Software Requirements for Prototype Land Use Model

Preface

This document describes the requirements of the proposed software system for metropolitan land use modeling. We distinguish between the model and the software that implements the simulation of this model. The working of the model system and additional functional requirements form the requirements of the software. Since the feasibility of the model depends on the tractability of software methods, there will have to be iterative development for feedback of feasibility.

Overview

The proposed software-based system is to be used for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. It is based on the model explained in this document. The working of the model is part of the requirements of the software as the purpose is to simulate the model. In addition, there are features desired in the software that make it flexible and easy to use.

The model is characterized by:

- Land use and transportation models are based on individual choice theory (utility/profit maximization) and using logit models for implementation;
- Key decision makers and choices impacting urban development are modeled; in particular, the mobility and location choices of households and businesses, and the development choices of developers.
- Urban development is modeled as a dynamic process over time and space - as opposed to a cross-sectional or equilibrium approach;
- Land market is modeled as interaction of demand and supply, with prices adjusting to clear market;
- Governmental policy assumptions are explicitly represented. Impacts are evaluated by modeling market responses.

The software should:

- Implement the model system as a robust, efficient, and flexible software with an intuitive user-interface;
- Integrate database management and Geographic Information Systems;
- Allow the user flexibility in defining and specifying the model equations;
- Be able to run on Microsoft Windows 95 and NT platforms, and be portable to Unix platforms.

The model is based on the recognition that the process of urban development as it evolves over time and space is the composite outcome of the interactions of individual choices and actions taken by households, businesses, developers, and governments. The choices made by households are assumed to be those that maximize the household's utility, or perceived benefit, and the choices of businesses are assumed to be taken to maximize profits. Table 1 summarizes some of the key decision-makers and their decisions that pertain to urban development in general and to land use and transportation in particular. The proposed software presently focuses on the behavior of households, businesses, developers, and local governments.

Environmental issues may be added later. It should be able to incorporate other decision makers, choices, and relationships as needed.

TABLE 1 Decision-makers and Choices Affecting Urban Development

Decision-makers	**Choices/Actions (* proposed for inclusion in prototype model)**	
Market Decisions (Endogenous)		
Household	Mobility (move or stay) * Location (where to move) * Housing Type (single/multi; quality) *	Housing Tenure (rent/own) * Housing Rent/Cost (willingness to pay) * Auto Ownership *
Worker	Labor Force Participation Job Change Full-time/Part-time Multiple Jobs	Workplace Choice Wage to Accept Mode of Transport to Work Trip Linking
Business	Number of Employees * Wages to Offer Type of Space (office, campus, retail) * Tenure (rent/own)	Lease/Purchase Cost (willingness to pay) Mobility (move or stay) * Location (where to move) *
Developer	Land Purchase Infrastructure Investment New Development *	Redevelopment * Land Use * Density * Disposition (sell/lease)
Public Policy Decisions (Exogenous)		
Municipality	Tax Rate Tax Abatement/Incentives Zoning Land Use Plan * Urban Design	Development Fees Amenities (Parks) Services (Fire, Police) Infrastructure (Transportation, Water, Sewer)
Transit Agency	Transit Infrastructure *	Levels of Service * Transit Fares *
Lender	Loans for Mortgages	Development Loans Interest Rates
School District	Tax Rates	School Quality
Other Local, State, Federal Agencies	Fees, Laws, Regulations governing land use, transportation, environment *	Highway, Rail, Ports, Airports *

The Model

The flowchart in Figure 1 outlines the proposed model structure. It includes components reflecting the behavior of households, businesses, developers, and governments, all interfaced through the land market. The model draws on individual choice, or random utility theory, for its theoretical foundation, and builds on techniques of disaggregate choice modeling widely employed in models of mode choice. In extending the discrete choice modeling framework to households and businesses, we are able to develop a model framework that is both intuitive and transparent to the user, as well as theoretically sound and computationally tractable.

We assume as exogenous inputs that we have base year land use and population and employment (activity) data, as well as regional economic forecasts that will be used as control totals. The focus of the analytical tools is on modeling the dynamics over time and geographic space of the locations of businesses and households, the residential and commercial developments constructed by developers, and the policy interventions made by governments that influence or constrain the market. Two modules, demographic and economic transition, address the evolution over time of the distribution of households and business by type (e.g. age, income, businesses by industry) at the regional level.

In the household mobility and location module, households make decisions about whether to move or remain in their current residences, and if they choose to move, the select a home and neighborhood from the available options. These choices can be modeled in much the same way as mode choices of commuters, using multinomial or nested logit estimation techniques. In the business mobility and location module, businesses make similar choices regarding mobility and location choice. Household and business characteristics influence choices, as do locational attributes such as accessibility or prices.

