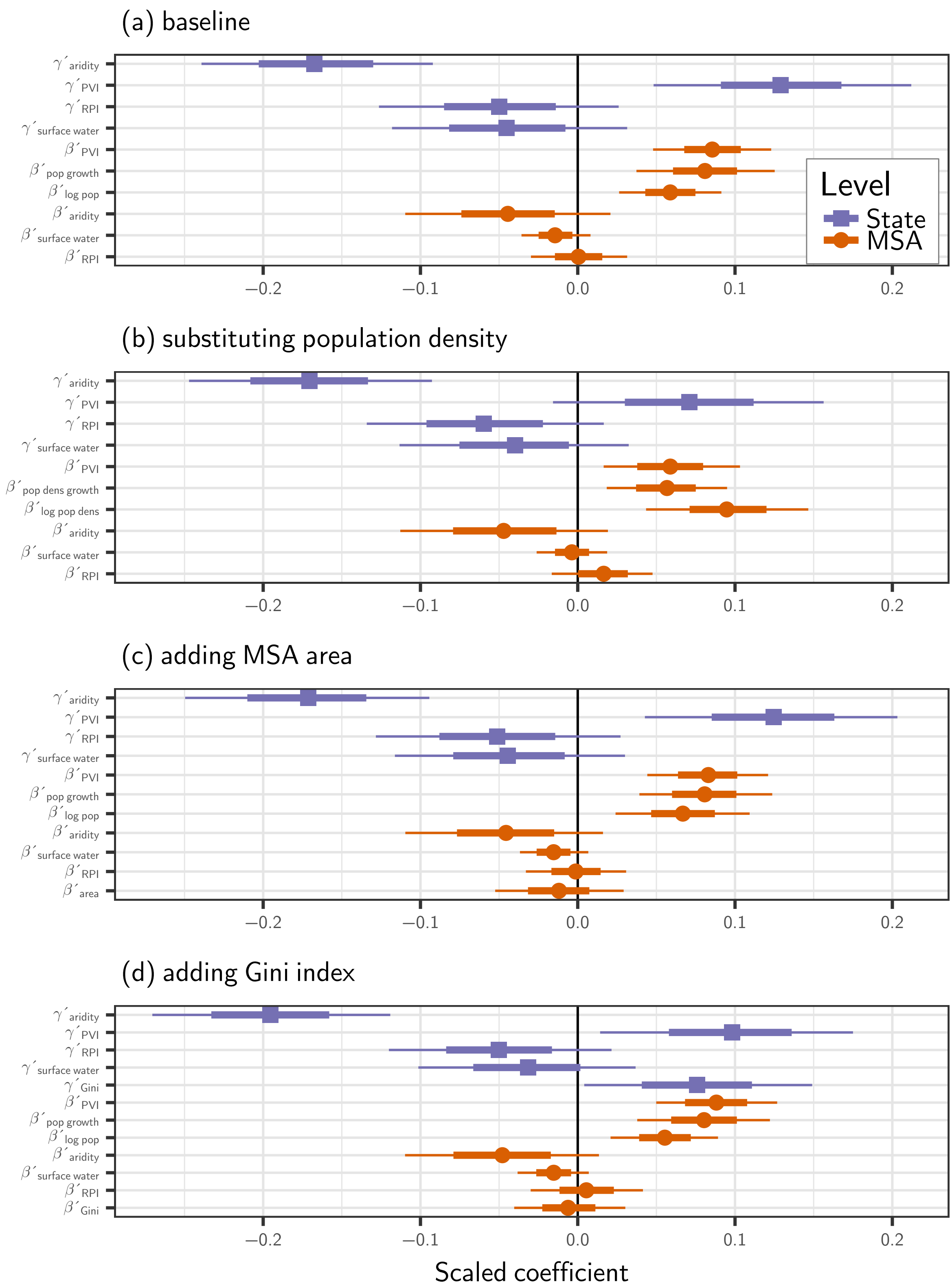

Figure 6. Regression coefficients for VWCI, with different covariates.

