

Stability and Change in the Freedom House Political Rights and Civil Liberties Measures

David A. Armstrong II
University of Wisconsin - Milwaukee
Department of Politics Science
PO Box 413
Milwaukee, WI 53201-0413

e: armstrod@uwm.edu
t: 414-229-4239

February 11, 2011

Word Count: 5645

Abstract

The Freedom House data on political rights and civil liberties are some of the most used by researchers and policy organizations as measures of democracy and human rights. Despite their popularity, little is known about the measurement properties of these variables. In 2006, Freedom House began releasing the sub-category indicators used to generate the seven-point political rights and civil liberties scales as well as the freedom status indicator. I investigate the measurement properties of these scales using Bayesian measurement models (factor analysis and latent class analysis) that explicitly incorporate the goals of the Freedom House organization (which I argue are, rigor, appropriate precision, and temporal stability). I find that a) there is considerable variation hiding in the seven point political rights and civil liberties classifications (e.g., many countries that are coded one by freedom house are interestingly different from each other), b) some countries coded in different categories by Freedom House are not interestingly different from each other (e.g., some twos are not different from some threes), and c) the sub-indicators are all reliable indicators of political rights and civil liberties. I show that these differences can be substantively and statistically consequential when using the Freedom House data if the researcher's goals include comparison across space, comparison over time or predictive statistical modeling.

Introduction

Scholars have long attempted to make distinctions among states based on regime type. Two of the most commonly used measures by both scholars and policymakers are the Freedom House *Freedom in the World* political rights and civil liberties measures. While there has been little empirical investigation of the measurement properties of these variables, the organization has put considerable thought into the what it wants to accomplish with the data. Freedom House has tried to ensure the data are appropriately precise (only indicating differences among regimes that are substantially different in the ‘real world’), temporally stable, and in recent years, as rigorous and transparent as possible.¹ While these are reasonable goals, I argue the current measures can be improved on all counts. I propose a new measurement strategy using existing Freedom House data that explicitly meets all of these goals while providing a considerable amount of additional information regarding the regime classifications.

To do this, I use the latest in measurement model technology classify to regimes. I use a dynamic Bayesian factor analysis to produce measures of political rights and civil liberties using the seven sub-indicators. I find that a) there is considerable variation in regimes *within* the same original Freedom House categories (e.g., some sevens are different from other sevens) and b) that a number of countries that are classified as different regime-types are not significantly different from each other (e.g., some twos are not different from some threes).²

Below, I briefly discuss the methods currently used to generate the political

¹I provide evidence of the prominence of these goals in the on-line appendix to this article.

²While space does not permit such an investigation in the print version of the article, in the on-line appendix to the article I use a Bayesian latent class analysis to classify regimes on the three-fold freedom status scale (‘Free’, ‘Partly Free’, ‘Not Free’) that is also a product of the Freedom House measurement exercise. There, I find that there is more movement between the categories over time than Freedom House’s classification would suggest. When the two classifications differ significantly, it is always in the direction of Freedom House indicating a country is more free than the latent class analysis would suggest.

rights, civil liberties and freedom status indicators. I provide an example of the potential consequences of this measurement strategy for predictive statistical models. After the methodological investigation, I conclude with some potential suggestions for increasing the rigor and transparency in Freedom House’s regime classification.

Political Rights and Civil Liberties

The political rights (PR) and civil liberties (CL) indicators proposed by Freedom House are seven-fold classification schemes. They start by asking a set of experts ten questions about PR and fifteen questions about CL.³ These questions are grouped into three sub-categories for PR and four sub-categories for CL, as in Table I.⁴

[Table 1 about here.]

The *Freedom in the World* report acknowledges that these might be closely related ideas, but they remain distinct concepts and should be treated as such (i.e., they should be considered as two ‘latent variables’). Respondents score the countries on a scale from zero (low) to four (high) on each question. The scores are then summed and an overall sub-category score is given. The sub-category scores range from zero to $4k$ where k is the number of questions in the subsection. For example, the ‘electoral process’ subcategory has three questions. Thus, the summed rankings on each of those variables ranges from zero to 12 (4×3). There are three questions asked in the areas of electoral process, functioning of government and associational rights and there are four questions asked in the remaining four areas. These seven sub-category scores are the primary data used in the measurement model discussed below.

³The number of checklist questions has changed over time, but is constant over the years under investigation here.

⁴A more thorough description can be found at http://www.freedomhouse.org/template.cfm?page=351&ana_page=341&year=2008

To generate the seven-point scales, Freedom House sums the sub-indicators creating a scales for PR (ranging from zero to 40) and CL (ranging from zero to 60). Freedom House then splits each scale into seven intervals of equal width.⁵ This scoring mechanism is meant to smooth over ‘insignificant’ changes in rights and liberties, reducing the unwarranted precision in the summed scores. The implication is that differences in category membership over time, or across countries signal interesting differences in the level of rights and liberties in those countries. The flaw in this logic is apparent. A country with a summed score of 30 on political rights is coded two. With an increases in the country’s summed score in the next year to 34, (10% of the scale) it remains in category two. However, if that country’s summed score decreases by a single point, to 29, in the next year (2.5% of the scale’s range) it will change from category two to category three. Thus, big changes in the summed score can result in no change in the coded category and small scale changes can result in changes in the coded category. In spirit, the idea makes perfect sense - highlight ‘real’ differences and downplay those due to chance. Unfortunately, this broad categorization seems ill suited to the task.⁶

Bayesian Model for Political Rights and Civil Liberties

There has been a long tradition of latent variable modeling of regime type (See for example Bollen & Paxton 2000). Much of the recent research on modeling democracy is Bayesian in nature due to its attractive features relative to its

⁵While summing is a reasonable idea (see Jacoby 1991), the categorization of summed scores is not well-supported from a statistical point-of-view.

