Title
BASS II: A New Generation of Metropolitan Simulation Models

Author
Landis, John

Publication Date
1992-04-01

Working Paper 573

BASS II: A New Generation of Metropolitan Simulation Models

John Landis

April 1992

University of California at Berkeley

$5.50

Working Paper 573

BASS II: A New Generation of Metropolitan Simulation Models

John Landis

BASS II was developed at the Institute of Urban and Regional Development, University of California at Berkeley, by Ted Bradshaw, Ted Egan, Peter Hall, John Landis, Ayse Pamuk, David Simpson, Qing Shen, and Michael Teitz. Support for the development of BASS II was provided by the California Policy Seminar.

University of California at Berkeley
Institute of Urban and Regional Development

I. INTRODUCTION

This paper summarizes the development of the Bay Area Simulation System (BASS II), a prototype of a new generation of metropolitan area forecasting and simulation models. The primary purpose of BASS II is to provide a framework for simulating how realistic growth and development policies, applied either regionally or locally, might alter the pattern and intensity of urban development in the ten-county San Francisco Bay Region (see Figure 1).[1]

BASS II is both similar to *and* different from other regional forecasting models. In BASS II, as in most other regional forecasting models, the process of forecasting *how much* population or economic growth (growth forecasting) will occur is separated from the process of determining *where* that growth will occur (growth allocation). A second similarity is that BASS II is designed to be used iteratively; that is, growth projections developed for 1995 are later used as inputs into the growth projections developed for the year 2000.

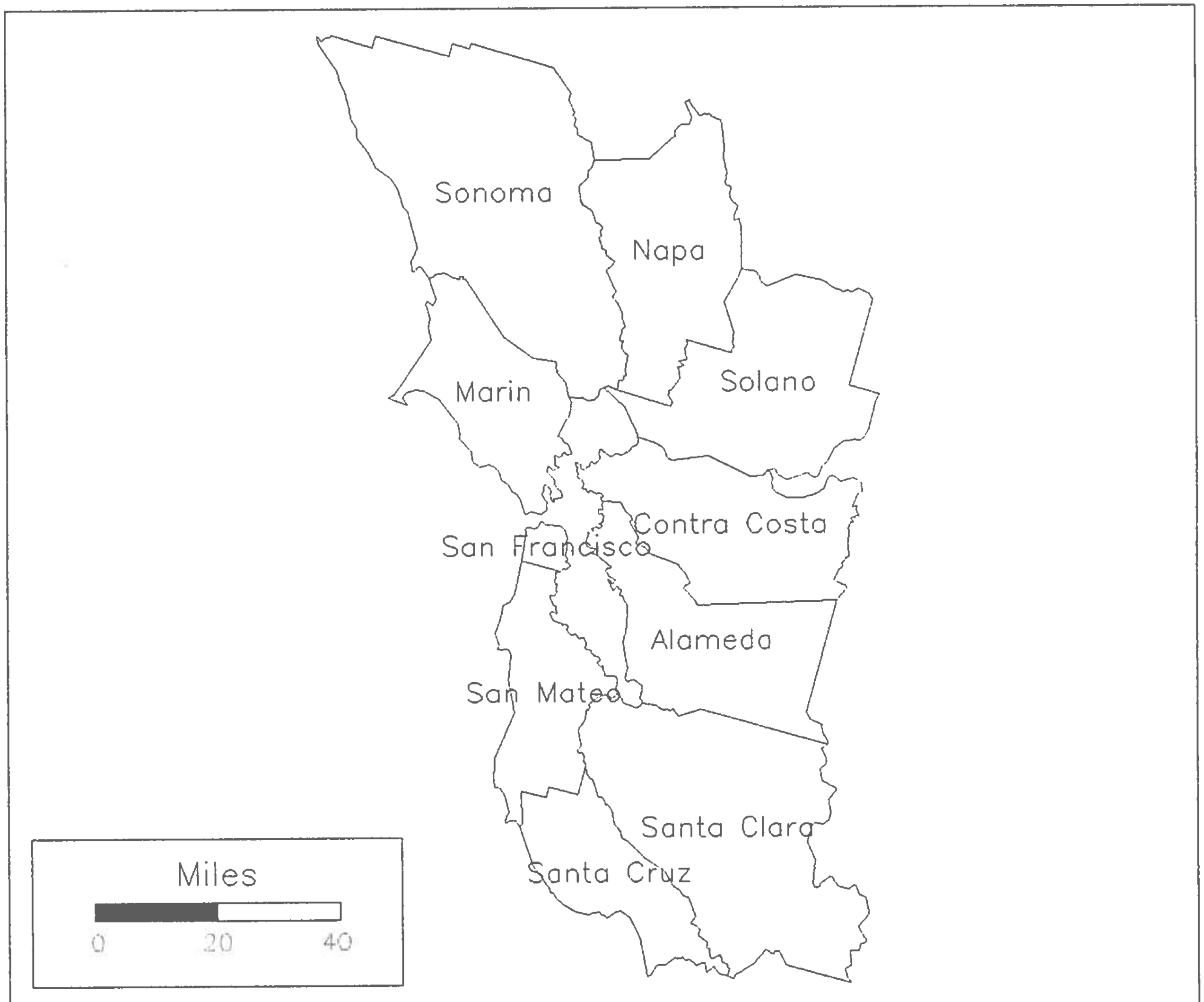
A major operational difference between BASS II and other regional forecasting models, however, is that BASS II relies on highly accurate computer map layers as a source of data and as a mechanism for analysis. These map layers are encoded, updated, and incorporated into the analytical and decision-making components of BASS II through the use of ARC/INFO, a Geographic Information System (GIS). To our knowledge, BASS II is the first regional model that directly incorporates GIS.

There are also theoretical differences between BASS II and other regional forecasting models. Previous forecasting models have tended to incorporate measures of transportation access (usually measured in terms of zone-to-zone highway travel times) as the *primary* determinant of new urban development patterns. BASS II, in contrast, incorporates transportation access (measured several different ways) as one of many variables that affect the allocation of growth to different parts of the region.

