International Trusteeship: External Authority in Areas of Limited Statehood

David A. Lake[*] Christopher J. Fariss[†]

[*]University of California San Diego, Dept. of Political Science, dlake@ucsd.edu
[†]Pennsylvania State University, Dept. of Political Science, cjf20@psu.edu

Supplementary Appendix

In this appendix, we provide additional details about the latent variable model and matching procedure that we use to examine the relationship between service provision and the presence of an international trusteeship. The latent variable model allows us to combine many indicators of service provision into one unidimensional dependent variable, which is available for the full time period of our study (1990-2010). The matching procedure allows us to test for the causal effect of a UN peace keeping mission on the service provision latent variable. The data and R code necessary to replicate the procedures reported here and in the main article are available at a Dataverse archive here: `http://hdl.handle.net/1902.1/22441`

Summary of the Matching Procedure

The matching procedure we use approximates the random assignment of units to a treatment or control group.[1] The procedure produces a group of country-year units in which the treatment variable is coded 1 and a control group is coded 0. The goal of this procedure is to produce two groups of country-year units that are equivalent in terms of a set of control variables. We use several matching algorithms to create the treatment and control groups.

To generate the treatment and control groups we use the nearest neighbor, optimal and CEM matching procedures, which are all run using the Matchit package developed by Ho et al. (2008) in R.[2] For the neighest neighbor algorithm a propensity score is calculated which is defined as the probability of receiving the treatment given a set of covariates. Next, the propensity score is used to select each control unit for comparison with each treated unit, one selection at a time. For each selection a control unit is selected that is not yet matched, but is closest to the treated unit on the distance measure (i.e., the propensity score). The optimal matching procedure is quite similar to the nearest neighbor matching procedure. The optimal procedure minimizes the average distance measure across all matched pairs, which reduces the influence of difficult to match units.[3]

We also use coarsened exact matching (CEM), which does not use a propensity score but instead

[1] See Ho et al. (2007); Imai, King and Stuart (2008); Rubin (1973, 1990, 2006).
[2] R Development Core Team (2009); see also Gu and Rosenbaum (1993); Hansen (2004).
[3] See Gu and Rosenbaum (1993); Imai, King and Lau (2007).

finds the treatment and control units within a multi-dimensional grid and weights the observations based on the number of both types of units within each grid cell. We use the matched data produced by the CEM algorithm; however, the matched data produced by the nearest neighbor matching and optimal matching algorithms produced similar results. The CEM algorithm is preferable in our case because we have a relatively small number of treatment units, approximately 30 from 1990-2010. We discuss the covariates that we include in the matching procedure in the main article.

In using a dichotomous measure of United Nations peacekeeping we violate the stable unit treatment value assumption (SUTVA) of treatment effects models. The assumption holds that there is only one version of the treatment (i.e., it does not vary in magnitude across units) and that no interference between units exist (i.e., the treatment does not spread through some mechanism of diffusion). We violate at least the first part of the SUTVA assumption given that UN peacekeeping missions are not all created equal. We attempt to address this issue by looking a several subsets of our treatment unit, believing these subsets to be less heterogeneous than the full set of treatment units. It is important to note that nearly all observational studies violate SUTVA with rare acknowledgement.

Finally, note that, after matching, we use a linear model to test for the relationship between the treatment variable and our measure of service provision. Imai, King and Stuart (2008) state that by matching, analysts can generate "doubly robust" estimates because the estimates will be consistent if at least the matching analysis or the model is correctly specified.[4]

Indicator Variables

The latent variable model rests on the assumption that the observed indicators for each country-year are each a function of the same underlying unidimensional latent variable. The model makes use of a combination of observable indicators of service provision, all of which are defined by Walter-Drop, and Wissel (in this volume) for the years they are available. The model uses this

[4]We preprocess the data using multiple imputation and matching before then estimating a simple linear model. We run the statistical models in R (R Development Core Team, 2009) using the Zelig library (Imai, King and Lau, 2007). Missing values are imputed using Amelia II (King et al., 2001).

information to estimate a single, united measure of service provision throughout the period 1990 to 2010. The indicator variables include (a) four public health indicators, including the proportion of maternal deaths during pregnancy, infant deaths, neonatal deaths, and deaths under 5 years of age, (b) four indicators of basic infrastructure, including the proportion of households with access to "improved" water resources, per capita electricity consumption, per capita kilometers of roads, and per capita kilometers of rail, and (c) five educational indicators, specifically the literacy rate, proportion of school age children that finish grade 5, proportion of enrolled 1st grade students, proportion of primary school students that enroll in secondary school, and the total enrollment of the school aged population. Each of these variables are taken from the United Nations data page (United Nations, 2012).