1 Tables S1–S9

Table S1 Caption

Table 1. Conservation scores and covariates for cities: VWCI = Vanderbilt Water Conservation Index (total # of conservation measures), Req. = # requirements, Reb. = # rebates, PVI = Cook Partisan Voting Index, Aridity = Köppen aridity index, RPI = per-capita real personal income (thousands of regionally adjusted chained 2009 dollars), Pop. = population (thousands), Growth = population growth rate (2010–2014), Surf. W. = surface-water fraction.

Table S2 Caption

Table 2. State-level covariates: PVI = Cook Partisan Voting Index, RPI = per-capita real personal income (thousands of regionally-adjusted chained 2009 dollars), Aridity = the Köppen aridity index, Surf. W. = the surface-water fraction.

Table S3

Model	LOO-IC	s.c. LOO-IC	$ELPD_{LOO}$	s.e. $ELPD_{LOO}$
hierarchical beta–binomial: 1985–2014	1247.3	20.4	-623.7	10.2
hierarchical beta–binomial: 1970–2014	1248.2	20.5	-624.1	10.2
hierarchical beta–binomial: 1995–2014	1249.6	20.4	-624.8	10.2
hierarchical beta–binomial: 2005–2014	1249.8	20.5	-624.9	10.3
hierarchical binomial: 1985–2014	1330.3	44.3	-665.2	22.1
hierarchical binomial: 2005–2014	1331.0	44.1	-665.5	22.0
hierarchical binomial: 1995–2014	1332.2	44.3	-666.1	22.1
hierarchical binomial: 1970–2014	1332.8	44.4	-666.4	22.2
single-level binomial: 2005–2014	2056.9	142.1	-1028.4	71.1
single-level binomial: 1970–2014	2060.4	143.2	-1030.2	71.6
single-level binomial: 1985–2014	2060.7	143.3	-1030.3	71.7
single-level binomial: 1995–2014	2063.4	143.5	-1031.7	71.8
single-level beta–binomial: 2005–2014	2429.2	182.0	-1214.6	91.0
single-level beta–binomial: 1970–2014	2439.8	182.2	-1219.9	91.1
single-level beta–binomial: 1985–2014	2440.3	182.3	-1220.1	91.1
single-level beta–binomial: 1995–2014	2440.8	181.4	-1220.4	90.7

Table 3. Comparing different climatological averaging periods and models: LOO = leave-one-out cross-validation, LOO-IC = LOO information criterion, ELPD = expected log pointwise predictive density, and s.e. indicates the standard error of estimates of quantities. Lower values of the information criteria and greater (less negative) values of ELPD indicate superior model performance. Models are labelled by the time-period for averaging aridity and the structure of the statistical mode.

Table S4

Model	LOO-IC	s.c. LOO-IC	$ELPD_{LOO}$	s.e. $ELPD_{LOO}$
hierarchical beta–binomial: baseline	1247.3	20.4	-623.7	10.2
hierarchical beta–binomial: with area,	1251.8	20.6	-625.9	10.3
hierarchical beta–binomial: with Gini index,	1252.5	20.3	-626.3	10.1
hierarchical beta–binomial: population density,	1261.8	23.1	-630.9	11.6
hierarchical binomial: baseline	1330.3	44.3	-665.2	22.1
hierarchical binomial: with area,	1336.6	44.3	-668.3	22.2
hierarchical binomial: with Gini index,	1337.2	44.2	-668.6	22.1
hierarchical binomial: population density,	1371.9	54.3	-686.0	27.1
single-level binomial: population density,	2037.6	112.3	-1018.8	56.1
single-level binomial: with area,	2059.4	145.4	-1029.7	72.7
single-level binomial: with Gini index,	2060.4	139.9	-1030.2	70.0
single-level binomial: baseline	2060.7	143.3	-1030.3	71.7
single-level beta–binomial: population density,	2364.0	136.9	-1182.0	68.4
single-level beta–binomial: baseline	2440.3	182.3	-1220.1	91.1
single-level beta–binomial: with Gini index,	2500.4	185.9	-1250.2	92.9
single-level beta–binomial: with area,	2514.5	192.8	-1257.3	96.4

Table 4. Comparing different models and sets of covariates with LOO-IC. Models are labelled by the covariates that differ from the baseline case and the baseline and by the structure of the statistical mode.