A final point of difference between BASS II and other regional models is the way in which population growth is projected. Most other regional forecasting models project population growth at the regional scale, and then allocate it "downward" to individual cities or sub-areas. BASS II takes the "bottom-up" approach of forecasting population growth at the city or sub-area level, and then aggregating it "upward" to yield total regional growth.

Experienced regional modelers will recognize that the BASS name is not new. BASS was the name given to a pioneer of the first generation of regional growth models. BASS was developed

Figure 1: 10–County San Francisco Bay Area



by William Goldner and the staff of the Association of Bay Area Governments in 1965-66 to generate twenty-year population and economic projections for the then nine-county San Francisco Bay Area. Based on the well-known Lowry Model, the original BASS Model involved relatively little *new* theory. Rather, by taking advantage of the capabilities of modern computers to process thousands of calculations, BASS made the theory useful as a tool for regional planning. Although computers and regional development theory have advanced considerably since the mid-1960s, we like to think that BASS II expands and carries on this most worthwhile tradition.

The remainder of this paper is organized into nine parts: Part II summarizes the four design principles that guided the development of BASS II; Part III schematically presents the overall structure of the BASS II model; Part IV presents the regression equations which form the core of the "Bottom-Up" Population Forecasting Model; Part V explains the organization of the Spatial Database, the GIS part of BASS II; and Parts VI and VII introduce the structure of the decision rules used to allocate projected population growth to specific areas and to annex newly developed areas to existing cities; Part VIII explains the several ways in which BASS II can be used to simulate development policy changes; Part IX offers an agenda for additional model development; and Part X reiterates some of the unique features of BASS II.

II. BASS II DESIGN GUIDELINES

Four principles guided the development of BASS II. These principles reflect a series of compromises between the desire for theoretical consistency and immediate policy relevance. The four principles include:

1. ***A Spatial Simulation System, Not a Regional Forecasting Model.*** First, and foremost, the BASS II had to be capable of simulating the spatial growth of the metropolitan area as it actually occurs— site-by-site, parcel-by-parcel, block-by-block, and city-by-city. County-level and zonal growth totals, such as those produced by most current regional forecasting models, were judged to be too aggregate to provide a clear picture of the spatial processes of urban growth. In this context, being able to simulate the specific locations ***where*** growth might occur was viewed as being as important as projecting how ***much*** growth might occur.

   Our insistence that BASS II be spatially accurate magnified both the complexity of the model, and the volume of data required to build it. This, in turn, required incorporating elements of a Geographic Information System (GIS) into BASS II. Given the current state of the GIS technology, it soon became clear that the only GIS system capable of handling the variety of geo-spatial transformations required, as well as the volume of data that would be incorporated into the model, was ARC/INFO.

2. ***A Policy-Relevant Approach.*** As the term "simulation" suggests, BASS II is designed to simulate alternative regional development futures as a function of specific policy changes, rather than produce a single best-guess forecast. As a result, the model must be usable and reliable over a wide range of real policy proposals. Such proposals might be regulatory in nature (e.g., significant down-zoning of undeveloped, unincorporated areas) or investment-oriented (e.g., construction of specific new transportation facilities or wastewater facilities). Moreover, BASS II would have to be capable of simulating the impacts of policies undertaken by various governmental units, including state government, local government, special districts, and (potentially) regional government. Finally, BASS II would have to be capable of simulating a complex of policy proposals; that is, simultaneously incorporating different policy initiatives as adopted in different jurisdictions.

   The requirement of policy relevance mandates that political jurisdictions—cities and counties —form one of the basic units of analysis of BASS II. Under California law, almost all development policies are adopted and implemented by individual cities, counties, or special districts. Current efforts to "regionalize" development planning in California notwithstanding, we expect that most decisions about how much new development can go where will primarily remain the province of local governments.

   The decision to use cities and counties as one of the basic units of analysis had both positive and negative ramifications. On the positive side, because most census and state population data (and some economic data) are reported at the jurisdictional level, the task of collecting some types of data is simplified. Complications arise, however, when trying to simulate the generation and expansion of unincorporated population centers— places that may have a popular identity but have not been formally incorporated, and thus for which data is usually unavailable.[2]

   The requirement of policy-relevance mandates that BASS II have a "bottom-up" structure. Traditional urban development models generally have a "top-down" structure: population and economic growth increments are projected for large areas (typically regions or counties) and then allocated to smaller units (typically traffic analysis zones). Under such a structure, local policy initiatives affect only the differential allocation of the growth increment, not the size or nature of the growth increment. In a "bottom-up" model, economic and population growth is projected for each unit of analysis, and then aggregated. Thus, in a bottom-up model, local policy initiatives affect not only the location of population and employment growth, but also its size and quality.[3]

3. ***A Tool Useable by and Understandable to Planners and Policy Analysts, Not Just Technicians.*** Most regional forecasting models make sense to the analysts and technicians who develop them, but not necessarily to the policy-makers and planners who try to use them. To avoid this problem, the growth-allocation mechanisms in BASS II are designed around a series of transparent and changeable decision rules (e.g., "limit development densities

to four units per acre in this city") rather than mathematical algorithms. Thus, it is possible to trace how specific policy changes will affect the pattern and level of population growth locally and regionally.

4. ***An Expandable System.*** Most regional forecasting models are constructed "all-of-a-piece"; that is, as a series of sequential and inter-related mathematical relationships. As a result, estimation and projection error which is introduced early in the model can propagate throughout the model, often in ways that are difficult to follow. This feature makes many regional models quite unstable (prone to over- and under-prediction for certain areas), and requires that they be subsequently fine-tuned by human judgement. The results of this type of compromise are model forecasts that are difficult to replicate or use in a policy context.

   To sidestep this problem, BASS II is designed in modular fashion, as a system of related but independent models. Thus, as improved forecasting procedures, better spatial data, or better allocation decision rules are developed, they can be smoothly integrated into BASS II.

   Taking a modular approach also allows BASS II to make use of *appropriate* theory. For example, while a trend-based approach may be the most appropriate way to forecast population growth, it is certainly not the most appropriate way to allocate growth to particular areas. By separating the growth forecasting and growth allocating functions of BASS II, it is possible to utilize a trend-model for growth forecasting, and a decision-rule based model for growth allocation.

