ANALYSIS OF OUTPUT AND POLICY APPLICATIONS OF AN URBAN SIMULATION MODEL

Michael A. Goldberg, University of British Columbia; and
Jeffrey M. Stander, Oregon State University

This paper reports the progress made in developing a regional simulation model. It is essentially a case study of a model-building process. The model is interactive and linked with employment, population, land use, and transportation components. The models have undergone considerable refinement, calibration, and elaboration during this period. In particular, emphasis has been placed on developing a housing model that not only forecasts total supply and demand for the region (i.e., a macrospatial model) and allocates these totals to subareas (i.e., a microspatial model) but also uses these microspatial data to generate successive housing forecasts at the macrolevel. The failure of earlier work to identify such micro-macro gaps and develop suitable linkages is one of the significant shortcomings of simulation modeling, and this paper discusses approaches that have been taken to develop such linkages. The response of the output variables to the inclusion of such links is compared with the output of the model without these links. Some observations about programming and debugging procedures are made. Output under various scenarios is presented, and a few general conclusions are drawn from this output.

•THIS paper presents the results of and approaches to the continued development and refinement of a regional simulation begun nearly 5 years ago. The details of this work are discussed in Goldberg (4, 5) and Goldberg and Davis (7); reviews of other work can be found in Brown et al. (3) and Sweet (13). Similarly, earlier works describing the present model can also be referred to for discussions of model-building strategy and philosophy. It is our intent to focus on the progress made in debugging the model and making it more useful for policy application. This is an ongoing process, and the current model can only be refined through continued use and modification. Of particular interest is the interaction between microspatial magnitudes, such as housing and land use in small areas, and macrospatial phenomena, such as regional migration and economic and employment growth.

MODEL-BUILDING STRUCTURE AND STRATEGY

The module structure is shown in Figure 1 and includes three principal component models: population, economic-employment, and land use. Figure 1 is an idealization because it shows all of the components connected by two links; although this is our ultimate goal, it does not describe our present level of achievement. Figure 2 shows the linkages between the module and the transportation model. The transportation-land use interaction is similar to that described in Wendt and Goldberg (15). Currently, the transportation model is being developed independently and in parallel with the three-component module, and the generation and distribution elements are operational. A modal-split model is being developed. Figure 3 shows both the present state of our ef-

---

Publication of this paper sponsored by Committee on Urban Activity Systems.

Figure 1. Module.

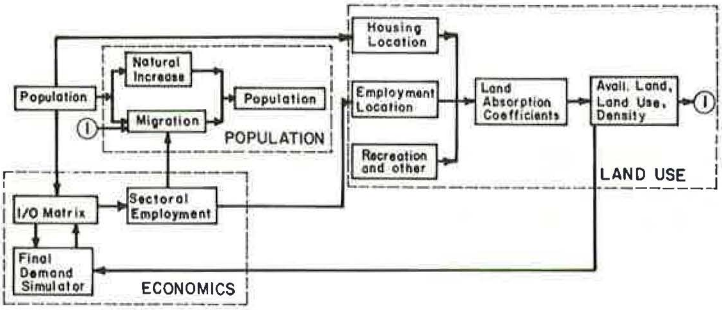


Figure 2. Transportation and the module.

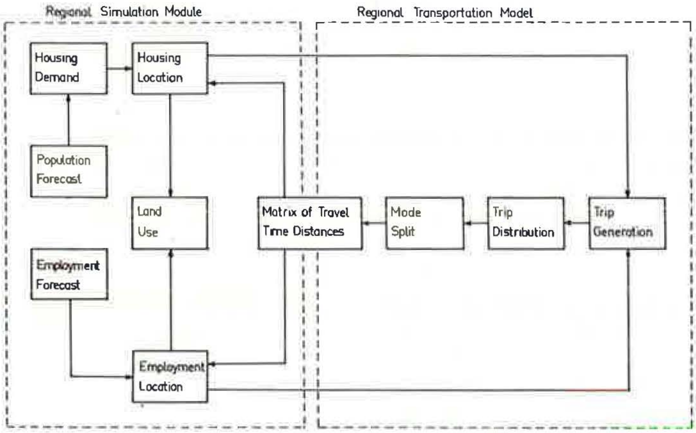
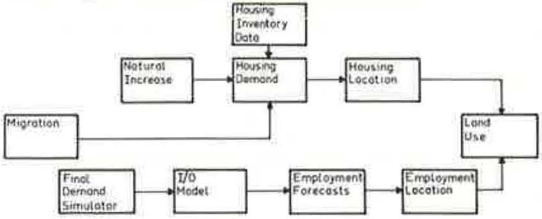
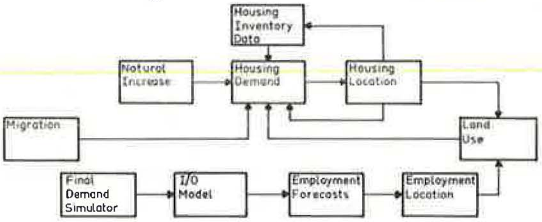