⁶Gastil (1985) suggests that changes of two or more points are interesting, but changes of one point could be due to arbitrary decisions or change. The Freedom House Survey Team is silent on what a ‘significant’ difference is. Below I use the measures I develop to address this question empirically.

frequentist counterparts (Pemstein et al. 2010). Before discussing the specific application of this technology, let me provide a bit of background.⁷

Bayesian inference rests on a different set of assumptions than its better-known frequentist counterpart. The foundations lie in Bayes' Law which suggests:

$$p(\theta|D) \propto p(\theta)p(D|\theta) \tag{1}$$

where we assume that D refers to our observed data and θ refers to our model (e.g., regression coefficients). As Gill (2007) suggests, on the left-hand side of the equation is the quantity about which inferences are to be made - the probability of obtaining certain parameters given a fixed set of data.⁸ You can think about $p(\theta|D)$ as the distribution of regression coefficients (generally referred to as the posterior distribution). The second term on the right-hand side is the probability of the data given a fixed set of parameters (aka the likelihood). And the only remaining part of the equation is $p(\theta)$ the unconditional probability of the model parameters - generally referred to as the prior distribution. In general, priors allow us to specify our prior knowledge/ignorance about the likely and unlikely values of the model parameters. In latent variable models, they can serve the additional purpose of model identification (more on this later). The posterior distribution, $p(\theta|D)$, is a compromise between our prior beliefs and the likelihood. The weight of the prior relative to the likelihood in this compromise is inversely proportional to n , the number of data points. In the latent variable models discussed below, the latent variable scores (i.e., the country-year scores of political rights and civil liberties) and other model parameters comprise θ . Thus, the models allow us to address the question – given a set of indicators of rights and liberties, what is the posterior

⁷See Gill (2007), Jackman (2009) for an introduction to Bayesian inference in the Social Sciences.

⁸Notice, that one of the main differences between frequentist and Bayesian inference is that Bayesians assume the data are fixed and parameters variable and frequentists assume data come from some infinitely repeatable generating process with constant, fixed parameters.

distribution of the latent variables political rights and civil liberties.

Bayesian inference is being used more frequently in the measurement models for many reasons. These models acknowledge the various levels of measurement of the data, a benefit also realized by some frequentist techniques implemented by boutique software like MPlus. These techniques do not rely on correlation structures (i.e., multivariate normality) to generate estimates. More importantly, these models provide direct measures of uncertainty for all model parameters, including latent variable estimates (political rights and civil liberties in this case). This permits inference on latent variable scores. Thus, it is possible to assess the statistical difference in political rights scores across time or space. The Bayesian model can also incorporate temporal information if it exists, as I do with the Freedom House data (see Armstrong 2009, for another example).

This model treats the *Freedom in the World* sub-categories as continuous data. It models each sub-category for each year as a function of the country-year-specific latent variable, as in equation 2 for political rights and equation 3 for civil liberties.

$$\text{PR}_{ijt} = \alpha_{tj} + \beta_{tj}\theta_{it}^{(\text{PR})} + \varepsilon_{ijt} \quad (2)$$

$$\text{CL}_{ijt} = \delta_{tj} + \gamma_{tj}\theta_{it}^{(\text{CL})} + \nu_{ijt} \quad (3)$$

where i indexes observations (countries), $t = \{2006, 2007, 2008, 2009\}$ and j indexes sub-category (three for political rights and four for civil liberties).⁹ Each observed indicator (PR_{ijt} or CL_{ijt}) is a linear function of its respective latent variable ($\theta_{it}^{(\text{PR})}$ or $\theta_{it}^{(\text{CL})}$). Notice that the θ 's have no j subscript. The model assumes that all three political rights sub-indicators are caused by the same underlying latent variable. However, the way in which that underlying latent variable relates to each observed

⁹Note, that in the models, the PR and CL variables have been centered in each year.

indicator changes (through the β and γ terms). I use Markov Chain Monte Carlo simulation to get estimates of all model parameters.

