DRAM Residential Location and Land Use Model: 40 Years of Development and Application

Stephen H. Putman

Abstract The DRAM residential model was one of the earliest to be developed and applied, with work beginning in 1971 and with applications in planning agencies continuing to this day. It was developed with the expectation that it would be applied together with an employment model (EMPAL), and with both being linked to a suite of transportation models. This chapter describes the development path of DRAM as well as those of related issues of model calibration and links to other models. The author concludes with the argument that while continued theory development is essential for models such as these, their use as forecasting and policy analysis tools depends as much upon ease of implementation for agency users as it does on any improvement in model formulation.

1 Introduction

In the U.S. the DRAM and EMPAL models of household and employment location and land use, including their successor GIS based model systems METROPILUS and TELUM, are the most widely applied models of these phenomena ever to be developed. They have seen use for public agency forecasting and policy analysis purposes in nearly 30 different metropolitan regions, including eight out of the country's ten largest cities. The development of these models began in the 1970s and continues to the present, via continued interplay between theory and practice. It would be silly to claim that these were perfect models. They do produce reliable estimates of long term regional patterns and have the ability to give sensible responses to many, though obviously not all, policy inputs. Of course other models have been developed and applied, some of them, too, began as long ago as the 1970s. In addition, it is inevitable that other operational models will be developed in

S.H. Putman
Department of City and Regional Planning, University of Pennsylvania, Philadelphia, PA, USA

F. Pagliara et al. (eds.), *Residential Location Choice*, Advances in Spatial Science,
DOI 10.1007/978-3-642-12788-5_3,

years to come. Some of these alternative approaches, both developed, and now being developed, are also presented in this book. DRAM and EMPAL, versions of which are still in active agency use more than three decades after their initial development, helped open the way.

In the following pages I provide a description of DRAM and the procedures for estimation of its parameters. The discussion begins with a description of its derivation from the principles of household location surplus maximization. The use of location surplus as an output indicator of policy effects is also described. This is followed by a discussion of the LANCON submodel's procedures for calculating land consumption by locators.

Following this is a description of CALIB, a constrained gradient search procedure, which is used to estimate the equation coefficients in both DRAM and EMPAL. This procedure calculates maximum likelihood estimates of the equation coefficients, goodness-of-fit statistics, asymptotic t-tests of the coefficients' statistical significance, and point elasticities which provide sensitivity information.

This is followed by a section on the several types of constraint procedures which are incorporated in the model structures, and then by a discussion of inter-model linkages as well as linkages to transportation models. Finally I give some concluding thoughts on the underutilization of these models in agency practice, and a new model system I have developed to address that issue.

2 The Residential Location Model: DRAM

I first began work on these models about 40 years ago. After experimenting with several functional forms, a spatial interaction structure (SI) was selected. This form was just emerging at that time as a mathematically consistent structure derived from entropy maximizing principles. In adapting the model for actual application I added the notion of a multivariate attractiveness function and a multiparametric travel cost deterrence function. This idea fell beyond the then current thinking about SI models, and required the development of new methods of model calibration.

The SI approach did, at that time, provide a nicely structured view of location models, but it lacked an appropriate grounding in economic theories of activity location. Further, the use of the multivariate attractiveness term was clearly necessary for satisfactory model application, yet did not seem to have a satisfactory basis in the model derivations. It was not until several years later that a satisfactory theoretical underpinning was developed.

It is difficult to sort out which came first, as at that time there was a great flurry of work, and many informal paper drafts were in circulation. An important paper was published by Cochrane (1975), in which a "location surplus" notion was developed. The derivation begins with the assertion that the trips which provide the trip-maker with the greatest *net* benefit are the trips that are chosen. The observed trip distribution pattern is thus indicative of the overall probability of trips being chosen on that basis. The approach taken to the subsequent derivation of a singly constrained SI model involves assuming that the probability of a particular trip-maker taking a trip from zone A to some other zone B, is the probability that a trip to zone B offers a surplus, or *net* benefit, greater than that which could be had from a trip to any other zone. It is then hypothesized that the probability that a trip to zone B for any individual trip-maker is the optimal trip (i.e., the trip which will actually be taken) increases with the number of opportunities for trip satisfaction in zone B, and decreases with trip cost between zones A and B, since the net benefit is reduced by greater trip cost.

In proceeding through the derivation of the functional form of the model it is assumed that the number of zones is large, say 100 or more. It is assumed that the underlying probability distribution is approximately exponential in the upper tail. It is shown that it is not necessary to know the actual number of trip possibilities represented by any trip attracting activity, but rather it is only necessary to assume that the number of trip possibilities is proportional to some measure of attraction.