Figure 3. Module evolution.

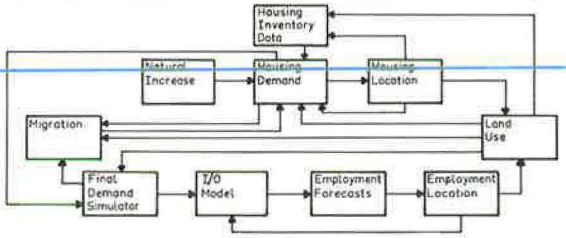
1st Stage in Evolution of Module



Present Stage in Evolution of Module



Completing the Module's Feedback Loops



forts and the strategy we are following in developing each of these models and the module simultaneously.

Our basic approach is that the models shown be flexible and that they evolve as our understanding of their behavior and of the world they simulate improves with continued research. Each of the models is carefully documented so that the model building remains process-oriented and we avoid, at any cost, creating a black box. Our emphasis on process instead of product helps by keeping us mindful that it is the usefulness and actual use of our models that will prove their worth. To be used they must be accessible, understandable, and subject to careful scrutiny, criticism, and most importantly, change. By viewing model building as an evolutionary process, we avoid creating a black box and, what is worse and quite common, justifying it. We are not interested in developing a product and then selling it, but rather in the continual evolution of our system of models and their continual documentation and criticism by users and others. This paper is therefore seen as part of that process. An integral part of that process is the linkage with the transportation planning process. Housing and employment location forecasts are key inputs to forecasting travel demand. The land use models set out below are intended to provide better inputs to transportation models. Figure 3 shows the development strategy being followed here and in our subsequent work.

The land use component of the module is important in allocating jobs and houses spatially. There are 27 economic sectors that have been aggregated into four groups for location purposes: 10 kinds of manufacturing in one group, retail trade in another location group, six sectors in the office-commercial group, and six sectors in the personal and business services group. Construction and three primary sectors are not allocated to subareas in these models.

The land use models are shown in Figure 4 in their simplest detail. The right side of the figure represents various supply elements, and the left side represents the principal demand element of land for jobs and houses. The conceptual structure is simple in these models, in that activities (jobs and houses) are allocated to each of the 82 subareas in the Vancouver region by using a number of algorithms (one algorithm associated with each activity). These allocations are then converted to land use by using a land absorption coefficient that represents the amount of land a unit of each activity requires. If there is sufficient land for the activity, then it is considered to be allocated to that zone, and the land use, employment, housing, and population files are all updated for that zone. If there is insufficient supply, the excesses are cumulated across all subareas and relocated, by using the initial algorithms, to subareas with excess capacity. When all jobs and households are allocated to subareas, for year t, a new set of forecasts from the population and economic models are read in for t+1 and allocated as above until the module reaches the terminal year for its forecasting horizon.

The housing model shown in Figure 5 represents its initial stage of development. The present stage of evolution is essentially identical, with the exception of a simple feedback mechanism that relates the microspatial allocation functions and land availability to the macrospatial supply-demand housing model.