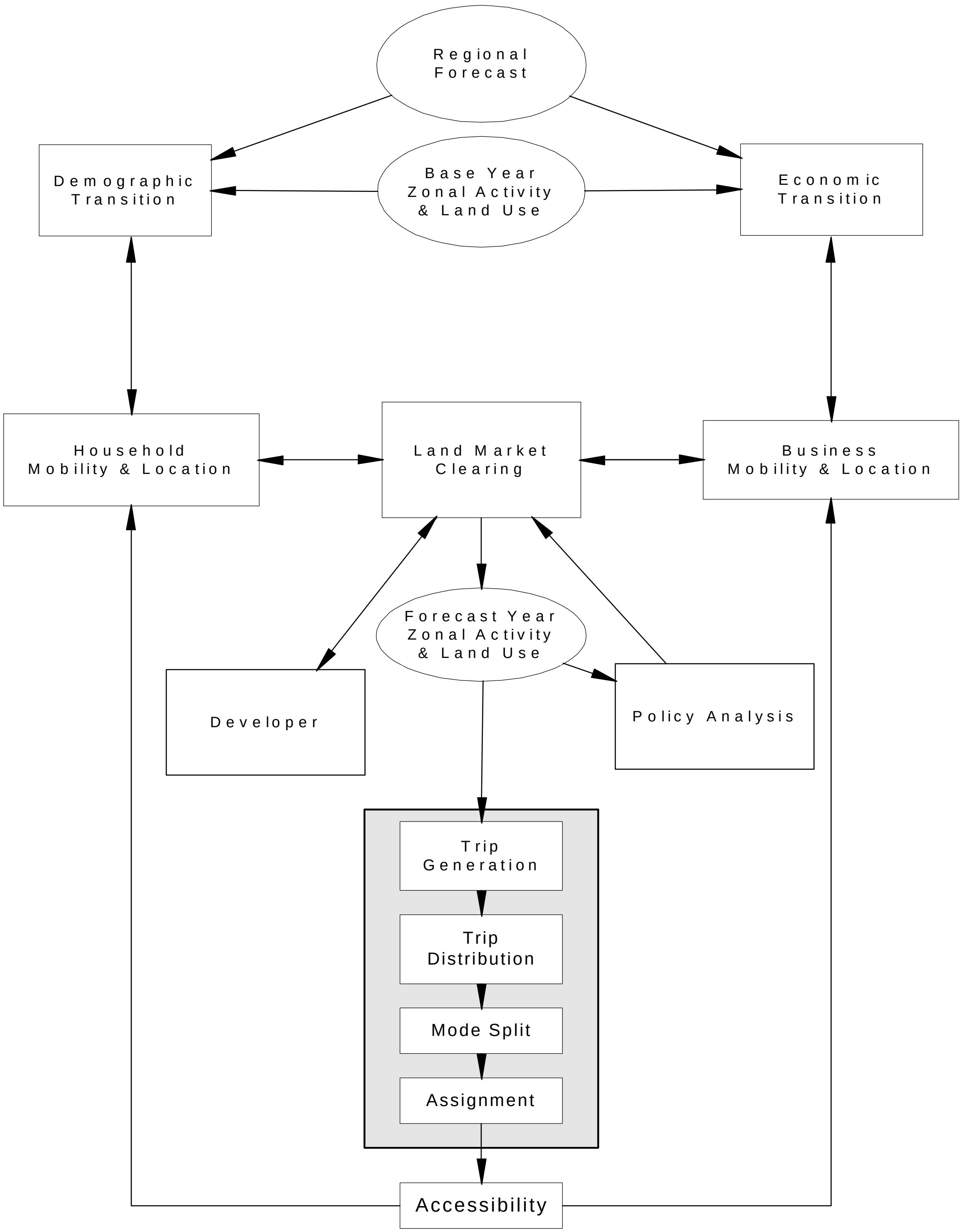
Developers make strategic choices about converting vacant land to developed uses, including the type of improvements, and density, based on their profitability expectations, and subject to constraints imposed by governmental policies such as zoning and infrastructure availability. These profitability expectations are in turn influenced by prior prices and revealed demand. We model these choices in the developer module.

Planners and analysts need to interact with the models to impose constraints or introduce policy assumptions, such as development impact fees, zoning, and land use plans. In the policy interface, we provide a structure for compiling these policy inputs from the user and allowing them to influence the choices made by households, businesses, and developers.

In the land market clearing module, we reconcile the competing demands for locations and structures among households and businesses, against the supply of space that has been created. The ratio of demand to supply at each location for each type of space (housing and commercial structures by type) will cause the prices of these structures to adjust to reflect market demand, and thereby influence their attractiveness for households and businesses as potential locations. An iterative market equilibration procedure alternates between price adjustment and the locational probabilities of moving households and businesses until the market clears and all movers are located (assuming initially no homeless).