Lee Walter-Drop, and Wissel also include several variables that are designed to approximate the monopoly of force a state holds and the amount of security it is able to provide. We do not include these violence variables in our latent variable model because we instead use these indicators in the matching algorithm to make sure we have a similar set of treatment and control cases. By matching on these covariates, we compare county-year treatment and control groups that have experienced statistically similar levels of violence.

Lee Walter-Drop, and Wissel also consider several environmental indicators in their analysis. We decided ti exclude these variables from our latent variable as well because we wish to include only those variables that are likely to improve because of the provision of an international trustee into a state with limited sovereignty. Including the environmental variables would likely bias the results in favor of the main argument in the article.

Summary of the Latent Variable Model

We present a brief description of the latent variable model in the main article. Here we present the formal description of the model and a longer discussion.

The latent variable model used here is similar to those used to study the ideology of members of Congress, the ideology of Supreme Court Justices, the level of democracy and the level of respect for human rights (Clinton, Jackman and Rivers, 2004; Martin and Quinn, 2002; Treier and

Jackman, 2008; Schnakenberg and Fariss, Forthcoming). Though the latent variable approach is computationally complex, the resulting estimates are easy to interpret because we have assumed that the latent variable is a normally distributed random variable with mean 0 and standard deviation 1. We could assume another distributional form for the latent variable, practically however, the relative differences between any two units would be similar.

Overall, the latent variable model formalizes the relationship between the underlying construct and the various indicators we have included in the model. Interested researchers can use our model to include additional indicators of service provision and compare it to ours using a number of model comparison techniques, which are easy to implement.[5] For face validity, we include plots of the latent variable for 1990, 2000, and 2010 in Figures 1, 2 and 3. We also present pairwise correlation coefficients between the latent variable and the indicator variables in Table 3.

Table 1: Latent Variable Indicators

Public Health Indicators
proportion of maternal deaths during pregnancy
infant deaths
neonatal deaths
deaths under 5 years of age
Basic Infrastructure Indicators
proportion of households with access to “improved” water resources
per capita electricity consumption
per capita kilometers of roads
per capita kilometers of rail
Educational Indicators
literacy rate
proportion of school age children that finish grade 5
proportion of enrolled 1st grade students
proportion of primary school students that enroll in secondary school
total enrollment of the school aged population

To specify the latent variable model, let $i = 1, \ldots, N$, index cross-sectional units and $t = 1, \ldots T$, index time periods. In each period, observe $y_{i,t,j}$ for each of $j = 1, \ldots, J$ indicators for each country-year unit. In our application, each indicator is continuous, however the model can also

[5]See Gelman and Hill (2007) for a discussion and Schnakenberg and Fariss (Forthcoming) for an application of some of these techniques.

include categorical values with no added difficulty (Quinn, 2004). A mixture of different responses requires only to specify a different link function $F(.)$ for the various indicators in the likelihood. However, since all of our indicator variables are continuous we need only specify $F(.)$ as a Guassian link function with a unique error term τ_j for each of the indicator variables included. Note that τ_j is an estimate of model level uncertainty, which links the latent variable to an observed indicator. Whereas, σ is the uncertainty estimate of the latent variable itself. The model assumes that each of the indicator variables depend on a single latent variable $\theta_{i,t}$ that varies between countries and across time. The prior distributions are summarized in Table 2. The likelihood function is simply:

$$\mathcal{L}(\beta, \alpha, \tau, \theta | y) = \prod_{i=1}^{N} \prod_{t=1}^{T} \prod_{j=1}^{J} F(\alpha_j + \theta_{it}\beta_j, \tau_j) \tag{1}$$

Note that we estimate a dynamic latent variable so that the prior on the latent variable estimate is $\theta_{it} \sim N(\theta_{it-1}, \sigma)$ for all i and t except when $t = 1$. The first year a country enters the dataset $t = 1$ and the prior is $\theta_{i1} \sim N(0, 1)$. This method for incorporating dynamics was first implemented for binary judicial decision data by Martin and Quinn (2002) and then extended to ordinal responses by Schnakenberg and Fariss (Forthcoming).

