What Explains Union Density? A Replication with Updated Bayesian Approaches

POSC 3410 – Quantitative Methods in Political Science

Steven V. Miller

Department of Political Science

Goal(s) for Today

- *Show students implementation of Bayesian methods.*
- *Introduce students to replicating an article they actually (hopefully) read.*

What Bayesian Want and Why

What Bayesian want:

- To know about the population parameter, given the data.
- Extend Bayes' theorem as a means to answering that question.
- To embrace/state outright the implications of subjective probability.

Why Bayesians want this:

- Frequentist inference gives a backdoor answer to the motivating question.
- Data-generating processes assumed by central limit theorem may not hold in real world.

The Benefits and Limitations

Benefits:

- Actually answers the question that interests us in inferential statistics.
- Declares/models outright prior beliefs/distributions.
- Uncertainty distributions come free in model output.
- Great for debugging/diagnosing/correcting problems in models.

Drawbacks:

- Computationally demanding
- Potential for "deck-stacking" (which is more of a strawman critique)

What We'll Do Today

1. Walk through Western and Jackman (1994) as one excellent introduction to Bayesian inference.
2. Update/replicate their findings with newer Bayesian methods (via Stan/`brms`)

See my blog for more detail:
`http://svmiller.com/blog/2019/08/what-explains-union-density-brms-replication/`

R Packages We'll Be Using Today

```r
library(tidyverse) # for most things
library(stevemisc) # for formatting and r2sd()
library(stevedata) # for data(uniondensity)
library(ggrepel) # for repelling labels
library(kableExtra) # for prettying up tables

# source("1-replicate-westernjackman1994bicr.R")
# ^ requires brms, tidybayes
con <- DBI::dbConnect(RSQLite::SQLite(), "westernjackman1994bicr.db")
# ^ Contains all model summaries and draws.
```

Western and Jackman (1994)

Nonstochastic and Weak Data

Two properties of comparative research violate foundations for frequentist inference.

1. Nonstochastic data (i.e. non-random DGP)
2. Weak data

Nonstochastic Data

Frequentist inference assumes data are generated by a repeated mechanism like a coin flip (hence: RDGP).

- A sample statistic is just one possible result from a draw of a probability distribution of the population.

Nonstochastic Data

However, political scientists can define the sample on the population. Examples:

- OECD countries
- Militarized interstate disputes
- Supreme Court decisions

You know what this is. We called this a **census**.

Nonstochastic Data

Frequentist inference is inapplicable to the nonstochastic setting.

- If we took another random draw, we'd get the exact same data.
- "Updating" the data doesn't generate a new random sample.
- Appeals to a "superpopulation" don't help either.

Weak Data

This takes on two forms in political science research.

1. Small n
2. Collinearity

If the population of interest is “advanced industrial societies”, our n is limited to about two dozen observations.

- We run out of degrees of freedom quickly when adding controls.

Weak Data

The issue of **multicollinearity** also arises in weak data with small n.

- This is when two predictors are so highly correlated that their estimated partial effects are uninformative.

Weak Data

This is relevant to a debate Western and Jackman address: what accounts for the percentage of the work force that is unionized?

- Wallerstein: size of civilian labor force (-).
- Stephens: industrial concentration (+).

Both agree that left-wing governments (see: Wilensky's (1981) index) matter as a control variable, but disagree about these two variables.

Table 1: The Data at the Heart of this Academic Dispute

Country	Union Density	Left Government	Logged Labor Force Size	Industrial Concentration
Sweden	82.4	111.84	8.28	1.55
Israel	80.0	73.17	6.90	1.71
Iceland	74.3	17.25	4.39	2.06
Finland	73.3	59.33	7.62	1.56
Belgium	71.9	43.25	8.12	1.52
Denmark	69.8	90.24	7.71	1.52
Ireland	68.1	0.00	6.79	1.75
Austria	65.6	48.67	7.81	1.53
NZ	59.4	60.00	6.96	1.64
Norway	58.9	83.08	7.41	1.58
Australia	51.4	33.74	8.60	1.37
Italy	50.6	0.00	9.67	0.86
UK	48.0	43.67	10.16	1.13
Germany	39.6	35.33	10.04	0.92
Netherlands	37.7	31.50	8.41	1.25
Switzerland	35.4	11.87	7.81	1.68
Canada	31.2	0.00	9.26	1.35
Japan	31.0	1.92	10.59	1.11
France	28.2	8.67	9.84	0.95
USA	24.5	0.00	11.44	1.00

The Problem of Weak Data

Problem: both are highly collinear (r = -.92).