The surplus then becomes the difference between the *probabilistic* utility u, which is the *gross* benefit of taking the trip, and a *deterministic* cost c_{ij} of actually taking the trip. Thus the *net* benefit, or surplus, from taking the trip is $s_{ij} = u_{ij} - c_{ij}$, and the probability that the surplus will be of some particular value s, given all the preceding assumptions, is

$$\Phi_{ij}(s) = exp\left[-h A_j\, e^{-\lambda(s-m+c_{ij})}\right] \tag{1}$$

where $\Phi_{ij}(s)$ is the cumulative distribution function of the "location surplus" accruing from the *optimal* trip between zone i and zone j;

h, λ, m = constants;
A_j = a measure of the attractiveness of zone j.

With this, we maintain the assumption that the trip chosen by a trip-maker will be the trip that maximizes his/her personal surplus. The probability that the trip chosen will be a trip from origin i to a particular zone j is the probability that the maximum surplus offered by a trip terminating is j is greater than the maximum surplus offered by a trip terminating in any other zone.

Continuing through the derivation, Cochrane winds up, given O_i trips originating from zone i, with the expected number of trips from zone i to zone j being given by

$$T_{ij} = \frac{O_i\, A_j\, e^{-\lambda c_{ij}}}{\sum_j A_j\, e^{-\lambda c_{ij}}} \tag{2}$$

which is the usual equation of the singly constrained SI model. Further, the total surplus for all trips actually made is

$$S_T = \frac{1}{\lambda} \sum_i O_i \left[0.577 + ln\left(h\, e^{\lambda m} \sum_j A_j\, e^{-\lambda c_{ij}}\right)\right] \tag{3}$$

Thus beginning with rather innocuous assumptions regarding a utility maximizing basis for trip-making behaviour a robust economic underpinning can be developed for a SI model formulation. The relationship of this derivation, which is also called group-surplus maximization, to a derivation from random utility theory and probabilistic choice models is described in (Wilson et al. 1981).

The actual formulation of DRAM differs from (2) in that the attractiveness variable A_j is replaced by a multivariate formulation with a Cobb–Douglas functional form, i.e., a product form with each term being an independent variable raised to a parameter power, as in,

$$A_j = X_{1j}^{\alpha} X_{2j}^{\beta} X_{3j}^{\gamma} X_{4j}^{\delta} \tag{4}$$

where the X_{1j}, X_{2j}, X_{3j}, X_{4j} are attributes of zone j such as available land, household income, etc., and where α, β, γ, and δ are empirically estimated parameters.

Thus, in application DRAM is an aggregate form of a multinomial logit model of location choice. In computational form this yields the functional equivalent of a modified singly-constrained spatial interaction model. There are two modifications (1) a multivariate, multiparametric attractiveness function is used, (2) a consistent balanced constraint procedure is included in the model, allowing zone and/or locator specific constraints. The model is normally used for up to eight household categories, defined in terms of income, whose parameters are individually estimated (Putman 1983). The equation structure currently in use also makes provision for a lag term which adds stability to the model. This structure is given here:

$$N_i^n = \eta^n \sum_j Q_j^n B_j^n W_i^n c_{i,j}^{\alpha^n} \exp(\beta^n c_{i,j}) + (1.0 - \eta^n) N_{i,t-1}^T \tag{5}$$

where

$$Q_j^n = \sum_k a_{k,n} E_j^k \tag{6}$$

and

$$B_j^n = \left[\sum_i W_i^n c_{i,j}^{\alpha^n} \exp(\beta^n c_{i,j}) \right]^{-1} \tag{7}$$

and

$$W_i^n = \left(L_i^v\right)^{q^n} (x_i)^{r^n} \left(L_i^r\right)^{s^n} \prod_{n'} \left(1 + \frac{N_i^{n'}}{\sum_n N_i^n}\right)^{b_{n'}^n} \tag{8}$$

where

E_j^k = employment of type k (place of work) in zone j
N_i^n = households of type n residing in zone i
$N_{i,t-1}^T$ = total households residing in zone i at time t − 1
L_i^v = vacant developable land in zone i
x_i = 1.0 plus the percentage of developable land already developed in zone i
L_i^r = residential land in zone i
$a_{k,n}$ = regional coefficient of type n households per type k employee
$c_{i,j}$ = impedance (travel time or cost) between zones i and j
$\eta^n, \alpha^n, \beta^n, q^n, r^n, s^n, b_{n'}^n$ = empirically derived parameters

DRAM is also capable of including additional attractiveness variables in the spatial potential term, (8), of the model. There has been rather little use of this option in practice, as *the inclusion of such variables does require the subsequent development of a means for their updating* in forecast runs of the model. We have, for example, explored the merits of including residential land value as an additional attractiveness variable. We found, using rather reliable data, that the inclusion of land value in addition to household incomes made rather little contribution to the model's overall reliability, and brought with it the not insignificant prospect of having to develop an extension of the model to update the land value variable as a part of the long term forecasting procedure. In every case, when considering the addition of variables to this sort of model, one must weigh the possible improvements in model performance to be had from such additions against the cost of updating those new variables over a 30 year forecast horizon.

