On-line Appendix to: Stability and Change in the Freedom House Political Rights and Civil Liberties Measures

January 10, 2011

The History of *Freedom in the World*

Nearly every year since 1973, Freedom House has published political rights, civil liberties and freedom status indicators for all of the countries and a number of dependent territories in the world. These indicators are used by scholars as measures of democracy or human rights, by non-governmental organizations to track the status of government's respect of the rights of citizens and, perhaps most importantly, by policy makers to adjudicate the allocation and effectiveness of aid and other funding to developing countries. Freedom House's advocacy and involvement in government policy making give its indicators of political rights, civil liberties and freedom status unparalleled importance in the political realm. The political importance of these data only makes a thorough understanding of their measurement properties more important. If decisions to intervene in a country's affairs, either financially or otherwise, are being made based on these data, then the ability to distinguish among different types of regimes is of more than academic importance.

To be fair, the political rights, civil liberties and freedom status indicators were never intended to be the product of a rigorous measurement exercise. Raymond Gastil, the Survey's founder, did not want to make unwarranted or unjustified claims to precision. In his introduction to the 1984-85 Freedom in the World survey, Gastil (1985) derides the work of Bollen (1980) for producing a measure on a "thousand-point scale" with data that cannot possibly justify or produce that level of precision. He even sees a rank-ordering of countries as difficult to justify and execute. The current incarnation of the survey would permit greater claims to rigor and precision, though I suspect they would not stray especially far from their founder's sentiment. As Gastil designed the Survey, it was intended to orient discussions of freedom, not to supply a precise numerical summary.

The first incarnation of the current rating system was devised and implemented by Raymond Gastil (1973). In his introduction to the volume published in 1979, Gastil describes the Survey as follows: "Since 1972 Freedom House has rated the level of freedom in each country in the world by means of a Comparative Survey of political and civil rights.

The Surveys provide an objective reference for judging political and civil rights within the maelstrom of emerging governments and changing standards" (Gastil 1979, 3). In 1991, looking back on his tenure as editor of the Freedom in the World Survey, he describes the intention of the enterprise: "The original intention was to produce, with relatively few manhours, an orienting discussion of variation in levels of freedom" Gastil (1991, 21), to provide a heuristic for describing freedom around the world.

While the Survey has explicitly disavowed any goals of quantitative precision,[1] it does lay claim to being scientific and rigorous. In 1991, Gastil said the extent to which the Survey is "scientific," is due to, "the author's determination not to let current international opinion, the interests of American foreign policy, Freedom House, or personal prejudices affect survey ratings" (Gastil 1991, 22). The survey continues to promote its ratings as scientific and rigorous as the Survey Methodology section of the more recent volumes suggest, "The survey process emphasizes intellectual rigor and balanced and unbiased judgements" (Piano & Puddington 2004, 711). Gastil thought that placing countries on a seven-point scale would reduce the temptation to make distinctions between countries that are unsupported by the data. While this is certainly a temptation worth avoiding, the eye remains drawn toward distinctions between categories - that is to say, countries with different category memberships are treated differently when, in fact, a country that is a "high" three and a country that is a "low" four are unlikely to be interestingly different from each other.

The *Freedom in the World* Survey had policy implications from the beginning. Again, in the introduction to the 1979 volume, Gastil describes what would become a continued American foreign policy goal,

> ... a recurrent American policy has been to go to the aid of other countries because their governments represented democratic systems similar to our own, or were struggling against forces hostile to democracy. But this policy can

[1] As Gastil said in his introduction to the volume published in 1981, "The object of the Surveys, now in their tenth year, is not quantitative. We hope only that the presentation allows a reader to better evaluate the changing systems of the world as daily events bring to his consciousness one or another nation or national leader" (Gastil 1981, 4).

> hardly be implemented if we cannot distinguish convincingly the more free from the less free states, particularly in the Third World (Gastil 1979, 3).