Step 1

Initially, for computational convenience, it was assumed that supply equalled demand for the region as a whole. However, it was not assumed that supply and demand had to be equal in any of the 82 subareas, nor did we even constrain regional totals for demand by structure type (single-family and high-rise) or by value class (four value classes) to be equal to the equivalent regional totals for supply. All we assumed was that the total number of units demanded equalled the total number of units supplied each period. Regional demand by structure type and value class is a function of information on family size and age distribution of household heads (from the population model) and of the income distribution (from the economic model) for each annual forecast increment. In addition, demand derives from households whose housing units have been demolished during the previous period. Finally, the initial model kept track of households that were forced into units other than those they desired, and these dissatisfied households also entered into the calculation of demand. Given these regional totals for demand by value

Figure 4. Land use models.

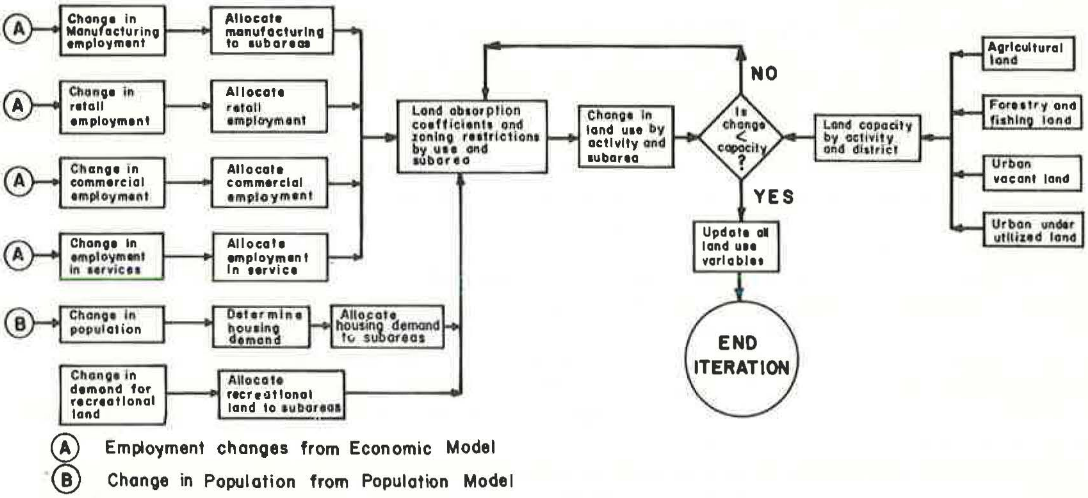
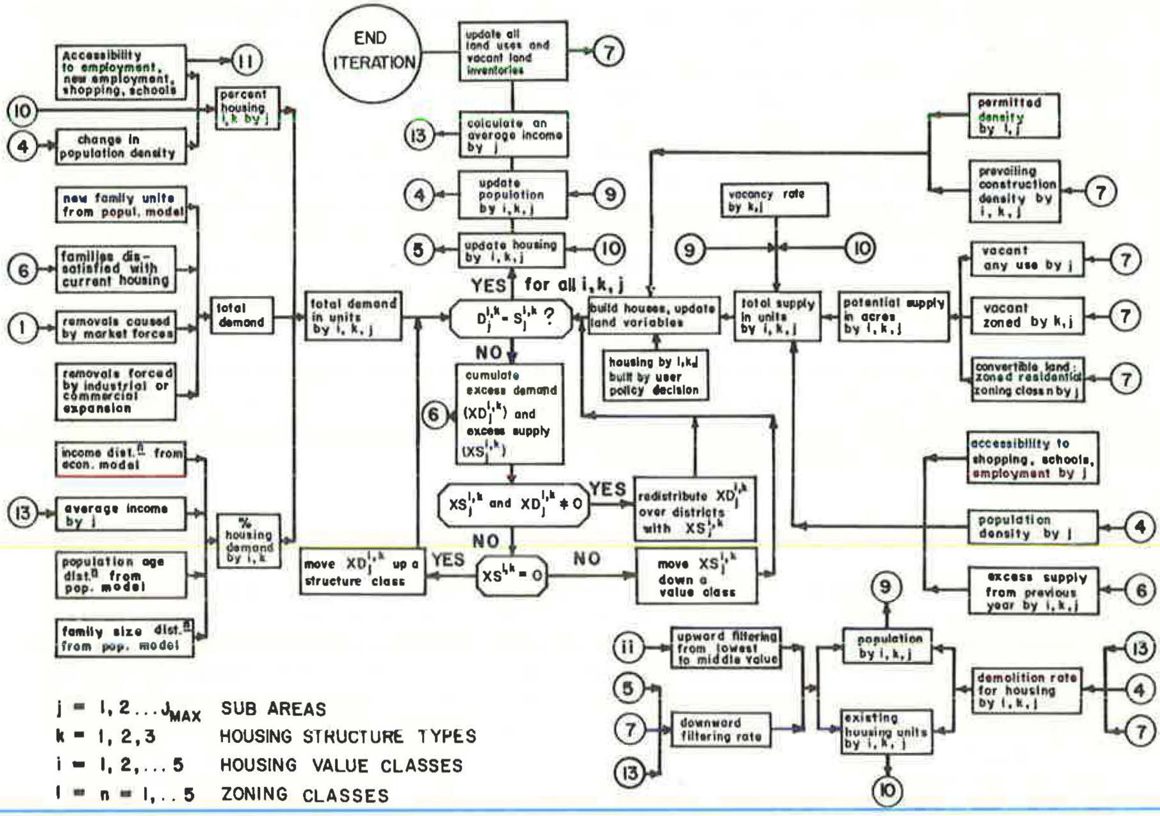