Table S5

Model	WAIC	s.c. WAIC	$ELPD_{WAIC}$	s.e. $ELPD_{WAIC}$
hierarchical beta–binomial: 1985–2014	1243.7	20.1	-621.9	10.1
hierarchical beta–binomial: 1970–2014	1244.4	20.2	-622.2	10.1
hierarchical beta–binomial: 2005–2014	1244.9	20.0	-622.4	10.0
hierarchical beta–binomial: 1995–2014	1245.5	20.1	-622.8	10.0
hierarchical binomial: 1985–2014	1317.3	43.4	-658.7	21.7
hierarchical binomial: 2005–2014	1318.7	43.4	-659.3	21.7
hierarchical binomial: 1970–2014	1318.8	43.5	-659.4	21.7
hierarchical binomial: 1995–2014	1319.6	43.6	-659.8	21.8
single-level binomial: 2005–2014	2056.0	142.0	-1028.0	71.0
single-level binomial: 1970–2014	2058.9	143.1	-1029.5	71.6
single-level binomial: 1985–2014	2059.1	143.1	-1029.6	71.6
single-level binomial: 1995–2014	2061.9	143.3	-1030.9	71.7
single-level beta–binomial: 2005–2014	2381.3	177.0	-1190.6	88.5
single-level beta–binomial: 1970–2014	2387.3	177.1	-1193.6	88.6
single-level beta–binomial: 1985–2014	2387.4	177.4	-1193.7	88.7
single-level beta–binomial: 1995–2014	2392.3	177.7	-1196.1	88.8

Table 5. Model comparison: WAIC = widely applicable information criterion (also known as the Watanabe-Aikake Information Criterion), ELPD = expected log-probability density, and s.e. indicates the standard error of estimates of quantities. Lower values of the information criteria and greater (less negative) values of ELPD indicate superior model performance. Models are labelled by the time-period for averaging aridity and the structure of the statistical mode.

Table S6

Model	WAIC	s.c. WAIC	$ELPD_{WAIC}$	s.e. $ELPD_{WAIC}$
hierarchical beta–binomial: baseline	1243.7	20.1	-621.9	10.1
hierarchical beta–binomial: with area,	1247.1	20.2	-623.5	10.1
hierarchical beta–binomial: with Gini index,	1248.6	19.9	-624.3	10.0
hierarchical beta–binomial: population density,	1259.1	23.0	-629.5	11.5
hierarchical binomial: baseline	1317.3	43.4	-658.7	21.7
hierarchical binomial: with area,	1322.5	43.5	-661.2	21.8
hierarchical binomial: with Gini index,	1324.3	43.5	-662.2	21.7
hierarchical binomial: population density,	1361.1	53.7	-680.5	26.9
single-level binomial: population density,	2037.5	112.3	-1018.7	56.2
single-level binomial: with Gini index,	2057.8	139.6	-1028.9	69.8
single-level binomial: with area,	2058.3	145.4	-1029.2	72.7
single-level binomial: baseline	2059.1	143.1	-1029.6	71.6
single-level beta–binomial: population density,	2314.6	133.1	-1157.3	66.6
single-level beta–binomial: baseline	2387.4	177.4	-1193.7	88.7
single-level beta–binomial: with Gini index,	2430.9	179.1	-1215.4	89.5
single-level beta–binomial: with area,	2438.5	185.6	-1219.2	92.8

Table 6. Model comparison with WAIC. Models are labelled by the covariates that differ from the baseline case and by the structure of the statistical mode.

Table S7 Caption

Table 7. Posterior probability distributions of regression coefficients for VWCI: mean, standard error of the mean, standard deviation of the posterior, quantiles of the posterior, and the Gelman-Rubin potential scale-reduction factor $\hat{R}$. γ coefficients correspond to state-level effects, β coefficients to MSA-level effects, δ coefficients represent state-level intercepts, α_0 is the overall intercept, and ϕ characterizes the overdispersion of the beta-binomial distribution. For more detail, see Materials and Methods.

