Supporting Information (Online)

Bureaucratic Responsiveness to LGBT Americans

Kenneth Lowande and Andrew Proctor

Study Diagnostics

Power

We simulated design power for each hypothesis with the observed N and distributions of each interactive variable.[11] These simulations are more conservative than the analyses presented in the paper, as they assume simple random assignment (rather than blocking) and exclude additional pre-treatment covariates. Figure A1 plots proportion of 1,000 simulations which return significant (i.e. $p < 0.05$) differences in the expected direction by effect size. The mean for the dichotomous outcome is set to 0.71, the value observed in the data. By conventional power thresholds, this design was well-powered to detect substantively moderate effects (+/-4%) for the overall discrimination hypothesis, large effects (+/-8%) for the legacy hypothesis, and under-powered for the final subgroup effects. Accordingly, we note that our experiment likely cannot detect substantively small effects in the theoretically-informed subgroups.

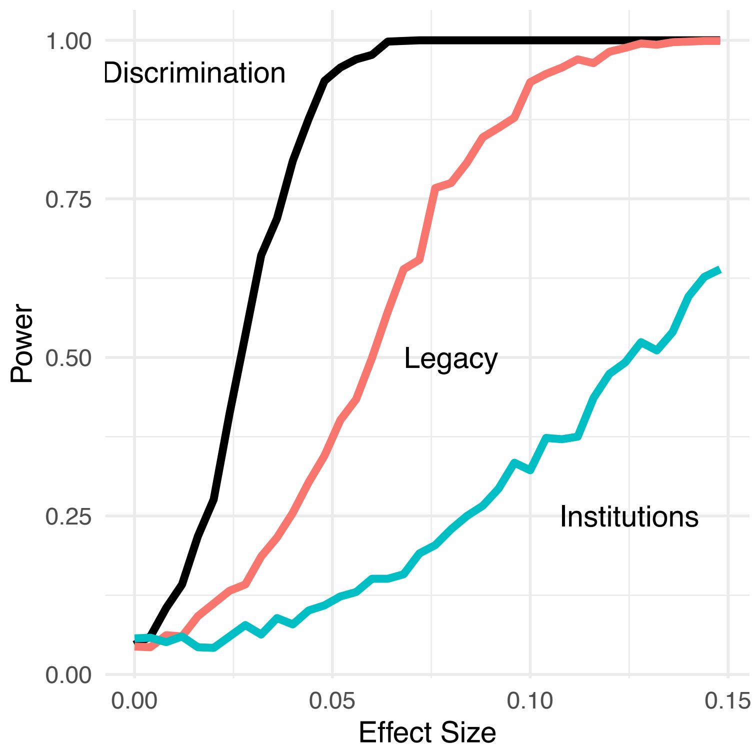

Figure A1 – Power Analysis

Missing Jurisdictions

What jurisdictions were less likely to be included in our sample? Figure A2 plots missing counties and the proportion of missing cities in municipal-granting states. In total, there were roughly 704 jurisdictions without online contact information. The exact number is unknown for reasons we

[11]The code for this simulation was adapted from "Power Analysis Simulations in R," url: `http://egap.org/content/power-analysis-simulations-r`. Last accessed August 1, 2019.

elaborate on in the Discussion. While there appears to be no relationship between population and inclusion, excluded jurisdictions tended to be more conservative. In models controlling for population, a standard-deviation increase in two-party Democratic presidential vote share from 2008 and 2012 is associated with a 3.5% increase in the probability of inclusion. (The unconditional probability of inclusion was 86%.) Note also, conservatism is correlated with census designations for "rural" geographic areas.

Thus, to address the concern that the exclusion of these cases may bias observed differences, we re-estimated the results with inverse probability weights. Specifically, in a first-stage logistic regression, we modeled the probability of inclusion in sample as a function of population, two-party Democratic presidential vote share from 2008 and 2012, selection method, official level, and state legal history. We then repeat the method for simulating responsiveness differences across sexual orientation in outcome with the inverse predicted probabilities from the first stage as regression weights. We report results for overall response and congratulatory language in A3. As the Figure indicates, typically, this procedure results in more precise estimates without appreciable differences in our results.