These interactions of households, businesses, developers, and governments produce outcomes representing the distribution of population and employment, as well as the prices, uses, and density of land development. These are fed into the traditional four-step travel models to produce new travel times, costs and patterns by mode. The accessibility module uses the travel time matrices and zonal activity from the prior time period to develop measures of accessibility that, in turn, influence the attractiveness of locations for households and businesses as well as developers.

Land Use-Transportation Model Flowchart



In this manner, we can develop a robust and realistic assessment of the relative impact of specific transportation improvements or policies on land use in its broadest context: land use, density, price, and population and employment location.

Two approaches to handling households/businesses are possible:

1. Aggregate:

Using an aggregate distribution of households by type and projecting and scaling this distribution to match an exogenous forecast of households. This could incorporate cohort survival techniques to age population.

2. Disaggregate:

Using a sample households with household sample weights and adjusting the weights to match an exogenous forecast of the distribution of households by type.

Software:

In addition to these modules in the flowchart, we identify software modules that will be needed to implement the models for use by planners and analysts. These are: a user-interface that is intuitive and powerful, an interface to GIS systems such as Arc/Info, Intergraph, and Mapinfo to allow the editing and viewing of inputs and results, a module for assisting users in the translation and loading of data from existing databases into the system, and interfaces to the predominant transportation planning tools available commercially, such as TRANPLAN, MINUTP, and EMME/2. We need a software environment that is flexible and robust enough to allow portability to multiple agencies and environments.

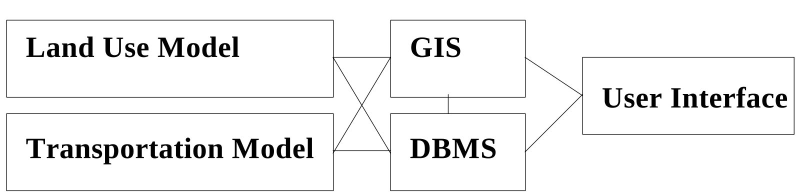
The user should be able to define household types as a joint distribution of household characteristics, such as number of persons, number of workers, income etc. The user should have the flexibility to define categories to group households and choose characteristics to use in defining the household types.

Zonal system should be equivalent between land use and travel models, although some land use modeling operations could be extended to higher resolution (smaller zones) to take advantage of more detailed data where it is available (such as parcel databases), that could be aggregated to the transportation zone level detail. This could be particularly useful in extending the models to deal with micro-locational issued such as urban design issues related to transit accessibility.

GIS Requirements:

The software should provide the users an ability to display maps with input data and results such as thematic maps and bandwidth plots of network values. Users should be able to edit input data using a GIS interface. The GIS functionality should not impose unnecessary overhead in terms of processing time or memory requirements. Conversion of GIS data to and from primary formats such as Arc/Info and Intergraph should be supported.

The software components can be visualized as interacting modules representing model components for land use and transportation, the database management system, the GIS, and the user interface:



Model components

Demographic Transition Model

The first of the land use models is designed to 'age' households and predict the transition of households from one type to another to reflect changes occurring over a given time increment between each application of the model. Since the proposed land use model is not a disaggregate microsimulation model, but deals instead with groups of households and employment of particular types, it is less straightforward to apply microsimulation techniques such as the ageing of individuals, probabilities of birth or death, divorce, marriage, job change, retirement, and so forth. Rather, an approximation of these individual outcomes must be developed by ageing households probabilistically, and predicting the probability that the households moving from one age group to another will change status in any of their other characteristics.

Initially, the proposed method for implementing this model is based on developing joint distributions of households by each dimension of household characteristic to be used in the model system (income, household size, number of workers, presence of children, etc.) for two or three historical periods, using Public Use Microdata as the principal source. A transitional probability matrix for moving from one cell to another can be estimated from these historical joint distributions, using an Iterative Proportional Fitting (IPF) technique to scale the transitional probabilities so that the product of the base year joint distribution and the transitional probabilities produce the aggregate (marginal) household totals by type provided exogenously in the regional forecast[1]. We separate the transitional probabilities into three components: transition from one household type to another based on demographic and economic changes such as ageing, household sizes, and income; inmigration to the region; and outmigration from the region.

$$H_{ht} = \sum_{h'} H_{ht-t} p_{hh'} + H_{ht-t} p_{hi} + H_{ht-t} p_{ho}$$

where

H_{ht} is the regional total of households of type h in time t

$p_{hh'}$ is the transitional probability of household type h to household type h' from time t-1 to time t

P_{hi} is the inmigration rate for households of type h

p_{ho} is the outmigration rate for households of type h.

