THE NOMINATE MODEL

Legislator i's (i=1,…,p) utility for the Yea outcome on roll call j (j=1,…,q) is:

$$U_{ijy} = u_{ijy} + \varepsilon_{ijy} = \beta e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijky}^2\right)} + \varepsilon_{ijy} \qquad (1)$$

where u_{ijy} is the deterministic portion of the utility function, ε_{ijy} is the stochastic portion, and d_{ijky}^2 is the distance of legislator i to the Yea outcome on the k^{th} (k=1,…,s) dimension for roll call j:

$$d_{ijky}^2 = (X_{ik} - O_{jky})^2 \qquad (2)$$

and X_{ik} is the i^{th} legislator's ideal point on dimension k, and O_{jky} is the "Yea" outcome location for the j^{th} roll call on the k^{th} dimension.

The same level of utility, U_{ijy} , can be produced in two different ways. First, fixing the scale of the deterministic utility, u_{ijy} , and varying the standard deviation of the error, ε_{ijy} , that is:

$$\varepsilon_{ijy} \sim N(0, \frac{\sigma^2}{2}) \quad \text{so that} \quad \varepsilon_{ijn} - \varepsilon_{ijy} \sim N(0, \sigma^2)$$

Second, σ^2 can be fixed and the relative weight of u_{ijy} in the overall utility, U_{ijy} , can be adjusted by increasing/decreasing β. In other words, without loss of generality we can assume that:

$$\varepsilon_{ijn} - \varepsilon_{ijy} \sim N(0, 1) \tag{5}$$

This implies that the distribution of the difference between the latent utilities for Yea and Nay is normal with mean $u_{ijy} - u_{ijn}$ and variance 1; that is

$$y_{ij}^{*} = U_{ijy} - U_{ijn} = u_{ijy} - u_{ijn} + \varepsilon_{ijn} - \varepsilon_{ijy} = \beta\left\{e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijky}^2\right)} - e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijkn}^2\right)}\right\} + \varepsilon_{ijn} - \varepsilon_{ijy}$$

$$\sim N(u_{ijy} - u_{ijn} , 1) \tag{6}$$

where y_{ij}^{*} is the difference between the latent utilities.

The probability that legislator i votes Yea on the j^{th} roll call is:

$P_{ijy} = P(U_{ijy} > U_{ijn}) = P(\varepsilon_{ijn} - \varepsilon_{ijy} < u_{ijy} - u_{ijn}) =$

$\Phi(u_{ijy} - u_{ijn}) =$

$$\Phi\left[\beta\left\{e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijky}^2\right)} - e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijkn}^2\right)}\right\}\right] \tag{7}$$

Let **Y** be the p by q matrix of observed Yea/Nay choices and let $\mathbf{Y}^*$ be the p by q matrix of unobserved latent utility differences. From a classical perspective the ***joint probability distribution of the sample*** is:

$$\mathbf{f}(\mathbf{Y}^* \mid u_{ijy}-u_{ijn}) = \prod_{i=1}^{p}\prod_{j=1}^{q}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y_{ij}^*-(u_{ijy}-u_{ijn}))^2} \qquad (8)$$

Where $\mathbf{Y}^*$ is the p by q matrix of latent utility differences. Note that equation (8) is not a typical joint p.d. of the sample. Technically, a random sample is a set of independent and identically distributed random variables so that the joint p.d. of the sample is (DeGroot, 1986, p. 316):

$$f_n(\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n} \mid \boldsymbol{\theta}) = f(\mathbf{X_1} \mid \boldsymbol{\theta})\, f(\mathbf{X_2} \mid \boldsymbol{\theta}) \ldots f(\mathbf{X_n} \mid \boldsymbol{\theta})$$

where $f(\mathbf{X} \mid \boldsymbol{\theta})$ is the distribution from which the random sample, $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}$, is drawn. In contrast each of the pq elements of $\mathbf{Y}^*$ is a random sample of size one from the corresponding $N(u_{ijy} - u_{ijn}, 1)$ distribution. The joint p.d. is a pq-variate normal distribution with variance-covariance matrix $\mathbf{I_{pq}}$:

$$\mathbf{f}(\mathbf{Y}^* \mid u_{ijy}-u_{ijn}) =$$

$$\frac{1}{(2\pi)^{\frac{pq}{2}}}e^{-\frac{1}{2}\{(y_{11}^*-(u_{11y}-u_{11n}))^2+(y_{12}^*-(u_{12y}-u_{12n}))^2+\ldots+(y_{pq}^*-(u_{pqy}-u_{pqn}))^2\}} \qquad \mathbf{(8)}$$