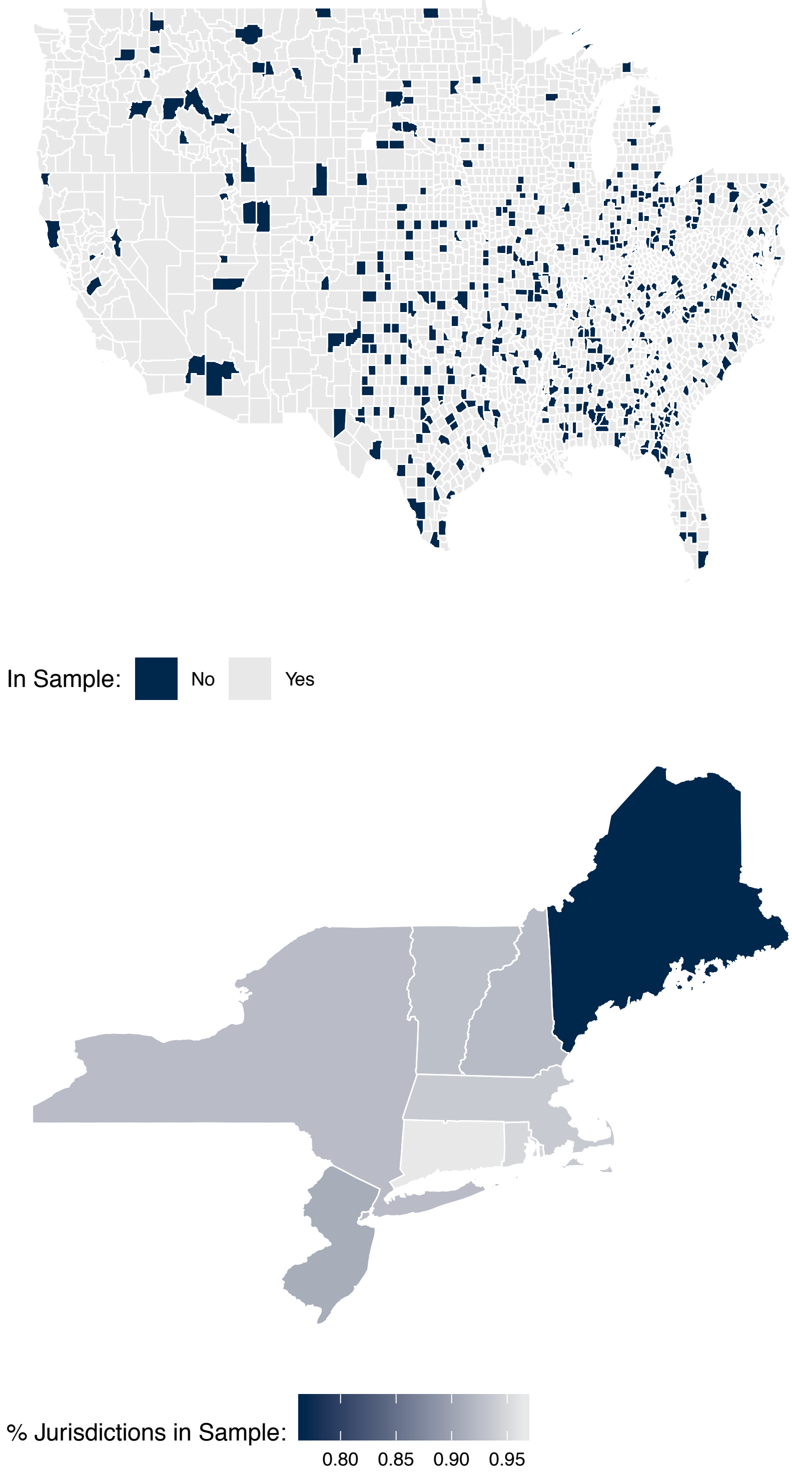

Figure A2 – Missing Jurisdictions

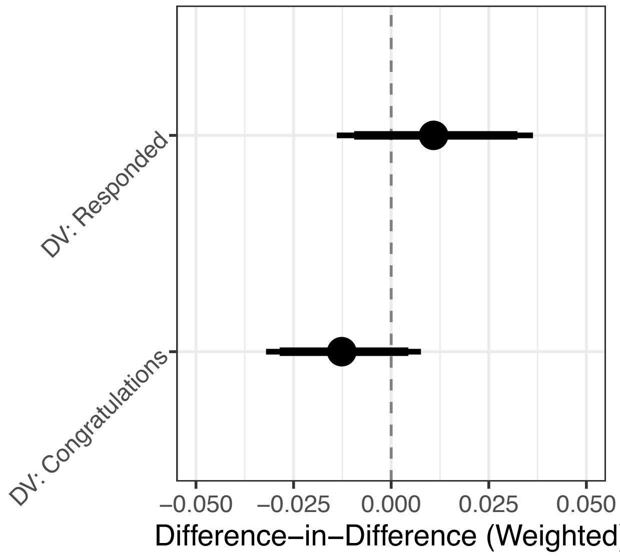

Figure A3 – Simulated Differences Robust to Inverse Probability Weighting. Plots marginal change in predicted probability of response quality, where negative values indicate a lower rate for same-sex couples. Weights are the inverse probability of inclusion in sample based on a logistic regression that includes population, Obama voteshare, official level, selection method, and state legal history as independent variables.

Table A1 – Imbalance Across Treatment Conditions

	Male, Male	Female, Female	Male, Female	Female, Male
Population	54825.78	54184.71	66317.08	57256.64
Democratic Voteshare	0.45	0.45	0.46	0.45
Appointed	278.00	278.00	278.00	278.00
Contact Forms	56.00	98.00	102.00	79.00
Errors	81.00	117.00	122.00	97.00
Bounces	118.00	150.00	158.00	143.00
N	1067.00	1094.00	1088.00	1072.00

Response Coding Procedures

Cost, waiting period, and time valid can vary by unit. Many webpages did not contain the necessary information to gauge accuracy. In addition, some respondents indicated that their webpages were out of date. Therefore, we simply code whether each question was answered. We used single-blind coding assess whether the response addressed the questions posed by the inquiry. Specifically, we followed the procedure below:

1. Removed treatment condition identifying information, and all names and partner synonyms from the text of replies,

2. Randomized the order of response coding by rater.

3. Coded each of the 4736 observations according to the following rules:

Cost. "Did the official answer the question about the cost of a marriage license?" An exact amount had to be provided. If the official indicated that the cost is dependent on residency or premarital counseling, they must provide costs for with/without each status to be counted as responsive.