Figure 5. Housing model.



class and structure type, the microspatial model then allocated demand to each of the 82 subareas, again by value class and structural type. The resulting demand $D_j^{i,k}$ is demand in subarea j, for housing type k, and value class i. Prices are not included in either the demand or the supply equations. This was a conscious decision derived from the great difficulty involved in forecasting subsequent period prices that would be needed if prices played a significant role in the model. It was also felt that the presence of prices created an illusion of precision that simulation does not warrant, and this could lead to misuse and misunderstanding of the models.

Since supply was constrained to equal demand in this initial phase, the supply model took the number of units demanded for the region as its starting point. Thus, the supply model proceeded, from the total regional supply in number of units, to disaggregate this supply into structure types, value classes, and subareas. The principal data inputs to this disaggregation procedure are actual and allowable densities; available land; accessibility; and excess supply by value class, structure type, and subarea from the previous iteration of the model. The result is supply $S_j^{i,k}$ by subarea j, structure type k, and value class i.

Differences between supply and demand by structure type and value class for each subarea are reconciled by cumulating excess demand and redistributing it to areas with excess supply until there is no excess demand or excess supply in any subarea, structure type, or value class of housing. Excess demand is allocated first to other subareas with similar housing (by type and class). If no similar housing is available, demand is allocated to those areas that have housing of the same value class but any structure type. If there is no such housing available, the excess is allocated to subareas with the originally desired structure type but the next lower value class. Such housing is then reclassified and raised one value class. The model generates a kind of upward filtering of houses and neighborhoods in this way. This process continues until all excess demands are allocated. If, on the other hand, there are excess supplies in certain subareas, the excess housing is assigned to the next lower value class. In this way excess supply moves down through value classes; this is what happens in practice where high vacancy rates lead to price cutting. Excess demand, however, moves across structure types within the same value class, unless no housing exists in any subarea in the desired value class. In this case, demand moves down one value class and then across the structure types again if necessary. This phenomenon has been observed in Vancouver (14) and in such renewal schemes as Society Hill in Philadelphia, Cobble and Boerum Hills in Brooklyn, Russian Hill and Jackson Heights in San Francisco, and Capitol Hill in Seattle.

Excess demand and excess supply are both kept in memory for one period: the former to measure dissatisfaction, the latter to introduce a dynamic lag into the supply determination process.

Step 2

In step 2 the assumption was dropped that demand and supply had to be equal for the region as a whole. Accordingly, we adopted a very simple multiplier-accelerator type of model from macroeconomics. Demand in this step was merely set equal to forecast population divided by the number of persons per household PPH to yield an estimate of the number of households forecast for each time period. New supply, on the other hand, was assumed equal to this change in the number of households TNHH plus a demand for vacancies DV to allow for equilibration of short-term disturbances (i.e., some inventories for short-run adjustments) minus the housing supply TH. Equation 1 sets out the supply relationship as follows:

$$NS_t = TNHH_t - TH_t + VACRAT_t \cdot TNHH_t \qquad (1)$$

The demand for vacancies in turn was assumed to be a function of vacancy rates in the preceding 3 years [there is 3-year planning horizon for developers in our region (9)].