To see that equation (8) is indeed a legal probability distribution note that:

$\mathbf{f}(\mathbf{Y}^* \mid u_{ijy}-u_{ijn}) \geq 0$ for all y_{ij}^*, $-\infty < y_{ij}^* < +\infty$

and

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(\mathrm{Y}^* \mid \mathrm{u}_{ijy} - \mathrm{u}_{ijn})dy_{11}^* dy_{21}^* \ldots dy_{pq}^* =$$

$$\left\{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y_{11}^*-(u_{11y}-u_{11n}))^2}dy_{11}^*\right\}\left\{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y_{21}^*-(u_{21y}-u_{21n}))^2}dy_{21}^*\right\}\ldots$$
$$\left\{\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y_{pq}^*-(u_{pqy}-u_{pqn}))^2}dy_{pq}^*\right\} = 1*1*\ldots*1 = 1$$

In the joint p.d. of the sample, $\mathbf{f}(\mathbf{Y}^* \mid u_{ijy}-u_{ijn})$, the y_{ij}^* are the random variables and the ps+2qs+1 parameters -- X_{i1} , X_{i2} , …, X_{is} , the qs Yea outcome coordinates – O_{j1y} , O_{j2y} ,…, O_{jsy} , the qs Nay outcome coordinates -- O_{j1n} , O_{j2n} ,…, O_{jsn} , and β – are fixed constants. Following DeGroot (1986, p. 317), if we regard $\mathbf{f}(\mathbf{Y}^* \mid u_{ijy}-u_{ijn})$ as a function of the parameters for given values of the y_{ij}^* then it is a ***likelihood function***; that is

$\mathbf{L}^*(u_{ijy}-u_{ijn} \mid \mathbf{Y}^*) =$

$$\frac{1}{(2\pi)^{\frac{pq}{2}}}e^{-\frac{1}{2}\{(y_{11}^*-(u_{11y}-u_{11n}))^2+(y_{12}^*-(u_{12y}-u_{12n}))^2+\ldots+(y_{pq}^*-(u_{pqy}-u_{pqn}))^2\}} \qquad \textbf{(9)}$$

Which is identical to equation (8) only now the pqy_{ij}^* are ***observed*** and the ps+2qs+1 parameters are ***variables*** (but ***not*** random variables), and the problem is to find values of the parameters that maximize equation (9).

Equation (8) is a probability distribution over the pq dimensional hyperplane with dimensions corresponding to the y_{ij}^*. Equation (9) is a function defined over the ps+2qs+1 dimensional hyperplane with dimensions corresponding to the ps legislator coordinates -- X_{i1} , X_{i2} , …, X_{is} , the qs Yea outcome coordinates – O_{j1y} , O_{j2y} ,…, O_{jsy} , the qs Nay outcome coordinates -- O_{j1n} , O_{j2n} ,…, O_{jsn} , and β. Although equation (9) is not a probability distribution it is the case that it is above zero over the ps+2qs+1 hyperplane; that is:

$$\mathbf{L}^*(u_{ijy}-u_{ijn} \mid \mathbf{Y}^*) \geq 0, \quad 0 < \beta < +\infty,$$

$$-\infty < X_{ik} , O_{jky} , O_{jkn} < +\infty$$

In addition, the hypervolume underneath the function is almost certainly finite:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{0}^{+\infty} L^*(\beta\left\{e^{\left(-\frac{1}{2}\sum_{k=1}^{s}d_{ijky}^2\right)} - e^{\left(-\frac{1}{2}\sum_{k=1}^{s}d_{ijkn}^2\right)}\right\}|\mathrm{Y}^*)d\beta dX_{11}\ldots dX_{pk}dO_{11y}\ldots dO_{qky}\ldots dO_{11n}\ldots dO_{qkn} =$$