A brief discussion of the specification of the model is required. Latent variable models are inherently unidentified and require some parameters to be set deterministically to ensure identification. To this end, I set $\beta_{t1} = \gamma_{t1} = 1$ and $\alpha_{t1} = \delta_{t1} = 0$ for each time-period.¹⁰ As the model is Bayesian, it requires the specification of prior distributions for all free model parameters. The free elements of α_{tj} , β_{tj} , δ_{tj} and γ_{tj} are given normal priors with mean zero and variance of ten. The residuals (ε_{itj} and ν_{itj}) are assumed to be normal with mean zero and standard deviations $\sigma_{tj}^{(\varepsilon)}$ and $\sigma_{tj}^{(\nu)}$, the squares of which (i.e., the variances) are given inverse gamma priors with rate and shape parameters equal to one. For reference, Figure 1 shows the probability density function of the prior on the standard deviations (i.e., an inverse gamma with rate=1 and shape=1). In the first period (2006), both latent variables ($\theta_{i1}^{(PR)}$ and $\theta_{i1}^{(CL)}$) are drawn from a normal distribution with mean zero and standard deviations, $\sigma_1^{(PR)}$ and $\sigma_1^{(CL)}$, the squares of which are given inverse gamma priors with rate and shape parameters equal to one. These standard deviations set the scales of the latent variables. In periods two through four, the latent variables are given normal priors with mean equal to that country's latent variable value from the previous period and standard deviations, $\sigma_2^{(PR)}$ and $\sigma_2^{(CL)}$, with the same priors as those in the first period. These second standard deviation parameters for each latent variable give the standard deviation of the temporal changes. Including temporal information this way will result in smoothing out changes over time, thus meeting one of the three goals attributed to Freedom House's measurement exercise.

[Figure 1 about here.]

Random starting values were chosen from their respective distributions for each

¹⁰Technically, the model could also be identified by constraining the variance of the latent variables, but this makes the inclusion of temporal information a bit more difficult.

of two chains. The model showed signs of convergence almost immediately and the results presented below are summary statistics of 2,500 chain values (1,250 for each of two chains) after 10,000 burn-in iterations. All model diagnostics (density plots and trace-lines as well as the Brooks, Gelman and Rubin diagnostic) indicated convergence.¹¹

Factor Model Results

There are two potentially interesting aspects of the model. The first, are the latent variable scores themselves which describe the level of political rights and civil liberties in each country-year. The second is the set of coefficients relating the latent variable to the observed indicators. This piece of information indicates the extent to which the observed data are good indicators of the latent construct. I confine a more thorough discussion of these parameters to the on-line appendix, but suffice it to say that all indicators are reliable. Figure 2 provide the mean and 95% credible interval¹² of the posterior distributions of political rights and civil liberties scores for the 50 most populous states in the world.¹³ In general, observations with non-overlapping credible intervals, of which there are many, will be statistically different from each other. The vertical line at zero is simply for reference. The graphs suggest that a considerable number of countries all have scores indistinguishable from the highest score and that there are only a few that attain the lowest value. Most importantly, notice the variance in the point estimates the values of PR and CL. A comparable graph of the original political rights and civil liberties would have seven groups of points with no variance estimates. The importance of the ability to make fine distinctions between countries can hardly be overstated.

¹¹This model was run in JAGS version 1.0.3 on a Mac running OS X (10.5.8), as were all of the others in this manuscript.

¹²These are similar to a frequentist confidence interval, but they can be interpreted differently. Here, the interpretation is more natural - 95% of the density of the latent variable lies between the lower and upper bounds.

¹³2007 population data come from Penn World Tables(Heston et al. 2009).

[Figure 2 about here.]

While these numbers are interesting in their own right, the analysis would be incomplete without a comparison to the classification scheme used by Freedom House - the seven-fold typology. First, the correlations between my numbers and the Freedom House counterparts are quite strong, ranging from -0.97 or -0.98.¹⁴ Despite high correlations, it is possible to see real differences in both statistical significance and effect size in statistical models, as the example below will demonstrate. Figure 3 shows violin plots of the new scores (x-axis) by the original Freedom House political rights and civil liberties scores (y-axis) (Kastellec & Leoni 2007). The strong negative relationship between the two types of variables is clear - the medians (indicated by dots in each violin) increase as the original Freedom House scores decrease. However, there are certainly some overlaps across the distributions. Take for example the violins for political rights categories three and four. The modes of the two distributions are almost identical, though the distributions are skewed in different directions. That means that some countries originally coded three by Freedom House (indicating more political rights) actually have fewer political rights than some countries that were originally coded as four.

[Figure 3 about here.]

Table II provides the probability of significant differences within and across the original categories of the Freedom House variables. For example, the upper-left entry in the table is 0.14. This indicates that 14% of countries scored one on Political Rights (across the four years under consideration) are significantly different from each other on my political rights measure.¹⁵ The second-row, first column of

¹⁴The scores on the sub-indicators are decreasing in freedom, so the smallest scores indicate most free countries. The latent variable were scored to increase in freedom, thus the negative correlations.

¹⁵Of the 772 observations over the four years under consideration, 235 of them are coded as one on Political Rights. This results in 27,495 possible undirected pairwise comparisons (1 → 2 and 2 → 1 are considered the same). Of these, 3,777 comparisons have a probability of being different of greater than 0.95.

the left-hand table has a value of 0.86. This indicates that in all possible pairwise comparisons between the original ones and twos on Political Rights, 86% of them show statistically significant differences. So 14% of pairwise comparisons of ones and twos are statistically insignificant. On the other hand, look at the entry of 0.73 for the sixth row and sixth column of the right-hand sub-table in Table II. This indicates that of all the observations originally coded as six on civil liberties by Freedom House, 73% of the pairwise comparisons among those observations are statistically significant. This suggests that the original categorization is hiding a lot of information.¹⁶

[Table 2 about here.]