If NS_t is negative, then a small number of units are still built; this reflects the fact that construction does not cease even when there are high excess housing stocks.

$$VACRAT_t = (VAC_t) \Bigg/ \left[\left(\sum_{i=1}^{3} VAC_{t-i} \right) \Bigg/ 3 \right] \tag{2}$$

The demand for vacancies is now to be equal to the ratio between last year's vacancy rate and the average over the last 3 years. Demand $TNHH_t$ and new supply NS_t are then disaggregated by structure type, value class, and subarea as before.

Step 3

Step 3 is the stage at which we are currently running the model. This step builds on step 2 and continues to assume that regional demand for units and regional supply are not necessarily equal. Supply is again calculated as in step 2. Demand, however, is now calculated by using feedback from the housing location (microspatial) submodel. This is accomplished by using an index of available land in each subarea weighted by the amount of development already existing in that subarea. Thus, demand is now calculated as follows:

$$D_t = TNNH_t \cdot \frac{CINDEX_t}{CINDEX_0} \tag{3}$$

where $TNHH_t$ is demand as calculated previously, and $CINDEX_t$ is the weighted capacity or land-availability index from the housing location (microspatial) model. CINDEX, the weighted land-availability index, provides the first micro-macro link that we have developed in accordance with our modeling strategy of identifying and then closing these micro-macro gaps. CINDEX itself is calculated as follows:

$$CINDEX_t = \frac{\sum_j TH_{jt}(AVLAND_{jt})}{\sum_j TH_{jt}} \tag{4}$$

where TH_{jt} is the total housing currently existing in the subarea j (=1, . . . , 82) at time t, and $AVLAND_{jt}$ is the land available for development in subarea j at time t. In equation 3, $CINDEX_t$ is divided by $CINDEX_0$ (base year for the run) to give a ratio; this serves as a simple valve to slowly shut off demand as the region runs out of land and as $CINDEX_t$ gets smaller relative to $CINDEX_0$.

MODEL DEVELOPMENT TECHNIQUES

In a large-scale simulation model like this one, there is much resilience. Very often, programming and data updates do not produce great changes in the model output. This is reasonable since the model is complex and is highly buffered. Thus, the gross picture of model behavior has remained the same: Vacant land is used up, and the region goes to capacity somewhere between 35 and 45 years. In the last year of work on the model many changes have been made, as described earlier, and many programming errors have been found; in a complex model like this, it is suspected that there are programming errors yet to be found (it is hoped they will not be major). The development process has thus taken much time and has required the attention of a full-time programmer-modeler because complete familiarity is the only way in which bugs can be identified when anomalous results are identified. In fact, the anomalies that signal bugs are usually imperceptible to someone not totally familiar with the program. This points out one of the major advantages of working with a very small group of researchers, programmers, and modelers. Although this model was initially developed in a large-scale program involving dozens of people, we were able to continue its development independently and had infinitely better success than an interdisciplinary team with much overhead and organizational difficulties.

There are two basic kinds of errors we ran across: conceptual and programming errors. Programming design should be such that conceptual errors may be corrected and conceptual changes made without a great deal of reworking. That is, there should be an inherent flexibility in the program, and this can occur by the hierarchic design of modules and submodules of the simulation model, allowing a subwhole to be changed without reworking the matrix in which it is set. Programming errors (bugs) show up as anomalies appear. The program should be writtern to encourage finding anomalous situations.

The first debugging tactic adopted was keeping current on all variables. This means that all changes to components of some totalizer variable should be immediately reflected in the totalizer itself. For instance, if some houses are demolished, the change should be reflected in all the relevant variables: for example, current density, vacancies, and total houses. The reason for this tactic is that the programmer may account for the change later on, but, perhaps in the future, some other programmer will accurately modify one part of the program and forget that the updating was not explicitly done in the segment where the change occurred. This creates a bug. The possible waste of CPU time by accessing the totalizer variables more than one time only at the end of a run is made up for by the programmer's time saved in tracking down an elusive bug (like negative vacant houses). The general rule is that program segments must represent some relatively complete or decomposible subwhole of the model and that updating be done either in that segment or by some integrator routine designed for the job; in any case, it must be done currently when the changes are made.