Table 2: Prior Distribution for Latent Variable and Model Level Parameter Estimates

Parameter			
Country i latent variable estimate in 1990 (or first year in system)	$\theta_{i,t=1}$	$\sim$	$\mathcal{N}(0, 1)$
Country i latent variable estimate in all other years	$\theta_{i,t}$	$\sim$	$\mathcal{N}(\theta_{t-1}, \sigma)$
Latent variable uncertainty	σ	$\sim$	$U(0, 1)$
Model j intercept "difficulty parameter"	α_j	$\sim$	$\mathcal{N}(0, 1)$
Model j slope "discrimination parameter"	β_j	$\sim$	$\mathcal{G}(4, 3)$
Model j uncertainty	τ_j	$\sim$	$\mathcal{G}(0.001, 0.001))$

Table 3: Pairwise Correlation Coefficients Between Latent Variable Estimate and Individual Indicator Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	θ **(Latent Variable)**	1.00	0.85	0.99	0.95	0.99	0.81	0.75	0.35	0.37	0.81	0.74	0.70	0.74	0.81
2	mortality_maternal	0.85	1.00	0.83	0.79	0.84	0.70	0.85	0.24	0.27	0.64	0.69	0.45	0.58	0.64
3	mortality_infant	0.99	0.83	1.00	0.97	0.98	0.79	0.76	0.35	0.35	0.78	0.75	0.71	0.73	0.79
4	mortaility_neonatal	0.95	0.79	0.97	1.00	0.94	0.74	0.78	0.39	0.36	0.77	0.73	0.71	0.70	0.74
5	mortality_under5	0.99	0.84	0.98	0.94	1.00	0.78	0.75	0.33	0.33	0.83	0.72	0.70	0.74	0.81
6	water	0.81	0.70	0.79	0.74	0.78	1.00	0.54	0.33	0.37	0.63	0.70	0.47	0.58	0.62
7	electrictiy_access_pc	0.75	0.85	0.76	0.78	0.75	0.54	1.00	-0.03	-0.13	0.61	0.78	0.49	0.73	0.44
8	roads_km_pc	0.35	0.24	0.35	0.39	0.33	0.33	-0.03	1.00	0.73	0.25	0.31	0.22	0.28	0.25
9	rail_km_pc	0.37	0.27	0.35	0.36	0.33	0.37	-0.13	0.73	1.00	0.41	0.26	0.05	0.33	0.23
10	literacy_rate	0.81	0.64	0.78	0.77	0.83	0.63	0.61	0.25	0.41	1.00	0.60	0.58	0.65	0.84
11	persistence_to_grade5	0.74	0.69	0.75	0.73	0.72	0.70	0.78	0.31	0.26	0.60	1.00	0.52	0.62	0.61
12	net_intake_rate_grade1	0.70	0.45	0.71	0.71	0.70	0.47	0.49	0.22	0.05	0.58	0.52	1.00	0.53	0.78
13	progression_to_secondary	0.74	0.58	0.73	0.70	0.74	0.58	0.73	0.28	0.33	0.65	0.62	0.53	1.00	0.63
14	total_primary_enrollment_rate	0.81	0.64	0.79	0.74	0.81	0.62	0.44	0.25	0.23	0.84	0.61	0.78	0.63	1.00

Table 4: Summary Statistics for Individual Indicator Variables

	mean	sd
θ **(Latent Variable)**	-0.068	0.537
mortality_maternal	-0.500	0.001
mortality_infant	-0.510	0.009
mortaility_neonatal	-0.505	0.004
mortality_under5	-0.514	0.014
water	0.680	0.051
electrictiy_access_pc	0.624	0.337
roads_km_pc	0.007	0.009
rail_km_pc	0.000	0.000
literacy_rate	0.705	0.039
persistence_to_grade5	0.689	0.042
net_intake_rate_grade1	0.659	0.050
progression_to_secondary	0.695	0.045
total_primary_enrollment_rate	0.705	0.037

Time Horizon of 4 Years

Table 5: Estimated Effects of Trusteeship on service provision, from 3 Models

Nearest Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.393	0.150	2.625
2	**Trustee Treatment**	-0.008	0.020	-0.378
3	Homicide Rate	-0.000	0.001	-0.202
4	One Sided Killing	-0.055	0.023	-2.353
5	Battle Death	-0.002	0.033	-0.061
6	Democracy	0.004	0.021	0.173
7	Population	-0.003	0.009	-0.390
8	GDP per capita	-0.034	0.013	-2.537

Optimal Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.389	0.148	2.637
2	**Trustee Treatment**	-0.001	0.020	-0.053
3	Homicide Rate	0.000	0.001	0.165
4	One Sided Killing	-0.045	0.029	-1.519
5	Battle Death	0.010	0.037	0.272
6	Democracy	0.009	0.022	0.412
7	Population	-0.002	0.009	-0.240
8	GDP per capita	-0.038	0.015	-2.598

CEM Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.269	0.211	1.275
2	**Trustee Treatment**	-0.011	0.017	-0.636
3	Homicide Rate	-0.000	0.001	-0.135
4	One Sided Killing	-0.093	0.042	-2.208
5	Battle Death	0.022	0.039	0.567
6	Democracy	-0.014	0.037	-0.378
7	Population	0.003	0.015	0.200
8	GDP per capita	-0.027	0.018	-1.547