This simple approach could potentially be extended by integrating the cohort survival technique as an initial step, providing a preliminary transitional matrix that would be subsequently scaled using IPF. The cohort survival technique applies death probabilities to age-sex cohorts (typically single or 5-year cohorts), then applies a probability of childbearing using fertility rates for women by age cohort. In the final step, net migration probabilities are applied to account for in and out migration. Integrating cohort survival into the household joint household distribution scaling method could prove complex, however, since cohort survival is based on counts of individuals by age and sex, and does not easily map onto distributions of households.

---

[1] IPF is an iterative procedure that scales a multidimensional matrix to user-supplied totals for some of the margins of the dimensions. For example, a simple joint distribution of households by income and household size would potentially be scaled to fit a distribution of households by income, or a distribution of households by household size, or the independent distributions of both.

Once the transitional probabilities are computed, we make the assumption that these transitional probabilities do not vary geographically within the region, allowing us to apply them at the zonal level to the base year distribution of households by type by zone. We apply initially only two components of the transition probabilities: the probability of transition from one household type to another, and the probability of outmigrating from the region. The transition from one household to another alters the classification of households within a zone, but does not affect the total household count within the zone. The application of theoutmigration rates by type of household at the zonal level will be accompanied by the removal of the predicted number of households, and the reclassification of housing units from occcupied to vacant to account for this loss of households. The third component of the transition probabilities, theinmigration rate, is applied in the aggregate only, and used to compute the initial set of movers by household type that will be located in the residential location model, along with other households predicted to move within the region in the residential mobility model.

Household Mobility Model

Once the Demographic Transition Model has accounted for transition between household types and outmigration from the region, the second step is to predict how many households of each type in each zone will move over the next five years and how many will stay in their current residence. The prediction and application of these mobility probabilities is the function of the Residential Mobility Model.

For households, mobility probabilities can be estimated for each type of household using a household survey, or if necessary, from the Public Use Microdata Sample file from 1990. This will reflect differential mobility rates for renters and owners, and households with and without children, etc. Logit models predicting probability of household mobility over a given span can be calibrated with independent variables based on the household characteristics such as age group, children present, number of workers, and housing tenure, and the ratio of housing cost to household income, each interacted with a dummy variable representing the choice to stay.

$$p_h(m) = \frac{1}{1 + e^{(V_m)}} + \varepsilon$$

where p(m) is the probability of a household of type h moving

V_m is the systematic component of the utility of moving.

A potential specification for the mobility equation is:

$$V_m = \$_1 + \$_2 a + \$_3 c + \$_4 w + \$_5 t + \$_6 c$$

where a is the age group of the head of the household

c is a dummy variable for whether children are present in the household

w is the number of workers in the household

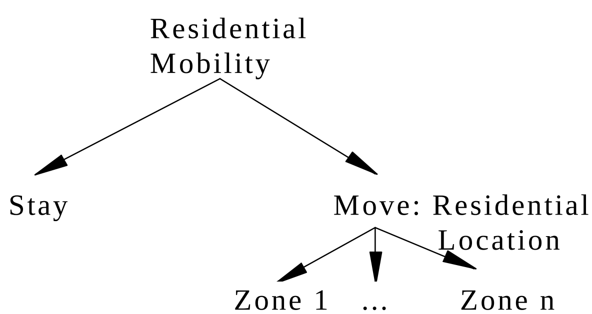
t is the housing tenure of the household

c is the ratio of housing cost to income

Application of the residential mobility model requires subtracting households by type probabilistically from the land use file by zone, and adding equivalent household counts by type to the unallocated inmigrant households by type calculated in the Demographic Transition Model. The combination of new and moving households serve as a pool of households to be located in the residential location choice model. Housing vacancy will be updated as movers are subtracted, making the housing available for allocation in the residential location choice model.

Since the relative attractiveness (utility), cost, and availability of housing in other locations compared to a household's current location may influence their decision to move, we could structure the mobility model as the higher level in a nested logit model with the lower level being residential location choice. In this way, we can use information about the relative utility of alternative locations, compared to the utility of the current location, in predicting whether households will move. Essentially, this represents a hierarchical choice of current residence vs. all other residences, and within the all other residences branch we further disaggregate the alternatives to zones and housing types in the residential location model.