The tension in Gastil's own statements is palpable. It seems that Gastil wanted to produce a measure that was precise enough to distinguish among regimes, but no more. This is further emphasized by Gastil's comments in 1981:

> [The Survey] quietly demonstrates the differences between contending systems, but also suggests that we should not imagine the distinctions are as clear and unambiguous as some may like. The Survey helps to orient its readers in the world; it may help to preserve balance in the struggle for that world; and it may in the end play its part in reducing the oppression it evaluates. (Gastil 1981, 11)

Another feature of the data is that some inertia is built into the scores. The scores for the previous year are used as a basis for the current year's scores. Piano and Puddington explain, "[t]he ratings were compared to the previous year's findings, and any major proposed numerical shifts or category changes were subject to more intense scrutiny" (Piano & Puddington 2008). Changes in category membership from year-to-year are taken quite seriously and only happen if there are real-world changes in the level of political rights and civil liberties. Again, there is an effort to ensure that changes are not the result of chance or whim, rather that they result from real, observable changes in government practices.

All of this suggests three primary motivations of the Freedom House organization in crafting the Freedom in the World studies.

1. Beware Unwarranted Precision - scores generated from the Freedom House data should not make claims to precision that are not warranted by the data.

2. Rigor in Measurement - the scores generated should be as transparent as possible and stand up to scrutiny.

3. Temporal Stability - the scores generated should not only use the data on hand this year, but should also use data from previous years to ensure trends show relative stability over time.

The methods I discuss in the print article meet all of these criteria. Further, they provide more information, even when taking uncertainty into consideration, than the the seven-point scales provided by Freedom House. The procedures used to generate scores are more justifiable from a measurement perspective and more effectively used to distinguish between levels of freedom across countries and time.

Freedom Status

Along with political rights and civil liberties, Freedom House also produces a freedom status variable that categorizes countries as free, partly free or not free. This was intended to be a rough categorization that permits the production of a 'map of freedom'. The current method for generating these results is to average the seven-point political rights and civil liberties scales and then split the averaged scores into three categories. Below, I discuss a Bayesian latent class analysis of the data that produces a three-fold classification and compare this to Freedom House's classification.

Dynamic Bayesian Latent Class Model for Freedom Status

The Bayesian latent class model used here has two advantages over Freedom House's method. First, measures of uncertainty exist for all latent point estimates. Second, it does not use arbitrary cut-off values to identify status membership. This model is a categorical analog of the model discussed above (Congdon 2003). As Freedom House does, I assume that there are three latent states (e.g., free, partly free and not free).[2] Each observation is

[2]While it would be possible to estimate a different number of latent states or even to make the number of states a parameter in the model, my objective is to provide a more appropriate model for Freedom House's three-state classification.

probabilistically a member of each state, though for some the probability of being in one particular category will approach one. The model takes the following form.

$$Y_{ijt} = \beta_{jt1}(\text{Status}_{it} = F) + \beta_{jt2}(\text{Status}_{it} = PF) + \beta_{jt3}(\text{Status}_{it} = NF) + \varepsilon_{ijt} \quad (1)$$

Y_{ijt} refers to the observed variables. These are the seven issues mentioned above in Table **??**. Since Freedom House uses both the political rights and civil liberties variables to construct the freedom status variable, I use all seven sub-category items as observed data. As above, i indexes observations (countries), j indexes the seven sub-categories and $t = \{2006, 2007, 2008, 2009\}$. The independent variables in the equation above are dummy variables coded one for each of the three possible levels of the freedom status variable at each time-point. For each period, the Status variable has the following prior:[3]

$$\text{Status}_{it} \sim \text{Categorical}(\boldsymbol{p}_{t,1:3}) \quad (2)$$

where $\boldsymbol{p}_{t,1:3}$ is simply the probability of being in each of the three categories at each of the four time-points. $\boldsymbol{p}_{t,1:3}$ is given a non-informative Dirichlet prior with α equal to a vector of ones for each time-point (See **?**, for a thorough discussion of the Dirichlet distribution). Each of the β_{jtk} coefficients is given a normal prior with mean zero and variance 10. The model is identified by forcing $\beta_{jt1} > \beta_{jt2} > \beta_{jt3}$. Each β_{tjk} gives the mean of variable j at time t for freedom status k. The only other parameters in the model are the residual variances ($\sigma^{(\varepsilon)}_{tj}$), the squares of which are given inverse gamma priors with rate and shape parameters equal to one.