Time Horizon of 3 Years

Table 6: Estimated Effects of Trusteeship on service provision, from 3 Models

Nearest Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.329	0.120	2.736
2	**Trustee Treatment**	-0.005	0.016	-0.305
3	Homicide Rate	0.000	0.000	0.036
4	One Sided Killing	-0.025	0.022	-1.129
5	Battle Death	-0.004	0.024	-0.176
6	Democracy	0.017	0.024	0.714
7	Population	-0.005	0.006	-0.836
8	GDP per capita	-0.027	0.012	-2.242

Optimal Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.244	0.126	1.939
2	**Trustee Treatment**	0.001	0.016	0.065
3	Homicide Rate	0.000	0.001	0.599
4	One Sided Killing	-0.031	0.018	-1.696
5	Battle Death	0.004	0.025	0.171
6	Democracy	-0.001	0.020	-0.054
7	Population	-0.002	0.007	-0.298
8	GDP per capita	-0.023	0.013	-1.764

CEM Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.231	0.177	1.306
2	**Trustee Treatment**	-0.005	0.024	-0.227
3	Homicide Rate	-0.000	0.001	-0.462
4	One Sided Killing	-0.064	0.026	-2.404
5	Battle Death	0.011	0.020	0.557
6	Democracy	-0.002	0.019	-0.106
7	Population	0.004	0.009	0.488
8	GDP per capita	-0.026	0.015	-1.717

Time Horizon of 2 Years

Table 7: Estimated Effects of Trusteeship on service provision, from 3 Models

Nearest Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.183	0.071	2.593
2	**Trustee Treatment**	0.004	0.012	0.294
3	Homicide Rate	-0.000	0.000	-0.298
4	One Sided Killing	-0.017	0.017	-1.047
5	Battle Death	-0.001	0.016	-0.083
6	Democracy	0.015	0.015	0.974
7	Population	-0.002	0.004	-0.528
8	GDP per capita	-0.015	0.007	-2.104

Optimal Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.130	0.092	1.422
2	**Trustee Treatment**	0.001	0.013	0.047
3	Homicide Rate	0.000	0.000	0.329
4	One Sided Killing	-0.016	0.013	-1.279
5	Battle Death	0.001	0.015	0.069
6	Democracy	0.008	0.014	0.538
7	Population	0.001	0.004	0.288
8	GDP per capita	-0.013	0.010	-1.318

CEM Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.075	0.118	0.642
2	**Trustee Treatment**	-0.006	0.010	-0.630
3	Homicide Rate	-0.000	0.000	-0.096
4	One Sided Killing	-0.044	0.022	-2.014
5	Battle Death	0.013	0.016	0.799
6	Democracy	-0.002	0.023	-0.102
7	Population	0.003	0.006	0.498
8	GDP per capita	-0.008	0.010	-0.767

Time Horizon of 1 Years

Table 8: Estimated Effects of Trusteeship on service provision, from 3 Models

Nearest Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.097	0.047	2.087
2	**Trustee Treatment**	0.001	0.005	0.160
3	Homicide Rate	0.000	0.000	0.331
4	One Sided Killing	-0.009	0.010	-0.927
5	Battle Death	-0.001	0.008	-0.111
6	Democracy	0.007	0.006	1.025
7	Population	-0.001	0.003	-0.345
8	GDP per capita	-0.009	0.004	-2.157

Optimal Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.071	0.040	1.772
2	**Trustee Treatment**	0.001	0.006	0.241
3	Homicide Rate	-0.000	0.000	-0.255
4	One Sided Killing	-0.007	0.007	-0.965
5	Battle Death	-0.001	0.011	-0.137
6	Democracy	0.008	0.007	1.105
7	Population	0.001	0.002	0.439
8	GDP per capita	-0.007	0.005	-1.615

CEM Neighbor Matching Algorithm

	Variables	Coefficients	SE	t
1	Intercept	0.020	0.080	0.248
2	**Trustee Treatment**	-0.001	0.006	-0.128
3	Homicide Rate	0.000	0.000	0.230
4	One Sided Killing	-0.021	0.010	-2.174
5	Battle Death	0.009	0.015	0.604
6	Democracy	0.001	0.019	0.076
7	Population	0.002	0.005	0.420
8	GDP per capita	-0.003	0.006	-0.468