3 Location Surplus as an Output Measure from DRAM

Location surplus is a measure of the aggregate benefit households receive from the attributes of their chosen residential zone. Because household utility can only be measured on an ordinal scale (i.e., it is not possible to determine the monetary value of utility), the location surplus measures are interpreted as index numbers. The larger the value of location surplus, the more utility households receive from their choices of residential location.

The location surplus measures used in DRAM can be derived by using either of two different methods. Both methods produce the same location surplus measures and are based on the assumption that households attempt to maximize utility when choosing residential locations. For the first method, the DRAM model is interpreted as a multinomial logit model and the location surplus measure is found by calculating aggregate indirect utility (McFadden 1974; Ben-Akiva and Lerman 1985; Freeman 1993). In the second approach, the location surplus measure is found by directly integrating the DRAM travel demand function (Neuburger 1971; Cochrane 1975; Williams 1976).

Consider first the calculation of *location surplus from indirect utility*. In DRAM residential attractiveness is defined in (8). From this, indirect household utility is defined as:

$$V_{ij}^{n} = \ln\left[\left(c_{ij}\right)^{\alpha n} \exp\left(\beta^{n} c_{ij}\right) W_{j}^{n}\right] \tag{9}$$

where

V_{ij}^{n} = the indirect utility of a type n household that resides in zone i with a head-of-household employed in zone j,

c_{ij} = the travel time between zone i and zone j, and

α^{n}, β^{n} = empirically derived parameters.

This definition of indirect utility is for a single household. To find aggregate location surplus, it is necessary to sum the values of indirect utility for all type n households.

$$LS^{n} = \sum_{j}\sum_{k}\left(a_{kn}E_{j}^{k}\right)\ln\left[\sum_{i}\left(c_{ij}\right)^{\alpha n}\exp\left(\beta^{n} c_{ij}\right)W_{i}^{n}\right] \tag{10}$$

where

LS^{n} = the location surplus for type n households,

a_{kn} = regional coefficient of type n households per type k employee

E_{j}^{k} = employment of type k (place-of-work) in zone j.

4 Procedures for Calculation of Land Consumption: LANCON

In the combined model EMPAL and DRAM structure the use of land by locating activities is calculated after completing the calculation of total location demand. EMPAL calculates location demanded by employers, followed by DRAM's calculation of location demanded by households. LANCON takes both these calculated demands and estimates the actual change in the amount of land, by zone, that will be used by each of the demand categories. If there has been a decrease in demand by a particular demand category, then land currently in use by that category is released into a "pool" of land available for any use. If there has been an increase in demand by a particular demand category, then the addition of land to use by that category is calculated. After the calculations are done for each demand category, the sum of land used is adjusted, by an increase in densities, to match the land available for such uses.

The land used by each demand category is estimated in terms of the rate of land use by a locator in that specific demand category. The calculation, for example, of the rate of residential land use by new household locators in a specific zone is given by the following equation:

$$\frac{L_{i}^{r}}{N_{i}^{T}} = k_{0}\left(\frac{L_{i}^{d}}{L_{i}^{d}+L_{i}^{v}}\right)^{k_1}\left(\frac{L_{i}^{B}}{L_{i}}\right)^{k_2}\left(\frac{L_{i}^{C}}{L_{i}}\right)^{k_3}\left(\frac{N_{i}^{1}}{N_{i}^{T}}\right)^{k_4}\left(\frac{N_{i}^{4}}{N_{i}^{T}}\right)^{k_5}\left(L_{i}^{d}+L_{i}^{v}\right)^{k_6} \tag{11}$$

where

L_i^d = total developed land area of zone i;
L_i^v = vacant developable land in zone i;
L_i^B = "Basic" employment land in zone i;
L_i^C = "Commercial" employment land in zone i;
N_i^1 = Number of low income households in zone i;
N_i^4 = Number of high income households in zone i;
$k_0, k_1, k_2, k_3, k_4, k_5$ = empirically derived parameters.

