

n DRAM Calibration

The DRAM model of household location has been successfully calibrated for more than forty metropolitan regions. These regions vary tremendously in geographic and demographic size. The largest is the Los Angeles metropolitan region, with a population in excess of 14.5 million. The smallest is the Colorado Springs metropolitan region, with a population under 400,000. The number of zones, or subareas, used in previous calibration efforts range from 772 for Los Angeles, to the 100 used here for Portland, Oregon.

The quality, or success, of DRAM calibrations is directly related to the quality of the land use and household data which are used as inputs to the calibration process. Errors or inconsistencies in either the land use or household data can result in calibrations with poor goodness-of-fit and unreliable parameter estimates. One of the purposes of this study is to examine the sensitivity of outcomes in the transportation and land use modeling process to the inputs to the components of the process. As such, the DRAM calibrations were examined for three different urban area data sets.

The relationships between data quality and the accuracy of DRAM calibrations are illustrated here by examining actual data sets and calibration results for three regions: Detroit, Kansas City and Portland. These three regions were chosen for comparison because they have comparable zone systems, but are different in terms of population growth and spatial structure.

The calibration results for the three regions were quite different. For Detroit, good calibration results were obtained for all household types. For Kansas City, calibration of the *standard* DRAM model was not successful. However, with the inclusion of an *additive* lagged household variable to the standard DRAM formulation, successful calibration results were achieved. For Portland, calibration was unsuccessful at the census tract level of geographic detail both for the standard DRAM model and for a model with an additive lagged household variable. However, successful calibrations were obtained when a *multiplicative* lagged household variable was included in the attractiveness function of the DRAM model equations. Successful calibrations were also obtained when the zone system for Portland was spatially aggregated. The primary reason for the differences in the results obtained from one city to another seems to be the quality of the data inputs to the calibration process.

As with the discussion of EMPAL in the previous section, it is useful to specify the mathematical structure of DRAM and to characterize the attributes of a successful calibration before describing the calibration experiments for the three regions.

The Mathematical Formulation of DRAM

Like EMPAL, DRAM is a modified version of the standard singly-constrained spatial interaction model. The two modifications are: 1) a multivariate, multiparametric attractiveness function is used; and 2) a consistent balanced constraint procedure is included in the model, allowing zone and/or sector specific constraints. In algebraic terms, this model may also be thought of as a transform of an aggregate logit model of residential choice. The model is described in more detail in Putman, S. *Integrated Urban Models* (1983, Pion Limited, London). The equation structure of DRAM is:

$$N_i^n = \sum_j Q_j^n B_j^n W_i^n c_{i,j}^{\alpha^n} \quad (1)$$

where

$$Q_j^n = \sum_k a_{k,n} E_j^k \quad (2)$$

and

$$B_j^n = \left[\sum_i W_i^n c_{i,j}^{\alpha^n} \right]^{-1} \quad (3)$$

and

$$W_i^n = (L_i^v)^{q^n} (x_i)^{r^n} (L_i^r)^{s^n} \prod_{n'} \left[\left(1 + N_i^{n'} / \sum_n N_i^n \right)^{b_{n'}} \right] \quad (4)$$

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

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

$\alpha^n, q^n, r^n, s^n, b_{n'}$ = empirically derived parameters

Since the standard implementation of the DRAM model is a cross-sectional model, all of the variables have the same time subscript (this will be one of the areas of investigation described later in this section).

The Calibration Procedure

Due to the nonlinear structure of DRAM, it is necessary to use a specialized parameter estimation procedure for its calibration. The goal of the calibration procedure is to develop estimates of the parameters of the DRAM model equations which best fit the general model structure to a specific data set. The calibration procedure is analogous to regression analysis but uses different mathematics.

In multiple regression analysis, the parameters of the regression equation are estimated by minimizing the sum of squared errors (i.e., the squares of the differences between the observed values and estimated values of the dependent variable.) For the DRAM model, the parameters of the model equations are estimated by maximizing the *likelihood criterion*. The likelihood criterion, ϕ , is a nonlinear function of the observed numbers and estimated numbers of households in each zone. It is analogous to the R^2 statistic used to measure goodness-of-fit in multiple regression analysis. The value of the likelihood criterion is always between zero and one, with higher values representing better goodness-of-fit. For a perfect calibration (i.e., the estimated numbers of households are exactly equal to the observed numbers of households in every zone), the likelihood criterion would be equal to one. It should be noted that it is *never* possible to obtain a perfect fit of socioeconomic data to a model. The behavior of individuals and of groups of humans is not perfectly predictable. Depending upon the phenomenon being analyzed, explanation of observed variation ranging anywhere from 10 or 20 percent on up may be considered “good”. In the sorts of models being described here, explanation of variance of 75 percent or more is generally considered to be acceptable. Here we apply a somewhat more stringent standard.

In multiple regression analysis, t-tests are used to evaluate the statistical significance of the estimated parameters. The statistical validity of these tests depends upon assumptions regarding the algebraic structure of the regression equation, and on assumptions regarding the shape of the distributions of the variables in the analysis. These assumptions do not hold for DRAM's nonlinear structure, or for the distributions of the variables, and so the usual t-tests cannot be used. However, since the parameters of the DRAM model are estimated by maximizing a likelihood criterion, *asymptotic t-tests* can be used in lieu of the usual t-tests. The asymptotic t-tests can be interpreted in a manner similar to that which is used for the usual t-tests, but their effectiveness cannot be guaranteed. Smaller values of the asymptotic t-tests indicate that a coefficient is not statistically significant. Larger values of the asymptotic t-tests **may** indicate that a coefficient is statistically significant, but do not guarantee significance. These tests approach the effectiveness of the standard t-test as the sample size becomes large. For most of the data sets being used in DRAM and EMPAL model calibrations, the numbers of observations are large enough to make the assumption that the asymptotic tests are equivalent to the normal tests.

For a successful calibration of the DRAM model equations, the likelihood criterion should exceed 0.80 for each household type, and the asymptotic t-tests should exceed 20.0 for a majority of the estimated parameters. If the likelihood criterion is less than 0.80, then a significant proportion of the variation in household location cannot be explained by the DRAM model equations. In most cases, low values of the likelihood criterion suggest that there are errors in the household and land use data used as inputs to the calibration process.⁶ A low value of a specific asymptotic t-test indicates that the variable associated with that statistic is, given the other variables in the equation, uncorrelated with household location. This absence of correlation may reflect the actual location behavior of households, the presence of high correlations amongst the variables in the equation, or may be due to errors in the measurement of the variable.

Another aspect of the quality of calibration results can be evaluated by comparing the signs and magnitudes of the estimated parameters with previous parameter estimations for other metropolitan regions. In general, we expect the values of the DRAM model parameters to conform to our theoretical and empirical knowledge of household location. The value of the travel time parameter, α , should always be negative, since employed household members are adverse to making long home-to-work trips, and we expect the probability of a worktrip being made to decline with the length of the possible trip.⁷ The values of the land use parameters should be positive, although in certain instances the parameter for vacant land, r , may be close to zero or negative. The household-to-household parameters measure the attraction or repulsion of households of a particular income group to high intrazonal concentrations of all other income groups. In general, households are most strongly attracted to similar (in terms of income) households. So, the household-to-household parameter for the same income group should usually be the largest (e.g., for low income households, the household-to-household parameter for low income households should be the largest.) The remaining household-to-household parameters should be smaller, or even negative, if the two income groups are very different (e.g., for low income households, the household-to-household parameter for high income households will usually be negative.)⁸

The DRAM Data Sets for Detroit, Kansas City and Portland

The Detroit, Kansas City and Portland regions are all quite different in terms of population, growth and spatial structure. In 1990, Detroit was the sixth largest metropolitan area in the United States, with a population of 4,665,236. The Kansas City and Portland metropolitan regions had populations of 1,566,280 and 1,477,895, respectively. However, of the three regions, Portland grew the fastest between 1980 and 1990, with a population growth rate of 13.9%. The population of Kansas City also increased significantly (9.3%), while the population of Detroit actually decreased (-1.8%). The aggregate land use for each of the metropolitan regions is summarized in the following table.

-
- ^{6.} Low values of the likelihood criterion may also be found if the spatial variation in household location is low. This problem is especially acute when census tract data is used in the calibration process.
- ^{7.} If a Tanner function is used as the trip function in the DRAM model equations, then we expect the α parameter to be positive, and the β parameter to be negative.
- ^{8.} An example of “good” DRAM calibration results is given in Table 21: DRAM Parameters for the Detroit Region.