Figure 1: Country Rank by service provision Latent Variable 1990

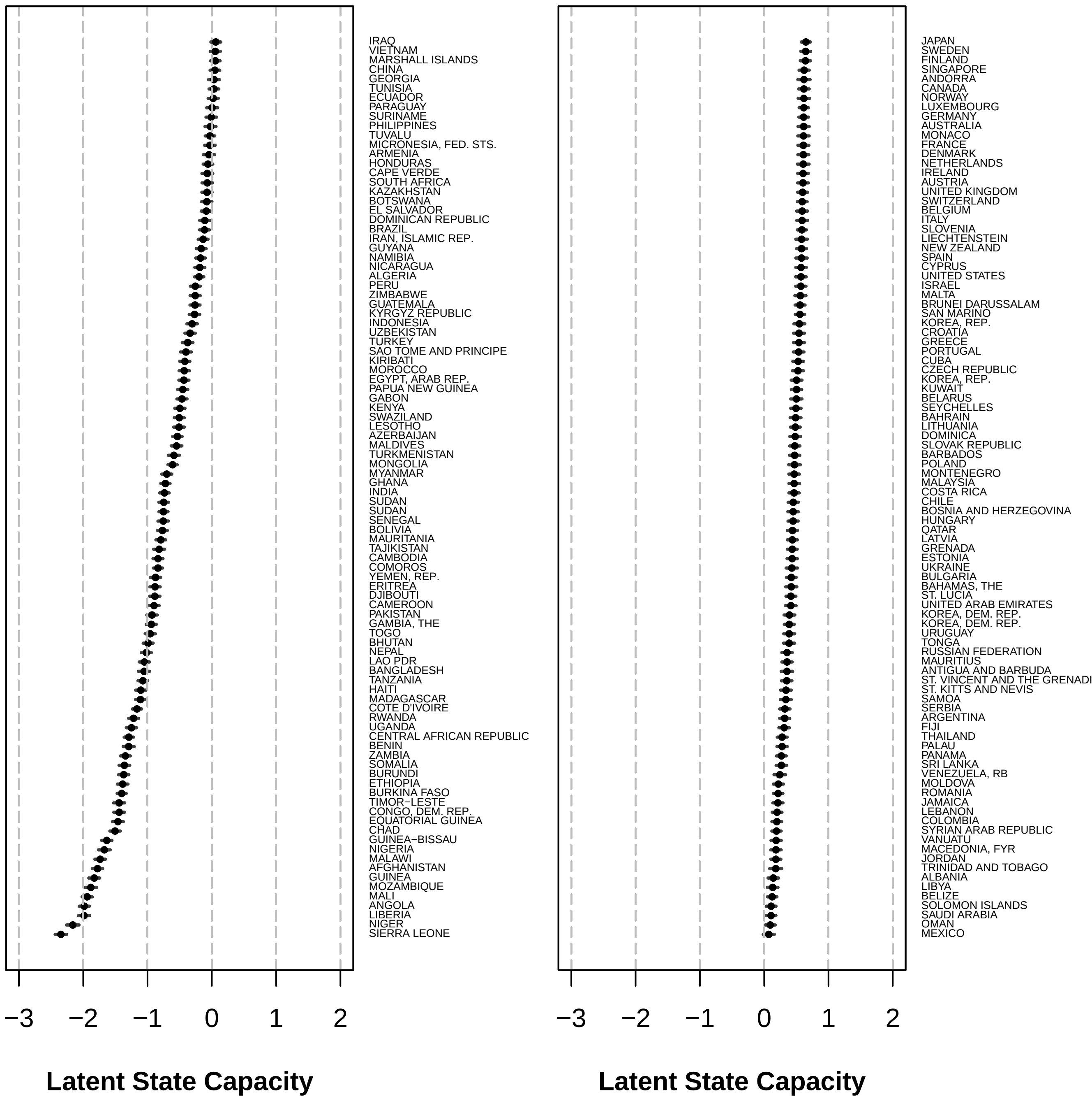


Figure 2: Country