5 Model Calibration

The calibration process involves "fitting" the equations of DRAM to the data which describe a particular region. When I first began to experiment with formulations of these models, I was unaware of the work done by others to calibrate spatial interaction models. Their efforts were devoted exclusively to making use of trip matrices (origin–destination trip data) for calibration of single parameter (usually the β in a travel function) spatial interaction model formulations. I assumed from the outset, that it would be necessary to have parameters not only in the travel function part of the model, but along with the attractiveness variables as well. Further, given the near complete unavailability of trip matrices for the cities in which I expected to do my initial model calibrations, it never would have occurred to me do use the procedures then in use by the spatial interaction modelers. Instead, I developed what we now call trip-end calibration as compared to the trip matrix based trip-interchange calibration. Beginning in about 1973 I became the scourge of various professional conferences as I pressed my colleagues to send me copies of their urban area data. Over the next 10 years I fit (statistically) the models to perhaps 40 urban areas while learning how best to do it (Putman 1977, 1980; Putman and Ducca 1978a, b). Somewhat later we were able to demonstrate that, for any specific region's dataset, the mean expected values of the parameters were identical for both trip-end and trip-interchange calibrations, though the variance, as would be expected, was somewhat higher for the trip-end procedure due to its having less information input (Putman and Kim 1984a, b).

To perform calibrations it is necessary to have one or more indicators of *Goodness-of-Fit* of the models to the data. The equation structure of DRAM is intrinsically nonlinear, and the data from which its parameters must be estimated are not normally distributed. Because of these factors it is not possible to use conventional regression techniques to calibrate DRAM. The procedure used for the estimation of parameters is gradient search. In effect, the partial derivatives of a goodness-of-fit criterion with respect to each specific parameter are calculated. The values of these derivatives determine the direction of parameter search (Putman 1983). The appropriate goodness-of-fit measure for calibration of DRAM is the likelihood function, a measure derived from the notion of maximum likelihood as developed in econometrics. This measure has the general form:

$$L = \sum_i N_i \ln \widehat{N}_i \tag{12}$$

where L is the computed likelihood measure, N_i is the observed value, and $\hat{N}i$ is the estimated value of the dependent variable. In DRAM the dependent variable would be households of a particular type located in a particular zone. It is important to note that in this equation form, the magnitude of L will be conditional on the magnitudes of the data being used. In a region with millions of households L will be larger than it will be in a region with hundreds of thousands of households. This means, unfortunately, that it is not possible to compare the results of analyses of different data sets, and thus not possible to evaluate the adequacy of one statistical analysis versus another.

The "Best Fit" is when the difference between the models' estimate of the dependent variable and the observed values in the calibration data set is as small as possible. A perfect fit would be obtained if, for each independent variable observation, i.e., locator type and zone, the estimated and the observed N_i were equal. This would give the following "Best Fit" value of likelihood:

$$L_b = \sum_i N_i \ln N_i \tag{13}$$

The "Worst Fit" would be when all values of the dependent variable are estimated by the mean of that variable. Thus, for example, if the region's total of Type 1 households were to be divided by the number of zones to get the mean of Type 1 households per zone, and all zones were assigned an amount of Type 1 households equal to the mean. This is also known as the uniform distribution assumption, where the estimated $\hat{N}i =$ the Zonal Mean $\bar{N}$, and gives the following "Worst Fit" value of likelihood:

$$L_w = \sum_i N_i \ln \bar{N} \tag{14}$$

From these two extreme values of likelihood we can, for a particular dataset, construct a relative measure of goodness-of-fit which is analogous to the more traditional R^2 measure, but which is appropriate to the nonlinear equations of DRAM and EMPAL, and to the non-normal distributions of the data. This measure of "Relative" goodness-of-fit is called a likelihood ratio, and takes the following equation form

$$\phi = \frac{L - L_w}{L_b - L_w} \tag{15}$$

The computed value of this Likelihood Ratio, ϕ, has a range such that for a perfect fit, $\phi = 1.00$, and for the worst fit, $\phi = 0.00$. Typical results obtained when fitting DRAM and EMPAL give $\phi = 0.80$–0.95. The values taken by ϕ are independent of the magnitude of the dependent variables and thus it is possible to compare the calibration results of one locator type to another, or from one region to another.

6 Asymptotic t-Statistics in DRAM Calibrations

In the estimation of nonlinear model parameters it is necessary to develop ways of assessing their statistical significance. The maximum likelihood estimator, when correctly calculated, is asymptotically normally distributed with its mean equal to the true parameter value, and with a covariance matrix which can be calculated by use of second order partial derivatives. These derivatives are calculated as part of the parameter estimation procedure, and allow the computation of asymptotic t-statistics which yield an indication of the statistical significance of the individual parameters in the models' equation structures.

7 Location Elasticities for DRAM and EMPAL

Location elasticities measure the sensitivity of household location to changes in the models' attractiveness variables. All of the location elasticities are defined for a single residential zone. For a 1% increase in an attractiveness variable in a specific zone, the location elasticity measures the resulting percentage change in the number of households in that zone. For example, suppose that for low-income households in zone 12 the DRAM location elasticity for residential land is equal to 0.2500. This means that a 1% increase in residential land in zone 12 will result in 0.25% increase in the numbers of low-income households in that zone.