Table 19. Aggregate Land Uses

	Total Land (Acres)	Residential Land (Acres)	Vacant Developable Land (Acres)
Detroit	3,169,048	579,592 (18.3%)	2,057,191 (64.9%)
Kansas City	1,271,673	179,826 (14.1%)	969,096 (76.2%)
Portland¹	2,382,789	133,019 (5.5%)	508,441 (21.3%)

^{1.} A large amount of land in the four counties containing the Portland region is national forest (approximately 130,000 acres). This land is not available for development.

In percentage terms, the amounts of vacant developable and residential land in Detroit and Kansas City are similar. For Portland, however, the percentages of residential land and vacant developable land are quite low.⁹

Since the measured quantity of residential land in Portland is relatively low, the average per zone residential density in Portland, as is shown in Table 20, is significantly higher than it is in Kansas City or Detroit.

Table 20. Average Residential Density per Zone

	Number of Zones in DRAM Data Set	Average Number of Households per Zone	Average Residential Density (HH/Acre)
Detroit	297	5908	3.02
Kansas City	377	1376	2.88
Portland	328	1688	6.14

Many of the zones in the Portland data set show no residential land, even though there are households present. For this reason, the amount of residential land in each zone was corrected by using the following procedure:

^{9.} In the Portland region, many households reside on agricultural or forest land. For this reason, the amount of residential land shown in the Table 19 is not completely representative of the amount of land used for residential purposes.

1. Compute the residential density for the zone. If the residential density is less than ten households per acre no correction is made. If the residential density is greater than ten households per acre go to the next step.
2. If the zone is urban no correction is made. If the zone is rural, the amount of residential land is corrected so that the residential density of the zone is equal to the average density of the surrounding zones. Newly created residential land is taken from agricultural land, so the total acreage of each zone remains constant.

The low percentage of vacant developable land in the Portland region data set is due to the existence of a legislatively defined Urban Growth Boundary. The Urban Growth Boundary prevents development beyond the urban fringe, effectively limiting the amount of vacant developable land. It would be a very substantial job to make the appropriate corrections to the data to attempt properly to represent the existence of the UGB. This task was beyond the scope of the current effort. As such, no corrections were made to the quantities of vacant developable land in each zone.

The limited amount of vacant developable land in the Portland region implies that the percentages of developable land developed in each zone will be higher than usual, and more uniform across zones. As we will see in the calibration results, the limits on developable land in the Portland region have a significant effect on estimated parameters for vacant land and the percentage of developable land developed. For Portland, vacant land and the extent of development are not strongly correlated with household location. (The magnitudes of the parameters for these two variables are small, and the asymptotic t-tests indicate that many of these parameters may not be statistically different from zero.) This contrasts with the calibration results for Detroit and Kansas City presented here, as well as with results for the many other urban areas for which DRAM model parameters have been estimated. For both the Detroit and Kansas City regions, the percentage of developable land developed is strongly correlated with household location. While for the Detroit region, the amount of vacant land is also strongly correlated with household location.

Calibration Results for the Standard DRAM Model

The calibration results for Detroit, Kansas City and Portland were quite different from one another. For Detroit (Table 21), good calibration results were obtained for all five household types. The likelihood criterion measures are all greater than 0.90, and only two parameters have asymptotic t-tests less than 20. Furthermore, the signs and magnitudes of most of the DRAM model parameters are comparable to parameter estimates for other metropolitan regions.

Table 21. DRAM Parameters for the Detroit Region¹
(297-Zone System)

	Low_Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	-0.6472 (94.2)	-0.8555 (123.0)	-0.9243 (149.8)	-0.9602 (156.0)	-0.9666 (143.8)
Vacant Land	0.0696 (74.2)	0.0669 (73.2)	0.0758 (82.6)	0.0520 (55.9)	0.0228 (23.2)
% Developed	2.8160 (264.5)	2.4294 (249.4)	2.3667 (254.5)	2.1350 (231.7)	1.9343 (196.7)
Residential Land	0.6768 (336.9)	0.7619 (332.3)	0.7760 (311.9)	0.8124 (309.9)	0.8367 (313.7)
LIHH	3.8114 (270.6)	1.0924 (63.9)	-0.0659 (3.3)	-0.4208 (19.3)	-0.7934 (34.6)
LMIHH	1.2372 (28.7)	5.2740 (130.7)	2.1610 (54.7)	0.6044 (15.9)	0.7598 (22.4)
MIHH	0.5191 (12.4)	0.3750 (7.73)	3.5443 (68.9)	1.0270 (20.3)	-0.7812 (19.0)
UMIHH	-1.0677 (37.4)	0.8481 (26.2)	1.9204 (54.8)	5.9217 (161.0)	4.1302 (110.9)
UIHH	-2.9773 (151.7)	-2.0737 (108.2)	-2.1550 (112.1)	-1.8209 (99.1)	1.9092 (131.8)
R²	0.8888	0.8907	0.9006	0.9078	0.8893
Likelihood Criterion - ϕ	0.9105	0.9013	0.9088	0.9201	0.9254
Avg. Trip Length (minutes)	34.36	33.84	34.19	34.43	34.01
MAPE	35.61%	30.76%	30.87%	29.21%	37.39%
MAPE (Smallest)	72.02%	60.94%	64.54%	55.03%	79.77%
MAPE (Largest)	17.01%	14.35%	14.07%	15.10%	19.04%

¹ Asymptotic t-tests in parentheses.

For Kansas City (Table 22), calibration of the *standard* DRAM model was not successful. The likelihood criterion measures were less than 0.80 for all four household types, and nine parameters had asymptotic t-tests that are less than 20.0. The signs and magnitudes of the land use and household-to-household parameters are as expected, but the estimated travel time parameter for low income households is positive. In fact, based on the asymptotic t-tests, for the three lowest income household groups, the travel time parameters may not be statistically different from zero.

Table 22. DRAM Parameters for the Kansas City Region¹
(377-Zone System)

Variable	Low Income	Low/Middle Income	Upper/Middle Income	Upper Income
Travel Time - α	0.04 (3.0)	-0.23 (18.3)	-0.10 (8.0)	-0.45 (31.6)
Vacant Acres	0.02 (19.9)	0.02 (18.7)	0.02 (27.6)	0.04 (40.9)
% Developed Acres	1.26 (83.6)	0.89 (64.0)	0.89 (67.3)	0.72 (57.8)
Residential Acres	0.55 (182.4)	0.58 (174.9)	0.51 (143.6)	0.55 (144.7)
% LI Households	3.71 (174.4)	0.46 (18.8)	-0.51 (17.7)	-2.33 (74.1)
% LMI Households	0.96 (16.2)	3.15 (63.4)	1.97 (35.9)	0.49 (10.4)
% UMI Households	0.93 (22.4)	1.33 (28.4)	4.96 (91.4)	1.16 (23.0)
% UI Households	-1.16 (43.0)	-0.61 (23.7)	0.53 (20.0)	3.15 (143.4)
R ²	0.53	0.54	0.65	0.70
Likelihood Criterion - ϕ	0.56	0.55	0.70	0.79
Average Trip Length (minutes)	26.6	27.1	28.4	27.8
MAPE	143.69%	83.01%	70.91%	82.10%
MAPE (Smallest)	510.72%	266.03%	209.57%	241.31%
MAPE (Largest)	26.61%	25.46%	21.83%	26.40%

¹ Asymptotic t-tests in parentheses.

For Portland (Table 23), after land use data corrections were made, all of the estimated travel time parameters were still positive, although two of these parameters may not be statistically different from zero. The signs of the land use parameters are acceptable, but the magnitudes of the parameters for vacant developable land and the percentage of developable land developed are significantly smaller than the estimates for the Detroit and Kansas City data sets. Acceptable goodness-of-fit is obtained only for upper income households, for which the likelihood criterion equals 0.8111. For all of the remaining household types, the likelihood criterion measures are less than 0.80.