Table S8 Caption

Table 8. Posterior probability distribution of regression coefficients for requirements

Table S9 Caption

Table 9. Posterior probability distribution of regression coefficients for rebates

Captions for Datasets S1–S4

Dataset S1: MSA-Level Data

This dataset contains MSA-level data: the FIPS (Federal Information Processing Standard) code for the MSA, the name of the MSA, the central city, state, latitude, longitude, VWCI, number of water-conservation requirements, number of rebate policies for water-conservation actions, the average annual precipitation (in millimeters) temperature (in Celsius), and Köppen aridity index, for the central city, the Cook Partisan Voting Index for the counties of the MSA, the 2014 population and average annual population growth rate from 2010–2014 for the MSA, the fraction of the municipal water supply derived from surface water, the BEA 2014 regional price parity and per-capita real personal income for the MSA (in chained regionally-adjusted 2009 dollars).

Dataset S2: MSA-Level Data Codebook

(Added: This dataset contains a codebook explaining the variable corresponding to each column in Dataset S1.)

Dataset S3: State-Level Data

This dataset contains stae-level data: the FIPS code for the state, the abbreviation and name of the state, the average annual precipitation (in millimeters), temperature (in Celsius), and Köppen aridity index for the state, the state-level Cook Partisan Voting Index, the fraction of the state water supply derived from surface water, and the BEA 2014 state-level regional price parity and per-capita real personal income (in chained regionally-adjusted 2009 dollars).

Dataset S4: State-Level Data Codebook

(Added: This dataset contains a codebook explaining the variable corresponding to each column in Dataset S3.)

Data Analysis Scripts S1

The zip file `scripts_S1.zip` contains R and Stan scripts to reproduce the regression analysis presented here. To reproduce the analysis, unzip the file with the scripts, copy Datasets S1 and S3 into the `data` subdirectory, and run the scripts `gilligan_vwci_ef_2017.R` and `gilligan_vwci_ef_si.R` in R.

This paper was produced with the following R packages:

- R version 3.4.3 (2017-11-30), `x86_64-pc-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF-8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF-8`, `LC_COLLATE=en_US.UTF-8`, `LC_MONETARY=en_US.UTF-8`, `LC_MESSAGES=en_US.UTF-8`, `LC_PAPER=en_US.UTF-8`, `LC_NAME=C`, `LC_ADDRESS=C`, `LC_TELEPHONE=C`, `LC_MEASUREMENT=en_US.UTF-8`, `LC_IDENTIFICATION=C`
- Running under: `Ubuntu 16.04.3 LTS`
- Matrix products: default
- BLAS: `/usr/lib/libblas/libblas.so.3.6.0`
- LAPACK: `/usr/lib/lapack/liblapack.so.3.6.0`
- Base packages: base, datasets, graphics, grDevices, methods, parallel, stats, utils
- Other packages: bindrcpp 0.2, bitops 1.0-6, broom 0.4.3, cowplot 0.9.1, dplyr 0.7.4, egg 0.2.0, extrafont 0.17, fontcm 1.1, forcats 0.2.0, geosphere 1.5-7, ggmap 2.7, ggplot2 2.2.1.9000, ggrepel 0.7.1, ggthemes 3.4.0, gridExtra 2.3, janitor 0.3.0, jgally 1.2.9.9999, jgmcmc 1.1.1, knitr 1.17, lazyeval 0.2.1, loo 1.1.0, lubridate 1.7.1, magrittr 1.5, maps 3.2.0, ncdf4 1.16, pacman 0.4.6, purrr 0.2.4, raster 2.6-7, RColorBrewer 1.1-2, RCurl 1.95-4.8, readr 1.1.1, readxl 1.0.0, rlang 0.1.4, rstan 2.16.2, shiny 1.0.5, shinystan 2.4.0, sp 1.2-5, StanHeaders 2.16.0-1, stringr 1.2.0, tibble 1.3.4, tidyr 0.7.2, tidyverse 1.2.1, viridis 0.4.0, viridisLite 0.2.0, xtable 1.8-2, zoo 1.8-0
- Loaded via a namespace (and not attached): assertthat 0.2.0, base64enc 0.1-3, bayesplot 1.4.0, bindr 0.1, cellranger 1.1.0, cli 1.0.0, colorspace 1.3-2,