Random starting values were chosen that conform to model constraints where they existed. 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

[3]The 'Categorical' distribution, as one astute reviewer pointed out, is an invention of the `bugs` community. This distribution assumes a pre-specified number of possible states, say $j = \{1, \ldots, J\}$. In any single random draw from this distribution, state j appears with probability p_j, where the p_j's sum to 1.

10,000 burn-in iterations. All model diagnostics (density plots and trace-lines as well as the Brooks, Gelman and Rubin diagnostic) showed evidence of convergence.

Latent Class Results

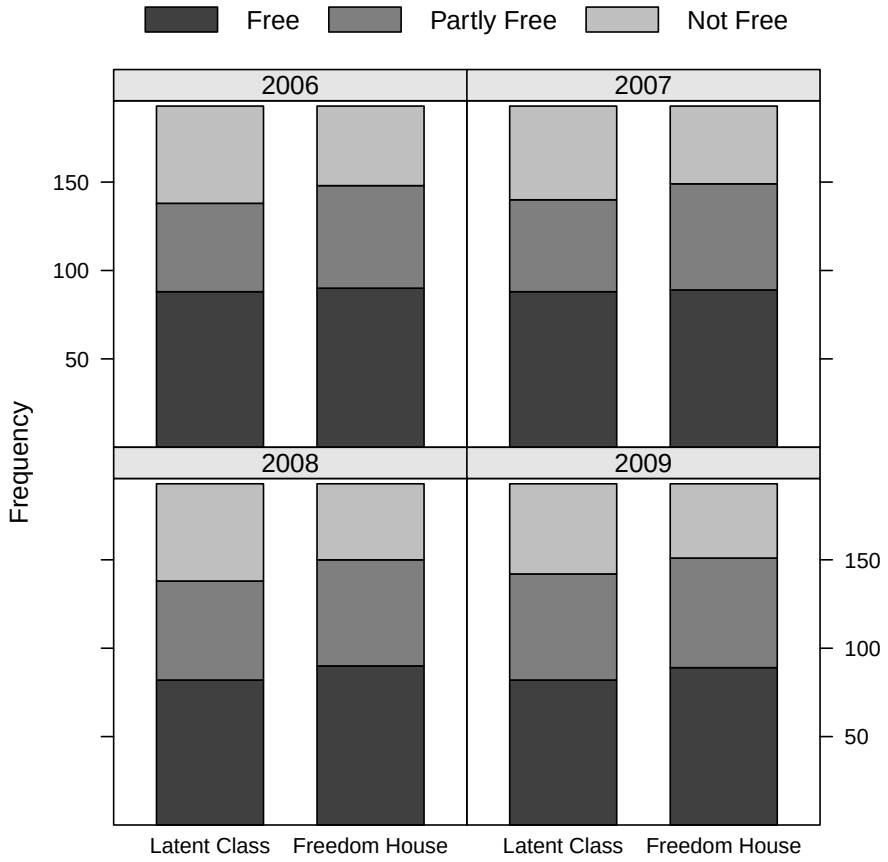
I present the information about freedom status estimates and model fit below. Figure A1 presents the distribution of Freedom House's Freedom Status variable as well as the posterior modes of the latent class result over the four years under consideration. Two things are clear from this graph. First the distributions are remarkably similar over time. I will discuss this in greater detail below. Second, the latent class model systematically identifies more countries in the not free category, though the differences are minor.

In many cases, the two estimation strategies produce the same result (i.e., Freedom House's classification is the same as the posterior mode from the latent class model). However, there is a clear bias in the differences. There are 61 cases for which the probability of being the same as the Freedom House designation is less than 0.05. For all of these cases, the latent class model places them in a less free category than Freedom House's coding. Thus, Freedom House's categorization scheme has a slight bias toward more freedom.