The second programming tactic evolved was that variables should be reasonably calculated: The program variables should represent easily understood quantities that are calculated as directly as possible. The example here is the current density calculation; last year's version had current density read in as data, and then the data were scaled up or down, depending on the relative changes in housing stocks versus new land acquired. This was initially done to get around a data problem. When the change was recently made to calculate current density directly, that is, by simply dividing housing by the land used, some very strange numbers turned up, specifically very large numbers. Some errors were indicated and were traced back to faulty data. The first programmers to work on the model were aware of this and chose to work around the problem. It is better strategically to acknowledge faults in the data base and even devise dummy data for the interim than to build into the model a device that obscures such problems. When the current densities were calculated directly, understandably, and clearly, many obscure bugs and misconceptions that were interfering with the simulation were uncovered. Model programming should always be done with the design that errors are likely to occur and that any error should stand out. A value less than zero for variables like housing units, vacancies, or land makes no sense in the real world, but the shrewd programmer will not use an if statement to alter a negative value to zero. That negative value indicates a bug somewhere. In the current version of the simulation model, where values such as these are set equal to zero, a message is printed indicating what happened; therefore, the simulation is allowed to continue to the end in a reasonable fashion, but the area where the problem might be is pinpointed. Often the negative value is a very small rounding error and can be ignored, but not always. The basic rule is as follows: When something is wrong, it should be obvious, and good program design will cause errors to be obvious.

ANALYSIS OF OUTPUT

There are two basic types of intervention that we can make with this model. The first type is quantitative intervention in which the magnitude of variables is changed, e.g., land may be frozen, densities changed through a zoning policy, and changes made in transportation facilities so that accessibilities are affected. At our current spatial scale [82 subareas for a region of roughly 600 miles^2 (1554 km^2)], other infrastructure (utility placement) is not meaningful because all 82 areas are serviced; however, subareas within these might not be. The second type is structural intervention in which an alternative model form is chosen; for instance, several feedbacks to migration from other system components have been identified, and any combination of these feedbacks may be switched on or off for a given model run. Figures 6 and 7 show three different simulations under which structural interventions were made. The output exhibits a degree of equifinality (1) defined as the process by which "the same final state or 'goal' may be reached from different initial conditions or in different ways." The backward link (backlink) from the microspatial housing model to the macrospatial housing model tends to make the construction more responsive, and to increase the frequency of the construction cycle as the amplitude is decreased. The addition of the feedback from the weighted land-availability index CINDEX to migration appears to cause an increased pressure on land resources, and migration is reinforced during partial development. However, as full development approaches, migration and construction decline.

These two types of interventions demonstrate the effects of different model structures on the model behavior. In all cases, about the same final point is reached. Full development occurs when all the land is gone. We specifically avoided building in dynamically adjusting density and redevelopment algorithms. We felt that these were the purview of decision makers and preferred to keep the models confined to if-then (if rezoning, then ... ?) types of runs. The terminal year of the simulation thus becomes of interest to decision makers in their attempts to provide more housing and jobs.

Figures 8 and 9 show the response of total regional housing to various scenarios. The university endowment lands (U.E.L.) scenario frees over 2,500 acres (1012 km^2) of land in West Point Grey for residential development at 10 units per acre (25 units/km^2). This is a policy that the provincial government has been contemplating. This large tract of land is located next to a region of predominantly single-family dwellings, much in demand. Figures 8 and 9 also show the basic, unconstrained land act run in which an agricultural land freeze is in effect and a grand up-zoning in which much of the city is up-zoned to higher density housing and multifamily dwellings. On regional level, the only noticeable differences are the smoothing effect of the up-zoning and the fact that the land act has removed much development land so that the region runs out of land much sooner and the maximum amount of units constructed is much less. Up-zoning appears to damp out the construction cycles possibly because it allows a higher single-family/multifamily ratio and because a greater proportion of the people moving into new housing are satisfied and do not reenter the market the next year.