III. THE STRUCTURE of BASS II

The purpose of BASS II is to predict the intensity, pattern, and location of population growth in the ten-county San Francisco Bay Region through the year 2020, as a function of regional *and* local development policy initiatives.

BASS II is built on two primary units of analysis: incorporated cities (and counties), and Developable Land Units (DLUs). Population growth, the demand side of BASS II, is projected on the basis of city population growth trends. Development potential, the supply side of BASS II, is calculated in terms of DLUs.

Under California law, control of development and land uses rests entirely in the hands of incorporated city and county governments. Villages, towns, municipal utility districts, regional authorities, and census-designated places lack control over land uses in California. As of January 1991, there were 121 city and county governments in the ten-county study area (San Francisco is both a city and a county) having direct control over local land uses and development.

Cities also have some measure of land use control over directly adjacent, unincorporated areas. Such areas, known as *spheres-of-influence,* are established and updated by county Local Agency Formation Commissions, or LAFCOs. Spheres-of-influence were originally intended as flexible urban limit lines; they were the areas into which growing cities would eventually expand, and to which cities could economically provide local public services. In recent years, the value of spheres-of-influence as a tool for coordinating inter-jurisdictional land use policies has been greatly diminished.[4]

Developable Land Units are the second primary unit of analysis in the BASS II system. DLUs are currently undeveloped or underdeveloped areas inside and outside cities which may be developed or redeveloped. DLUs are polygon constructs generated by the GIS component of the model and are described according to the geometric union and/or intersection of various environmental, market, and policy attributes. An example would be an undeveloped site with steep slopes, served by sewers, zoned for light industrial, and that is less than 500 meters from a major freeway. In more developed areas, DLUS may approximate collections of developable parcels.

In a nutshell, BASS II "grows" the ten-county San Francisco Bay Region by determining how much new development to allocate to each DLU during each model period as a function of population growth in each city and county; the characteristics of each DLU; and a series of user-specified decision-rules. Thus, the structure of BASS II incorporates four related models, shown schematically in Figure 2.

1. ***The Bottom-Up Population Growth Model.*** This model is the demand side of BASS II: it generates five-year population growth forecasts for each city, county, and major CDP in the study region.

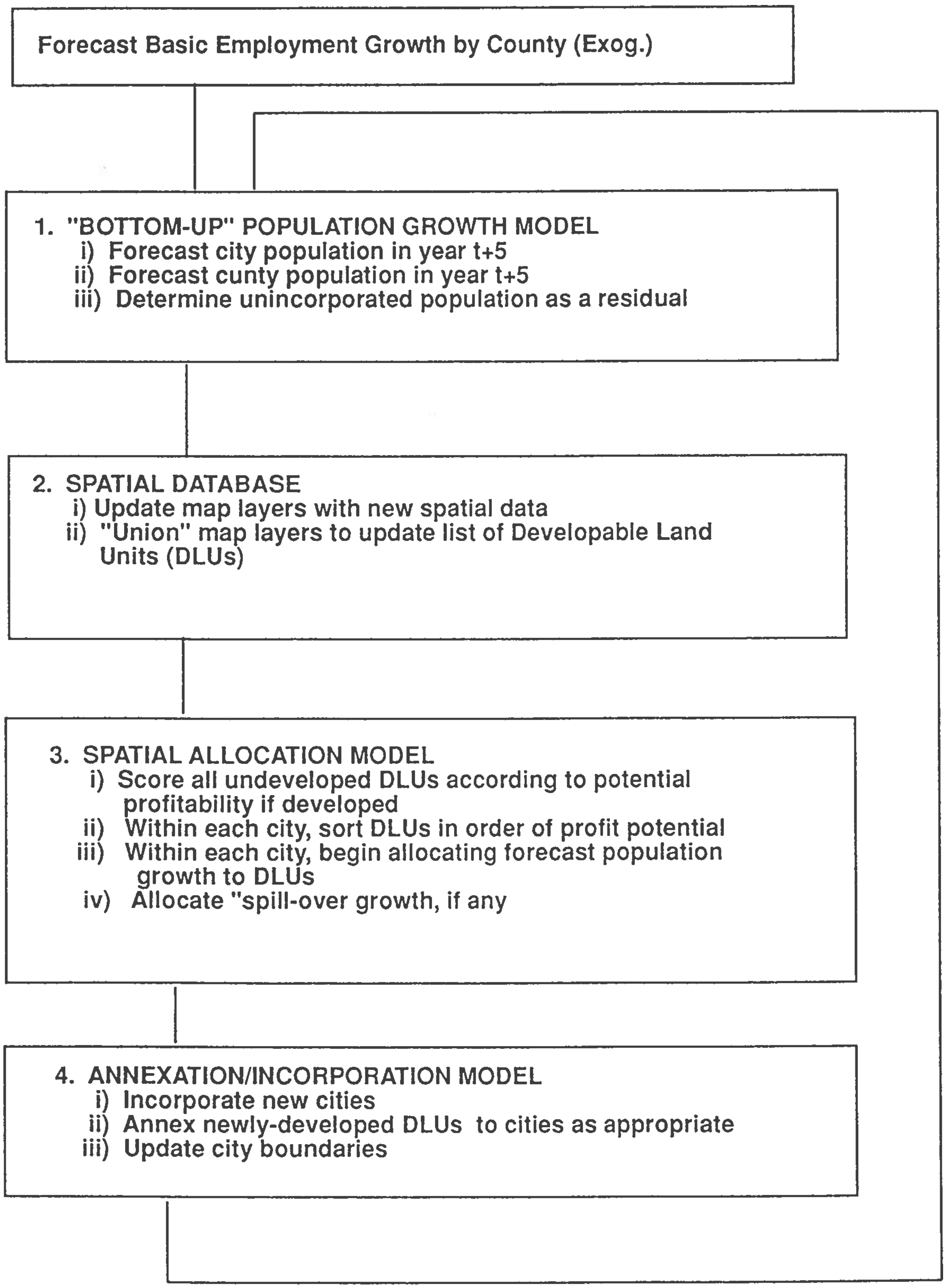
2. ***The Spatial Database.*** This GIS-based model is the supply side of BASS II: it generates and updates the geometry, location, and attributes of each Developable Land Unit (DLU). It is also the primary tool for displaying the spatial pattern of growth.