Figure 2
Residential Mobility/
Residential Location



Household Location Choice Model

The function of the household location choice model is to assign moving and new households to a specific zone and housing type, as well as choices regarding housing tenure. It represents the demand of households for locations and housing types, and must be balanced with the supply of housing and the competition for locations by employment. The model will be specified as a multinomial logit:

$$P_{ihtd} = \frac{\exp(V_{iht})}{\sum_{i} \exp(V_{jht})} + \varepsilon$$

where P_{iht} is the probability of choosing zone i, housing type h, and housing tenure t

V_{iht} is the systematic utility of location i, housing type h, and housing tenure t

ε is the error

The utility of a location is a function of household characteristics, zonal characteristics, tenure and housing type:

$$V_{iht} = f(hZ, hT, hH)$$

where h is an array of household characteristics

Z is an array of zone characteristics (e.g. accessibility)

T is a dummy variable for tenure = own

H is an array of dummy variables for housing types, plus housing quantity and cost

The household characteristics are interacted with zonal characteristics, tenure, and housing type in order to differentiate the attractiveness of these characteristics to different household types, such as the attractiveness of homeownership to households with children, or of the income distribution of the zone to households of different income levels. In addition, in a discrete choice econometric framework, household characteristics cannot be entered into the model except by interacting them with characteristics of the alternatives, since there would be no variation in these characteristics across alternatives. An initial set of variables proposed for inclusion are listed below, grouped into household, zonal, housing type and tenure:

Household characteristics:

a is the age group of the household head

c is the children present status of the household

w is the number of workers in the household

i is the income group of the household

Zonal characteristics:

$za_1...za_4$ is the zonal percent of households in age group 1...4

zc is the zonal percent of households with children

zt is the zonal percent of households who own their homes

$zi_1...zi_4$ is the zonal percent of households in income group 1...4

zha is the zonal average housing age

zba is the zonal accessibility to employment in the basic sector

zra is the zonal accessibility to employment in the retail sector

zsa is the zonal accessibility to employment in the service sector

zma is the zonal accessibility to employment in the military sector

zea is the zonal accessibility to elementary schools

zpa is the zonal accessibility to population

zcbd is the zonal accessibility to the CBD

zpd is the zonal population density

zed is the zonal employment density

zd is the distance of the zone from the zone of the prior residence

d is an array of district-specific dummies, to allow testing as a hierarchical model

Housing Type and Tenure

o is a dummy variable for tenure = own

d is a dummy variable for households that double up in occupying a housing unit

$zhq_1 \ldots zhq_t$ is the zonal housing quantity available (vacant) by type and tenure

$zhc_1 \ldots zhc_t$ is the zonal housing cost for housing by type and tenure (to be entered in the model as the ratio of housing cost to income)

Although all of these variables may be anticipated to influence the choice of location, tenure, and housing type, a final specification of the model will clearly depend on the results of calibration. The model calibration can be tested with data from a household survey. The model will be calibrated using a random sample of alternative locations, which has been shown to provide consistent estimates of the coefficients.

In application for forecasting, the logit equation will be applied to all alternatives for each household type being allocated, producing the demand for housing of each type and tenure in each zone. All of the independent variables will be updated in each forecast year increment, based on results of the prior period iteration of the model set. Total housing demand in each zone is computed by summing demand across household types.

Economic Transition Model

Businesses will be classified using two characteristics: industry and size. Users may aggregate or disaggregate the grouping of industries or size categories as needed. Possibly additional characteristics of business could be added.

Just as households may age and change status of children, income, number of workers, move out of the region, or be created, business establishments may increase or decrease in employment, go into or out of business, or move into or out of the region. There is evidence of very high levels of turnover (birth and death of businesses) in some industries, and especially among smaller establishments. It is possible that failure to account for this turnover, even with rough estimates, may significantly underestimate just how fluid employment location is in response to changing accessibility patterns and market forces such as land prices or the construction of new space.

$$B_{nst} = \sum_{s'} B_{nst-t}\, p_{nss'} + B_{nst-t}\, p_{nsi} + B_{nst-t}\, p_{nso}$$

where

B_{nst} is the regional total of businesses of industry n and size s in time t

$p_{nss'}$ is the transitional probability of a business in industry n from size s to s' from time t-1 to t

P_{nsi} is the inmigration and creation rate for businesses of industry n and size s

p_{nso} is the outmigration and destruction rate for businesses of industry n and size s.

As in the demographic transition model, the transition rates from one size to another can be applied to the distribution of businesses by zone, as can the outmigration/destruction probabilities, with appropriate accounting of the vacating of commercial space. Inmigration and creation of businesses, as in the demographic model, are accumulated by type and allocated in the business location model.

Business Mobility Model

In parallel with the residential mobility model, the business mobility model predicts the probability that establishments of each type will move from their current location or stay during a five year forecast period. If a business establishment file can be obtained for two years, preferably five years apart, mobility rates by industry and size of establishment will be computed. Otherwise, estimates of these rates will need to be developed from other sources.