colourpicker 1.0, compiler 3.4.3, crayon 1.3.4, crosstalk 1.0.0, digest 0.6.12, DT 0.2, dygraphs 1.1.1.4, evaluate 0.10.1, extrafontdb 1.0, foreign 0.8-69, GGally 1.3.2, glue 1.2.0, grid 3.4.3, gtable 0.2.0, gtools 3.5.0, haven 1.1.0, hms 0.4.0, htmltools 0.3.6, htmlwidgets 0.9, httpuv 1.3.5, httr 1.3.1, igraph 1.1.2, inline 0.3.14, jpeg 0.1-8, jsonlite 1.5, lattice 0.20-35, mapproj 1.2-5, markdown 0.8, matrixStats 0.52.2, mime 0.5, miniUI 0.1.1, mnormt 1.5-5, modelr 0.1.1, munsell 0.4.3, nlme 3.1-131, pkgconfig 2.0.1, plyr 1.8.4, png 0.1-7, proto 1.0.0, psych 1.7.8, R6 2.2.2, Rcpp 0.12.14, reshape 0.8.7, reshape2 1.4.2, RgoogleMaps 1.4.1, rjson 0.2.15, rsconnect 0.8.5, rstudioapi 0.7.0-9000, Rttf2pt1 1.3.4, rvest 0.3.2, scales 0.5.0.9000, shinyjs 0.9.1, shinythemes 1.1.1, stats4 3.4.3, stringi 1.1.6, threejs 0.3.1, tools 3.4.3, xml2 1.1.1, xts 0.10-0

References

Associated Press (2006). Many state legislative races are uncontested. Retrieved from `http://www.nbcnews.com/id/15446775/ns/politics/t/many-state-legislative-races-are-uncontested`.

Brown, K. P., & Hess, D. J. (2017). The politics of water conservation: Identifying and overcoming political barriers to successful policies. *Water Policy*, *19*, 304–321. `https://doi.org/10.2166/wp.2016.089`.

Carpenter, B., Gelman, A. Hoffman, M. Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*, 1–32. `https://doi.org/10.18637/jss.v076.i01`.

CQ Press (2016). Voting and elections collection. Retrieved from `http://library.cqpress.com/elections/`.

Fiorina, M. P. (2016). *The (re)nationalization of Congressional elections* (Essays on Contemporary Politics 7). Stanford, CA: Hoover Institute. Retrieved from `https://www.hoover.org/sites/default/files/research/docs/fiorina_renationalizationofcongressionalelections_7.pdf`.

Fleck, J. (2016). *Water is for Fighting Over and Other Myths about Water in the West*. Washington, DC: Island.

Gelman, A., & Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*, 460–465.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, *2*, 1360–1383. https:\\doi.org/10.1214/08-AOAS191.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis*, 3rd ed. Boca Raton, FL: CRC.

Gleick, P. H. (2002). Water management: Soft water paths. *Nature*, *418*, 373–373, https://doi.org/10.1038/418373a.

Hawkins, C. B., & B. A. Nosek (2012). Motivated independence? implicit party identity predicts political judgments among self-proclaimed independents. *Personality and Social Psychology Bulletin*, *38*, 1437–1452. https://doi.org/10.1177/0146167212452313.

Hess, D. J., Wold, C. A., Hunter, E., Nay, J. J., Worland, S., Gilligan, J., & Hornberger, G. M. (2016). Drought, risk, and institutional politics in the American Southwest. *Sociological Forum*, *31*, 807–827. https://doi.org/10.1111/socf.12274.

Hess, D. J., Wold, C. A., Worland, S. C., & Hornberger, G.M. (2017). Measuring urban water conservation policies: Toward a comprehensive index. *Journal of the American Water Resources Association*, *53*, 442–455. https://doi.org/10.1111/1752-1688.12506.

Hornberger, G. M., Hess, D. J., & Gilligan, J. (2015). Water conservation and hydrological transitions in cities in the United States. *Water Resources Research*, *51*, 4635–4649. https://doi.org/10.1002/2015WR016943.

Klarner, C. (2015). Competitiveness in state legislative elections: 1972–2014. *Ballotopedia*. Retrieved from https://ballotpedia.org/Competitiveness_in_State_Legislative_Elections:_1972-2014.