3. ***The Spatial Allocation Model.*** This model is a series of user-specified Functions and decision rules for allocating population growth to each DLU.

4. ***The Annexation-Incorporation Model.*** This model is a series of decision rules for annexing newly-developed DLUs to existing cities, or for incorporating clusters of DLUs into new cities.

The structure of each of these components is explained in greater detail in the following sections.

Figure 2: Outline of the BASS II Model Structure



IV. THE BOTTOM-UP POPULATION GROWTH MODEL

The Bottom-up Population Growth Model is the demand side of BASS II. It consists of two regression equations of population growth in the cities and unincorporated areas of the ten-county San Francisco Bay Area. Both equations are essentially *trend* models. That is, they predict current population levels (the dependent variable) primarily as a function of past population levels. Other independent variables are included in the models to account for place-specific differences from the overall trend-line.

The Bottom-Up Population Growth Model takes its name from Equation 1, shown in Table 1. Equation 1 is used to project city-by-city population growth levels, at five-year increments, for the 112 incorporated cities in the ten-county San Francisco Bay Area.

Equation 2 is used to forecast county-wide population growth (including cities), also at five-year increments, for the ten counties which comprise the San Francisco Bay Area. Projections of population growth in unincorporated county areas are then estimated by subtracting the sum of city population growth (from Equation 1) from county-wide population growth (from Equation 2). Both Equations 1 and 2 were estimated using ordinary least squares regression on a database which combines cross-sectional data (cities and counties) and time-series data (covering the periods 1970-75, 1975-80, 1980-85, and 1985-90). Coefficient estimates and relevant statistics for all three equations are discussed below.

Equation 1: The Bay Area City Population Growth Sub-Model

Statewide, the period 1970-90 was one of consistent, regular, and predictable population growth (Teitz, 1990; California Department of Finance). It was also a period of consistent population growth across the ten-county San Francisco Bay Area. Thus, it is not too difficult to estimate a simple (linear) trend model which explains the general pattern of population growth across all Bay Area cities during the 1970-90 period. Far more difficult is the task of explaining the specific population growth trajectories of 112 individual cities. This is because individual cities have grown at different rates according to their sizes, their potential for outward expansion, their densities, supplies of developable land, regional growth pressures, and a host of other factors. When used to forecast the growth ofindividual cities, simple trend-based models overestimate population growth in some cities while underestimating population growth in others. To develop a model that fits past trends *and* is stable enough to be used for forecasting, it is essential to: (1) identify the normal growth paths (or "regimes") of particular types of cities; and (2) identify the key factors which cause population growth to diverge from those normal paths. That is, it is important to identify those

TABLE 1:
Equation IA: Bay Area City Population Growth Model:
Regression Results

Dependent Variable: *CITYPOP(i,t):* Population of City i in time period t

Independent Variable:	Coef. Est.	t-stat	Beta
CITYPOP(t-5)	.9367	27.03	.89
City Size Dummy Variables:			
VERY SMALL	(omitted to guarantee a unique solution)		
SMALL	2597.3	4.39	.011
MEDIUM	6298.4	7.25	.022
MEDIUM LARGE	10286.0	8.112	.037
LARGE	15437.9	7.71	.032
OAKLAND	51706.4	7.86	.051
SAN FRANCISCO	162825.4	9.06	.160
SAN JOSE	116654.7	10.69	.114
GROWTH CONTROL(t-5)	-173.82	-2.51	-.00054
LANDLOCK(t-5)	-234.72	-6.31	-.021
DENSITY(t-5)	-00000105	-3.601	-.116
CA_WAVE(t)	.000000612	3.314	.029
Constant	977.26	2.47	

R-squared	.998
F-statistic	21496.9
Standard Error	3977.5
N	383

Definitions of Independent Variables:

CITYPOP(t-5): the population of the same city five years previously.
GROWTH CONTROL(t-5): a dummy variable indicating whether or not the city had adopted a population, housing, or development cap, weighted by th land area of the city in period t-5.
LANDLOCK(t-5): a dummy variable indicating whether a city is land-locked (or water-locked) by neighboring communities, weighted by the land area of the city in period t-5.
DENSITY(t-5): the gross population density of the city in year t-5, weighted by the population of the city in year t-5.
CA_WAVE(t): current statewide population weighted by the land area of each city in period t-5.

City Size Dummy Variables:

VERY SMALL:	population less than 10,000;
SMALL:	population between 10,000 and 29,999;
MEDIUM:	population between 30,000 and 49,999;
MEDIUM LARGE	population between 50,000 and 99,999;
LARGE:	population larger than 100,000.
OAKLAND	
SAN FRANCISCO	
SAN JOSE	

fundamental factors that serve to accelerate the growth of some cities while providing a growth "brake" in others.

Table 1 presents the results of Equation 1, a linear trend model of population growth for the 112 cities in the ten-county San Francisco Bay Area. The primary dependent variable is *CityPop(t)*, the current city population; the primary independent variable is *CityPop(t-5)*, the population of the same city five years previously. (So strong is the trend effect that the r-squared measure for this single independent variable by itself is .994.)[5]

Cities of different sizes tend to add new population in different increments. All else being equal, we observe that smaller cities tend to attract fewer new residents than larger ones.[6] To capture this effect, we classified the 112 cities in the sample into five size classes, according to population in year *t*: (1) *Very Small*: population less than 10,000; (2) *Small*: population between 10,000 and 29,999; (3) *Medium:* population between 30,000 and 49,999; (4) *Medium Large*: population between 50,000 and 99,999; and (5) *Large*: population larger than 100,000. City-size classes were updated every five years to account for population growth. Three separate city-size classes were generated for the region's three largest cities: *Oakland, San Francisco, and San Jose*. The city-size classes were entered into the model as dummy variables.