Two important conclusions can be drawn from Table II. First, there is a considerable amount of information being lost in the categorization implemented by Freedom House. The fact that the numbers on the table's diagonal are considerably different from zero (sometimes greater than 0.5), suggests that there are interesting differences within categories. The second conclusion, complementary to the first, is that not all countries with different scores in the original data are statistically different from each other. These results confirm Gastil's (1985) original intuition that the probability that two countries in adjacent categories are indistinguishable might be non-trivial (here shown empirically to be as high as 0.27 or [1-0.73], the difference between three and two on Political Rights), the probability that countries two or more units apart are not truly different from each other is very low (here only as high as 0.01)

This model also permits the assessment of cross-temporal differences within each country. Table III summarizes the agreement of over-time changes using Freedom House's original scores (columns) and the scores I generated above (rows). The

¹⁶For reference, a similar matrix for Freedom House's original data would have zeros on the diagonal and ones on all of the off-diagonal elements.

entries in the table are counts of the number of countries identified by both Freedom House and this study as experiencing significant changes over time.¹⁷ If the two methods agreed perfectly, all of the entries would fall on the diagonal of each four-by-four sub-table. The table indicates considerable disagreement about which countries are changing over time. Even though one of the stated goals of the organization is temporal stability, Freedom House seems to indicate many more over-time changes than are supported by the data.

[Table 3 about here.]

Despite high correlations between the variables generated here and those from Freedom House, there are some significant differences (both substantively and statistically) in the discrepancies. To the extent that differences either across time or across countries are important, the models presented here are on much firmer statistical ground to assess those differences.

Statistical Results of the Dynamic Bayesian Factor Model

The “dynamic” part of the model suggests that something can be learned about behavior over time. In the discussion of the parameters above there were two different variances for each set of latent variable estimates. The first standard deviation for each ($\sigma_1^{(PR)}$ and $\sigma_1^{(CL)}$) are the cross-sectional standard deviations in the first period - the variance that characterizes the spread of countries in 2006. For civil liberties, the first-period standard deviation is 4.26 and for political rights, it is 4.10. The second set of variances refer to the variance of the time-series innovations. They tell us the spread of observations from one time-period to the next. These standard deviations are much smaller at 0.25 and 0.41 for civil liberties and political rights, respectively. The standard deviations for the first period are an order of

¹⁷A single country can experience both positive and negative changes across the four years under investigation.

magnitude bigger than the over-time standard deviations. This is an indicator of the general stability in scores over time. This underscores the finding from above that only a few countries make statistically significant moves over the four years under consideration here.

The model estimated here also produces results akin to those from a traditional factor analysis. First, the model produces coefficients (the α , β , δ and γ parameters from equations 2 and 3) and an estimate of model fit. Figure 4 presents the results for β and γ .¹⁸ The points represent the coefficients relating the latent variables to the observed variables and the line segments represent the 95% credible intervals for their respective coefficients. The coefficients for ‘rule of law’ and ‘personal autonomy and individual rights’ are less than the coefficients for ‘freedom of expression and belief’ and ‘associational and organizational rights’. Indicators with smaller coefficients are not as reliable indicators of the underlying phenomenon, though these would still meet any reasonable cutoff for inclusion in a measurement model.

[Figure 4 about here.]

In general, the model results show that while the Freedom House procedure produces results that largely stand up to closer scrutiny, at least as much closer scrutiny as is possible with the data provided, the similarities among the indicators hide important differences in inferences that may be made about changes over time and space. Thus, the models presented here emerge as clearly superior on those grounds.

¹⁸The coefficients β and γ provide interesting information about the fit of the model to the indicators. The α and γ parameters are less interesting because they simply give the mean of the indicators when the latent variables are zero. These results, while not presented here, are available from the author upon request.

Example

Above, I demonstrated that comparison of countries over time and space can benefit greatly from the method proposed here. Proper inference can be done on these quantities, where it would have been impossible in the past (unless perfect measurement on the part of Freedom House was assumed). I also alluded to the fact that statistical results can change as well, depending on whether the method above was chosen over Freedom House’s original data. It is to this result that I turn now.

In her article, Whitten-Woodring (2009) suggests that among other things, an interaction between media freedom and regime type influences a state’s respect for physical integrity rights. While I am not replicating the findings in the article *per se*, I estimate a model using political rights (both Freedom House’s original method and from the Bayesian latent variable model) to show how results in statistical models can change even when substitutable variables of different operationalizations exhibit relatively high correlations. The model I estimate is as follows:

$$\begin{aligned} \text{Physical Integrity Rights}_{it} = & \beta_0 + \beta_1 \text{Partly Free Press}_{it} + \beta_2 \text{Not Free Press}_{it} \quad (4) \\ & + \beta_3 \text{Political Rights}_{it} + \beta_4 \text{Political Rights}_{it} \times \text{Partly Free Press}_{it} \\ & + \beta_5 \text{Political Rights}_{it} \times \text{Not Free Press}_{it} + \varepsilon_{it} \end{aligned}$$

Table IV shows the results of the estimation. The first column shows the results using Freedom House’s original variable. The second and third column show the results using the latent variable estimates (multiplied by negative one for comparison purposes). The second column uses the posterior means without propagating the measurement uncertainty through the model and the third column does propagate the measurement uncertainty through the model as suggested in Treier & Jackman (2008). Given the multiplicative nature of the model, rather than

presenting coefficients, I present conditional intercepts and coefficients (i.e., marginal effects of political rights) for each of the three press freedom regimes.