Maggioni, E. (2014). Water demand management in time of drought: What matters for water conservation. *Water Resources Research*, *51*, 215–139. https://doi.org/10.1002/2014wr016301.

Matsuura, K., & Willmott, C. J. (2015a). *Terrestrial Air Temperature: 1900–2014 Gridded Monthly Time Series (Version 4.01)* (Department of Geography, University of Delaware). Retrieved from http://climate.geog.udel.edu/~climate/

html_pages/Global2014/README.GlobalTsT2014.html.

Matsuura, K., and Willmott, C. J. (2015b). *Terrestrial Precipitation: 1900–2014 Gridded Monthly Time Series (Version 4.01)* (Department of Geography, University of Delaware). Retrieved from http://climate.geog.udel.edu/~climate/html_pages/Global2014/README.GlobalTsP2014.html.

Maupin, M., Kenny, J., Hutson, S., Lovelace, J., Barber, N., & Linsey, K. (2014). *Estimated use of water in the United States in 2010* (USGS Circular 1405). Washington, DC: U.S. Geological Survey. Retrieved from https://doi.org/10.3133/cir1405.

Melillo, J. M., Richmond, T. C., & Yohe G. W. (Eds.). (2014). *Climate Change Impacts in the United States: The Third National Climate Assessment*. Washington, DC: U.S. Global Change Research Program. Retrieved from https://doi.org/10.7930/J0Z31WJ2.

Quan, C., Shuang, H., Torsten, U., Chunhua, Z., and Liu, Y.-S. (2013). Validation of temperature—precipitation based aridity index: Paleoclimatic implications. *Palaeogeography Palaeoclimatology Palaeoecology*, *386*, 86–95. https://doi.org/10.1016/j.palaeo.2013.05.008.

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rodden, J. (2016). The geographic distribution of political preferences. *Annual Review of Political Science*, *13*, 331–340. https://doi.org/10.1146/annurev.polisci.12.031607.092945.

Saurí, D. (2013). Water conservation: Theory and evidence in urban areas of the developed world, *Annual Review of Environment and Resources*, *38*, 227–248. https://doi.org/10.1146/annurev-environ-013113-142651.

Stan Development Team (2016). *Stan Modeling Language: User's Guide and Reference Manual*, version 2.14.0. Retrieved from http://mc-stan.org/documentation/.

Svara, J. H. (2003). *Two Decades of Continuity and Change in American City Councils*, Washington, DC: National League of Cities. Retrieved from http://www.skidmore.edu/~bturner/Svara%20citycouncilrpt.pdf.

Switzer, D., and Vedlitz, A. (2016). Green colored lenses: Worldviews and motivated reasoning in the case of local water scarcity. *Environment and Behavior*, *49*, 719–744. https://doi.org/10.1177/0013916516669391.

U.S. Bureau of Economic Analysis (2016a). Real personal income and regional price parities. In *Regional Economic Accounts Methodology* (Report RPP2016). Retrieved from http://www.bea.gov/regional/pdf/RPP2016_methodology.pdf.

U.S. Bureau of Economic Analysis (2016b). *RPI1: Real per-capita personal income*. Retrieved from http://www.bea.gov/iTable/index.cfm.

U.S. Census Bureau (2016). *Metropolitan and micropolitan statistical areas*. Retrieved from https://www.census.gov/data/datasets/2016/demo/popest/total-metro-and-micro-statistical-areas.html/.

U.S. Census Bureau (2017). Gini index of income inequality. In *2014 American Community Survey 1-Year Estimates*. Retrieved from https://www.census.gov/data/datasets/2016/demo/popest/total-metro-and-micro-statistical-areas.html/.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, textit27, 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.

Vose, R. S., Applequist, S., Squires, M., Durre, I., Menne, M. J., Williams, C. N., Fenimore, C., Gleason, K., & Arndt, D. (2014). Improved historical temperature and precipitation time series for U.S. climate divisions, *Journal of Applied Meteorology and Climatology*, *53*, 1232–1251. https://doi.org/10.1175/JAMC-D-13-0248.1.

Wasserman, D. (2013). *Introducing the 2014 Cook Political Report Partisan Voter Index*. Cook Political Report. Retrieved from http://cookpolitical.com/story/5604.