The location elasticities are static measures of model sensitivity. This means that when a location elasticity is calculated for a specific attractiveness variable in a specific zone it is assumed that the values of all other attractiveness variables remain fixed. In the example above, the only variable that is allowed to change is the quantity of residential land in zone 12. All of the other attractiveness variables in zone 12 are assumed to be fixed, as are the attractiveness variables (including residential land) in all zones other than zone 12. Because they are static measures of model sensitivity, the location elasticities will change as the values of the DRAM attractiveness variables change (e.g., the location elasticities for forecast years will be different from the location elasticities for the base year).

The value of the location elasticity for a specific attractiveness variable and zone is a function of (1) the value of the calibrated parameter for the attractiveness variable, (2) the numbers of households or employees in the zone, (3) the magnitude of the attractiveness variable, and (4) the relative attractiveness of other zones in the region. Location elasticities will be larger when the calibrated parameter for the attractiveness variable is large (in absolute value), the number of households or employees is small (relative to other zones in the region), or the value of the attractiveness variable is small (relative to other zones in the region). For more a more detailed description of the derivation of location elasticities for residential location models see (Anas 1982) and (Anas and Chu 1984).

Except for travel time, all of the DRAM location elasticities have the same mathematical function definition. For the percentage of developable land developed and the household percentage variables, the location elasticities are defined for changes in one plus the value of the variable. (For example, if the percentage of developable land developed equals 66%, the DRAM attractiveness variable is equal to 1.66. A 1% increase in this variable is equal to 0.0166.)

Location elasticity for any attractiveness variable (shown for residential land) is:

$$\varepsilon^{n}_{L^{r}_{i}} = \frac{\partial N^{n}_{i}}{\partial L^{r}_{i}} \frac{L^{r}_{i}}{N^{n}_{i}} = \sum_{j} \left[\left(\sum_{k} a_{k,n} E^{k}_{j} \right) \left(\frac{s^{n}}{N^{n}_{i}} \right) \left(p^{n}_{i,j} \left(1 - p^{n}_{i,j}\right) \right) \right] \tag{16}$$

where

$\varepsilon^{n}_{L^{r}_{i}}$ = elasticity of type n households to changes in residential land in zone i,
$a_{k,n}$ = a matrix of conversion coefficients of type n households per type k employees,
E^{k}_{j} = employment of type k (place-of-work) in zone j,
s^{n} = the calibrated DRAM parameter for residential land,
L^{r}_{i} = residential land in zone i,
$p^{n}_{i,j}$ = the probability of a type n household, with an employed head-of-household in zone j, residing in zone i, and
N^{n}_{i} = households of type n residing in zone i.

For DRAM, the location elasticities for travel time are defined for a 1% increase in the travel time for trips from all employment zones to the specified residential zone. The equation for the location elasticity for travel time is as follows:

$$\varepsilon^{n}_{c_{j.}} = \frac{\partial N^{n}_{i}}{\partial c_{j.}} \frac{c_{j.}}{N^{n}_{i}} = \sum_{j} \left[\left(\sum_{k} a_{k,n} E^{k}_{j} \right) \left(\left(\frac{\alpha^{n}}{c_{ij}} \right) + \beta^{n} \right) \left(p^{n}_{i,j} \left(1 - p^{n}_{i,j}\right) \right) \left(\frac{c_{ij}}{N^{n}_{i}} \right) \right] \tag{17}$$

where

$\varepsilon^{n}_{c_{j.}}$ = elasticity of type n households to changes in travel times from all employment zones to residential zone i,
c_{ij} = travel time between zones i and j, and
α^{n}, β^{n} = the calibrated DRAM parameters for travel time.

The purpose of all this is to provide a means for assessing, without the need for innumerable model runs, the relative sensitivities of locators to the different independent variables in the model structure. This knowledge, in turn, provides a means for assessing the likely degree of impact of specific policy proposals on individual locator–zone combinations.

8 The Consistent Imposition of Constraints on Location

Once having calibrated DRAM and done preliminary tests of its forecasts, it is usually necessary to add additional user knowledge to the model structure. One way in which this may be done is by the imposition of constraints on location. It is of prime importance that when constraints are imposed on particular locators in particular zones, that this does not have the effect of swamping, or overwhelming, the forecasts of location by that locator type in other zones of the region being modeled. In DRAM this precaution is taken by use of a model which is a computational hybrid of "singly constrained" zones which are not constrained, and "doubly constrained" zones which are constrained. By this means various "out of the ordinary" locations can be represented, e.g., prohibiting residential location at or too near an airport, preventing decline of a locator type as a means for representation of policy incentives which it may otherwise not be possible to represent within the model structure, or forcing forecast numbers of residents on military bases to remain at exogenously specified levels. Judiciously applied, constraints inform the model structures of "anomalies" in reality.