Three variables were included in the model to provide a "brake" on population growth. The first, *Growth Control(t-5)*, is a dummy variable indicating whether or not the city had adopted a population, housing, or development cap; the dummy variable is weighted by the land area of the city in period t-5 to account for differences in geographic size.

A second "braking" variable, *Landlock(t-5)*, indicates whether a city is land-locked (or water-locked) by neighboring communities, and thus prevented from expanding; this dummy variable is also weighted by the land area of the city in period t-5.

The final "braking" variable, *Density(t-5)*, is the gross population density of the city in year t-5, weighted by the population of the city in year t-5. All else being equal, we observe that cities with higher densities tend to grow more slowly than cities with lower densities.

The final variable to enter the model, *CA-Wave(t)*, causes local population growth to accelerate during periods of high statewide population growth. This variable consists of current statewide population weighted by the land area of each city in period t-5.

Overall, Equation 1 explains the historical trend line of Bay Area city population growth exceptionally well (r-squared = .998). Equally important to the r-squared measure (statistical "goodness-of-fit") is the standard error of the estimate, which is extremely small. All of the coefficients are statistically significant, and have the expected signs. For example, cities that are land- or water-locked, and thus cannot annex undeveloped parcels, grow somewhat more slowly than cities that can expand. Cities that have adopted formal growth control ordinances also tend to grow more slowly, as do cities with higher residential densities. And the positive sign on coefficient of the *CA-Wave* variable reflects the fact that local population growth responds, albeit only slightly, to changes in statewide population growth.

The five city-size classification dummy variables capture the observed effect that population growth levels tend to be correlated with city size. For example, whereas small cities typically add about 2,600 new residents every five years, large cities add about 15,400 new residents every five years.

Equation 2: The County Population Growth Sub-Model

The original design of the BASS II model was entirely as a "bottom-up" model of population growth. County population growth was to be computed as the sum of population growth in cities and unincorporated places. Unfortunately, the data required to build models of population growth for unincorporated areas (of the type shown in Equation 1) are not generally available. Thus, it became necessary to forecast population growth in unincorporated county areas as a residual: that is, as the difference between total county population growth and the sum of population growth within incorporated cities.

Table 2 presents regression results for Equation 2, a trend model of five-year population growth during the 1970-90 period for nine of the ten San Francisco Bay Region counties (San Francisco, which is also a city, is omitted). Equation 2 is similar in form to Equation 1, the city forecasting model: *CntyPop(t)*, the current county population, is modeled primarily as a function of *CntyPop(t-5)*, the population of the same county five years previously.

Where Equation 1 differs from Equation 2 is in the variables that explain deviations from the historical trend line. Two variables, *ChBasic*, and *CA-Wave*, "accelerate" county population growth; another two variables, *Growth Control* and *CityLand*, provide a "brake" on county population growth. *ChBasic*, the numerical change in "basic"[7] employment in the county during the previous five years, is the one variable which must be projected exogenously to the BASS II Model. As Table 2 indicates, *ChBasic* is strongly and positively correlated with county population growth; for every increase in basic employment in a county over a five-year period, the county's population rises by .37 persons. *CA-Wave* is the other variable which "accelerates"

TABLE 2:
Equation 2: 15-County Population Growth Model: Regression Results

Dependent Variable: *CNTYPOP(i,t):* Population of County i in time period t

Independent Variable:	Coef. Est.	t-stat	Beta
CNTYPOP(t-5)	.67	5.34	.62
CHGBASIC	.371	3.42	.0175
GROWTH CONTROL	-157.75	-4.16	-.022
CITYLAND	-969.67	-7.22	-.065
CA_WAVE(t)	.128	3.89	.445

R-squared	.999
F-statistic	15368.01
Standard Error	16656
N	56

Definitions of Independent Variables:

CNTYPOP(t-5): the population of the same county five years previously.
CHBASIC: change in employment in county "basic" industries during the previous five years.
GROWTH CONTROL: a dummy variable indicating whether or not the county has adopted a population, housing, or development cap, weighted by the total land area of incorporated cities within the county.
CITYLAND: the ratio of land area within cities (squared) to total county land area.
CA_WAVE: county share of region-wide population, lagged five years, and

county population growth. *CA-Wave* is a county's share of region-wide population growth (lagged five years), and weighted by state population in the current year. *CA-Wave* measures the extent to which state-wide population growth filters down to the county level. As Table 2 shows, it is both positive and significant.

In the past, persistently high rates of population growth have led many unincorporated places to incorporate in order to gain control over land-use decisions and locally generated revenues. Upon incorporating, those same places then attempt to boost revenue-generating commercial development at the expense of revenue-using population growth. Thus, all else being equal, we would expect that population growth would decline as the ratio of incorporated land area to total county land area increases. This effect is captured in the variable *CityLand*, which, as expected, is negatively correlated with county population growth.

Several counties in Northern California have adopted county-wide growth control ordinances to slow growth, protect the natural environment, or preserve their agricultural base. Such development limits are captured in the variable *Growth Control*, a dummy variable (weighted by the total land area of incorporated cities within the county) that indicates whether or not a county has adopted a population, housing, or development cap. As expected, this variable is negatively correlated with county population growth.