[Table 4 about here.]

The dependent variable here is the same as in Whitten-Woodring's study - the CIRI physical integrity violations variable (Cingranelli & Richards 2004). The two independent variables are different than the original study. I use the Freedom House *Freedom of the Press* indicator which is a three-fold classification of presses, indicating whether they are 'free', 'partly free' or 'not free' (Freedom House 2008). In all cases, the political rights variable was standardized to have zero mean and unit variance.

Notice that there is an interesting change when moving from the original political rights variable to the latent variable estimates. First, using the original data, there was no significant effect of regime-type among countries with free presses. Using the latent variable estimate a significant effect appears. The greater ability to distinguish among regimes at the high-end of the scale has paid dividends in the form of a statistically significant relationship. Further, the results using the original data showed a significant effect of regime-type in places with partly free presses. This relationship has disappeared in both models using the latent variable. Finally, the relationship between regime-type and physical integrity violations among countries with not free presses remains largely the same in sign, significance and magnitude.

There is the opportunity for statistical results to change in interesting ways. Consider the third column of the table. This shows that even when the measurement error is propagated through the model, there is still a significant effect of regime-type in countries with free presses. The fact that the identification of relationships as statistically important or unimportant depends on measurement should encourage everyone to take matters like this seriously.

Conclusion

I have shown here that while the aggregate scores produced by Freedom House largely stand up to greater scrutiny, they fail to capitalize on a considerable amount of information. The model employed here, in an effort to take advantage of this information, had three goals. First, it estimated latent variables with measures of uncertainty for the latent point estimates. Second, it explicitly accounted for the temporal structure of the data, using the previous period's values as starting points for the current period's estimation. Third, it did not employ arbitrary cut-off points to categorize continuous data.

This model takes these data about as far as they can go. While there are potentially some spatial modeling techniques that could provide some further purchase, I suspect any improvement over the current model with these data will be marginal at best. This is not to say that there are no remaining problems with the Freedom House data. In fact, there are some potentially serious problems that simply cannot be addressed with the data released by Freedom House.

First, it appears that only one person, the person in charge of writing the country report, fills out the survey (provides numerical ratings for each country). However, the numerical ratings are debated by a number of individuals and the final numbers provided for each country are, from what I can gather, a result of considerable debate and compromise. The fact that there is likely variation in opinion does not get considered. If a number of individuals completed the survey, it would be possible for the debate and compromise to take place within the statistical model, which would permit a characterization of the uncertainty inherent in this process. The fact that the data available do not reflect this uncertainty means the numbers I produced above should be less precise than indicated.

Another smaller and soluble problem is that the organization is still providing aggregate scores. The estimates provided above would be better if the 25 questions

used to generate the seven sub-category scales were released. Perhaps this work, showing that the likely result of increased scrutiny will be better information that is in line with Freedom House's current data and goals, will provide the push required for the release of these data.

Having said all of this, it is important to note that the numbers produced by Freedom House are quite close to the ones proposed here. For political rights and civil liberties, the variables correlate around .97-.98 with the Freedom House variables. While the numbers are similar, I think the more complex model proposed here pays dividends in the continuous nature of the political rights and civil liberties measures and with the uncertainty estimates around all latent point estimates. For researchers interested in comparing countries, the benefit of these data as opposed to their original Freedom House counterparts is obvious.

Data replication: The dataset, codebook, and code files for the empirical analysis in this article can be found at <http://www.prio.no/jpr/datasets>.

References

- Armstrong, David A. (2009) Measuring the democracy-repression nexus. *Electoral Studies* 28(3): 403–412.
- Bollen, Kenneth & Pamela Paxton (2000) Subjective measures of liberal democracy. *Comparative Political Studies* 33: 58–86.
- Cingranelli, David L. & David L. Richards (2004) *The Cingranelli-Richards (CIRI) Human Rights Dataset*. [Http://www.humanrightsdata.org](http://www.humanrightsdata.org), accessed 12-1-2004.
- Freedom House (2008) *Freedom of the Press 2008*. Lanham, MD: Rowman & Littlefield.
- Gastil, Raymond D. (1985) *Freedom in the World: Political Rights and Civil Liberties 1984-1985*. Westport, CT: Greenwood Press.
- Gill, Jeff (2007) *Bayesian Methods: A Social and Behavioral Science Approach 2nd ed.* Boca Raton, FL: Chapman Hall/CRC.
- Heston, Alan; Robert Summers & Bettina Aten (2009) Penn world table version 6.3.
- Jackman, Simon (2009) *Bayesian Analysis for the Social Sciences*. Chicester, UK: John Wiley & Sons, Ltd.
- Jacoby, William G. (1991) *Data Theory and Dimensional Analysis*. Thousand Oaks: Sage.
- Kastellec, Jonathan P & Eduardo L. Leoni (2007) Using graphs instead of tables in political science. *Perspectives on Politics* 5(4): 755–771.
- Pemstein, Daniel; Stephen A. Meserve & James Melton (2010) Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis* 18(4): 426–449.
- Treier, Shawn & Simon Jackman (2008) Democracy as a latent variable. *American Journal of Political Science* 52(1): 201–217.
- Whitten-Woodring, Jennifer (2009) Watchdog or lapdog? media freedom, regime type and government respect for human rights. *International Studies Quarterly* 53: 595–625.