Rank by service provision Latent Variable 2000

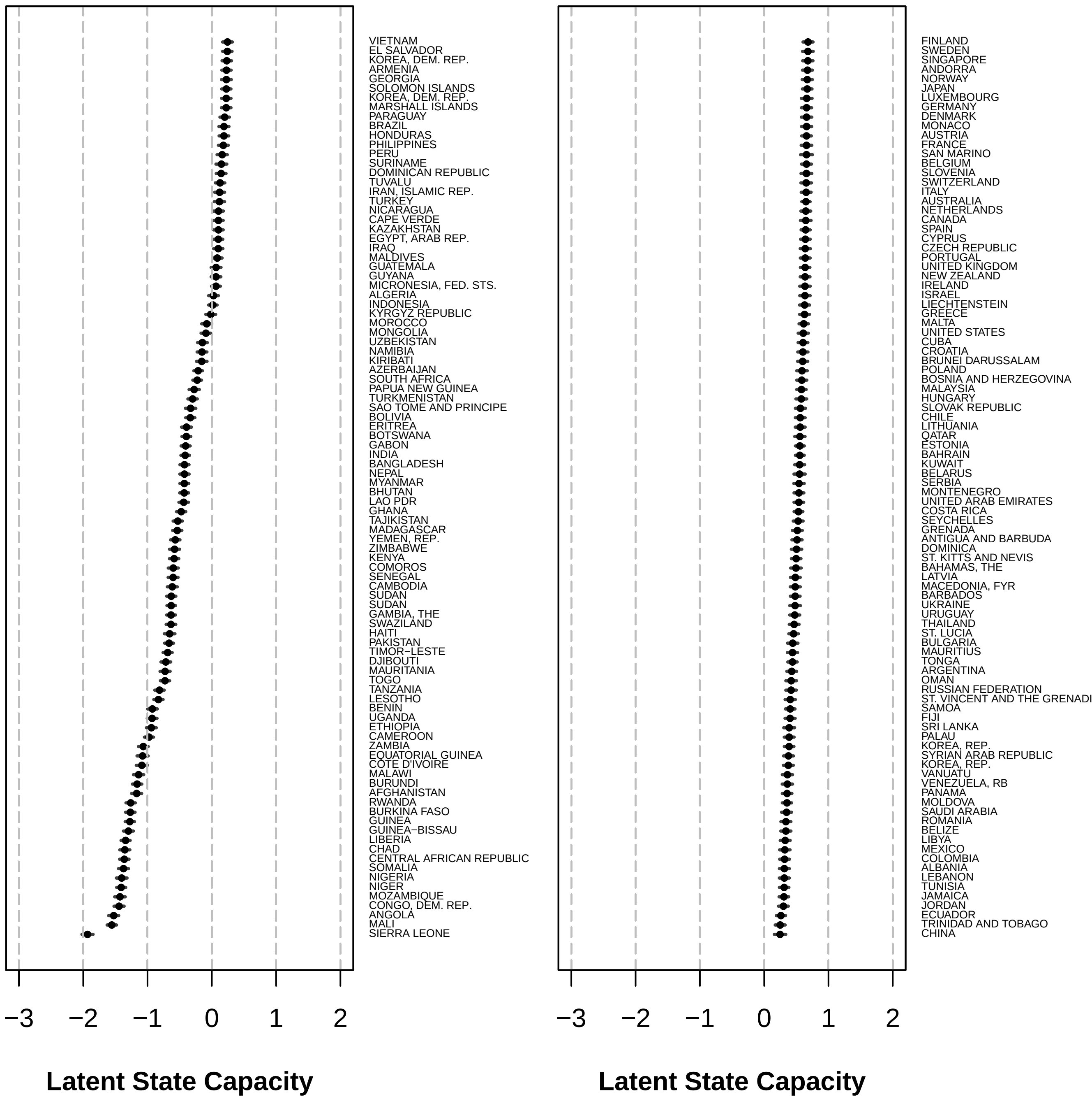
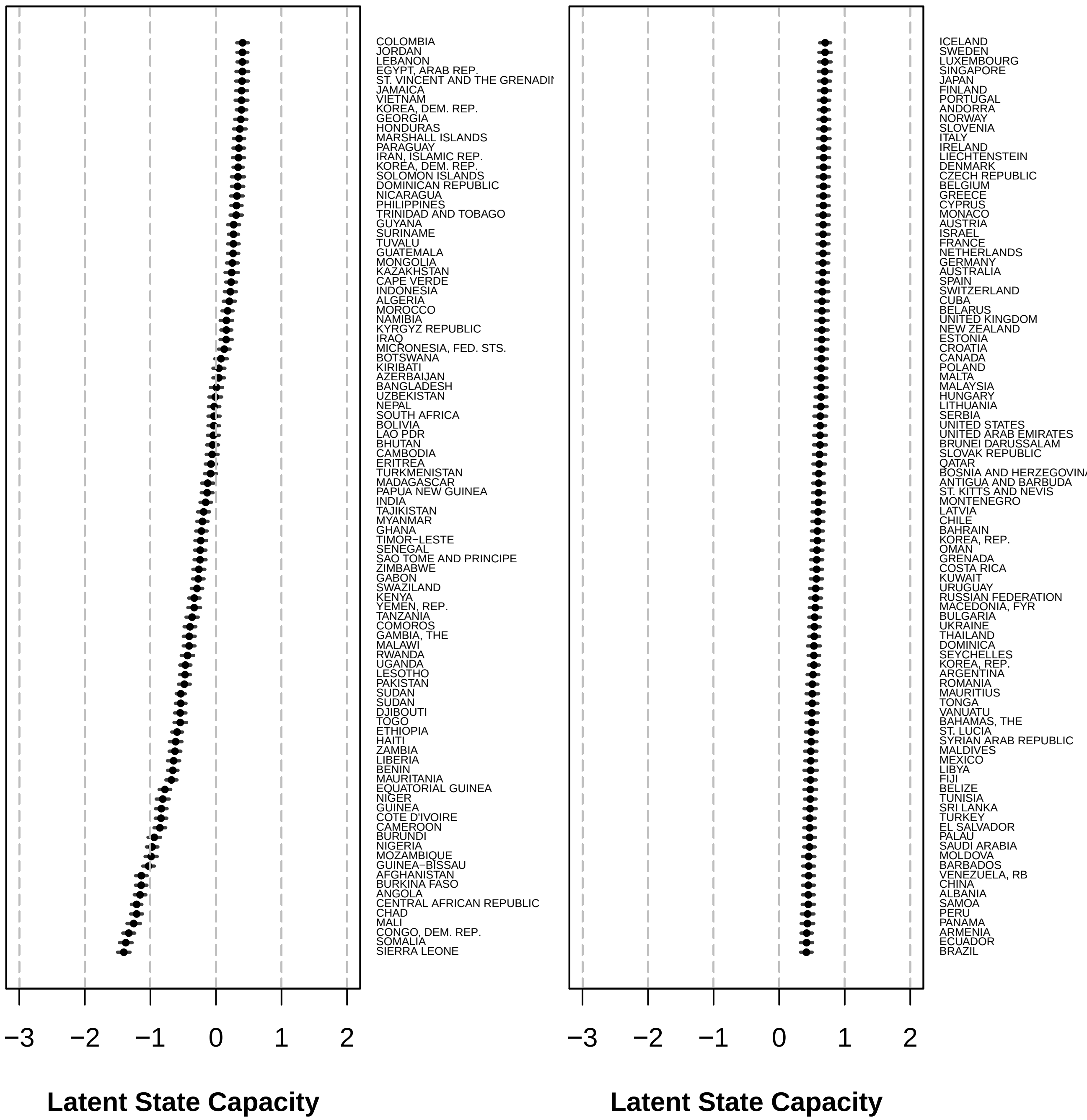