Table 23. DRAM Parameters for the Portland Region¹
(328-Zone System - Corrected Land Use Data)

	Low Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	0.2508 (24.4)	0.1727 (14.7)	0.1715 (14.3)	0.3331 (29.1)	0.2857 (26.8)
Vacant Land	-0.0302 (22.5)	-0.0064 (4.5)	0.0047 (3.2)	-0.0026 (1.9)	0.0259 (19.9)
% Developed	0.0434 (2.2)	0.1545 (7.5)	0.2761 (13.4)	0.2783 (14.8)	0.4984 (29.8)
Residential Land	0.4428 (122.7)	0.4753 (103.2)	0.4810 (94.9)	0.5262 (107.2)	0.4838 (105.6)
LIHH	3.6931 (134.4)	0.6245 (17.3)	0.4689 (11.7)	-0.1798 (4.6)	-1.6935 (44.7)
LMIHH	2.1659 (31.6)	8.2862 (115.0)	2.3792 (32.3)	2.0654 (30.5)	1.4939 (26.2)
MIHH	-0.8297 (10.5)	-0.2511 (2.7)	6.3458 (64.8)	0.5706 (6.5)	0.0012 (0.2)
UMIHH	-1.7073 (28.7)	-0.6726 (9.2)	-0.5645 (7.2)	3.4987 (50.1)	-1.0901 (15.8)
UIHH	-1.2336 (40.5)	0.1435 (4.4)	0.3642 (10.9)	0.2074 (6.7)	3.5546 (144.8)
R²	0.5781	0.6601	0.6507	0.6859	0.7717
Likelihood Criterion - ϕ	0.6637	0.6747	0.6686	0.7035	0.8111
Avg. Trip Length (minutes)	24.32	25.21	26.05	27.25	26.58
MAPE	34.97%	31.06%	30.51%	34.12%	40.19%
MAPE (Smallest)	70.25%	62.78%	64.05%	75.84%	89.06%
MAPE (Largest)	23.08%	21.40%	18.71%	18.22%	21.35%

¹ Asymptotic t-tests in parentheses.

Table 24 shows, for the purpose of comparison, the initial calibration results for the Portland region which were obtained when uncorrected residential land data was used. These results are similar to

those shown in Table 23. However, the corrections to the residential land data did improve the goodness-of-fit of the DRAM model equations, and produced larger estimates for the residential land parameters.

Table 24. DRAM Parameters for the Portland Region¹
(328-Zone System - Uncorrected Land Use Data)

	Low Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	0.8094 (81.1)	0.7375 (65.1)	0.7239 (61.9)	0.9350 (83.0)	0.8639 (81.2)
Vacant Land	-0.0897 (55.5)	-0.0515 (30.5)	-0.0398 (23.5)	-0.0427 (26.5)	-0.0137 (9.0)
% Developed	0.0354 (1.4)	0.0845 (3.3)	0.2031 (8.3)	0.2703 (11.9)	0.4427 (21.4)
Residential Land	0.2351 (105.6)	0.2066 (83.3)	0.1778 (75.8)	0.1922 (88.4)	0.1566 (79.1)
LIHH	2.8875 (104.7)	0.1094 (3.1)	-0.4161 (10.6)	-1.0261 (26.8)	-2.6138 (71.1)
LMIHH	1.3737 (20.3)	7.6789 (107.9)	2.3742 (32.9)	1.8369 (27.8)	1.2698 (22.7)
MIHH	-0.7621 (9.6)	0.1871 (2.0)	5.8616 (60.7)	1.5680 (18.2)	-0.2213 (3.1)
UMIHH	-1.7253 (29.3)	-0.1553 (2.2)	0.1099 (1.4)	3.0458 (45.7)	-1.2888 (19.4)
UIHH	-1.2973 (42.7)	0.2893 (8.9)	0.4944 (14.9)	0.6186 (20.2)	3.5971 (147.0)
R²	0.5841	0.5653	0.4942	0.5277	0.6863
Likelihood Criterion - ϕ	0.6177	0.5706	0.5306	0.5685	0.7360
Avg. Trip Length (minutes)	25.42	26.48	27.32	28.53	27.80
MAPE	41.59%	39.44%	38.90%	47.70%	55.55%
MAPE (Smallest)	89.91%	88.10%	86.02%	106.81%	133.66%
MAPE (Largest)	23.38%	23.95%	23.74%	24.38%	26.67%

¹. Asymptotic t-tests in parentheses.

One reason for the poor calibration results in Kansas City and Portland is that the data sets for these regions are spatially disaggregated by census tract. All previous work with EMPAL and DRAM

has been done with a zone system where the zones are aggregations of census tracts. The consequence of this customary aggregation is to smooth over tract-to-tract variation and thus obtain better fits of model to data. The zones used to divide the Detroit data set are aggregations of census tracts. We will discuss this matter further in the next section.

Spatial Variation: Using Census Tract Data

The DRAM calibration procedure, in fact any statistical estimation procedure, will produce better goodness-of-fit and parameter estimates when the independent variables of the model are strongly correlated with the dependent variable (i.e., in DRAM, the numbers of households in each zone). A necessary condition for strong correlation between the independent and dependent variables is variation in both types of variables. If the dependent variable is constant, or varies very little from zone-to-zone, then it will not be correlated with, i.e., will not vary along with, any of the independent variables. (Mathematically, the correlation coefficient is undefined if the standard deviation of the dependent variable is zero.) As a consequence, adequate zone-to-zone variation in the numbers of households is a necessary (but not sufficient) condition for finding statistically significant parameters that maximize the goodness-of-fit of the DRAM model to any particular urban area data set.

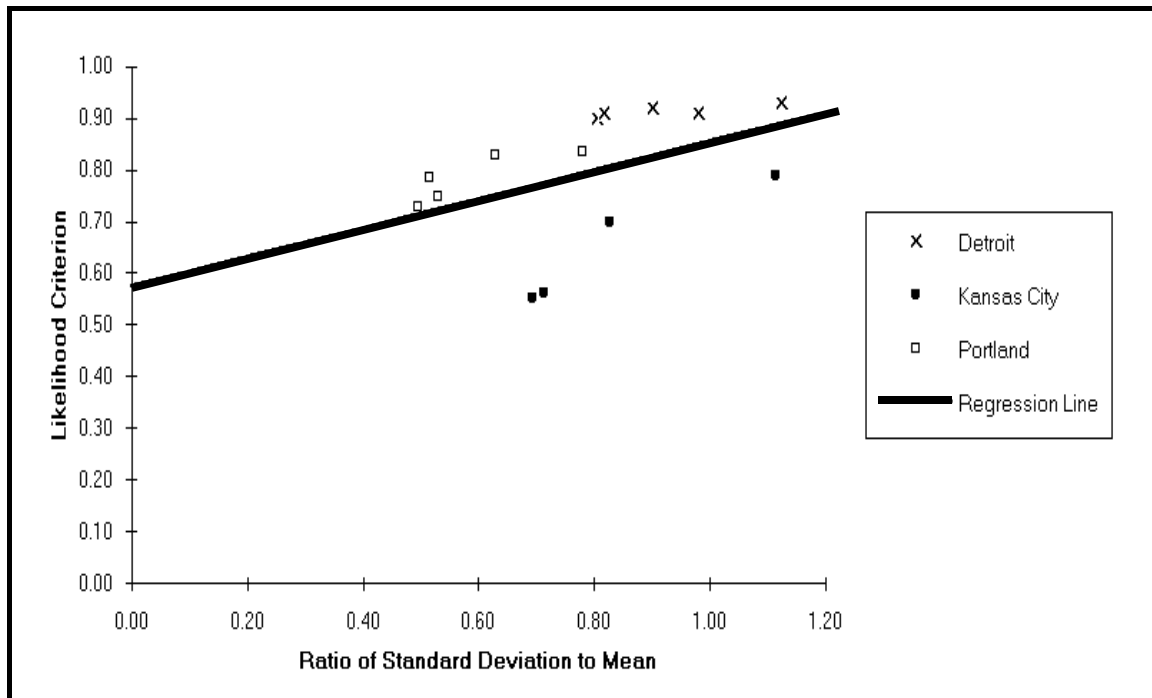
In general, the likelihood of finding parameters that maximize goodness-of-fit *increases* as the variation in the number of households in each zone *increases*. This is an important consideration when a metropolitan region is partitioned into zones and data is collected for the number of households in each zone. If the standard deviations of the numbers of households in each zone are low, then the calibrated DRAM model may not be capable of producing accurate estimates of household location (i.e., the goodness-of-fit may be low).

Because the zone systems for Kansas City and Portland are defined in terms of individual census tracts, the variation in the numbers of households per zone in these two regions is relatively low. Census tracts are defined by the Bureau of the Census so that there are approximately 4,000 people residing within each tract. This definition minimizes the variance of population across census tracts, and, as a consequence, reduces the standard deviation of the number of households in each tract (when compared with other definitions of residential zones). Therefore, using data defined at the census tract level of geographic detail for calibration of the *standard* DRAM model may result in reduced goodness-of-fit. Table 25 summarizes these statistics for the three regions.