Next, I considered the extent to which countries had a higher than .95 probability of being in a single category. Of the 772 country-years, 93% (718) had a higher than .95 probability of being in a single category. Of the remaining 54, 35 were originally coded by Freedom House as free and 19 were originally coded as partly free (in both cases, about 10% of the observations in each category). For my categorization scheme, about 11% of the observations estimated to be partly free had a probability of greater than 0.05 of being in another category. It was around 5% of observations in the other two categories that were uncertain in their group memberships.

Finally, I look at the stability over time. This can be captured by the transition probabilities – the probability that countries experience a change in category membership over time. In Table A I are the transition probabilities both from the latent class model

Figure A1: Distribution of Freedom Status



In the stacked bar graph above, each different colored section indicates a the number of observations in each of the three possible categories.

and Freedom House. The diagonal elements represent the probability of staying in the same category as in the pervious period. Both tables are dominated by the diagonal. However, the diagonal elements are significantly smaller in in the latent class model indicating movement among the different categories as significantly more fluid than Freedom House would suggest. One thing on which both models agree is that it is not possible to move directly from free to not free or vice versa. The two tables also agree on general patterns. The first of which is that a country is essentially equally likely to transfer into free from partly free as they are to transfer to partly free from free. Further, countries are more likely to transfer from not free to partly free than vice versa, but the latent class

analysis shows that this is not a significant difference.

Table A I: Transition Probabilities for Freedom Status

	Latent Class			Freedom House		
	Free_t	Partly Free_t	Not Free_t	Free_t	Partly Free_t	Not Free_t
Free_{t-1}	0.965 [0.950,0.981]	0.035 [0.019,0.050]	0.000 [0.000,0.000]	0.993	0.007	0.000
Partly Free_{t-1}	0.031 [0.006,0.068]	0.928 [0.891,0.955]	0.040 [0.030,0.058]	0.006	0.978	0.017
Not Free_{t-1}	0.000 [0.000,0.000]	0.061 [0.044,0.075]	0.939 [0.925,0.956]	0.000	0.045	0.955

Main entries on the left-half of the table are the posterior median transition probabilities from the latent class model with 95% credible intervals in brackets beneath. On the right-half of the table, the entries are transition probabilities from the Freedom House data. The off-diagonal elements in both table indicate the probability of transitioning from the row value to the column value.

Taken together, these results are similar in nature to the ones above. The overwhelming majority of cases are coded in a similar fashion. However, there are some interesting patterns lurking in the dissimilarities. Here, the dissimilarities indicate a more fluid system of transitions than Freedom House would suggest. Further, they indicate, on average, a less free world, than Freedom House would have us believe.

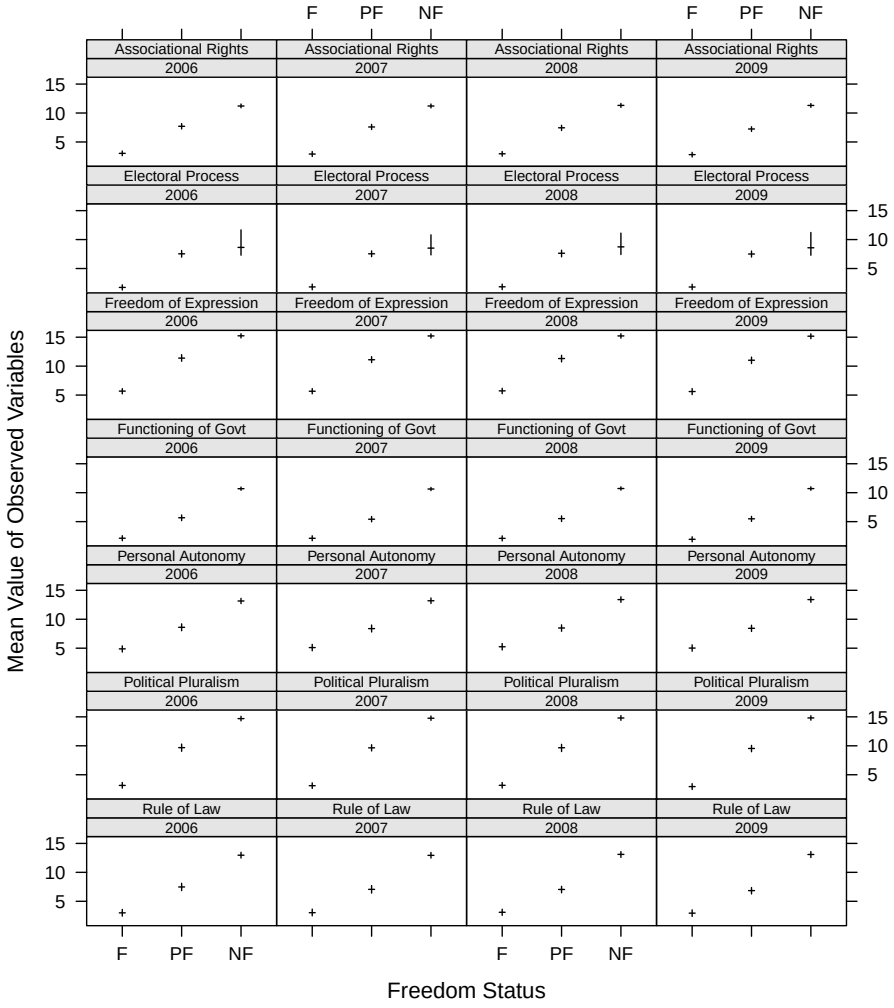
Statistical Results of the Dynamic Bayesian Latent Class Model