Table I: Description of Freedom House Sub-categories

Political Rights	Civil Liberties
Electoral Process	Freedom of Expression and Belief
Political Pluralism and Participation	Associational and Organizational Rights
Functioning of Government	Rule of Law
	Personal Autonomy and Individual Rights

Table II: Significant Differences in Latent Variable Scores

	Political Rights							Civil Liberties						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	0.14							0.26						
2	0.86	0.48						0.89	0.42					
3	1.00	0.85	0.56					1.00	0.78	0.72				
4	1.00	1.00	0.73	0.56				1.00	1.00	0.87	0.67			
5	1.00	1.00	0.99	0.82	0.47			1.00	1.00	1.00	0.92	0.68		
6	1.00	1.00	1.00	0.99	0.87	0.50		1.00	1.00	1.00	0.99	0.89	0.73	
7	1.00	1.00	1.00	1.00	0.99	0.88	0.41	1.00	1.00	1.00	1.00	1.00	0.88	0.62

Entries represent the percentage of significant pairwise differences in the latent variables for comparisons of the row and column categories of the Freedom House scale.

Table III: Significant Over-time Swings in Political Rights and Civil Liberties

		FH Scores							
		PR				CL			
		None	Positive	Negative	Both	None	Positive	Negative	Both
My Scores	None	144	5	17	7	159	6	17	4
	Positive	0	7	0	1	0	4	0	0
	Negative	2	0	7	2	1	0	2	0
	Both	0	0	0	1	0	0	0	0

Cell entries are the number of countries (total = 193) that saw positive and/or negative vs no change in freedom over time according to the values I generated and those generated by Freedom House.

Table IV: Models of Physical Integrity Rights (Conditional OLS Coefficients)

	Original	Latent 1 [†]	Latent 2 [‡]
Intercept			
Free Press (β_0)	5.960*	5.770*	5.858*
	(0.421)	(0.394)	(0.402)
Partly Free Press ($\beta_0 + \beta_1$)	4.052*	3.655*	3.635*
	(0.293)	(0.286)	(0.285)
Not Free Press ($\beta_0 + \beta_2$)	4.617*	4.596*	4.597*
	(0.131)	(0.133)	(0.136)
Conditional Coefficients of Political Rights			
Free Press (β_3)	0.785	1.010*	0.921*
	(0.420)	(0.402)	(0.408)
Partly Free Press ($\beta_3 + \beta_4$)	0.603*	0.252	0.238
	(0.233)	(0.224)	(0.227)
Not Free Press ($\beta_3 + \beta_5$)	0.577*	0.564*	0.558*
	(0.221)	(0.239)	(0.242)
SEs in parentheses, * $p < 0.05$, two-tailed			
† Measurement error not propagated			
‡ Measurement error propagated			