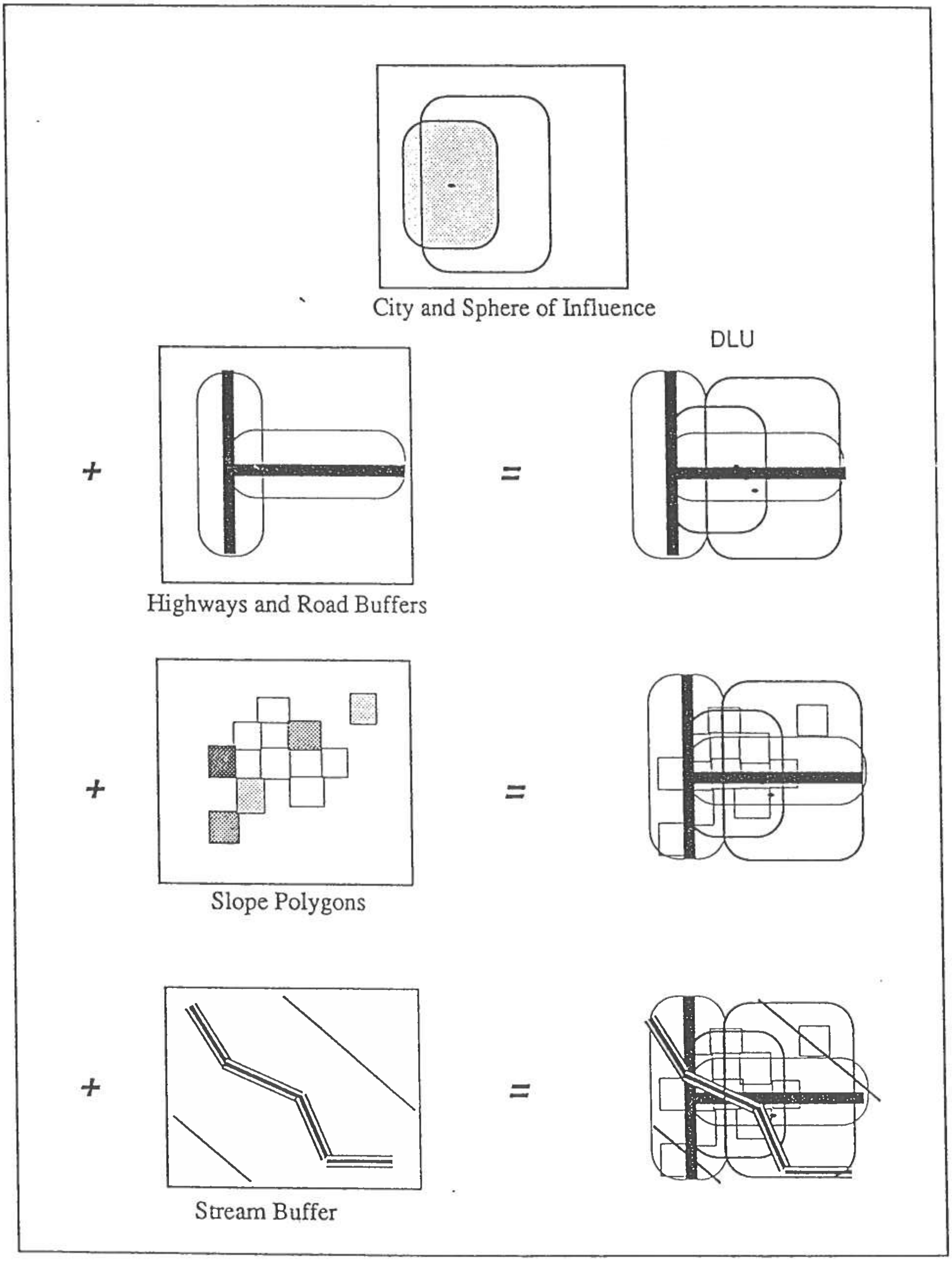
Overall, Equation 2 fits the data extraordinarily well, with a goodness-of-fit measure approaching unity, and an extremely small standard error.

V. THE SPATIAL DATABASE

The spatial database consists of a series of map layers that describe the environmental, land use, zoning, current density, and accessibility characteristics (or attributes) of all sites in the ten-county San Francisco Bay Region. These various layers can be analyzed individually, or merged into a single layer which includes all the relevant attribute information for each resulting polygon. The spatial database is maintained and managed through the use of ARC/INFO, a Geographic Information System (GIS) which incorporates a relational database and true map feature topology.[8]

The spatial database functions as the supply side of BASS II. It is a comprehensive list of the locations and attributes of currently undeveloped (or underdeveloped) sites that may be available to accommodate city and county forecast population growth. These sites are known as Developable Land Units, or DLUs. DLUs do not have regular shapes or sizes, but are generated as the geometric union of different map features and their attributes (Figure 3). Depending on how

Figure 3: Generation of Developable Land Units (DLUs)



the different map layers combine, DLUs can vary in size from the very small to the very large. DLUs in or adjacent to urbanized areas tend to be very small, typically a few acres. By contrast, DLUs in rural areas may exceed several hundred acres in size. The number of DLUs varies by county, but also tends to be very large—ranging from more than 25,000 in Santa Clara County to less than 10,000 in Solano County.

The spatial database currently includes the following map layers:

1. ***TIGER Roads:*** As part of the 1990 Census, the Census Bureau digitally encoded maps of major roads and highways in all metropolitan areas. These map files are known as TIGER (Topologically Integrated Geographic Encoded Reference) files. TIGER files can be referenced and projected through a variety of spatial referencing systems, including USGS, UTM, and State-Plane Coordinate. The TIGER roads file, as encoded in ARC/INFO, is the base map layer for BASS II. This map layer includes federal interstate highways, federal roads, state highways, local arterials, and neighborhood-serving roads.

2. ***TIGER Census Tracts:*** TIGER files also include the boundary lines of 1990 Census tracts. The boundaries were assembled into census tracts (polygons) using ARC/INFO.

3. ***TIGER City Boundaries:*** TIGER files also include the boundary lines of counties and other local governments—including both incorporated cities and unincorporated "census-designated-places." TIGER city boundaries for the San Francisco Bay Region were imported into ARC/INFO, then corrected using updated boundaries supplied by the cities themselves.

4. ***TIGER Hydrology:*** TIGER files also include the locations of major streams and water bodies. These were imported into ARC/INFO as a separate layer.

5. ***Other TIGER Features:*** including railroads and airports.

6. ***Spheres of Influence:*** Under California law, every incorporated municipality has a sphere of influence, which includes, in addition to the city itself, those surrounding unincorporated lands over which the city has some measure of land-use control. Originally, spheres of influence were intended to demarcate each city's ultimate "build-out" and public service area. Thus, they are essential for analyzing possible limits to growth. The size and extent of spheres of influence varies widely by city and county. A map layer incorporating every city's sphere of influence was digitally encoded.