Table 25. Standard Deviation of Households per Zone

Household Type	Detroit	Kansas City	Portland
Low Income	1160.98	243.91	218.94
Low-Middle Income	949.73	238.04	159.30
Middle Income	967.14	n.a.	146.37
Upper-Middle Income	1064.90	283.84	175.42
Upper Income	1326.96	386.41	313.97

The spatial variation/goodness-of-fit phenomenon described above is more easily seen in the graph below, which compares, for each of the three regions, the standard deviations of the numbers of households in each zone with the likelihood criterion measure obtained in calibration of DRAM for each household type. As can be seen in Figure 6, there is a

Figure 6. Standard Deviation of Households vs. Likelihood Criterion

positive relationship between the variation in household location (households per zone) and the goodness-of-fit of the calibrated DRAM model. This relationship is not a perfect one, but the general tendency is quite clear. The household data for Detroit has the greatest zone-to-zone spatial variation. In terms of goodness-of-fit, the calibration results for Detroit are better than the results

for either Kansas City or Portland. The household data for Portland has the least zone-to-zone spatial variation. However, the likelihood criterion measures for Portland are higher than the likelihood criterion measures for Kansas City. This may be due to differences in the spatial structure and average number of households per zone in Kansas City and Portland, as well as to other data set differences which were not identifiable. Within the two region's household types, it is still true that there is a positive relationship between likelihood criterion and the ratio of the standard deviation to the mean. In general, it seems that it is more difficult to obtain good calibration results for metropolitan regions with large rural populations and/or small numbers of household in each zone.

Extensions to the Standard DRAM Formulation

The original specification of the DRAM equation structure was couched in terms of changes in household location from one time period to another.¹⁰ Unfortunately, it was not long thereafter that it became clear that a dynamic model structure, much to be preferred from a theoretical point of view, was completely impractical when it came to actual agency application of the model. No agency had, at that time, the requisite time series of data for employment and household location, land use, and travel times or costs. Without these data resources, it was impossible to estimate the parameters of a dynamic model structure. As it was a principal goal of the effort to provide a *practical* model for agency use, the formulation was revised to a cross-sectional form. This form became the standard formulation of DRAM as was shown in equations 1-4.

There are several difficulties with a cross sectional formulation for DRAM. One of these is that some variables when examined cross-sectionally have exactly the opposite relationship to the dependent variable that they have when examined longitudinally. Take for example a vacant land variable. In a cross sectional analysis, one finds that the largest (geographically) zones in most regions are found at the outer edges of the region. This is also where one finds, generally speaking, the least amount of locating activities. From these two facts, one gets a negative relationship between a zone's quantity of vacant land and the number of households located in the zone. Yet, our expectation is that vacant land, all other things being equal, should be a positive factor in a zone's attractiveness to residential location. Another instance of this sort of problem can be seen in the fact that if a zone has 500 low income households, a cross-sectional formulation provides no information on whether these 500 low income households are the remnants of a prior population of 2000, or a substantial increase over a prior population of 100.

In 1994, most agencies were still unable to provide reliable time series data on *all* the variables necessary for a fully dynamic formulation for DRAM. Yet, with the current availability of the 1990 census results, and with the increased data resources which can be assembled by use of GIS, we begin to have some data on, at least, prior time periods' household location. Given this, several experiments were undertaken as part of this project in an attempt to remedy some of the inherent difficulties of the standard cross-sectional formulation of DRAM. These experiments are described in the next sections of this report.

¹⁰ Putman, S. "Calibrating a Disaggregated Residential Model - DRAM", *Alternative Frameworks for Analysis*, Batey & Massey, eds. Pion, London 1977.

Including an Additive Lag Term in the DRAM Model

The most promising approach for improving the DRAM formulation seemed to be in adding a lagged household variable. There are several ways in which this can be accomplished. The precise manner in which a lag term can be added depends upon what quality of lagged household data are expected to be available for calibration. If the lagged household data are only available, or only reliable, for zonal household totals, then the lag term can only be incorporated as an additional variable in the model's zonal attractiveness term. If the lagged data are reliably available by household type in each zone, then it is possible to include the variable as *either* a term in the zonal attractiveness, or as a separate additive lag term. The advantage of incorporating an additive lag term is that the resulting model formulation is quasi-dynamic in form. To include the additive lag term, we modify the original equation (1) as follows:

$$N_{i,t}^n = \lambda \sum_j Q_j^n B_j^n W_i^n c_{i,j}^{\alpha^n} + (1 - \lambda)N_{i,t-1}^n \quad (5)$$

where

$N_{i,t}^n$ = households of type n residing in zone i at time t

$N_{i,t-1}^n$ = households of type n residing in zone i at time $(t-1)$

and

λ = an empirically derived parameter

then the standard DRAM model is converted from a purely cross-sectional model into a hybrid cross-sectional/quasi-dynamic model. In fact, if we rearrange the terms of equation (5), this modification of the DRAM model produces a “change” model of household location.

$$\Delta N_i^n = N_{i,t}^n - N_{i,t-1}^n = \lambda \left[\sum_j Q_j^n B_j^n W_i^n c_{i,j}^{\alpha^n} - N_{i,t-1}^n \right] \quad (6)$$

Adding a lagged household variable to the DRAM model structure should always improve calibration results. However, as we will see, the degree of improvement in calibration results depends on the quality of the lagged household data.

Calibration Results for a DRAM Model with an Additive Lag Term

When an additive lag term is included in the DRAM model equations, the likelihood criterion measures increase for all household types in all three regions. In terms of goodness-of-fit, the calibration results for Detroit (Table 27) are nearly “perfect”, since all of the likelihood criterion measures are close to one. For Kansas City (Table 28), the inclusion of an additive lag term dramatically

improves goodness-of-fit, and produces negative estimates for all of the travel time parameters. Unfortunately, similar improvements in the calibration results for the Portland data set do not occur (Table 29). For three of the five household types in the Portland region, the likelihood criterion measures remain below 0.80 when an additive lag term is included in the DRAM model structure. Furthermore, although the magnitudes of the travel time parameters are reduced, they remain positive, and thus counter intuitive, for all household income groups.

After a good deal of rather careful checking, the lack of improvement in the calibration results for Portland were shown to stem from inconsistencies between the base year (1990) and lag year (1985) household data.

Inconsistencies between the Base Year and Lag Year Household Data

Tables 27, 28 and 29 demonstrate the improvements in goodness-of-fit that result when an *additive* lag term is included in the DRAM model structure. Note, however, the improvements in goodness-of-fit are not at all equal for all three regions. Table 26, below, shows the with and without goodness-of-fit results for the three regions. The DRAM calibrations for Kansas City show the greatest increase in goodness-of-fit as a result of the inclusion of an additive lag term. The improvement in goodness-of-fit for Detroit is small, but there was not much room for improvement, as the likelihood ratio was already 0.90 or greater for all five household types. The goodness-of-fit for Portland shows significant improvement too, but not by as much as was obtained for Kansas City. It is in the examination of why there are these different degrees of improvement, that we discover the problems with the lagged household data for Portland. The differences between the calibration results for Kansas City and Portland are due to inconsistencies between the base year and lag year household data for Portland.