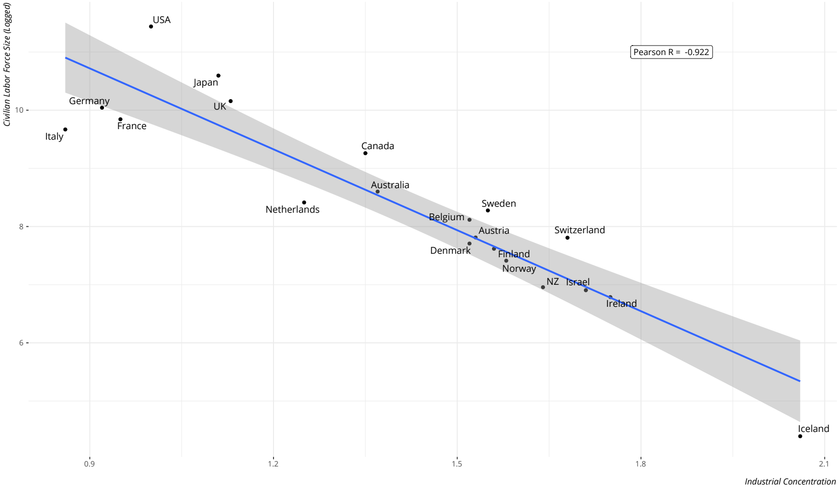
- In normal regression, one has to be dropped for a better model fit.
- This gets us no closer to settling an academic debate, though.

A Bayesian regression will have no problem with this.

- It's great for weak data situations.
- The prior distribution will exert more weight on the posterior distributions.

The Collinearity Between Industrial Concentration and Logged Labor Force Size

The correlation is -.922, which is almost a perfect negative correlation.



Modeling Prior Information

TABLE 1

Prior Means (Standard Deviations) and Their Substantive Interpretations for Bayesian Regression Analysis of Union Density

VARIABLE	PRIOR	SUBSTANTIVE INTERPRETATION
Wallerstein's priors		
Left government	.3 (.15)	One year of left-wing government increases union density by about 1 percentage point. A year of left-wing government may increase union density by as much as 2 percentage points of union density, but its effect is almost certainly not negative.
Logged labor-force size	−5 (2.5)	Doubling the size of the labor force would reduce union density about $\ln(2) \times 5 \approx 3.5$ percentage points. This increase in labor-force size may generate a union decline as big as 7 percentage points, but a growing labor force is unlikely to increase union density.
Economic concentration	0 (10^6)	The diffuse prior indicates that the researcher has no strong prior beliefs about the sign or magnitude of an effect. When the explanatory variables are uncorrelated, the diffuse prior yields posteriors that are approximately given by the sample data.
Stephens's priors		
Left government	.3 (.15)	Like Wallerstein's prior, one year of left-wing government increases union density by 1 percentage point.
Logged labor-force size	0 (10^6)	Diffuse prior.
Economic concentration	10 (5)	If economic concentration were to increase by 100% in relation to the United States, union density would increase by 10 percentage points. This increase in the concentration ratio may generate a density increase as large as 20 percentage points, but any increase in concentration is unlikely to decrease union density.

Note: Left government is measured by Wilensky's (1981) cumulative index of left-wing government; logged labor-force size is the natural log of the size (in thousands) of the dependent labor force in the year that union density is measured; and economic concentration is measured by the four-firm concentration ratio, in proportion to the United States.

Using Prior Information

Notice what's happening with our prior information.

- Wallerstein and Stephens agree on the effect of left governments.
- They disagree on the two other variables.

Using Uninformative Priors

Let's start with using uninformative priors.

- `M1`: A simple linear model (ostensibly used by Western and Jackman)
- `B0`: A Bayesian linear model with flat/undeclared priors

We'll compare the results with what Western and Jackman (1994) report in their paper (Table 2).

R Code

```
M1 <- lm(union ~ left + size + concen, data=uniondensity)

B0 <- brm(union ~ left + size + concen,
          data=uniondensity,
          seed = 8675309,
          chains = 4, cores = 4,
          family="gaussian")

tribble(
  ~term, ~estimate, ~`std.error`, ~lwr, ~upr,
  "Intercept", 97.59, 57.48, 3.04, 192.14,
  "Left Government", .27, .08, .15, .39,
  "Labor Force Size (logged)", -6.46, 3.79, -12.70, -.22,
  "Industrial Concentration", .35, 19.25,  -31.32, 32.02
) -> WJM1
```

Table 2: Comparing OLS, an Uninformative Bayesian Model, and Table 2 of Western and Jackman (1994)