The distinction between up-zoning and land-freeing policy is made clear in Figures 10 and 11, which show, respectively, the effect of a U.E.L. scenario for the total housing units in West Point Grey and an up-zoning scenario for the West End, the high-rise section of the city. Freeing land allows new development to take place, mere up-zoning is restricted because there are already buildings present, and demolition occurs only on initially substandard properties. A fault of the model is that there is no aging process to allow for redevelopment of depreciated properties. This is necessitated by the absence of any age distribution data for the standing stock of buildings. In either case there is a ceiling that, when reached, precludes further development without the freeing of more land in some way. Renewal and removal of standing stock can be read in interactively as policies. As noted earlier, however, they are not accounted for in the dynamics of the present model.

Figure 6. Housing stock—basic run, with backlink, and with backlink applied to housing and migration.

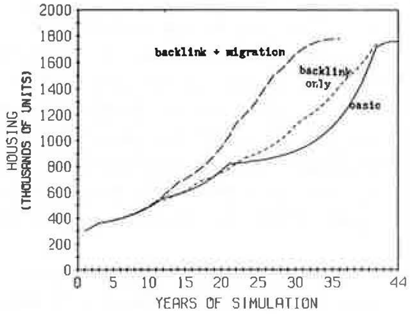


Figure 7. Housing starts—basic run, with backlink, and with backlink applied to starts and migration.

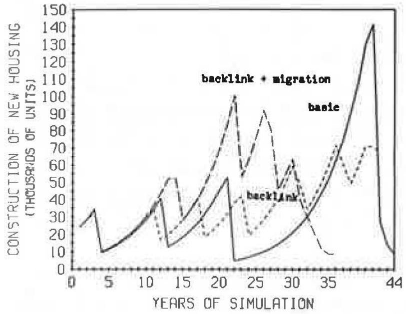


Figure 8. Housing stock—basic run, university endowment lands up-zoning, grand up-zoning in Vancouver, and farmland freeze (land act).

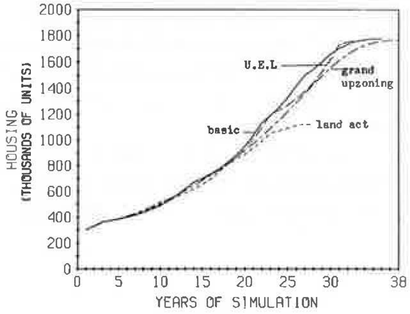


Figure 9. Housing starts—basic run, university endowment lands up-zoning, grand up-zoning in Vancouver, and farmland freeze (land act).

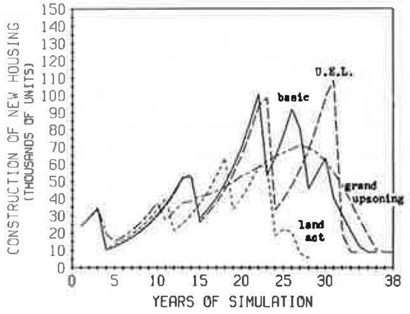


Figure 10. Housing stock—university endowment lands up-zoning and basic run.

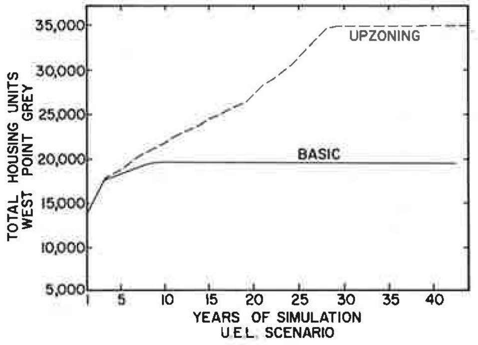
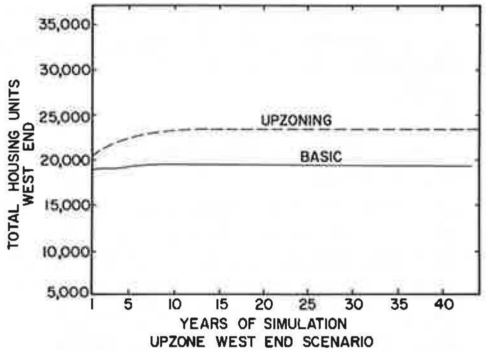


Figure 11. Housing stock—west-end up-zoning and basic run.