7. ***Slope Polygons:*** Slopes play a major role in determining site developability. Flat and gently sloped parcels are easily and inexpensively developed. As the slope of a site increases, so too does the difficulty and expense of developing it. Sites with average slopes of more than 5 percent and less than 25 percent can be developed, but at increasing cost. Sites with average slopes greater than

25 percent are usually unstable, and are thus rarely developed. To incorporate information on slopes, 500-meter-square slope grid cells were generated from U.S. Geological Survey topographic maps. Seven sets of grid cells were generated, subsuming the following slope categories: (i) 0 percent slope; (ii) 1-2 percent slope; (iii) 3-5 percent slope; (iv) 6-10 percent slope; (v) 11-15 percent slope; (vi) 16-25 percent slope; (vii) 26+ percent slope. Adjacent grid cells of similar slope were then merged into slope polygons.

8. ***Highway Buffers:*** Developers favor sites which are accessible through the existing transportation network. To identify relative accessibility, we generated 500-, 1,500-, and 5,000-meter polygon buffers around major state and federal highways.

9. ***Urban Buffers:*** Most new urban development occurs at the periphery of existing developments, not in entirely new areas. This is because the cost of extending essential urban services to new undeveloped areas usually outweighs any land cost savings. To capture this "adjacency-preference", we generated 1,000- and 2,500-meter polygon buffers around existing urbanized areas as a map layer.

10. ***Earthquake Faults:*** A California law, the Alquist-Priolo Act, stipulates that structures may not be built on top of a known earthquake fault line. The locations of known earthquake fault lines were obtained from the USGS.

11. ***Prime Agricultural Lands:*** Most non-urbanized lands in California currently have some use. In 1988, as part of the California State Farmland Mapping project, the state generated a base map of major agricultural and urban uses. Agricultural use designations are based on current use and soil quality. Agricultural lands are differentiated into: (1) prime agriculture; (2) grazing; (3) forest; (4) of unique state interest; and (5) of unique local interest.

12. ***Marsh and Wetlands:*** All else being equal, development on marsh and wetland areas tends to be costly. This is due both to the higher costs of site preparation (draining and filling) as well as to the added costs of any required environmental mitigation. Moreover, in many areas of California, depending on which agencies have jurisdiction, intense development may be altogether prohibited from marsh and wetland areas. Digitally encoded maps of wetland and marsh areas for the ten-county San Francisco Bay Area were obtained from the U.S. Fish and Wildlife Service.

13. ***Sewer and Water Utility Service Areas:*** The availability of sewer and water service is an important determinant of site developability. Sites without sewer and water service, and for which no such service is planned, cannot be intensely developed. Sites inside the service areas of existing water and sewer utility districts can be developed —usually for the cost of extending service to the site. To capture these differences, the boundaries of the major sewer and water utility districts in the ten-county Bay Area were digitally encoded.

Because these various map layers rarely have common polygon boundaries, the number of DLUs generated by merging the different layers for a single county can easily exceed 10,000. Figure 4 illustrates a portion of the merged DLU map for the city of Livermore in Alameda County; Table 3 lists a portion of the attributes associated with each DLU.

VI. THE SPATIAL ALLOCATION MODEL

The Spatial Allocation Model is a series of decision rules for allocating future population growth (as projected using the Bottom-Up Population Growth Model) to the thousands of developable land units (as generated through the Spatial Database). In economic terms, the function of the Spatial Allocation Model is to "clear the market"— to match the demand for developable sites (as manifest through city and county population growth) to the supply of developable sites (as described by the attributes, size, and location of DLUs).

Unlike most economic models of the development process, the spatial allocation model does not work by solving for the land and housing reservation prices that equilibrate supply and demand (Gore and Nicholson, 1991). Rather, the spatial allocation model seeks to mimic the way private sector developers screen potentially developable sites according to their likelihood of development and ultimate profit potential.

The primary assumption underlying the Spatial Allocation Model is that the location and timing of land development decisions are almost entirely in the hands of private-sector developers, but that such decisions are subject to the policy stipulations of state, regional, county, and local governments. We further presume that private housing developers will seek to develop or redevelop sites in order of expected profitability, subject to land use and environmental regulations as imposed by the public sector, and in accordance with prevailing or permitted development densities. This logic is incorporated into the Spatial Allocation Model in the following steps:

1. All undeveloped DLUs in a county are scored according to their potential profitability if developed.

2. Those DLUs that are unsuitable for development due to environmental, ownership, or public policy reasons are eliminated from consideration. Examples of DLUs that would not be considered for additional growth might include publicly owned parks and open-space, or steeply sloped DLUs having unstable soils.

3. Within each city and its sphere-of-influence, the remaining DLUs (those that could be developed) are sorted from high to low in order of their potential profitability.