Logit models predicting probability of business establishment mobility over a given span can be calibrated with independent variables based on the establishment characteristics of industry and size group, interacted with the alternative-specific dummy representing the choice to stay.

$$p_{ns}(m) = \frac{1}{1+e^{(V_m)}} + \varepsilon$$

where p(m) is the probability of a business of industry n and size s moving

V_m is the systematic component of the utility of moving.

A simple specification of this utility would be:

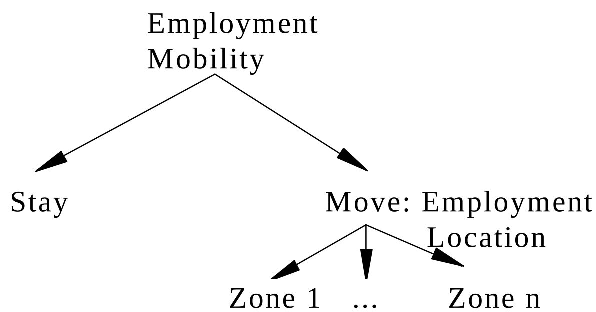
$V_m = \$_1 + \$_2 i + \$_3 s$

where i is the industry

s is the establishment size

As in the residential case, the relative attractiveness (utility), cost, and availability of commercial space in other locations compared to an establishment's current location may influence their decision to move, so we structure the mobility model as the higher level in a nested logit model with the lower level being employment location choice. In this way, we use information about the relative utility of alternative locations compared to the utility of the current location in predicting whether establishments will move. This represents a hierarchical choice of current location vs. all others, and within the all other branch we further disaggregate the alternatives to zones and space types in the employment location model.

Figure 3
Employment Mobility/
Employment Location



As in the household mobility model, the application of this model requires subtracting establishments by type probabilistically from the zonal file. These counts will be added to the unallocated new establishments by type calculated in the economic transition model. The combination of new and moving establishments serve as a pool to be located in the employment location choice model. Vacancy of nonresidential space will be updated, making space available for allocation in the employment location choice model. Average space per employee will be stored in a lookup table, and is an exogenous input in the initial model specification.

Business Location Choice Model

The employment location choice model is similar in concept to the residential counterpart. The form of the model remains multinomial logit:

$$P_{il} = \frac{e^{V_{il}}}{\sum_{il} e^{V_{jk}}}$$

where P_{il} is the probability of choosing zone i and land use type l (tenure choice is omitted from the employment model since we have no data on it)

V_{il} is the systematic utility of location i and land use type l

where V_{iht} = f(e*Z, e*S)

The utility can be defined as a function of establishment characteristics (industry and size)m zonal characteristics, and characteristics of commercial space:

$$V_{il} = f(eZ, eS)$$

where e is an array of establishment characteristics

Z is an array of zone characteristics (e.g. accessibility)

S is an array of dummy variables for space types, plus quantity and cost

The establishment characteristics are interacted with zonal characteristics and nonresidential space type in order to differentiate the attractiveness of these characteristics to different establishment types, such as the attractiveness of high-rise office in the CBD to large establishments in the service industry. The initial variables to be used are listed below, grouped into establishment, zonal, and nonresidential space type:

Establishment Characteristics:

i is the industry of the establishment

s is the employment size group of the establishment

Zonal Characteristics:

z is an array representing the mix of zoning within the zone

zb is the zonal employment in the basic sector

zr is the zonal employment in the retail sector

zs is the zonal employment in the service sector

zm is the zonal employment in the military sector

zh is the zonal employment in the hotel sector

zas is the zonal average age of nonresidential space of type l

zba is the zonal accessibility to employment in the basic sector

zra is the zonal accessibility to employment in the retail sector

zsa is the zonal accessibility to employment in the service sector

zma is the zonal accessibility to employment in the military sector

zpa is the zonal accessibility to population

$zai_1 ... zai_4$ is the zonal accessibility to households of income group 1...4

zcbd is the zonal accessibility to the CBD

zpd is the zonal population density

zed is the zonal employment density

zd is the distance of the zone from the zone from the prior location

d is an array of district-specific dummy variables, to allow testing as a hierarchical model

Nonresidential Space Characteristics:

zqs is the zonal quantity of available nonresidential space of type l

zsc is the zonal cost per square foot of nonresidential space of type l

Calibration of the model can be based on a geocoded establishment file (matched to the parcel file to link employment by type to land use by type). If movers can be identified in the file, location choice can be calibrated using establishments that moved within the model simulation period. As in the residential location choice model, the application of the model produces demand by each employment type for space of each nonresidential type, by zone. The total demand for space by zone by type is computed by summing demand across employment types and space types.

