Date: September 27, 2004
RE: Model implementation within ORNIM
BY: Kenneth Train and Wesley Wilson

The report “Shippers’ Responses to Changes in Transportation Costs and Times: The Mid-American Grain Survey” by Train and Wilson (hereafter referred to as the Train/Wilson report) describes models of shippers’ choice of mode and quantity. The current memo describes how these models can be implemented within the structure of existing Army Corps Planning Models. The memo is a working draft whose purpose is to initiate discussion on implementation procedures. The suggested procedure is based on our understanding of the planning models. However, our understanding is somewhat limited, and we expect the implementation procedure to be revised to be as consistent as possible with the current programming structure and data inputs of the models.

The unit of observation in ORNIM is the ODC, where O is the originating pool, D destination pool and C is the commodity. For each ODC, the current planning models have data on the quantity of the commodity that is shipped annually from the origin pool to the destination pool (Typically, these data are from the Waterborne Commerce Statistics Center). This quantity is the outcome of choices by numerous shippers for numerous shipments. The shipments generally originate at various locations that need not be on the river. Those shipped from locations that are not on the river are first shipped overland to a port on the origin pool and then shipped by barge from the origin pool. To differentiate these concepts, we refer to the location from which a shipment is sent as the “origin location” and the pool on the river at which the shipment is put onto a barge as the “origin pool.” Similarly, for the destinations of the shipments.

The mode choice model in the Train/Wilson report predicts the probability that a shipment will be sent by a given mode, given the rates and times for the alternative modes that are available from the origin location to the destination location. This probability has the following meaning: Consider a set of shipments that face the same costs and times for each mode, such as barge and rail. The probability for a given mode is the share of these shipments that are predicted to be sent by that mode. The model therefore operates at the level of sets of shipments that face the same costs and times for each available mode from an origin location to a destination location. The quantity choice model in the Train/Wilson report operates at the same level, predicting the average reduction in quantity shipped annually in response increases in transit times and costs, where the “average” is over a set of shippers who face the same times and costs from an origin location to a destination location.

These models can be implemented in the Army Corps planning models by assuming that the annual quantity for each ODC in Army planning models is the outcome of shippers’ choices for a set of shipments from a “representative” origin location and to a “representative” destination location, such that the set of shipments within each ODC face the same costs and times. As we understand, this assumption is already being made within these models. That is, for each ODC, we understand that the planning models specify the cost of shipping the commodity overland as well as by river. These costs are assumed within the planning models to be the same for all shipments that utilize the river from the O pool to the D pool for commodity C, which implicitly assumes that all the shipments within each ODC face the same "representative" costs by river and by overland mode. The models in the Train/Wilson report can be implemented within Army Planning models by utilitizing this existing concept, i.e., that each ODC represents a set of shipments that are assumed to face the same costs and times by barge and overland modes.

The following data are required for each ODC. Some of these data already exist within the current structure. For the variables that do not currently exist within ORNIM, there are several options. In some cases, it may be possible to calculate the new variables from existing data within ORNIM, using reasonable assumptions. If this is not possible, then perhaps new data can be generated and added to ORNIM. A third option is to assume that the variables are not materially important to the analysis and omit them. This last option might apply most readily to variables 8 and 9 listed below than to the more critical time and cost variables.

For each ODC, a representative origin location and a representative destination location are determined. The time and cost variables are calculated on the basis of these origin and destination locations. The complete set of variables is:

1. The quantity shipped annually by barge for the ODC (that is, the quantity of commodity C shipped by barge from origin pool O to destination pool D.) Label this variable Q. Note that this is the quantity shipped from the origin location to the destination location using barge rather than the overland mode.
2. The cost for shipments by barge. This cost is the full cost of sending the shipment from the representative origin location to the representative destination location using barge from the O pool to the D pool, including the cost of transit by rail or truck from the origin location to a port on the O pool and the cost of transit by rail or truck to the destination location from a port on the D pool. Label this variable C_b, where the subscript "b" refers to barge.
3. The transit time for shipments by barge. This time is the full transit time from the origin location to the destination location using barge from the O pool to the D pool, the same as for cost. Label this variable T_b.
4. A dummy for whether rail is utilized to access barge from the origin location or to reach the destination location from barge. Label this variable R_b.
5. The cost for shipments by the best alternative overland modes. As for variables 2 and 3, this cost is from the origin location to the destination location. Label this variable C_o, where the subscript "o" refers to overland.
6. The transit time for shipments by the best overland modes. As for item 2-4, it is from the origin location to the destination location. Label this variable T_o.
7. A dummy for whether rail is utilized on the overland mode from the origin location to the destination location. Label this variable R_o.
8. The average or typical number of years that shippers of this commodity from the origin location to the destination location have been at their current location. Label this variable Y.

9. Transportation costs from the origin location to the destination location expressed as a share of the value of the commodity. Label this variable H.

These variables are calculated for current conditions. Changes in conditions are represented as changes in these variables. The purpose of ORNIM is to predict the impact of changes in transit costs and times on the quantity shipped annually on the river. Let superscript “1” denote the value of a variable at the current conditions and superscript “2” denote the value at the changed conditions.

The models in the Train/Wilson report can be implemented through the following steps to predict the impact of changes in transit costs and times. These steps are undertaken for each ODC, and the description of each step is to be interpreted as being for a given ODC.