Table 26. ϕ - Goodness-of-Fit (Likelihood Criterion) for DRAM Model

Household Type	Detroit		Kansas City		Portland	
	Standard	w/ Lag Term	Standard	w/ Lag Term	Standard	w/ Lag Term
Low Income	0.91	0.99	0.56	0.95	0.66	0.83
Low-Middle Income	0.90	0.99	0.55	0.94	0.67	0.78
Middle Income	0.91	0.99	n.a.	n.a.	0.67	0.75
Upper-Middle Income	0.92	0.99	0.70	0.95	0.70	0.76
Upper Income	0.93	0.97	0.79	0.96	0.81	0.86

Table 27. DRAM Parameters for the Detroit Region¹
(297 Zones, Lagged Households Included in Additive Term)

	Low_Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income

Table 27. DRAM Parameters for the Detroit Region¹
(297 Zones, Lagged Households Included in Additive Term)

Alpha - α	-0.5879 (32.9)	-0.9031 (63.9)	-0.9962 (76.5)	-1.0287 (82.3)	-1.0706 (97.8)
Vacant Land	0.7775 (166.7)	0.2393 (92.0)	0.2682 (104.7)	0.2688 (105.5)	0.1361 (72.2)
% Developed	2.7830 (87.5)	2.4904 (108.4)	2.4007 (112.7)	2.1872 (108.8)	2.0175 (122.0)
Residential Land	1.1609 (115.6)	1.0798 (145.6)	1.1498 (156.7)	1.2479 (175.4)	0.9775 (194.8)
LIHH	3.7243 (33.9)	1.0346 (16.4)	-0.1580 (2.2)	-0.5477 (7.22)	-0.9573 (19.9)
LMIHH	1.2245 (9.3)	5.2300 (51.4)	2.1228 (22.0)	0.5205 (5.9)	0.6909 (11.8)
MIHH	0.5402 (2.3)	0.3889 (2.9)	3.5514 (28.9)	1.0084 (9.5)	-0.7828 (11.8)
UMIHH	-1.0293 (7.2)	0.8590 (8.1)	1.9617 (18.5)	5.9825 (60.6)	4.1931 (63.5)
UIHH	-2.9274 (40.0)	-2.0178 (42.7)	-2.0738 (45.8)	-1.6645 (41.0)	2.1036 (85.5)
Lambda - λ	0.0207 (23.6)	0.0996 (35.1)	0.1225 (45.5)	0.1519 (60.67)	0.3119 (85.5)
R²	0.9870	0.9907	0.9887	0.9863	0.9533
Likelihood Criterion - ϕ	0.9929	0.9919	0.9888	0.9867	0.9678
Avg. Trip Length (minutes)	48.47	36.33	36.63	36.99	39.28
MAPE	6.35%	6.67%	7.94%	9.19%	17.28%
MAPE (Smallest)	7.13%	9.35%	12.26%	16.55%	30.34%
MAPE (Largest)	5.69%	3.64%	4.00%	5.01%	9.51%

¹ Asymptotic t-tests in parentheses.

Table 28. DRAM Parameters for the Kansas City Region¹
 (377 Zones, Lagged Households Included in Additive Term)

Variable	Low Income	Low/Middle Income	Upper/Middle Income	Upper Income
Travel Time - α	-1.20 (41.9)	-1.65 (63.1)	-1.65 (62.9)	-2.25 (95.6)
Vacant Acres	0.14 (62.0)	0.18 (72.9)	0.24 (83.7)	0.29 (96.3)
% Developed Acres	1.42 (47.4)	0.80 (28.5)	0.20 (6.8)	-0.94 (34.3)
Residential Acres	0.78 (103.1)	0.62 (73.9)	0.42 (52.8)	0.30 (48.6)
% LI Households	3.12 (53.9)	-1.56 (18.8)	-2.41 (24.4)	-5.97 (59.6)
% LMI Households	1.96 (16.6)	3.46 (37.0)	1.67 (17.9)	0.40 (5.3)
% UMI Households	1.91 (16.6)	3.71 (29.8)	5.91 (51.4)	1.40 (17.9)
% UI Households	0.32 (6.0)	2.12 (41.9)	3.97 (80.8)	6.27 (153.8)
Lambda - λ	0.14 (38.9)	0.13 (42.5)	0.11 (42.7)	0.15 (72.6)
R ²	0.95	0.93	0.94	0.93
Likelihood Criterion - ϕ	0.95	0.94	0.95	0.96
Average Trip Length (minutes)	24.8	24.7	26.1	25.0
MAPE	22.96%	18.13%	18.94%	23.81%
MAPE (Smallest)	57.47%	40.04%	42.55%	51.70%
MAPE (Largest)	9.11%	9.72%	9.09%	11.64%

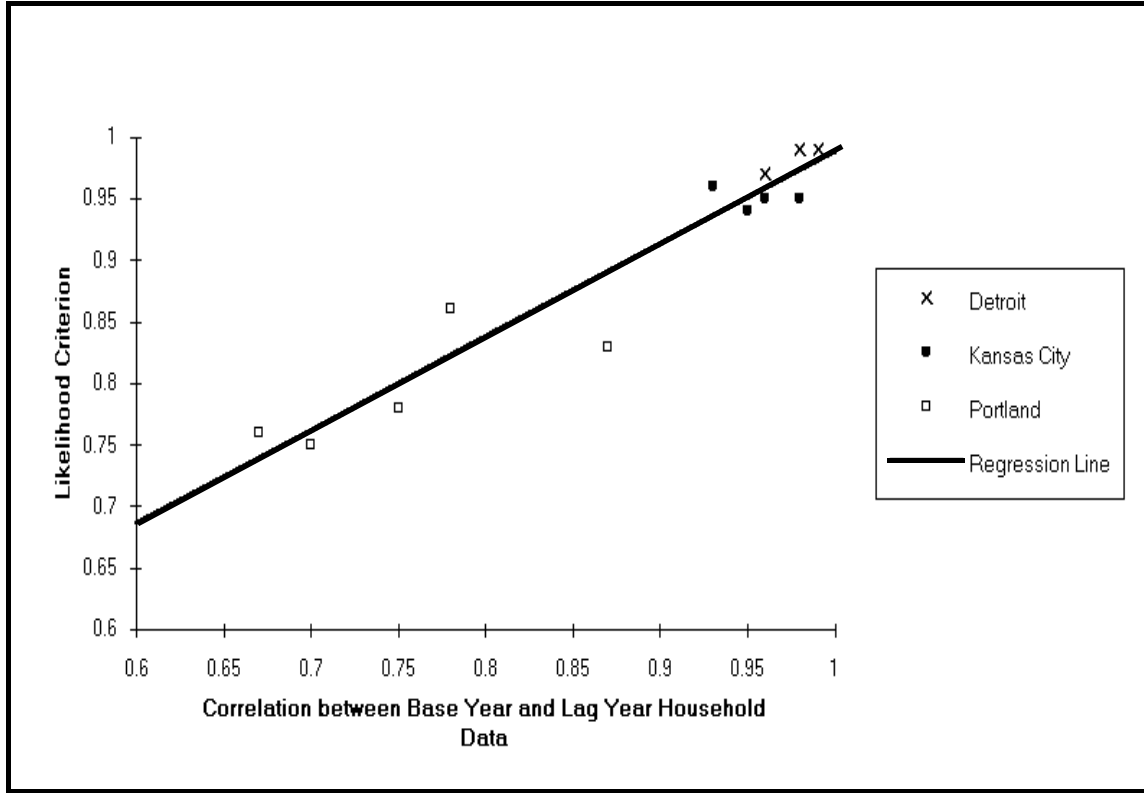
¹ Asymptotic t-tests in parentheses.

Table 29. DRAM Parameters for the Portland Region¹
(328 Zones, Lagged Households Included in Additive Term)

	Low_Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	0.3021 (19.1)	0.1926 (12.4)	0.0102 (0.7)	0.3695 (27.9)	0.3842 (31.4)
Vacant Land	-0.0035 (1.7)	0.0079 (4.2)	0.0199 (10.9)	-0.0099 (6.0)	0.0153 (10.2)
% Developed	0.2621 (9.0)	0.2488 (9.1)	0.2843 (10.7)	0.2546 (11.3)	0.4575 (23.6)
Residential Land	0.6440 (107.3)	0.5421 (88.5)	0.4942 (80.0)	0.5570 (96.0)	0.5020 (93.7)
LIHH	3.9512 (92.5)	0.6111 (12.5)	0.7340 (14.7)	0.0065 (0.1)	-1.8817 (42.0)
LMIHH	2.1651 (21.6)	10.5176 (111.7)	2.9120 (31.3)	2.5187 (30.9)	1.6670 (25.2)
MIHH	-0.9602 (7.9)	-1.2974 (10.4)	8.1446 (65.1)	0.8700 (8.3)	0.6508 (7.8)
UMIHH	-1.7641 (19.2)	-0.0372 (0.4)	-0.7759 (8.0)	4.5443 (56.5)	-0.2315 (2.9)
UIHH	-1.2722 (29.9)	0.3356 (7.8)	0.3106 (7.2)	-0.1936 (5.1)	3.6450 (126.1)
Lambda - λ	0.4381 (76.1)	0.5888 (82.4)	0.6437 (86.9)	0.7176 (119.3)	0.7397 (165.3)
R²	0.8162	0.7802	0.7295	0.7475	0.8361
Likelihood Criterion - ϕ	0.8283	0.7842	0.7463	0.7596	0.8578
Avg. Trip Length (minutes)	25.05	25.37	25.33	27.39	27.08
MAPE	25.45%	26.61%	25.55%	28.79%	35.10%
MAPE (Smallest)	52.34%	57.90%	51.65%	63.30%	73.76%
MAPE (Largest)	13.20%	16.72%	17.24%	16.21%	19.53%

¹ Asymptotic t-tests in parentheses.