Model	Parameter	Coef.	SD|SE	Lower Bound	Upper Bound
Standard OLS	Industrial Concentration	0.35	19.25	-31.32	32.02
Bayesian LM	Industrial Concentration	0.89	20.21	-31.74	33.60
Western and Jackman (Table 2)	Industrial Concentration	0.35	19.25	-31.32	32.02
Standard OLS	Intercept	97.59	57.48	3.04	192.15
Bayesian LM	Intercept	95.84	60.28	-3.46	193.11
Western and Jackman (Table 2)	Intercept	97.59	57.48	3.04	192.14
Standard OLS	Labor Force Size (logged)	-6.46	3.79	-12.70	-0.22
Bayesian LM	Labor Force Size (logged)	-6.34	3.99	-12.75	0.16
Western and Jackman (Table 2)	Labor Force Size (logged)	-6.46	3.79	-12.70	-0.22
Standard OLS	Left Government	0.27	0.08	0.15	0.39
Bayesian LM	Left Government	0.27	0.08	0.13	0.41
Western and Jackman (Table 2)	Left Government	0.27	0.08	0.15	0.39

Using Uninformative Priors

We see that the effects of left governments and logged labor force size are significant.

- Prima facie, Wallerstein seems to be correct (though the Stan estimates are bit more diffuse).
- The industrial concentration variable is insignificant in all three models.

The results do seem to suggest that perhaps what Western and Jackman call "Uninformative Priors" is really just OLS.

Using Informative Priors

In the interest of brevity, let's focus on analyses that comprise Table 3.

- We are looking at the regression results using both sets of prior information.

R Code

```
# Wallerstein's priors
# left: 3(1.5)
# size: -5(2.5) // This is what he's arguing
# concen: 0(10^6) // diffuse/"ignorance" prior
# Intercept: 0(10^6) // diffuse/"ignorance" prior
wall_priors <- c(set_prior("normal(3,1.5)", class = "b", coef= "left"),
                 set_prior("normal(-5,2.5)", class = "b", coef="size"),
                 set_prior("normal(0,10^6)", class="b", coef="concen"),
                 set_prior("normal(0,10^6)", class="Intercept"))

# Stephens priors
# left: 3(1.5) // they both agree about left governments
# size: 0(10^6) // diffuse/"ignorance" prior
# concen: 10(5) // This is what Stephens thinks it is.
# Intercept: 0(10^6) // diffuse/"ignorance" prior
stephens_priors <- c(set_prior("normal(3,1.5)", class = "b", coef= "left"),
                     set_prior("normal(0,10^6)", class = "b", coef="size"),
                     set_prior("normal(10,5)", class="b", coef="concen"),
                     set_prior("normal(0,10^6)", class="Intercept"))
```

R Code

```r
# Wallerstein's priors
B1 <- brm(union ~ left + size + concen,
          data = uniondensity,
          prior=wall_priors,
          seed = 8675309,
          chains = 4, cores = 4,
          family="gaussian")

B1 %>% gather_draws(b_Intercept, sigma, b_left, b_concen, b_size) -> tidyB1


# Stephens' priors
B2 <- brm(union ~ left + size + concen,
          data = uniondensity,
          prior=stephens_priors,
          seed = 8675309,
          chains = 4, cores = 4,
          family="gaussian")


B2 %>% gather_draws(b_Intercept, sigma, b_left, b_concen, b_size) -> tidyB2
```

Table 3: A Reproduction of Table 3 from Western and Jackman (1994)

Prior	**Parameter**	**Coef.**	**SD**	**Lower Bound**	**Upper Bound**
Wallerstein's Priors	Industrial Concentration	4.56	12.68	-15.95	25.91
Wallerstein's Priors	Intercept	82.45	32.87	28.01	135.81
Wallerstein's Priors	Left Government	0.28	0.08	0.15	0.41
Wallerstein's Priors	Labor Force Size (logged)	-5.40	2.08	-8.77	-1.94
Stephen's Priors	Industrial Concentration	9.41	4.79	1.58	17.20
Stephen's Priors	Intercept	70.51	20.50	37.51	104.17
Stephen's Priors	Left Government	0.27	0.08	0.14	0.40
Stephen's Priors	Labor Force Size (logged)	-4.76	1.86	-7.85	-1.73

Interpreting Table 3

Using Wallerstein's priors:

- Posterior estimates for left-wing governments remain precise.
  - Actually gain a little precision too.
- Prior information makes confidence interval for labor-force size much less diffuse.
- No effect of industrial concentration.

Interpreting Table 3

Using Stephens' priors:

- Same posterior estimates for left-wing governments.
- Labor-force size estimate still significant, though magnitude decreases.
- Significant effect of industrial concentration.
    - But notice: we had prior beliefs about that effect!

The data we ultimately observed don't discount the effect of industrial concentration if you build in the prior belief.

Conclusion

Bayesians highlight how many liberties we can take with our research design if we're not careful.

- A census (a non-random DGP) does not permit conventional statistical inference.
- Collinearity magnifies problems of weak data.

Importantly, why start agnostic of the population parameter if we do not have to do this?

- If you have prior information or plausible effects, use it.

Table of Contents