Valid. "Did the official answer the question about how long the marriage license was valid?" Official must indicate if the license expires, and when. Replying that the emailer can "apply anytime" is not sufficient.

When. "Did the official answer the question about when you could apply for a marriage license?" If they said "you can apply anytime now" (e.g. there was no statutorily mandated waiting period) it must include office hours to be counted as responsive. Thus, either notification of office hours or information about a waiting period is considered responsive.

Felicitations. "Did the official say 'congratulations' or wish 'good luck' at the wedding/marriage, etc.?" Email signatures and salutations like "best" or "have a nice day" are not counted as a felicitation.

4. Disagreements between raters were resolved by a third, double-blind coder—who was unaware of the study purpose.

Table B2 – Interrater Reliability for Response Quality

	Cost?	Valid?	When?	Felicitations
% Disagree	0.01	0.03	0.05	0.02
No. Disagree	47.00	143.00	226.00	112.00
Cohen's Kappa	0.98	0.94	0.89	0.88

Table B3 – Interrater Reliability for Response Quality (Non-Response Removed)

	Cost?	Valid?	When?	Felicitations
% Disagree	0.02	0.06	0.10	0.05
No. Disagree	47.00	143.00	226.00	112.00
Cohen's Kappa	0.89	0.76	0.80	0.87

In general, coding was consistent across raters to suggest that a third coder to resolve discrepancies was sufficient. Not surprisingly, cost was mostly clear for human coders, whereas the most disagreement arose for the more generic question of "when" the emailer could apply and whether the reply included some kind of felicitation.

Estimation Results

Table C4 – Full Estimation Results for Figures 3, 4, 5, 7, and 8

	Dependent variable:							
	DV: Response	DV: Congrats	DV: Response	DV: Congrats	DV: Cost	DV: Valid	DV: When	DV: Felicit.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female-Female	0.31	0.07	0.32	0.08	0.22	0.15	0.22	0.04
	(0.10)	(0.13)	(0.10)	(0.13)	(0.09)	(0.09)	(0.09)	(0.13)
Male-Female	0.01	0.13	0.01	0.13	−0.01	−0.02	0.09	0.07
	(0.09)	(0.13)	(0.09)	(0.13)	(0.09)	(0.09)	(0.09)	(0.13)
Female-Male	0.18	0.15	0.19	0.16	0.20	0.13	0.21	0.09
	(0.10)	(0.13)	(0.10)	(0.13)	(0.09)	(0.09)	(0.09)	(0.13)
No SSM Ban	0.46	0.59			0.24	0.22	0.31	0.60
	(0.09)	(0.14)			(0.08)	(0.08)	(0.09)	(0.14)
No Sodomy Ban			0.44	0.14				
			(0.09)	(0.14)				
County-level	−0.10	−0.46	−0.09	−0.61	−0.15	−0.15	0.10	−0.50
	(0.11)	(0.13)	(0.11)	(0.14)	(0.09)	(0.09)	(0.10)	(0.13)
Population	0.0000	−0.00	0.0000	0.0000	−0.0000	0.00	−0.00	−0.00
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Elected	0.11	0.27	0.06	0.25	0.29	0.32	0.13	0.28
	(0.09)	(0.12)	(0.09)	(0.12)	(0.08)	(0.08)	(0.09)	(0.12)
Unk. Selection	−0.08	−0.09	−0.11	−0.12	0.24	0.33	−0.69	−0.06
	(0.25)	(0.32)	(0.25)	(0.32)	(0.23)	(0.23)	(0.27)	(0.32)
Obama Vote	0.34	−0.27	0.52	0.002	−0.14	−0.21	−0.11	−0.44
	(0.27)	(0.37)	(0.26)	(0.36)	(0.24)	(0.24)	(0.25)	(0.37)
Contact Form	0.29	−0.15	0.29	−0.14	0.24	0.27	0.17	−0.11
	(0.14)	(0.17)	(0.14)	(0.17)	(0.12)	(0.12)	(0.12)	(0.17)
Constant	0.22	−2.24	0.16	−1.93	−0.32	−0.36	−1.02	−2.13
	(0.17)	(0.24)	(0.18)	(0.25)	(0.16)	(0.16)	(0.17)	(0.25)
Observations	4,321	4,321	4,321	4,321	4,321	4,321	4,321	4,321
Log Likelihood	−2,576.54	−1,669.26	−2,577.49	−1,678.04	−2,973.04	−2,974.98	−2,816.88	−1,658.02
Akaike Inf. Crit.	5,175.07	3,360.52	5,176.97	3,378.09	5,968.08	5,971.95	5,655.77	3,338.04