Figure 7. DRAM Model with Additive Lag Term



The degree of correlation between the base year and lag year household data is a prime determinant of the goodness-of-fit for the form of the DRAM model with an additive lag term. If household location remained relatively stable during the five year period between the base year and the lag year, then the correlation between the base year and lag year household data should be very high. This is the case for Detroit and Kansas City. For Kansas City, the correlation between the total number of households per zone in the base year and the lag year is 0.95, while for Detroit this correlation is 0.98. For Portland, the correlations between the 1980, 1985 and 1990 household data are shown in the following table:

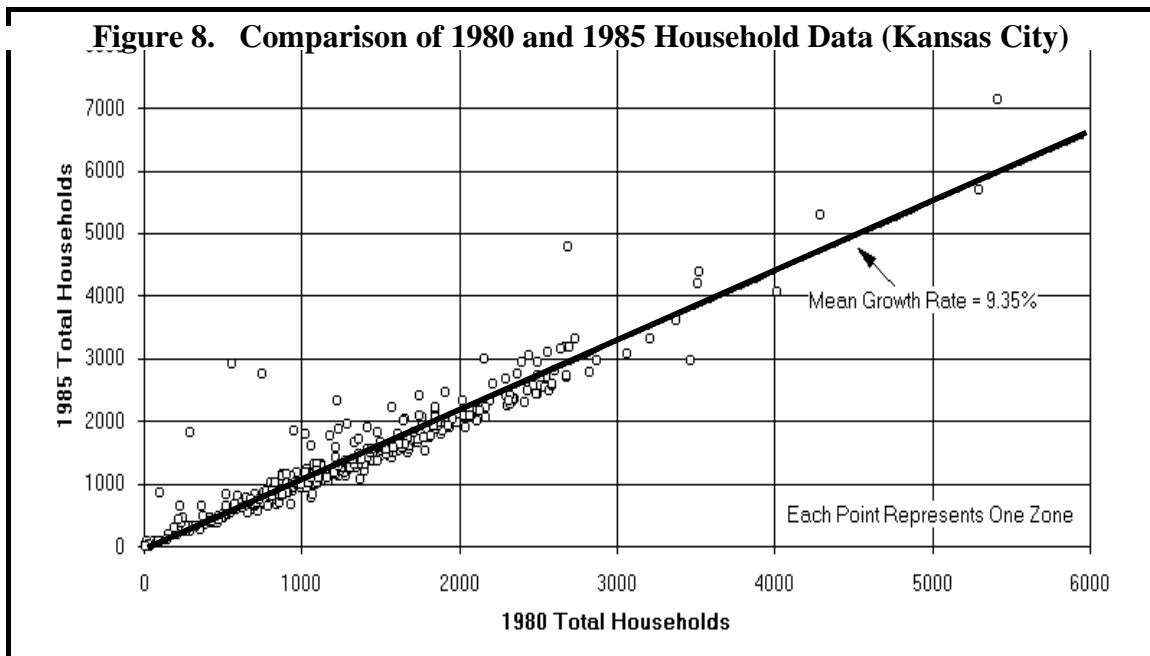
Table 30. Pearson Correlations for Portland Household Data

	1980	1985	1990
1980	1.00	*	*
1985	0.98	1.00	*
1990	0.77	0.81	1.00

The relatively low correlation in the base year and lag year household data for Portland is partly a result of the relatively rapid growth from 1985 to 1990. However, the principal cause is a set of inconsistencies in the 1985 household data.

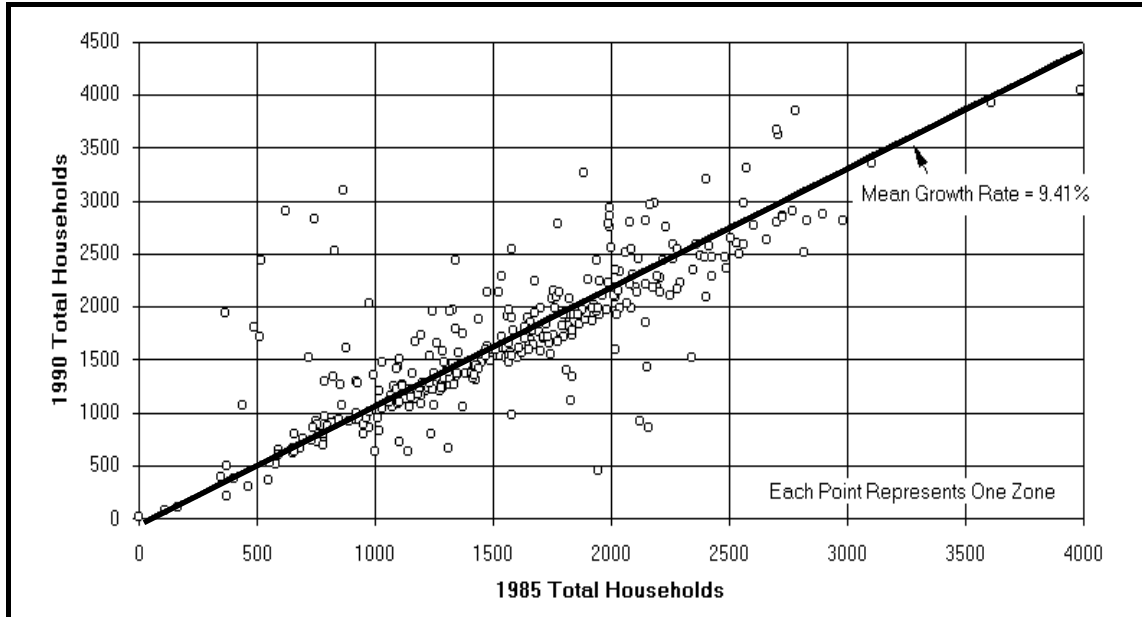
If we compare the 1990 and 1985 census tract household data for Portland, there are 89 tracts (27.1% of the total number of zones) in which the total number of households increased or decreased by more than 20%. While it is possible for the number of households in a single zone to change by this amount over a five year period, it is unlikely that so many zones would undergo such rapid change. In Kansas City, 77 zones (20.4%) grew or declined by more than 20%, while in Detroit 44 zones (14.8%) changed by more than 20%.

If we plot the number of households in each zone for the base year against the number of households in each zone for the lag year, then we can quickly identify zones for which there may be errors in either the base year or lag year household data. Zones with potential errors in the household data will appear farther from the line representing the average per zone growth rate. Note the relatively tight clustering of the points for Kansas City in Figure 8, below.



The greater dispersion of points in Figure 9 for Portland illustrates the relatively low correlation between the base year and the lag year household data.

Figure 9. Comparison of 1985 and 1990 Household Data (Portland)



The 1985 household data for Portland were calculated by interpolating between 1980 and 1990 data. The 1980 household data, however, were defined for 1980 census tracts, while the 1990 data were defined for 1990 census tracts. In order to make the interpolation between 1980 and 1990 data, the 1980 household data were reapportioned to 1990 census tracts on a simple acreage basis. Unfortunately, this reapportionment of 1980 households to 1990 tracts was not accurate. Of the 89 zones which grew or declined by more than 20%, 42 are zones which only exist as 1990 tracts (i.e., they are tracts that are divisions of 1980 tracts). The numbers of households in many of these zones probably did not really change by more than 20% between 1985 and 1990.

The reapportionment of 1980 households to 1990 census tracts reduces the accuracy of the interpolated estimates of 1985 household location. In addition, even after the problems of the interpolation of the total households by tract, the data were then disaggregated by income group, more or less by assuming a regionwide household income distribution. Further inaccuracies were obviously introduced to the 1985 data by this step of the process. As a result, the 1985 household data are not strongly correlated with the 1990 household data, and the calibration results for Portland, even with the inclusion of an additive lag term, are unsatisfactory when compared to the results obtained for a similar model configuration for Detroit or Kansas City.