Wilson, S. G. and Plane, D. A. and Mckun, P. J. and Fischetti, T. R. & Goworowska, J. (2016). *Patterns of Metropolitan and Micropolitan Population Change: 2000–2010* (Report No. C2010SR-01). Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/library/publications/2012/dec/c2010sr-01.html

List of Changes

Replaced: ~~S2~~ replaced with: S4, on page 1.

Replaced: ~~S5–S7~~ replaced with: S7–S9, on page 1.

Replaced: ~~sharply peaked~~ replaced with: skewed, with a sharp peak, on page 2.

Added: , and a long tail at higher populations, on page 2.

Added: Our Monte Carlo analysis sampled four Markov chains, allowing each chain to warm-up and tune sampling parameters for the first 1000 iterations and then sampling each chain for 1000 more iterations, yielding a total of 4000 samples. Each sample is a vector of length 57, with values for each of the parameters α_0, β_j, γ_k, δ_{state}, σ, and ϕ, where j indexes over the six MSA-level covariates, k indexes over the four state-level covariates, and *state* indexes over 44 of the 45 states (leaving one out for identifiability). The samples approximate random draws from the joint posterior probability distribution of the parameters, given the priors and the observed data. Thus, the statistics of the sampled values approximate the statistics of the joint posterior distribution. , on page 2.

Replaced: ~~Pairwise correlation plots of the posterior probability distributions of regressions parameters (Figures S3–S5) are smooth and show little correlation. The Hamiltonian Monte Carlo calculations proceeded without any divergences or exceessive tree depths, and the Gelman-Rubin $\hat{R}$ potential scale-reduction factor (Tabs S7–S9) converged to ≥ 0.999 for each parameter.~~ replaced with: Collinearity among the predictor variables is diagnosed by observing correlations in the joint posterior probability distributions of the regression coefficients [*Stan Development Team*, 2016, pp. 288–293]. Inefficient sampling due to varying curvature in the log-probability manifold or poorly chosen priors can be diagnosed by irregularities in joint posterior distributions [*Stan Development Team*, 2016, pp. 316–321]. Pairwise correlation plots of the Monte-Carlo samples for the regression coefficients in our models of VWCI, requirements, and rebates (Figures S2–S4) are smooth with little correlation and give no cause for concern. In addition, the Hamiltonian Monte Carlo calculations proceeded without any divergences or exceessive tree depths, and the Gelman-Rubin $\hat{R}$ potential scale-reduction factor converged to ≥ 0.999 for each parameter [*Stan Development Team*, 2016]. , on page 2.

Added: We used several model-selection criteria in deciding whether to model the VWCI, requirements, and rebates as binomial or beta-binomial processes. At each joint sample of the model parameters in the Monte-Carlo process, we both computed the log-likelihood of the observed data under the sampled parameters and also generated posterior predictions obtained by drawing simulated observations from binomial or beta-binomial distribution at each joint sample of the model parameters. , on page 3.

Added: Visual comparisons of distributions of posterior predictions to observed data and comparisons of the posterior predictions of mean, maximum, and minimum VWCI over the cities in our data set showed better agreement for the overdispersed β-binomial process than for a purely binomial one [*Gelman et al.*, 2014]. , on page 3.

Replaced: ~~Leave-one-out cross-validation (Table S3) and the Widely Applicable Information Criterion (Table S4) were used for model selection (overdispersed beta-binomial versus binomial and hierarchical versus single-level regressions).~~ replaced with: A separate test for overdispersion, which accounts for the danger of overfitting by introducing new free parameters, uses the Leave-One-Out cross-validation Information Criterion (LOO-IC) or the Widely Available Information Criterion (WAIC, also known as the Watanabe-Aikake Information Criterion), obtained by Pareto-smoothed importance sampling [*Vehtari et al.*, 2017]. Both information criteria favored the overdispersed beta-binomial distribution over a pure binomial, and also strongly favored hierarchical over single-level models (Tables S3–S6). Our choice to use very weakly informative priors in our model reduces the accuracy of our estimates of LOO-IC and WAIC [*Vehtari et al.*, 2017], but we do not worry overly about this potential inaccuracy both because the posterior prediction test yields the same results and because a pure binomial model gives very similar results to those presented here. , on page 3.