Figure 3: Country Rank by service provision Latent Variable 2010



Permutation Test

As reported in the article, we also conducted a permutation test. Recall that in the main analysis we used a matching algorithm to create a group of control units that were as statistically similar to the treatment units as possible. For the permutation test reported here, we wanted to determine if their was a treatment effect between the treatment group and some other random combination of control units not selected by the matching algorithm. To check for this possibility we ran 10,000 regressions using the same 44 treatment observations in each regression. The control units were sampled from all possible observations that did not receive the treatment. We sampled $2 * n$ control observations, where n is the number of treatment units. Only a tiny fraction of these random control groups yielded a positive t-statistic for the treatment effect at conventional levels of significance (0.0078 of the samples produced a p-value of 0.05 or smaller). Figure 4 displays the distribution of all 10,000 t-statstics generated for the treatment variable in these regressions. The results were also consistent with random samples of control observations of 44, 122, and 176 observations respectively (n, $3 * n$, and $4 * n$). These tests provide additional evidence for the lack of an effect reported in the article when comparing the treatment units with the matched sample of control units.

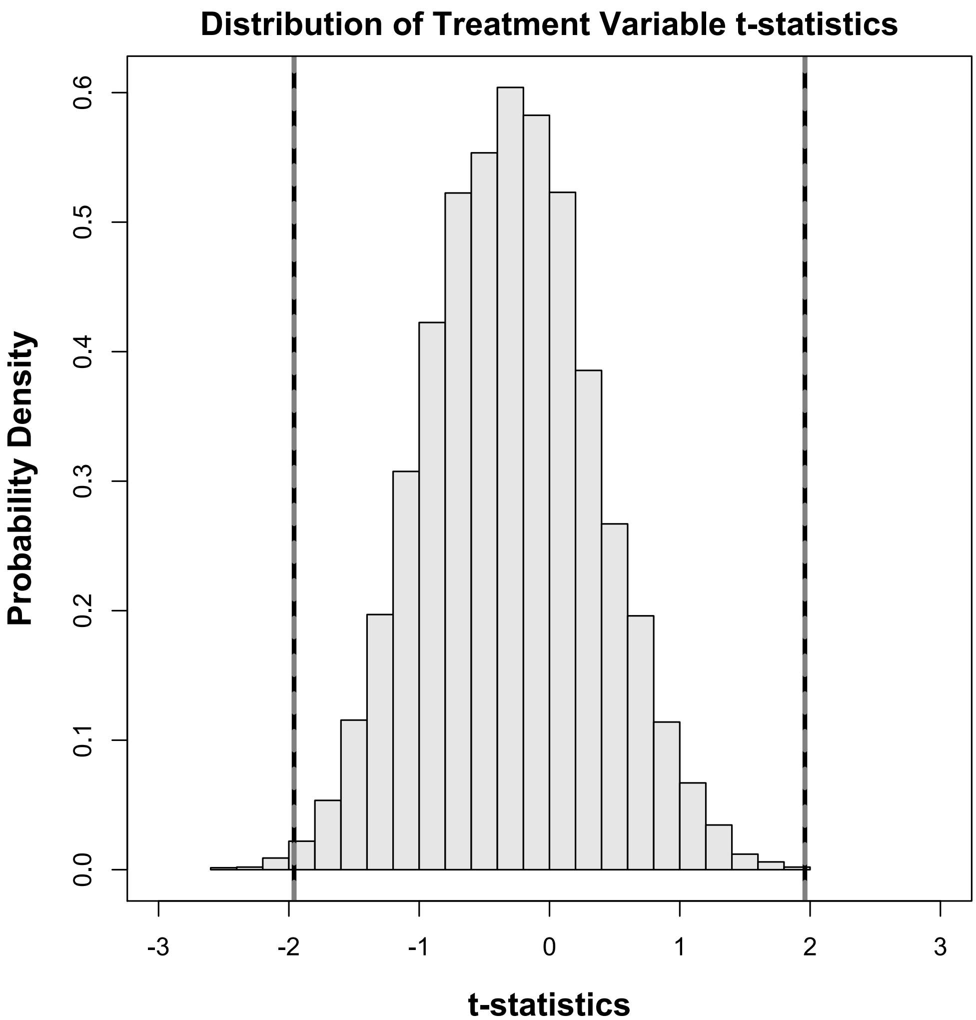


Figure 4: Distribution of all 10,000 t-statistics generated for the treatment variable compared to $2 * n$ randomly sampled control units. $n = 44$.

JAGS Code for Latent Variable Model

```
model{
 for(i in 1:n){# n is the number of records
   for(j in 1:J){ # j is the total number of observable indicators
     p[i,j] <- alpha[j] + beta[j]*x[i]
     y[i,j] ~ dnorm(p[i,j], tau[j])
 }

# redraw latent variable parameter from mu matrix because of unbalanced panels
   x[i] <- mu[country[i], year[i]]
 }

# draw percision for latent variable parameter estimate
 sigma ~ dunif(0,1)
 kappa <- pow(sigma, -1)

# draw dynamic latent variable parameter
 for(c in 1:n.country){
  mu[c, 1] ~ dnorm(0, 1)
  for(t in 2:n.year){ #n.year is number of years