$$K^* < +\infty \qquad (10)$$

Because as $\beta \to +\infty$ clearly $\mathbf{L}^*(u_{ijy}-u_{ijn} \mid \mathbf{Y}^*) \to 0$; and as the absolute value of any of the legislator and roll call parameters becomes large the likelihood function goes to zero; that is, as $|X_{ik}| \to +\infty$ clearly $\mathbf{L}^*(u_{ijy}-u_{ijn} \mid \mathbf{Y}^*) \to 0$. **L*** is shaped like a multivariate normal in that it is quasi-concave along each dimension and asymptotes towards zero fairly quickly. However, I have no formal proof that the hypervolume is finite.

Unfortunately, the latent utility differences are not observed and we do not have any simple expression for the joint probability distribution for the sample of discrete choices -- $f(\mathbf{Y} \mid u_{ijy}-u_{ijn})$. However, it is easy to write down the distribution corresponding to any particular choice, that is, $f_{ij}(y_{ij} \mid u_{ijy}-u_{ijn})$. The product of these pq distributions is proportional to the joint p.d. of the sample and the likelihood function. Specifically, let:

$$y_{ij} = \begin{cases} 1 \text{ (Yea) if } y_{ij}^* > 0 \\ 0 \text{ (Nay) if } y_{ij}^* \leq 0 \end{cases} \text{ so that } \begin{cases} P(y_{ij}^* > 0) = \mathbf{\Phi}(u_{ijy} - u_{ijn}) \\ P(y_{ij}^* \leq 0) = 1 - \mathbf{\Phi}(u_{ijy} - u_{ijn}) \end{cases}$$

If the y_{ij} are independent Bernoulli random variables, that is:

$$f_{ij}(y_{ij}|u_{ijy} - u_{ijn}) \sim \text{Bernoulli}(\mathbf{\Phi}(u_{ijy} - u_{ijn}))$$

then

$f(\mathbf{Y} \mid u_{ijy}-u_{ijn}) \propto \prod_{i=1}^{p}\prod_{j=1}^{q} f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn})$ **=** $L(u_{ijy}-u_{ijn} \mid \mathbf{Y})$ **=**

$$\prod_{i=1}^{p}\prod_{j=1}^{q}[\Phi(u_{ijy} - u_{ijn})]^{y_{ij}}[1-\Phi(u_{ijy} - u_{ijn})]^{(1-y_{ij})} = \prod_{i=1}^{p}\prod_{j=1}^{q}\prod_{\tau=1}^{2} P_{ij\tau}^{C_{ij\tau}} \qquad (11)$$

where τ is the index for Yea and Nay, $P_{ij\tau}$ is the probability of voting for choice τ, and $C_{ij\tau}$= 1 if the legislator's actual choice is τ and zero otherwise.

Note that $f(\mathbf{Y} \mid u_{ijy}-u_{ijn})$ is a ***discrete*** distribution with 2^{pq} possible outcomes. By definition, the ps+2qs+1 parameters are fixed constants and the $\mathbf{y}_{ij}$ are the random variables. Hence, we can apply standard probability theory to find the proportionality constant:

$$\sum_{y_{11}=0}^{1}\sum_{y_{12}=0}^{1}\cdots\sum_{y_{pq}=0}^{1}\left\{\prod_{i=1}^{p}\prod_{j=1}^{q} f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn})\right\} = K$$

So that

$$f(Y \mid u_{ijy} - u_{ijn}) = \frac{1}{K}\left\{\prod_{i=1}^{p}\prod_{j=1}^{q} f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn})\right\}$$

Fortunately, knowing the value of K is not important and does not affect the analysis of the likelihood function, $L(u_{ijy}-u_{ijn} \mid \mathbf{Y})$. Technically, the likelihood function has exactly the same expression as $f(\mathbf{Y} \mid u_{ijy}-u_{ijn})$. In this case the

1/K is missing. This has no effect because it is as if the true likelihood function were multiplied by K. When gradient methods are applied to the likelihood function all that matters is the relative heights of the function. In addition, when logs are taken the proportionality constant becomes an additive constant and plays no role in the estimation.