Four types of constraint may be applied to DRAM forecast outputs. The first type of constraint, Type I, is an absolute constraint on the number of households of a specified type in a specified zone. The second type of constraint, Type II, is an absolute constraint on the total households of all types in a specified zone. When a Type II constraint is imposed on a zone, the procedure scales the unconstrained households of each type in the zone to sum to the constraint total. If one of the household types in the zone has already been constrained to a particular value, the program attempts to maintain that value while scaling the remaining household types in the zone. The third type of constraint, Type III, is a maximum value for a particular household type in a specific zone, and only operates when the forecast of that household type in that zone exceeds the maximum. The fourth type of constraint, Type IV, is a minimum value for a particular household type in a specific zone, and only operates when the forecast of that household type in that zone is less than the minimum. Again, the program tries to avoid violating previously specified constraints. If, for example Type 1 households have been scaled with a Type I constraint in zone 14, and then are scaled as a part of a Type II constraint in zone 39, the imposition of a Type III constraint on Type 1 households in zone 73 could cause some interactions when the regional control totals were being enforced. The program constraint procedures attempt to maintain consistency throughout the various possible interactions which can arise with different combinations of constraints. Note that is possible to impose both maximum and minimum constraints on a particular locator type in a particular zone thus, in effect, providing the ability to constrain the location of that activity to fall within a predefined range. Finally, we may also impose constraints in terms of densities, by first calculating the number of households that would result in a particular density in a zone, and then applying a constraint on the zones households to limit them to the previously calculated value.

9 Linked Transportation and Land Use Model Runs

In most cases DRAM is run in conjunction with other models. That is, it is linked "behind" EMPAL, which produces forecasts of employment location, for *all* employment types, which are then used by DRAM as an input to its forecasts of household location. An important facet of the linkage between these models is the employment-to-household conversion process which provides a direct means for exogenously forecast changes in regional employment mix to produce a change in the region's income distribution. There are other socioeconomic links as well, involving regional unemployment rates, household size, and employees-per-household, which provide a means for incorporating some of these important phenomena as integral and consistent components of the forecasting process.

In addition to the linkage between EMPAL and DRAM, they are typically both used together as components of an integrated transportation and land use model system. In the initial development of these models a prototype set of transportation modeling procedures were developed for system testing purposes. In agency application other procedures for travel demand, mode split, and trip assignment are regularly used. These are typically one or another of the proprietary software packages. Various applications have combined agency developed models of travel demand and mode split, with such commercial software packages for trip assignment. While some additional work was necessary to produce seamless links from EMPAL and DRAM with these packages, my practical experience is that whenever the agency actually does wish to see the connections made, they are a relatively straight forward matter. A convex combinations procedure is added to solve the combined systems for an equilibrium solution (Putman 1991). This is an equilibrium between transportation and land use, involving both employment and residence location and land use, with travel demands (as well as mode split where the agency has the capability) and trip volumes on transportation network links, and subsequent loaded, or congested, transportation network characteristics. The procedure is straightforward and usually requires only a few iterations to converge. In practice, some agencies have found that a partial equilibrium between transportation and land use will suffice to improve forecasting reliability. The partial equilibrium is measured in terms of mean absolute percentage change in activity locations or network travel characteristics from one system iteration to the next. These days computation time is often of no importance in either case, but still, it has been possible to achieve quite good results with, say three iterations rather than the four or five that might be run to reach equilibrium. It is worth noting for the record that research by Shen (1995) has proven the existence and uniqueness of the solutions obtained (a reassuring result, even though in practice it had previously been demonstrated that the solutions were computationally stable and unique).

10 Concluding Thoughts

The development of DRAM, and its companion models was originally undertaken with the intention of bringing the best practical technology into regular agency use. It has taken a great deal longer than I expected for this to happen. Today, even though there have been a substantial number of agency users, and even though more seem to become interested with each passing year, the use of statistically valid models for producing the land use inputs to travel modeling still is not regular practice amongst Metropolitan Planning Organizations (MPOs) in the U.S. In part this is a political matter. Land developers are often closely connected to politicians. Thus, while transportation modeling is used by virtually all MPOs, the statistically valid modeling of land use, even though required as input to the transportation models, is often pushed aside by political considerations. This, however, is not the sole reason for the underutilization of these methods. Other reasons derive from both theory and practice.