Added: We had initially used the interval 1970–2014 for climatological averaging, but in response to reviewer comments, we investigated other averaging intervals and switched to the 1984–2014 interval because it performed slightly, but insignificantly better according to the information criteria. There are no important differences between between the two analyses (Figure S5). , on page 4.

Added: We chose our explanatory variables based on theoretical considerations, as described in *Hess et al.* [2016]. To test the robustness of our analysis, we compared the results described above to two kinds of alternate regression analyses for the VWCI. , on page 4.

Added: In the first series, we varied the interval over which we averaged the Köppen aridity index, considering the 45 year period 1970–2014, the 20 year period 1995–2014, and the 10-year period 2005–2014. This tests for sensitivity to recent extreme events versus the longer-term average climate. We observed very little difference in the regression coefficients and very little difference in the information criteria (Figure S5 and Tables S3 and S5). The 1985–2014 interval produced the best results (by a small and insignificant margin) and all of the intervals were statistically consistent with one another. , on page 4.

Added: Second, we performed regressions with additional or different explanatory variables (Figure S6 and Tables S4 and S6). Population density has been found to correlate well with voting patterns, and thus might affect water conservation policies [*Rodden*, 2016]. We substituted 2010 population-weighted population density and the rate of change of population density from 2000–2010 [*Wilson et al.*, 2012] for total population. The results were similar to those of our original analysis, and produced slightly and insignificantly inferior information criteria scores. , on page 4.

Added: We also considered that the area of an MSA might be important to collecting and distributing water, so we conducted regressions that included an additional explanatory variable representing the total area of the MSA as reported in the 2010 U.S. Census [*Wilson et al.*, 2012], the coefficient for area was consistent with zero and the information criteria scores were slightly and insignificantly inferior to our original analysis. , on page 4.

Added: Finally, we considered that in addition to mean personal income and relative purchasing power, the distribution of income might be important, so we performed regressions that included Gini indices of income inequality at both the MSA and the state levels, taken from the 2014 American Community Survey [*U.S. Census Bureau*, 2017]. The Gini index lies in the range zero (complete equality, with everyone receiving the same income) to one (complete inequality, with one person receiving all of the income and everyone else receiving nothing). In these regressions the coefficient for the state-level Gini index was positive and of comparable magnitude to the state-PVI coefficient, and the coefficient for the MSA-level Gini index was very small and consistent with zero. The information criteria scores were slightly and insignificantly inferior to our original model. , on page 5.

Added: All of these different analyses of VWCI consistently found that at the state level, the largest coefficients were for aridity and PVI, and at the MSA-level PVI, population (or population density), and population (or population-density) growth rates were positive and of comparable magnitude. The variations in coefficients across all of the alternate analyses were well within the 95% highest-density intervals of the posterior probability distribution. , on page 5.

Added: Similarly, in all of the analyses, the MSA-level PVI, population, and population growth coefficients were positive and distinct from zero. The values and the ranking of these three coefficients changed, but by amounts that were well within the posterior probability distributions. The remaining MSA-level variables were consistent with zero. , on page 5.

Added: The posterior distributions were considerably narrower than the prior distributions and lay well within those prior distributions, which indicates that they are not constrained by the priors. We tested this by varying the scales of the priors and by replacing the Cauchy priors on α_0, β, and γ with normal priors. The results were very similar to and consistent with the original analysis. , on page 5.

Added: We conclude from this that the results of our analysis are robust against many changes of time-spans, explanatory variables, and assumptions about priors. , on page 6.

Added: There are myriad other potential explanatory variables, but our concern that further exploration of alternative models might unintentionally become an exercise in "p-hacking" due to "garden of forking paths" effects [*Gelman and Loken*, 2014] led us to confine this analysis to our original set of variables, which we had previously chosen for theoretical reasons [*Hess et al.*, 2016]. , on page 6.

Added: This dataset contains a codebook explaining the variable corresponding to each column in Dataset S1. , on page 17.

Added: This dataset contains a codebook explaining the variable corresponding to each column in Dataset S3. , on page 18.