Figure 1: Probability Density Function of Inverse Gamma Distribution

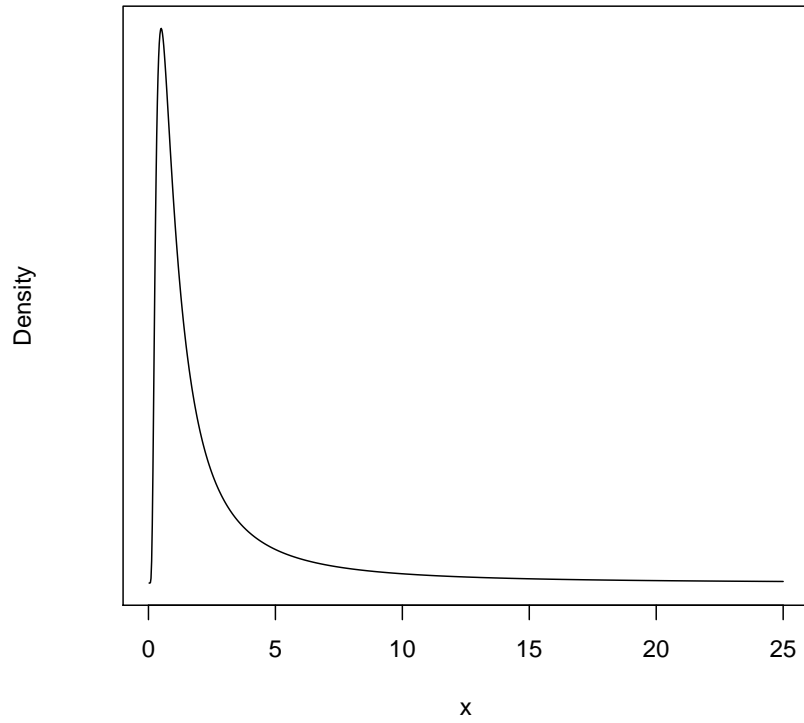
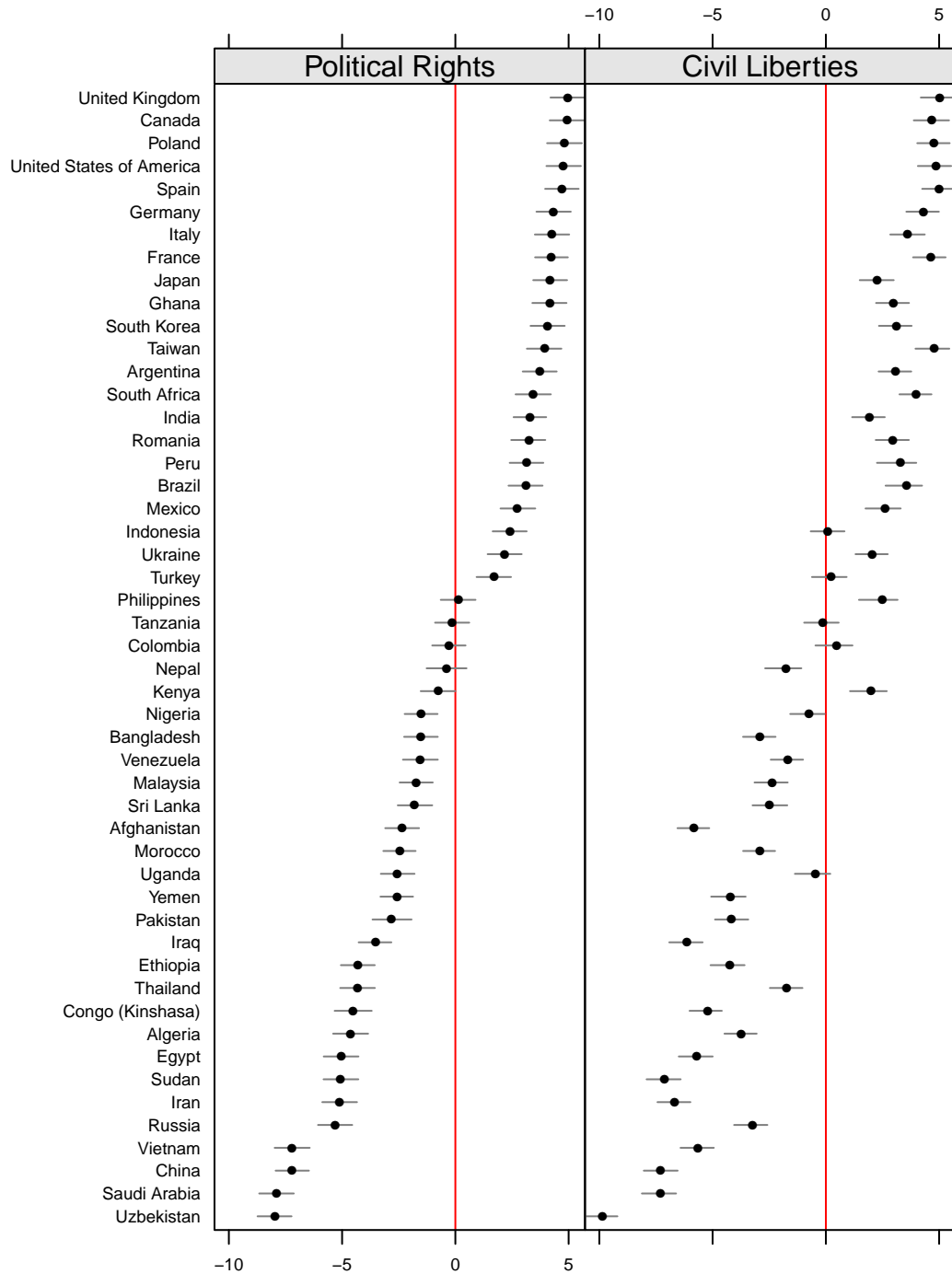
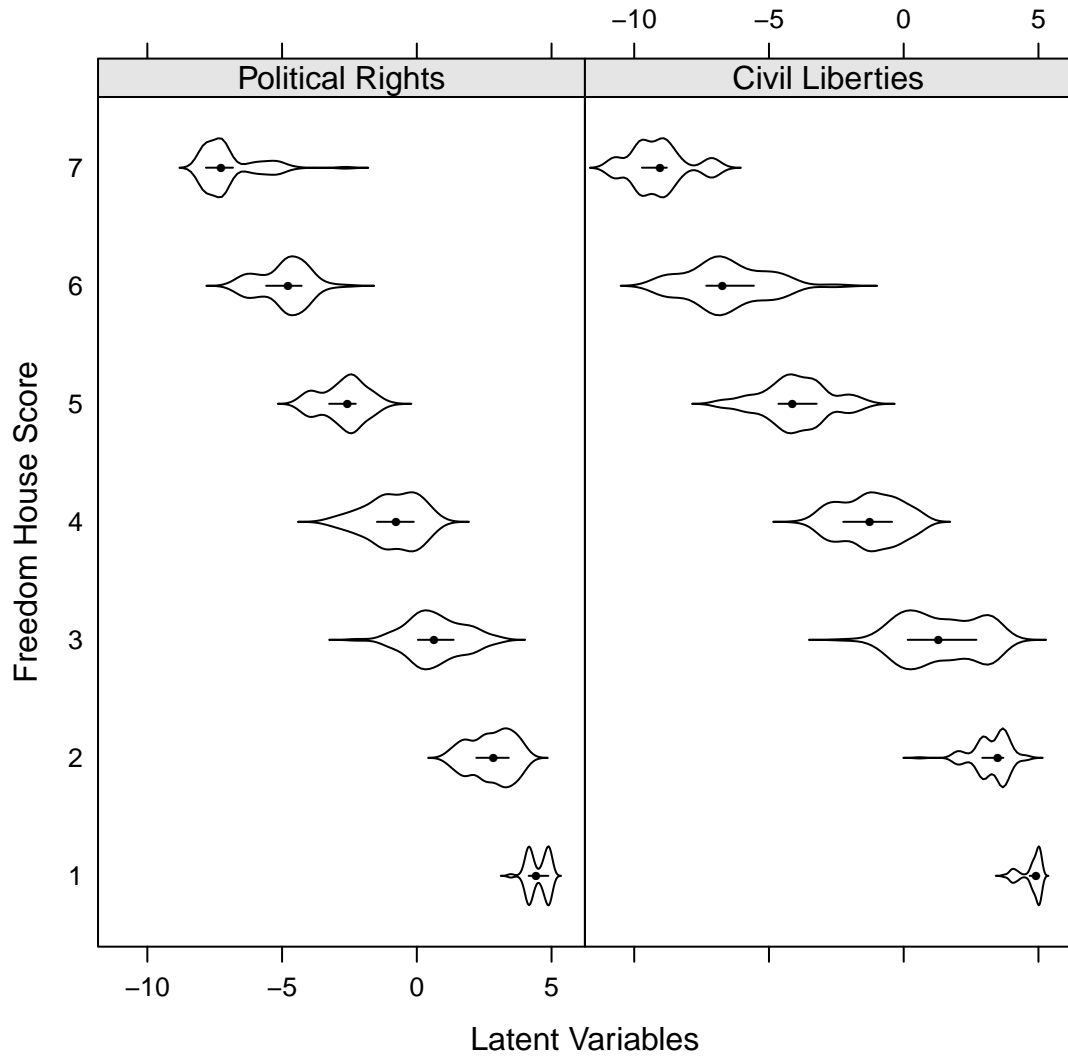