First, no matter how good or bad the theoretical underpinnings, what is important in practice is whether a model system is implementable and understandable. It helps if the underlying theory, albeit often incomplete, is comprehensible to the users. Often, however, model users in operating agencies are rather less concerned with the model's theoretical basis than they are with whether or not it can readily be calibrated and its outputs adjusted where necessary to meet agencies' political constraints. Only after a substantial user community evolves, do the model developers find it possible to address questions of model improvement. Even then, the disparities between different agencies' priorities for model developments and improvements can sometimes be insurmountable with limited budgets. The size of the user community in transportation modeling is probably two orders of magnitude greater than the size of the land use modeling user community. This makes a big difference in what can and should be done, and the ways in which it may happen.

Second, one of the apparent determinants of the success of a model application is the extent to which the work is being done by agency staff or by consultants. Rarely do agency staff have the training necessary to do this work, yet if agency staff are not intimately involved, there will be issues of credibility of the results within the agency. At the same time, agency staff who may be inexperienced in the use of these sorts of models have a difficult adjustment to make between spreadsheet programs and complex models of socioeconomic phenomena. The fact that it is not possible to, in effect, simply press a button and get a valid result means that new users will always have an initial period of frustration. Some simply quit at that point, with projects going unfinished and the blame being placed on the difficulty of use and/or inadequate reliability of results. Not too many years ago I was involved in a project for a city in Florida. The entire effort was plagued by serious data availability and reliability issues. It seemed that after considerable struggling, both the agency staff and their politically motivated external advisory committee had understood the data issues and the effects on forecasting and policy analysis which data problems might be expected to have. Then, at the very first public presentation which I saw of the modeling project results, the opening speaker (an agency staff member worried about intra-agency political conflicts) walked to the podium and began his presentation with the sentence: “It doesn’t work.”

After more than 30 years’ experience of straddling the model development – model application fence, I became convinced that major improvement in planning practice was not to be had from making the land use models better. After all, existing models such as DRAM are often able to account for 90% or more of the variation in household location patterns. Instead of trying to add a few percent more to the explained variation, it seemed that it would be more effective to try to decrease the vast number (the majority) of agencies that used no statistically sound models at all. This led to the development of a new model system, called METROPILUS, (Putman and Chan 2001) which contains DRAM and EMPAL, and other programs including calibration procedures, embedded in a GIS environment, operating behind an extensive graphical user interface (GUI). The principal aim in its development was to achieve a dramatic reduction in the difficulty ordinarily encountered in land use model application. This system was quite successful in allowing agency users to do their own model runs. The system was usually set up with consultant assistance. Preliminary calibrations were sometimes also done with consultant assistance. After the initial system had been set up and calibrated they were able to do this with minimal additional assistance from consultants. METROPILUS, incorporating DRAM, continues in use by several regional agencies to this time.

Even so, it still was not “user friendly” enough for some users. The world had become used to general spreadsheet and database programs, as well as more specialized applications such as income tax preparation programs, touch screen computers in convenience stores for ordering sandwiches, and to restaurant waiters or waitresses placing patron’s meal orders with touch screens. Thus we developed an even more user friendly system, named TELUM. With this, after the user enters a few preliminary numbers, a spreadsheet, dimensioned to the user’s region, is opened. The task of data collection is presented as a structured set of steps, with software assistance, towards populating this spreadsheet with the region’s data. Various consistency checks are performed automatically during the user’s progress towards assembling the required data. The program also checks the consistency of the links between the database and the geographical data by which it is embedded in the GIS. On the completion of the data assembly work, statistical analyses of the data are performed and evaluated by the model system, followed by the running of a fully automatic procedure which performs the calibration of the models. After this, the model software rearranges the data along with the statistical analyses, and notifies the user of the results, while at the same time completing the preparations necessary for making forecasts using the models. Then, literally at the press of a button, the system runs the forecasts for the user’s region.

We designed and constructed this new model system around a knowledge based systems approach (Pozoukidou 2005). Implementing this kind of automation for a complex land use modelling system had not been done before, and we had some work to do to understand the function and use of software wizards and other artificial intelligence notions. As such, the early version of this model system had several simplifications over the full form of METROPILUS, including some limits on numbers of analysis zones and numbers of locator categories. Also, certain optional procedures available in METROPILUS were temporarily disabled in TELUM. Even so, the new system, now operational, does a remarkable job of assisting agency staff, often inexperienced in location modelling work, in performing this rather complex planning/ analysis activity (Pozoukidou 2006, 2007).