Including a Multiplicative Lag Term in the DRAM Model

An alternative method for incorporating information about past household location into the DRAM model structure is to include a *multiplicative* lag term in the residential attractiveness function. An advantage of selecting this approach is that it is not necessary to use the lagged households disaggregated by type. The disadvantage is that the quasi-dynamic model structure is no longer possible. In the case of Portland there was reason to believe that while the lagged total households by zone were not as good data as one would like, the same data when disaggregated by income group within zones were very poor indeed. It was hoped that by aggregating the households by zone, some information would still be included on the prior “size” of the zone, while reducing the “noise” in the data due to difficult estimation. The resulting form of the equation is as follows:

$$W_i^n = (L_i^v)^{q^n} (x_i)^{r^n} (L_i^r)^{s^n} (N_{i,t-1}^T)^{t^n} \prod_{n'} \left[\left(1 + N_i^{n'} / \sum_n N_i^n \right)^{b_{n'}^n} \right] \quad (7)$$

where

$N_{i,t-1}^T$ = total households residing in zone i at time $(t - 1)$, and

t^n = an empirically derived parameter.

When a multiplicative lag term is included in the DRAM model equations, successful calibration results are obtained for the Portland data set (see Table 31).¹¹ The likelihood criterion measures exceed 0.80 for four of the household types, and most of the estimated parameters have asymptotic t -tests that are greater than 20.0. More importantly, negative estimates of β are obtained for all five household types; the expected result, verifying that households are adverse to long home-to-work trips. Similar results were obtained for Kansas City calibration using the extended *multiplicative* version of DRAM in equation (7), but the improvement over the results obtained for the extended *additive* version were insignificant. This is undoubtedly due to the fact that the Kansas City lagged household data were of somewhat better quality than that for Portland.

¹¹ For these calibrations, the Tanner function is used as the trip function. A power function was used for all of the other calibration experiments. The Tanner function is defined as: $c_{i,j}^\alpha \exp(\beta c_{i,j})$. For many metropolitan regions, the Tanner function more accurately represents observed trip distributions.

Table 31. DRAM Parameters for the Portland Region¹
 (328 Zones, Lagged Households Included in Potential Term)

	Low Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	1.0591 (98.1)	0.9196 (75.1)	0.8589 (68.1)	1.0832 (87.8)	0.9930 (86.4)
Beta - β	-0.3331 (76.2)	-0.2895 (60.5)	-0.2707 (56.1)	-0.3273 (70.3)	-0.3110 (69.8)
Vacant Land	-0.0112 (8.5)	-0.0042 (3.0)	0.0032 (2.2)	0.0024 (1.7)	0.0291 (22.7)
% Developed	0.0403 (2.1)	0.0470 (2.3)	0.1123 (5.5)	0.1767 (9.5)	0.4183 (25.3)
Residential Land	0.1286 (35.6)	0.2089 (45.1)	0.2616 (51.2)	0.2895 (58.2)	0.2319 (50.1)
Lagged Total Households	0.7595 (102.8)	0.6061 (74.5)	0.5296 (64.2)	0.4620 (63.0)	0.5277 (75.9)
LIHH	3.6893 (133.9)	0.1871 (5.2)	0.2550 (6.4)	-0.6015 (15.3)	-2.0369 (53.9)
LMIHH	2.0011 (29.3)	8.0098 (111.6)	2.2469 (30.5)	1.6370 (23.9)	1.2084 (21.2)
MIHH	-0.4355 (5.5)	-0.2491 (2.6)	6.3917 (65.0)	0.2159 (2.4)	0.1733 (2.4)
UMIHH	-0.7499 (12.4)	-0.3713 (5.0)	-0.3284 (4.1)	4.5480 (59.3)	-0.5039 (7.0)
UIHH	-0.8972 (29.0)	0.2416 (7.4)	0.4639 (13.8)	0.0122 (0.4)	3.6879 (149.5)
R²	0.8432	0.8227	0.7814	0.7738	0.8410
Likelihood Criterion	0.8418	0.8165	0.7912	0.8064	0.8670
Avg. Trip Length (minutes)	24.62	25.57	26.32	27.27	26.64
MAPE	22.08%	19.92%	20.53%	20.83%	26.43%
MAPE (Smallest)	43.21%	37.62%	37.99%	35.71%	50.12%
MAPE (Largest)	11.80%	14.32%	15.12%	15.64%	16.52%

¹. Asymptotic t-tests in parentheses.

Spatial Aggregation Tests for Portland

There are two factors that are responsible for the poor calibration results for Portland when a multiplicative lag term is *not* used in the DRAM model equations: 1) relatively low spatial variation in the base year household data, and 2) inconsistencies between the base year and lag year household data. When the 328-zone Portland data set is aggregated to 100 zones, these factors are mitigated, and good calibration results can be obtained for both the *standard* DRAM model and an *extended* DRAM model with an *additive* lag term. In part, this is because the aggregation procedure for Portland, as distinct from the census tract definitions, does not create zones of equal size (in terms of numbers of households). The spatial aggregation of the household data increases the amount of zone-to-zone spatial variation in household location for all income groups. The extent of this aggregation effect can be seen in the following table.

Table 32. Standard Deviation of Households per Zone

Household Type	Portland (100-Zones)	Portland (328-Zones)
Low Income	1345.50	218.94
Low-Middle Income	1024.32	159.30
Middle Income	866.36	146.37
Upper-Middle Income	889.72	175.42
Upper Income	1121.63	313.97

Spatial aggregation also reduces the inconsistency between the 1985 household data and the 1990 household data. This is because the aggregation of census tracts eliminates many of the inconsistencies created when the 1980 household data (which is used to interpolate 1985 household data) is reapportioned to 1990 zones. If a 1980 census tract was split into two or more 1990 census tracts, the reapportionment of 1980 households at the 328-zone level of spatial disaggregation was likely to be inaccurate. However, if all of the tracts produced by a census tract split are contained within one zone at the 100-zone level of spatial disaggregation, then the split has no effect on the consistency of the household data.

The improved consistency of the 100-zone household data is apparent if we compare the correlations between the 1985 and 1990 household data for both the 328-zone and the 100-zone data sets.

Table 33. Correlation between 1985 Households per Zone and 1990 Households per Zone

Household Type	Portland (100-Zones)	Portland (328-Zones)
----------------	----------------------	----------------------

Table 33. Correlation between 1985 Households per Zone and 1990 Households per Zone

Low Income	0.98	0.87
Low-Middle Income	0.97	0.75
Middle Income	0.95	0.70
Upper-Middle Income	0.92	0.67
Upper Income	0.89	0.78

Since spatial aggregation mitigates the errors that were present in the 328-zone data set, it is not surprising that the calibration results for the 100-zone Portland data set are similar to those for the 297-zone Detroit data set. Note, too, that these results at the 100 zone level of detail are substantially improved over the previous ones. For the standard DRAM calibration (Table 34), all of the likelihood criterion measures are greater than 0.90. Twenty-three of the estimated parameters, however, have asymptotic t-tests less than 20.0, which indicates that some of the parameters cannot be precisely estimated, i.e., they are probably not statistically significant. The signs and magnitudes of the majority of the estimated parameters are consistent with previous calibration results for other metropolitan regions. It is especially important to note that the estimated travel time parameters are properly negative for all household types.

Table 34. DRAM Parameters for the Portland Region
(100-Zone System)¹

	Low Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	-0.3423 (28.3)	-0.2408 (18.2)	-0.2345 (18.0)	-0.1260 (10.3)	-0.2293 (20.7)
Vacant Land	-0.0961 (43.2)	-0.0904 (38.5)	-0.0819 (34.5)	-0.0708 (32.0)	-0.0598 (29.4)
% Developed	0.1284 (6.2)	0.0920 (4.3)	0.0793 (3.7)	0.1732 (8.7)	0.2101 (11.9)
Residential Land	0.7130 (196.9)	0.7448 (178.4)	0.7430 (177.2)	0.7400 (189.4)	0.7039 (202.1)
LIHH	5.2822 (136.1)	1.4374 (31.0)	1.3097 (26.6)	0.6739 (14.2)	-0.4364 (10.1)
LMIHH	0.2971 (3.0)	7.3024 (71.0)	1.4530 (13.9)	1.0958 (11.4)	0.1063 (1.3)
MIHH	-2.9328 (25.3)	-1.9638 (14.8)	4.8754 (36.2)	-0.8822 (7.2)	-1.3267 (13.1)
UMIHH	-1.2799 (15.2)	-0.4346 (4.8)	-0.3648 (3.6)	4.0672 (42.6)	-1.3890 (15.6)
UIHH	-1.8071 (46.0)	-0.2714 (6.7)	-0.0783 (1.9)	-0.3601 (9.4)	3.1583 (100.3)
R²	0.9375	0.9343	0.9224	0.8974	0.8834
Likelihood Criterion - ϕ	0.9322	0.9273	0.9113	0.8912	0.8806
Avg. Trip Length (minutes)	22.69	24.03	24.80	25.86	25.06
MAPE	30.41%	29.39%	29.87%	30.35%	33.85%
MAPE (Smallest)	49.02%	57.27%	53.45%	58.71%	63.38%
MAPE (Largest)	17.60%	14.87%	12.50%	12.24%	15.02%

¹. Asymptotic t-tests in parentheses.