The user may want to identify "major businesses" that will be excluded from the mobility location choice models. These can be perhaps identified as non movers. The user can input a database of major establishments in addition to aggregate distribution of businesses by type used in models (geocoded by zone). The user may add into a forecast a known business relocation.

Land Market Model

The land market model reconciles demand for land from households and establishments with each other, and with the available land supply in every forecast interval. It adjusts land prices according to the ratio of demand to supply in each zone. Since prices enter the locational utilities for businesses and households, an adjustment in prices will alter their locational preferences. Similarly, any adjustment in land prices alters the preferences of developers to build new construction by type of space, and the density of the construction. We assume the developers respond to price stimuli in the previous time period, since the time scale for physical construction is considerably longer than the time scale for residential and business relocation. The price adjustment mechanism therefore iterates with the location choice components, causing an adjustment in the locational preferences of businesses and households, until the market clears. The developers respond to price changes in the next time period.

The supply of housing and commercial space used in any iteration is drawn from existing vacant structures plus the new construction and redevelopment of structures that occurred over the most recent period. The new construction in a forecast interval could include committed, proposed, and potential development projects known to the planner or analyst operating the model. Any developments added by the analyst represent an exogenous policy input.

Access Model

An accessibility procedure reads the congested zonal impedance matrix (from the travel model in Tranplan, Emme/2 orMinuTP) and the forecast land use file, and creates accessibility indices for use in the household and business location choice models. The general framework is to summarize the accessibility from each zone to various urban activities for which accessibility is considered important in residential or employment location choice. The accessibility indices referenced in the utility functions of the residential and employment location logit models were:

zba is the zonal accessibility to employment in the basic sector

zra is the zonal accessibility to employment in the retail sector zsa is the zonal accessibility to employment in the service sector

zma is the zonal accessibility to employment in the military sector

zpa is the zonal accessibility to population

$zai_1 ... zai_4$ is the zonal accessibility to households of income group 1...4

zcbd is the zonal accessibility to the CBD

These accessibility indices could each be measured using the gravity formulation:

$$Access_i = \sum_{1}^{J} \frac{Activity_j}{(Impedance_{ij})^2}$$

where impedance is generalized cost, or travel time.

The user should be able to change exponent (2 can be changed) and substitute variables and functions for impedance or could replace denominator function.

Accessibility will normally be calculated from a zonal impedance matrix but optionally the loaded network could be used for this purpose, to develop detailed accessibility indications. A GIS could be used to compute theseaccessibilities.

Separate accessibility indices could be evaluated for transit and highway, to test land use sensitivity to separate effects. For example, low income households may be more sensitive in their residential location to the accessibility of jobs and shopping by transit, whereas high income households might be more sensitive to auto accessibility to jobs and shopping. Transit accessibility needs to be considered separately, and may require additional and very different measures than auto accessibility. Common measures of transit accessibility, for example, focus on the non-motorized access to the transit station, which is a measure requiring substantially different kinds of information and geographic resolution than auto access measures.

An alternative specification of these accessibility indices would be to measure the percentage of the regional total of each activity that is within a given commuting time. While this approach is perhaps easier to interpret, it may have a scaling problem. As the region grows over time, the percentage of the regional total accessible to each zone will tend to diminish even if accessibility remains constant, since the denominator is increasing. However, this scale effect may fall out of the residential and employment logit models, since the relativeaccessibilities are what influence choice probabilities. This and other specifications of the accessibility indices will also be evaluated such as number of activities of different types available within a specifies distance. Substantial effort will be needed in testing alternative accessibility measures, related to auto and transit, at regional and local scales, to appropriately specify the most effective combination of measures.

Developer Model

We will model the activity of the developers such as building new housing and non-residential buildings and re-developing existing development. Developers may be assumed to be myopic in their expectations, making predictions about market trends and expected profits from alternative development plans based on recent trends. The assumption of imperfect information based on recent history is entirely consistent with the familiar cycle of real estate speculation, overbuilding, and bust. Over longer forecasting horizons, we would expect the model to predict behavior without the short-run volatility present in actual real estate markets.

We can specify the developer decision as one of converting land from existing status to one of a finite set of alternative land uses at various densities:

$$p_i(ud) = f(P_{lt-1}, P_{lt-2}, C_u, D_u, Z_i)$$

where $p_i(ud)$ is the probability that a developer will convert a unit of vacant land in zone i to use u at density d

P_{lt} is the price of land at time t in zone i

C is the cost of constructing one unit of structure of type u

D is the cost of demolishing one unit of structure of type u

Z is an array of zonal characteristics, including the land use mix and density, and accessibility measures.