We developed TELUM under contract to the New Jersey Institute of Technology (NJIT), who subsequently have overseen its deployment. Development of the TELUM model system, incorporating DRAM, was sponsored by the USDoT, and TELUM has since been distributed to every MPO in the US. The software and documentation are available from an NJIT website (NJIT 2009). In pre-release tests by regional agencies participating in "beta testing," staff from MPOs for several mid-sized regions were able to produce statistically reliable, replicable, forecasts for their agencies' use, without the need for consulting assistance. Since then several more agencies have downloaded the software and produced their own forecasts (Casper et al. 2009). In order to achieve this level of automation a modest reduction in the models' flexibility of use has been necessary, but otherwise these agencies have done, largely on their own, work that has, hitherto, always required outside consulting assistance and a major budgetary commitment. All told, a half dozen or so agencies have downloaded and applied the TELUM system in areas such as Colorado Springs, Little Rock, and Des Moines. The results have been mixed, some quite satisfactory to their users, some not so. A major issue seems to be the availability of appropriate data, and the agency staff's ability to manage data issues as they arise. Even so, agencies continue to show interest in at least attempting to use the model for their specific local forecasting and policy analysis needs. We shall see whether having thus greatly reduced the difficulty of use, land use models will be more commonly implemented as a part of planning agencies' forecasting and policy evaluation procedures. This question will be settled not, as some are wont to say, on whether one model provides a few percent better goodness-of-fit than does another, but more on a mix of internal agency politics, regional politics, and the extent to which the whole affair is run by the region's moneyed interests.

References

Anas A (1982) Residential location markets and urban transportation: economic theory, econometrics and public policy analysis. Academic Press, New York

Anas A, Chu C (1984) Discrete choice models and the housing price and travel to work elasticities of location demand. J Urban Econ 15:107–123

Ben-Akiva M, Lerman SR (1985) Discrete choice analysis: theory and application to travel demand. MIT, Cambridge
Casper CT et al (2009) Application of TELUM at Pikes Peak Area Council of Government Metropolitan Planning Organization: Lessons Learned, Paper 09-2697, presented at Transportation Research Board Annual Meeting, Washington, DC. http://www.trb.org/am/ip
Cochrane RA (1975) A possible economic basis for the gravity model. J Transp Econ Policy 9:34–49
Freeman AM (1993) The measurement of environmental and resource values: theory and method. Resources for the Future, Washington DC
McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York, pp 105–142
Neuburger H (1971) User benefit in the evaluation of transport and land use plans. J Transp Econ Policy 5:52–75
NJIT (2009) http://www.telus-national.org/documentationpaper/htm
Pozoukidou G (2005) Increased usability of urban and land use models. The role of knowledge based systems in facilitating land use forecasting to planning agencies, Ph.D. dissertation viewable at. http://repository.upenn.edu/dissertations/AAI3179791
Pozoukidou G (2006) Planning support systems' application bottlenecks, European Regional Science Association, ERSA conference papers, No. ersa06p769 http://www-sre.wu-wien.ac.at/ersa/ersaconfs/ersa06/papers/769.pdf
Pozoukidou G (2007) Facilitating land use forecasting in planning agencies. In: Kungolos A, Brebbia CA, Beriatos E (eds) Sustainable development III, vol 1. WIT, Southampton, pp 67–80
Putman S (1977) Calibrating a disaggregated residential model – DRAM. In: Batey P, Massey D (eds) Alternative frameworks for analysis, vol 7. Pion, London, pp 108–124
Putman S (1980) Calibrating urban residential location models 3: empirical results for non-U.S. cities. Environ Plann A 12:813–827
Putman S (1983) Integrated urban models: policy analysis of transportation and land use. Pion, London
Putman S (1991) Integrated urban models 2: new research and applications of optimization and dynamics. Pion, London
Putman S, Chan S-L (2001) The METROPILUS planning support system: urban models and GIS. In: Brail R, Klosterman R (eds) Planning support systems: integrating geographic information systems, models, and visualization tools. ESRI, Redlands, California, pp 99–128
Putman S, Ducca FW (1978a) Calibrating urban residential models 1: procedures and strategies. Environ Plann A 10:633–650
Putman S, Ducca FW (1978b) Calibrating urban residential models 2: empirical results. Environ Plann A 10:1001–1014
Putman S, Kim Y-S (1984a) Calibrating urban residential location models 4: effects of log-collinearity on model calibration and formulation. Environ Plann A 16:95–106
Putman S, Kim Y-S (1984b) Calibrating urban residential location models 5: a comparison of trip end and trip interchange calibration methods. Environ Plann A 16:1649–1664
Shen P-N (1995) Optimized network equilibrium models of combined travel and residential location choices. Ph.D. Dissertation, Department of City and Regional Planning, University of Pennsylvania, Philadelphia
Williams HCWL (1976) Travel demand models, duality relations and user benefit analysis. J Reg Sci 16:147–166
Wilson AG et al (1981) Optimization in Location and Transport Analysis. Wiley, Chichester, Sussex