   mu[c, t] ~ dnorm(mu[c, t-1], kappa)
  }
 }

# prior distribution for model level parameters
 for(j in 1:J){
  beta[j] ~ dnorm(0, .25)
  alpha[j] ~ dnorm(0, .25)
  tau[j] ~ dgamma(0.001, 0.001)
 }
}
```

References

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.

Gelman, Andrew and Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge, MA: Cambridge University Press.

Gu, X. and P. R Rosenbaum. 1993. "Comparison of multivariate matching methods: structures, distances, and algorithms." *Journal of Computational and Graphical Statistics* 2(405–420).

Hansen, B. B. 2004. "Full Matching in an Observational Study of Coaching for the SAT." *Journal of the American Statistical Association* 99:609–618.

Ho, Daniel E., Kosuke Imai, Gary King and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3):199–236.

Ho, Daniel, Kosuke Imai, Gary King and Elizabeth Stuart. 2008. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." *Journal of Statistical Software* 171:481–502.

Imai, Kosuke, Gary King and Elizabeth A. Stuart. 2008. "Misunderstandings between experimentalists and observationalists about causal inference." *Journal of the Royal Statistical Society, Series A* 171(Part 2):481–502.

Imai, Kosuke, Gary King and Olivia Lau. 2007. "Zelig: Everyone's Statistical Software.".
URL: *http://GKing.Harvard.Edu/zelig*

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

Martin, Andrew D. and Keven M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.

Quinn, Keven M. 2004. "Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses." *Political Analysis* 12(4):338–353.

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
URL: *http://www.R-project.org*

Rubin, Donald B. 1973. "Matching to remove bias in observational studies." *Biometrics* 29(1):159–184.

Rubin, Donald B. 1990. "Formal Modes of Statistical Inference For Causal Effects." *Journal of Statistical Planning and Inference* 25:279–292.

Rubin, Donald B. 2006. *Matched Sampling for Causal Effects*. Cambridge, MA: Cambridge University Press.

Schnakenberg, Keith E. and Christopher J. Fariss. Forthcoming. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* .

Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–217.

United Nations. 2012. *UN Data A World of Information.*
URL: *http://data.un.org/*