The 100 zone Portland calibration results for the *extended* DRAM model with an *additive* lag term (Table 35) are similar to the results for the standard DRAM model. The goodness-of-fit, as measured by the likelihood criterion, is slightly improved for all household types. The estimated parameters and their associated t-tests are nearly the same as those for the standard DRAM model. The mean absolute percent error (MAPE) measures of the estimates are significantly reduced however, implying that forecast from this extended version of the model will be substantially more reliable.

Table 35. DRAM Parameters for the Portland Region¹
(100-Zone System with Additive Lag Term)

	Low Income	Lower-Middle Income	Middle Income	Upper-Middle Income	Upper Income
Alpha - α	-0.1253 (6.1)	-0.3541 (16.7)	-0.5704 (27.3)	-0.1119 (6.9)	-0.1193 (8.6)
Vacant Land	-0.0364 (9.9)	-0.0564 (15.9)	-0.0976 (27.2)	-0.0820 (26.8)	-0.0742 (29.3)
% Developed	0.3099 (9.4)	0.0934 (2.9)	0.0452 (1.4)	0.0792 (2.9)	0.0379 (1.7)
Residential Land	0.8404 (132.1)	0.7871 (122.5)	0.7120 (114.1)	0.7865 (144.7)	0.7467 (170.4)
LIHH	5.0400 (78.3)	1.3285 (19.4)	1.4755 (21.2)	0.8109 (12.2)	-0.7901 (13.7)
LMIHH	0.1869 (1.2)	7.5286 (49.5)	1.6261 (10.5)	0.9485 (7.0)	-0.0265 (0.3)
MIHH	-2.9422 (15.3)	-1.9428 (9.8)	5.1073 (25.6)	-0.4528 (2.6)	-0.7261 (5.7)
UMIHH	-1.2764 (9.1)	-0.4221 (3.0)	-0.3976 (2.8)	6.5993 (52.0)	0.2549 (2.2)
UIHH	-1.4545 (25.4)	-0.3587 (6.0)	-0.5638 (9.1)	-1.4012 (25.5)	2.9239 (72.3)
Lambda - λ	0.3624 (51.2)	0.4529 (47.8)	0.4719 (54.5)	0.5403 (71.2)	0.6223 (108.9)
R²	0.9839	0.9676	0.9564	0.9405	0.9217
Likelihood Criterion	0.9765	0.9548	0.9521	0.9322	0.9243
Avg. Trip Length (minutes)	23.77	23.55	22.98	25.97	25.75
MAPE	17.03%	19.37%	20.39%	22.54%	28.61%
MAPE (Smallest)	26.93%	39.31%	34.90%	44.66%	59.78%
MAPE (Largest)	7.53%	11.25%	9.39%	9.52%	13.46%

¹ Asymptotic t-tests in parentheses.

Including Land Value as a Variable in the DRAM Model

It has long been thought that models of household location could be improved in their performance if it were possible to include some aspect of the urban/regional land market in their formulations. For the 100-zone Portland data set, estimates of total residential land value for each zone were available. It was hoped that the inclusion of a land value variable in the DRAM model equations would improve goodness-of-fit. Were this to be the case, then it would ultimately become necessary to develop a method for updating this variable for recursive forecasting, but in any case the potential for improving model performance with the introduction of a land market representation would have been examined. Unfortunately, as Table 36 shows, used as an additional attractiveness variable in a test of an extended version of DRAM, residential land value is not a statistically significant determinant of the location of low income households. (Because of the insignificant results for low income households, experiments with other household types were not performed.)

The calibration results for an extended model with a residential land value variable are nearly identical to the results for the standard DRAM model. Furthermore, the estimated parameter for residential land value may not be statistically different from zero, since its asymptotic t-test is only 11.0.

In the few other tests which have been done involving the use of land value variables in the DRAM formulation, similar results have been obtained. Here the insignificance of the estimated residential land value parameter may be due to inaccuracies in the land value data for the Portland region. On the other hand, it may well be that since DRAM already uses the income distribution of households in a zone as a part of its constructed attractiveness measure, that the land value is already rather well represented. Income distribution of residents is known to be rather highly correlated with land values. If however it is of importance to a prospective model user to have estimates of house value or land value, then a substantial investigation will be necessary to determine whether it can be done in this model framework. This examination falls beyond the scope of work for the current study.

Table 36. DRAM Parameters for the Portland Region
(Low Income Households, 100-Zone System, Three Tests)

	Standard Model	Additive Lag Term	Land Value
Alpha - α	-0.3423 (28.3)	-0.1253 (6.1)	-0.2758 (22.9)
Vacant Land	-0.0961 (43.2)	-0.0364 (9.9)	-0.1058 (47.8)
% Developed	0.1284 (6.2)	0.3099 (9.4)	0.1139 (5.5)
Residential Land	0.7130 (196.9)	0.8404 (132.1)	0.6984 (192.1)
Residential Land Value per HH			0.0178 (11.0)
LIHH	5.2822 (136.1)	5.0400 (78.3)	5.2437 (135.3)
LMIHH	0.2971 (3.0)	0.1869 (1.2)	0.3161 (3.2)
MIHH	-2.9328 (25.3)	-2.9422 (15.3)	-2.8874 (24.9)
UMIHH	-1.2799 (15.2)	-1.2764 (9.1)	-1.2874 (15.2)
UIHH	-1.8071 (46.0)	-1.4545 (25.4)	-1.8338 (46.8)
Lambda - λ		0.3624 (51.2)	
R²	0.9375	0.9839	0.9365
Likelihood Criterion - ϕ	0.9322	0.9765	0.9327
Avg. Trip Length (minutes)	22.69	23.77	22.86
MAPE	30.41%	17.03%	30.52%
MAPE (Smallest)	49.02%	26.93%	50.01%
MAPE (Largest)	17.60%	7.53%	17.95%

Conclusions

The accuracy of DRAM forecasts of household location depend on the accuracy of the parameter estimates produced by the calibration procedure. If the goodness-of-fit of an estimated DRAM model equation is poor, then it is not possible to obtain accurate forecasts of household location. If individual parameters are imprecisely estimated, then the factors which determine household location may be misidentified.

For the Detroit region, very good calibration results are obtained for the standard DRAM model and an extended DRAM model that includes an additive lag term. The calibration procedures for each of these models produced excellent goodness-of-fit and precise parameter estimates. Therefore, either one of these model specifications could be used for forecasting household location.

For the Kansas City and Portland regions, the calibration procedures are unsuccessful for the standard formulation of the DRAM model. For both regions, the goodness-of-fit statistics (R^2 and the likelihood criterion ϕ) indicate accurate forecasts of household location cannot be made with the standard model. Of more concern, however, are the signs and asymptotic t-test statistics of the travel time parameters. For the standard DRAM formulation at the 328 census tract level of detail, all of the estimated travel time parameters for Portland are positive. One of the estimated travel time parameters for Kansas City is positive, while another may not be statistically significantly different from zero. These estimates are contrary to our theoretical and empirical knowledge of the location decisions of households, as they imply that households prefer longer home-to-work trips. Obviously calibrations which yield such results are not acceptable.

If an *additive* lag term is included in an extended version of the DRAM model equations, then very good calibration results are obtained for the Kansas City data set. If a *multiplicative* lag term is incorporated into the DRAM model structure, along with a Tanner function to measure travel deterrence, then good calibration results are obtained for the Portland and Kansas City data sets. These formulations of DRAM should produce accurate forecasts of household location.

Successful calibrations for the Portland region are also obtained when the 328-zone data set is spatially aggregated to 100 zones, for both the standard DRAM model and an extended model with an additive lag term. The *most accurate forecasts* of household location should be produced by either of these two formulations when the *100-zone* Portland data set is used.

It should be recalled that similar results were obtained for the employment location model (EMPAL), and were described in the first part of this chapter.