CONCLUSIONS

The initial development and subsequent evolution of the housing model contained in the larger simulation module have served as a convenient focus to illustrate the approach to modeling as well as the response of the model to various changes in its structure. The output presented will improve in quality with continued use, calibration, and specification of more appropriate initial conditions. It is presented in its current form primarily to illustrate the process and is not intended to be illustrative of the kind of output needed to aid detailed policy making. Such high-quality output can only result from the development process discussed. Ultimately, it is the usefulness of the module for specific land use and transportation policy testing that will prove its worth. Its refinement to a suitable high degree of realism and accuracy is the goal of the development process presented.

The general conclusions from the present simulation are that available land determines the stopping place for development; zoning to free or freeze-vacant land will have the biggest effect on development, and rezoning of land under existing use for housing or business will have a lesser overall effect. Most of the effect of zoning existing high-use land will be temporal. Full urban development cannot be halted by changing the land use patterns. The implications are that feedbacks to population growth and life-styles will ultimately prove to be the only effective means of limiting the growth of cities.

That the models are highly buffered with respect to policy changes is, as far as we are concerned, an indication of the usefulness of the models. Cities do not exhibit radical departures from their recent history, and the present output exhibits similar characteristics. Clearly different policies will have different impacts on the subareas of the region, although at the regional level (and this is consistent with equifinality) these impacts will be much less apparent. These relationships between subregional impacts and regional impacts are the micro-macro link presented.

By illuminating this link, we hope to develop more meaningful models. Most important, we hope to provide insight to decision makers about how this region works so that they can better plan for its alternative futures and ensure that we have alternatives in the future.

REFERENCES

1. L. von Bertalanffy. Robots, Men, and Minds: Psychology in the Modern World. Braziller, New York, 1967.
2. D. Boyce, N. Day, and C. McDonald. Metropolitan Plan Making. Regional Science Research Institute, Philadelphia, 1972.
3. H. J. Brown et al. Empirical Models of Urban Land Use: Suggestions on Research Objectives and Organization. National Bureau of Economic Research, New York, 1972.
4. M. A. Goldberg. Bay Area Simulation Study: Employment Location Models. Annals of Regional Science, Vol. 2, 1968, pp. 161-176.
5. M. A. Goldberg. Simulation, Synthesis and Urban Public Decision-Making. Management Science, Vol. 20, No. 4, part 2, 1973, pp. 629-643.
6. M. A. Goldberg. Quantitative Approaches to Land Management: A Survey Critique and Exposition of Recent Work, Monograph Series. Urban Land Economic Program, Faculty of Commerce, Univ. of British Columbia, Vancouver, 1973.
7. M. A. Goldberg and H. C. Davis. An Approach to Modeling Urban Growth and Spatial Structure. Highway Research Record 435, 1973, pp. 42-55.
8. M. A. Goldberg. Housing, Employment, Land Use, and Transportation: A Regional Simulation Model. Urban Land Economics, Faculty of Commerce, Univ. of British Columbia, Vancouver, Rept. 2, 1974.
9. ICO Real Estate Management Ltd. The Land Development Process. Real Estate Board of Greater Vancouver, 1974.
10. D. T. Kresge and P. O. Roberts. Systems Analysis and Simulation Models.

Brookings Institution, Washington, D.C., 1971.
11. D. B. Lee. Requiem for Large-Scale Models. Journal of the American Institute of Planners, Vol. 39, 1973, pp. 163-178.
12. S. H. Putman. Intraurban Employment Forecasting Models: A Review and a Suggested New Model Construct. Journal of the American Institute of Planners, Vol. 38, 1972, pp. 216-230.
13. D. C. Sweet, ed. Models of Urban Structure. D.C. Heath and Co., Lexington, Mass., 1972.
14. Vancouver Urban Renewal Study. Vancouver Planning Department, 1969.
15. P. F. Wendt and M. A. Goldberg. The Use of Land Development Simulation Models in Transportation Planning. Highway Research Record 285, 1969, pp. 82-91.