Since land prices enter the profit function of developers by affecting their costs and expected profit from development, it influences their development probabilities. In addition, the changes in price over the past two periods form a trend on which the developer bases the expectation of future price appreciation. As noted earlier, the time scale of physical development is such that it would be inappropriate to equilibrate development within a single time period. Rather, developers are modeled as responding to lagged price information and forming myopic expectations about the future.

The probability predictions can be applied at the zonal level and scaled to meet regional totals for the regional quantity of new construction by type for each time interval. If regional constraints for new construction are not given, the zonal probabilities would be scaled to meet the predicted space requirements of the regional forecast business and household growth.

Policy Analysis Module

There will be a policy analysis module that allows users to specify policy inputs and assumptions, generate scenarios and compare them, compute evaluation measures, and query the database of results. Users may edit policy variables imposing constraints (e.g. land use plans) or influencing costs (development fees). The software should provide the user reports, charts and maps useful in evaluating policy outcomes or objectives.

Calibration

The user should be able to calibrate the entire model system using longitudinally observed data from prior census data and other sources. The calibration procedure should produce a ‘maximum likelihood’ estimate of the coefficients in all of the model equations.

Validation

The software should produce detailed reporting on errors in the predictions generated during the calibration process and the user should be able to specify alternative summaries for these error reports.

Description of the Data

The data files used in the land use model system are described in the following sections. The data files include:

1. Regional Forecast

2. Land Use
3. Committed/Proposed Development Projects
4. Major business establishments.

1. Regional Forecast

This input file is an exogenous input to the land use forecasting models, which are designed to allocate regional changes in households and employment. The data items required as inputs to the land use models are employment by industry and total households. Employment will contain counts of establishments by industry by size of employment, and total employment by industry. Households will be stored as a distribution of households types for example age of head, whether children are present, household income group, and total population.

2. Land Use

This database is the primary database that maintains the accounting of households, housing, employment, nonresidential space, land use, and prices. In order to capture information about household characteristics needed to predict location, mobility, and travel behavior, households are stratified for example by age of head, whether children are present, number of workers, income, and housing tenure.

Age of Head:	Under 35, 35-49, 50-64, 65 or over
Children:	Yes, No
Workers:	None, One, Two or more
Income:	Under $20K, $20-40K, $40-60K, Over $60K
Tenure:	Rent, Own

These characteristics are each important in predicting mobility, location, or travel behavior. Age of head captures one aspect of life cycle, and income, tenure, probability of having children, number of workers, and travel behavior are all likely to be influenced by age. Children present in a household influence choice of residential location and travel behavior. Number of workers influence income, mobility, location choice, and travel behavior. Tenure affects mobility, location choice, and travel behavior. In addition to these disaggregate effects, the proportion of households in a neighborhood that are of a given age group, or income group, or that have children, or are homeowners, all affect the attractiveness of the neighborhood to other households. Keeping track of these household characteristics enables us to provide a rich neighborhood context that adds realism to the location choice model.

Employment will be stratified primarily by industry and establishment size. Example of Industry groupings are as follows:

Military
Government
Hotel
Agricultural
Transportation, Communications, Utilities
Industrial (Manufacturing, Wholesale)
Finance, Insurance, & Real Estate
Service (other services)
Retail
Construction

Establishment size categories corresponding to county business patterns: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500+, which could be condensed by the user.

The development of the base year land use file will require integration of several data sources and a procedure to estimate their joint distribution. Household information is available at a small area level in the 1990 (also available for 1970 and 1980) Census STF1A and STF3A files, but insufficient information is available about joint distributions of household income and structure, and the allocation of households to housing by type. The household survey being conducted for calibration of travel models will provide the primary source for calibrating the residential models. Household records provide full descriptions of the demographic and economic structure of the household, as well as their housing characteristics and location. GIS techniques will be used to augment the data available directly from the survey with detailed spatial descriptions of the neighborhoods and locations occupied by the households. Depending on the sample distribution, some aggregation of the initial target stratification of households will be required.

Creation of the employment/space/zone allocation will use an establishment file, ideally address-matched to the parcel file. This will provide the joint distribution of employment by industry, nonresidential space, and zone needed as input for the land use models. The employment data will be maintained in the form of establishments by industry by size category, since this more closely reflects the decision-making units, and size of establishment is important in predicting mobility, location choice, trip attractions, and potentially, policy responsiveness to TDM (Travel Demand Management) measures.

3. Committed/Proposed Development Projects

A database of known or anticipated development projects by zone may be used to add to the supply of new construction in any forecast time interval. It would be treated as an exogenous policy input.

4. Major business establishments

The user may want to identify "major businesses" that will be excluded from the business mobility and location choice model. These can be perhaps identified as non movers. The user can input a database of major establishments in addition to aggregate distribution of businesses by type used in models. (would be geocoded) The user may add into a forecast known business relocations.