The joint p.d. of the sample, f(**Y** | u_{ijy}-u_{ijn}), is a ***discrete*** probability distribution with 2^{pq} possible outcomes. The likelihood function, L(u_{ijy}-u_{ijn} | **Y**), is a continuous distribution over the ps+2qs+1 dimensional hyperplane with dimensions corresponding to the ps legislator coordinates, the 2qs outcome coordinates, and β. Although equation (11) is not a probability distribution it is the case that it is above zero over the ps+2qs+1 hyperplane; that is:

$$\mathbf{L}(\, u_{ijy}-u_{ijn} \mid \mathbf{Y}\,) \geq 0, \quad 0 < \beta < +\infty,$$

$$-\infty < X_{ik}\ ,\ O_{jky}\ ,\ O_{jkn} < +\infty$$

Unfortunately the hypervolume underneath the function is ***not*** finite; that is:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{0}^{+\infty} L\left(\beta\left\{e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijky}^{2}\right)} - e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijkn}^{2}\right)}\right\}\middle|\mathrm{Y}\right) d\beta\, dX_{11}\ldots dX_{pk}\, dO_{11y}\ldots dO_{qky}\ldots dO_{11n}\ldots dO_{qkn} =$$

$$+\infty \qquad (12)$$

The likelihood function in equation (11) is the product of the pq probabilities of the observed choices. The value of the function is a maximum of 1 and a minimum of zero. Note that if all the legislators are voting correctly, that is, $\mathbf{P}_{ijc} > .5$ (or $u_{ijc} > u_{ijb}$) for all i and j where "c" means "correct choice" and "b" means "incorrect choice", then as $\beta \rightarrow +\infty$ clearly $\mathbf{L}(u_{ijc}-u_{ijb} \mid \mathbf{Y}) \rightarrow 1$. Conversely, if for at least one choice a legislator votes "incorrectly", $u_{ijc} < u_{ijb}$ so that $\mathbf{P}_{ijc} < .5$, then as $\beta \rightarrow +\infty$ clearly $\mathbf{L}(u_{ijc}-u_{ijb} \mid \mathbf{Y}) \rightarrow 0$. With voting error $\mathbf{L}(u_{ijy}-u_{ijn} \mid \mathbf{Y})$ asymptotes very quickly to zero because $\Phi(u_{ijc}-u_{ijb})$ goes to zero very quickly as β increases.

Now consider the effect of the legislator and outcome coordinates. Suppose $|X_{ik}| \rightarrow +\infty$ then a simple inspection of equation (7) shows that:

$$\Phi\left[\beta\left\{e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijkc}^2\right)} - e^{\left(-\frac{1}{2}\sum_{k=1}^{s} d_{ijkb}^2\right)}\right\}\right] \rightarrow \Phi\left[\beta\left\{e^{(-\infty)} - e^{(-\infty)}\right\}\right] = \Phi\left[\beta\{0\}\right] = .5$$

The portion of the likelihood function corresponding to legislator i is:

$$\prod_{j=1}^{q} \Phi(u_{ijc} - u_{ijb}) \qquad (13)$$

So that this converges to $.5^q$ as $|X_{ik}| \rightarrow +\infty$. This shows that

L(u_{ijy}-u_{ijn} **|** **Y**) does not asymptote to zero along the dimensions corresponding to legislator coordinates so that the hypervolume underneath **L**(u_{ijy}-u_{ijn} **|** **Y**) is infinite.

The fact that **L**(u_{ijy}-u_{ijn} **|** **Y**) has an infinite hypervolume has no practical effect on a standard maximum likelihood analysis. This is so because at a great distance from the center of the space defined by the legislator and outcome points the likelihood function is a flat, featureless vista. That is, the maxima are towards the interior of the function and are easily found by conventional gradient and quasi-gradient methods. However, the use of simulation methods is inappropriate because **L**(u_{ijy}-u_{ijn} **|** **Y**) cannot be treated as if it were a probability distribution.