Just as with the factor model presented above, there are some statistical results other than category membership that are interesting. The coefficients and their 95% credible intervals are provided in Figure A2. These coefficients are essentially the means of the observed variables for the countries in each of the latent classes. These are very tightly bound and increase in a roughly linear fashion across the categories moving from free to partly free.[4] The only exception to this general pattern is Electoral Process, which is has rather large bounds for the not-free group. This suggests that the distinction made by electoral process

[4]Freedom House codes observations with higher values on the sub-indicators as less free.

is really one between the free countries and the others.

Figure A2: Parameters of the Latent Class Model



In the latent class model, I am estimating the mean of each observed variable for each latent class. The small horizontal ticks represent the mean for each of the three possibilities on each of the seven observed variables. The vertical lines represent the 95% credible interval for the mean estimates.

The main difference between the method proposed above and the one proposed by Freedom House is that the one here does not rely on arbitrary cut-off points to group observations in categories. The fact that the two variables are not very different provides some *post hoc* justification for the Freedom House cut-off points, but to the extent that these variables differ, the one proposed here stands on firmer statistical ground.[5]

[5]The "optimal" cutoffs for the current problem are 2.5 and 4.5 which generate a probability of agreement

References

Bollen, Kenneth (1980) Issues in the comparative measurement of political democracy. *American Sociological Review* 45(2): 370–390.

Congdon, Peter (2003) *Applied Bayesian Modeling*. West Sussex: Wiley.

Gastil, Raymond D. (1973) The new criteria of freedom. *Freedom at Issue* 17: 2–5, 20–23.

Gastil, Raymond D. (1979) The comparative study of freedom: Nature and purposes. In: , 3–14. New York: Freedom House.

Gastil, Raymond D. (1981) The comparative study of freedom: Nature and purposes. In: , 3–12. Oxford: Clio Press.

Gastil, Raymond D. (1985) *Freedom in the World: Political Rights and Civil Liberties 1984-1985*. Westport, CT: Greenwood Press.

Gastil, Raymond D. (1991) The comparative survey of freedom: Experiences and suggestions. In: . New Brunswick, NJ: Transaction Publishers.

Piano, Aili & Arch Puddington (eds.) (2004) *Freedom in the World 2004: The Annual Survey of Political Rights and Civil Liberties*. Lanham, MD: Rowman & Littlefield Publishers, Inc.

Piano, Aili & Arch Puddington (eds.) (2008) *Freedom in the World 2008: The Annual Survey of Political Rights and Civil Liberties*. Lanham, MD: Rowman & Littlefield Publishers, Inc.

with the latent class estimate of 0.96 as opposed to 0.92 for the cutoffs currently used by Freedom House.