1. Calculate the share of shipments from the origin location to the destination location that go by barge under current conditions. This share is calculated from the model in Table 8 of the Train/Wilson report. The procedure for calculating the predicted share from the model is given in appendix 1 to this memo. Label this share as S^1.
2. Calculate the share of shipments from the origin location to the destination location that go by barge under changed conditions. This calculation is the same as in step 1 except that the changed values of costs and times are used. Label this share S^2.
3. Calculate the share using barge under the changed conditions as a proportion of the share under the original consiitons: S^2/S^1.
4. Calculate the percent reduction in annual volumes as a result of increased transit costs. This reduction is calculated from the model in Table 12, assuming that C_b^2 is greater than C_b^1. The steps for implementing this model are given in appendix 2. Label the predicted reduction M, where the percent is given in decimal form (e.g., a 10% reduction is expressed as 0.10.) Note that the volume under changed conditions is (1-M) times the volume under the original conditions.
5. Calculate the percent reduction in annual volumes as a result of increased transit times. This reduction is calculated from the model in Table 13, assuming that T_b^2 is greater than T_b^1. Label this reduction L, expressed in decimal form.
6. Using the changes calculated in steps 3-5, calculate the predicted annual quantity for the ODC under the changed conditions as $Q^2=(S^2/S^1)*(1-M)*(1-L)*Q^1$.

Note that S^1 is the same for any scenario for changes in transit times and costs, since it represents the share under current conditions. This variable can therefore be calculated once, and the calculated value of S^1 can be retained as an input to ORNIM. Thereafter, when ORNIM is used to predict the impact of changes in conditions, the value of S^1 for each ODC is taken as input; only S^2, M, and L are calculated each time ORNIM is run for changes in transit times and costs.

In addition to predicting the change in quantities for each ODC, ORNIM calculates the change in surplus that results from a change in transti times and costs. There are are a variety of ways that changes in surplus can be calculated using the models from the Train/Wilson report.. The various options have implications for the computer time required to run the model. We need to discuss these options to arrive at the best procedure taking computer time into account.

Appendix 1: Procedure for Implementing the Mode Share Model of Table 8

The model contains random coefficients for cost and time. Let w^r, $r=1,\ldots,R$ be R independent draws from a standard normal distribution, which will be used in constructing draws of the cost coefficient. Let u^r, $r=1,\ldots,R$, be R independent draws from a standard normal distribution, which will be used in constructing draws of the time coefficient. We will provide the ORNIM team with "intelligently draw" values for w^r and u^r with $R=100$. The values can be held in a file as input to ORNIM and reused in each run.

The following steps are repeated for each *r,* for a total of *R* times.

1. Calculate the cost coefficient: $\beta_c^r = -\exp(1.1767 + 0.6329 * w^r)$. Note that 1.1767 and 0.6329 are the mean and standard deviation, respectively, of a normal deviate that, when exponentiated, constitutes a lognormal deviate with a median of 3.2436 and mean of 3.9629 as given in Table 8.
2. Calculate the time coefficient: $\beta_t^r = \exp(0.5846 + 0.3726 * u^r + 0.7972K)$, where K is a dummy that indicates that the commodity is not corn, wheat or soy. That is, $K=0$ is the commodity is corn, wheat or soy and $K=1$ otherwise. Note that this K is identified on the basis of the commodity in the ODC.
3. Calculate the representative utility of mode b: $V_b^r = \beta_c^r C_b + \beta_t^r T_b + 3.7036 R_b$ + 4.7048. Recall that R_b is a dummy that identifies whether rail is used as access to or egress from barge. The term 4.7048 always enters because barge is necessarily used for this mode. Note that this equation assumes that the origin and destination locations are the same for both modes, such that distance is the same and hence does not affect the relative representative utilities. (If a different destination location is specified for the overland mode than for the barge mode, then 3.3566 times the distance is added here for mode b and in step 4 for mode o.)
4. Calculate the representative utility of mode o: $V_o^r = \beta_c^r C_o + \beta_t^r T_o + 3.7036 R_o$. Recall that R_o identifies whether rail is used on the overland mode.
5. Calculate the share that is predicted for mode b under these draws of the cost and time coefficients: $S^r = \exp(V_b^r) / [\exp(V_b^r) + \exp(V_o^r)]$.

The predicted share for mode b is then calculated as the average of S^r over all r: $S = (1/R)\sum_r S^r$. The share can be calculate at any values for the cost and time by each mode. The share under current conditions, labeled S_1 in the body of the memo, is calculated using C_b^1,

T_b^1, C_o^1, and T_o^1. The share under changed conditions, labeled S_2, is calculated using C_b^2, T_b^2, C_o^2, and T_o^2.

Appendix 2: Procedure for Implementing the Quantity Reduction Models of Tables 12 and 13

Let ε^r, *r=1,…,R,* be *R* independent draws from a standard normal distribution. We will provide intelligently drawn values, which can held in an input file and reused each time ORNIM is run. The reduction in annual volumes from increased transit costs is calculated from Table 12 as follows for each value of *r*:

1. Calculate $y^r = 0.8813 * ((C_b^2 / C_b^1) - 1) + .7246 * H^1 - .00171 * Y + 0.0906 - .4933 + \varepsilon^r$. Recall that H^1 is the transportation costs as a share of product value (calculated with the current costs) and Y is the years at current location. Note that the term 0.0906 always enters since all of the shipments under consideration are by barge.
2. Censor y^r from above at 1 and from below at 0: $\tilde{y}^r = \max(0, \min(1, y^r))$.

The predicted average reduction is the average of $\tilde{y}^r$ over all *r*: $M = (1/R)\sum_r \tilde{y}^r$.

This M is the predicted reduction from an increase in costs. The reduction due to a increase in transit times, L, is calculated analogously, using the coefficients from Table 13 instead of those from Table 12 and using the ratio T_b^2 / T_b^1 instead of C_b^2 / C_b^1.