The Bayesian approach avoids the problem of infinite volume through the judicious choice of prior distributions that when multiplied through the likelihood function produce a distribution that is proportional to a probability distribution. In a standard Bayesian approach the prior distribution for a legislator ideal point is a multivariate normal distribution with variance-covariance matrix $\sigma^2\mathbf{I_s}$:

$$\boldsymbol{\xi}(\mathbf{X_i}) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2}(X_{i1}^2+X_{i2}^2+\ldots+X_{is}^2)} \qquad (14)$$

Similarly, assume that the prior distributions for the outcome points are also multivariate normal distributions with variance-covariance matrices $\sigma^2\mathbf{I_s}$:

$$\xi(\mathrm{O}_{\mathrm{jy}}) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2}(O_{j1y}^2 + O_{j2y}^2 + \ldots + O_{jsy}^2)}$$

and (15)

$$\xi(\mathrm{O}_{\mathrm{jn}}) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2}(O_{j1n}^2 + O_{j2n}^2 + \ldots + O_{jsn}^2)}$$

The posterior distribution for the NOMINATE model is:

$$\boldsymbol{\xi}(\beta, \mathbf{O_y}, \mathbf{O_n}, \mathbf{X} \mid \mathbf{Y}) \propto \prod_{i=1}^{p}\prod_{j=1}^{q}\left\{ f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn})\xi(X_i)\xi(O_{jy})\xi(O_{jn})\right\} =$$

$$\prod_{i=1}^{p}\prod_{j=1}^{q}\left\{[\Phi(u_{ijy} - u_{ijn})]^{y_{ij}}[1-\Phi(u_{ijy} - u_{ijn})]^{(1-y_{ij})}\xi(X_i)\xi(O_{jy})\xi(O_{jn})\right\} =$$

$$\prod_{i=1}^{p}\prod_{j=1}^{q}\xi(X_i)\xi(O_{jy})\xi(O_{jn})\prod_{\tau=1}^{2}P_{ij\tau}^{C_{ij\tau}} \quad (16)$$

By definition

$\boldsymbol{\xi}(\beta, \mathbf{O_y}, \mathbf{O_n}, \mathbf{X} \mid \mathbf{Y}) \geq 0, \quad 0 < \beta < +\infty,$

$-\infty < X_{ik}, O_{jky}, O_{jkn} < +\infty$

and the hypervolume underneath $\boldsymbol{\xi}(\beta, \mathbf{O_y}, \mathbf{O_n}, \mathbf{X} \mid \mathbf{Y})$ is finite. Specifically, as the legislator ideal points and/or the outcome points go to $\pm\infty$ then $\boldsymbol{\xi}(\beta, \mathbf{O_y}, \mathbf{O_n}, \mathbf{X} \mid \mathbf{Y})$ goes to zero. For example, as $|X_{ik}| \rightarrow +\infty$ then $\boldsymbol{\xi}(\mathbf{X_i}) \rightarrow 0$ so that

$\boldsymbol{\xi}(\beta, \mathbf{O_y}, \mathbf{O_n}, \mathbf{X} \mid \mathbf{Y}) \rightarrow 0.$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{0}^{+\infty} \xi(\beta, \mathrm{O_y}, \mathrm{O_n}, \mathrm{X} \mid \mathrm{Y}) d\beta dX_{11}...dX_{pk} dO_{11y}...dO_{qky}...dO_{11n}...dO_{qkn} = K < +\infty$$

Theoretically, simulation methods can be applied to the posterior distribution expressed in equation (16) but not to the likelihood distribution expressed in equation (11). The irony here is that if "believable" standard errors are to be obtained for the basic model either the parametric bootstrap (L&P) or the simulation approach are the best methods to do so because inverting the information matrix is problematic. I say "ironic" because one could favor the use of (16) over (11) on these practical grounds rather than the philosophical grounds of the "Bayesians" (non-"Frequentists").