Figure 2: Dot-plot of Political Rights and Civil Liberties 2009



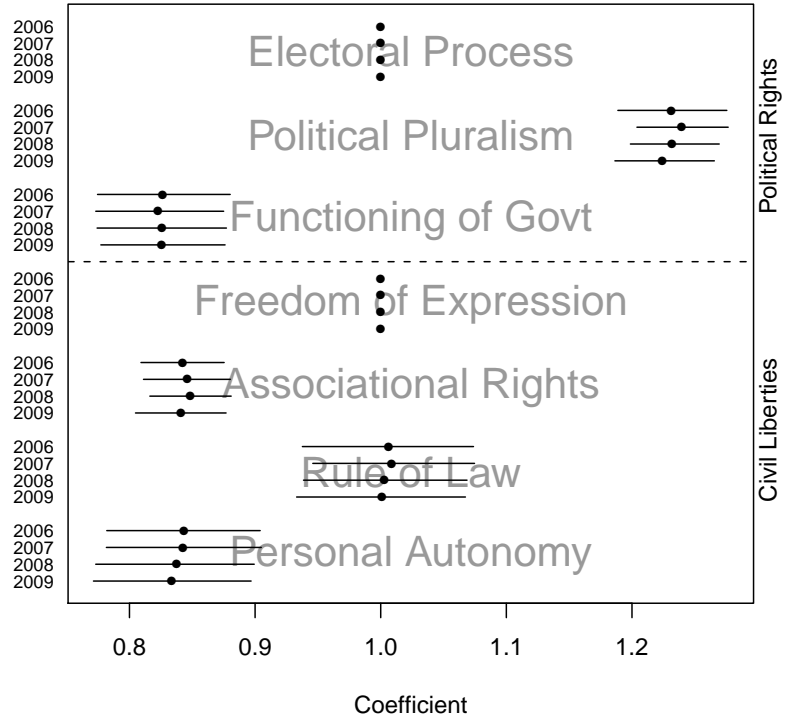
The points represent the posterior means and the line segments represent the 95% credible interval for each estimate. The points included are those for the 50 most populous countries as of 2007 according to Penn World Tables 6.3 (Heston et al. 2009). Similar figures for all countries and for other years are available from the author upon request.

Figure 3: Violin Plots of New and Old Political Rights and Civil Liberties Scores



The violin plots represent the distribution of the posterior means of the latent variables for all periods (2006-09) organized by their original Freedom House score.

Figure 4: Statistical Results from Dynamic Bayesian Factor Model



The points represent coefficients relating the latent variables to the observed variables indicated with text. The coefficients for Electoral Process (PR) and Freedom of Expression (CL) were set to 1 for identification purposes. The line segments represent the 95% credible intervals for each coefficient.