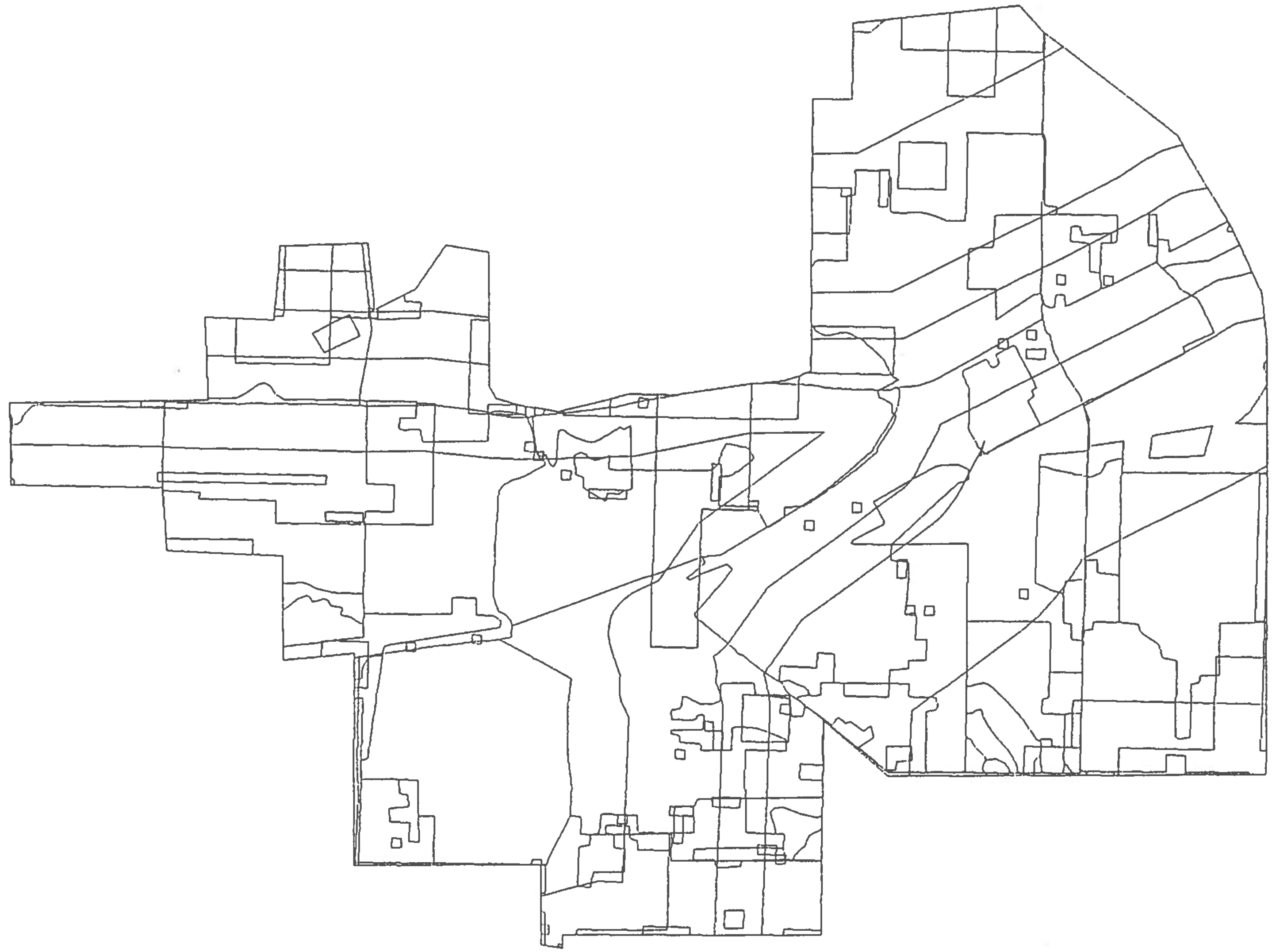


Figure 4: Developable Land Units (DLUs) in the City of Livermore

TABLE 3: Partial Listing of Developable Land Unit (DLU) Attributes for Livermore, California

DLU#	DLU		Wetlands	Buffers						Publicly	Land Type						Slope	Census Tract			
	Area (sqM)	Perimeter		500m Hwy	1000m Hwy	2500m Hwy	1000m Urban	2500m Urban	5000m Urban	Owned Land	Devel.	Grazing	Locally Import.	Prime Agric.	State Import.	Unique	Class	Tract	Population 1980	Population 1990	Density 1990
2	425	176	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
3	211646	1824	0	0	0	0	1	1	1	0	0	1	0	0	0	0	7	4512	4511	8697	176
4	223757	1899	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
5	191152	1769	0	0	0	0	1	1	1	0	0	1	0	0	0	0	7	4512	4511	8697	176
6	1294310	5710	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
7	29708	859	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
8	110760	1951	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
9	27652	777	0	0	0	0	1	1	1	0	0	1	0	0	0	0	0	4511	3589	3569	6
10	1424574	5151	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4511	3589	3569	6
11	231314	1894	0	0	0	0	1	1	1	0	0	1	0	0	0	0	3	4512	4511	8697	176
12	1332692	10199	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
13	18816	709	0	0	0	1	1	1	1	0	0	1	0	0	0	0	3	4512	4511	8697	176
14	1969604	9654	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	4512	4511	8697	176
15	23846	1990	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	4511	3589	3569	6
16	250104	2000	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
17	611168	3962	0	0	1	1	1	1	1	0	0	1	0	0	0	0	0	4511	3589	3569	6
18	158	63	0	0	1	1	1	1	1	0	0	1	0	0	0	0	7	4511	3589	3569	6
19	737	215	0	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
20	39281	986	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	4512	4511	8697	176
21	176658	1732	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
22	9298	389	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	4512	4511	8697	176
23	604631	5188	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	4511	3589	3569	6
24	703	124	1	0	0	1	1	1	1	0	1	0	0	0	0	0	0	4512	4511	8697	176
25	17373	887	1	0	0	1	1	1	1	0	1	0	0	0	0	0	0	4512	4511	8697	176
26	3403	304	0	0	1	1	1	1	1	0	1	0	0	0	0	0	0	4511	3589	3569	6
27	73161	1349	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	4512	4511	8697	176
28	553	122	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	4511	3589	3569	6
29	257600	3021	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	4511	3589	3569	6
30	31259	857	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	4511	3589	3569	6
31	403336	2725	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	4512	4511	8697	176
32	208933	2462	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	4511	3589	3569	6

4. Forecast population growth for each city is allocated to the DLUs within each city sphere-of-influence in order of DLU profit potential (high to low); and at population densities consistent either with current zoning and general plan requirements, or at "up-zoned" population densities comparable to other developed areas in the city. After it has allocated as much population growth as will "fit" into the DLU with the highest profit potential, the model moves to the next most profitable DLU, and so on.

   The allocation process within a city is complete either when: (1) all forecast population growth is allocated, or (2) when there is insufficient undeveloped land in the city to accommodate all forecast population growth. Unallocated population growth, if any remains, is then accumulated for re-allocation (or spill-over) into unincorporated county areas.

5. The same logic as above is used to allocate forecast county population growth (plus any unallocated spill-over growth from individual cities) to unincorporated county DLUs.

   The allocation process within a county is complete either when: (1) all forecast and spill-over population growth is allocated, or (2) when there is insufficient undeveloped land in the county to accommodate forecast population growth. Unallocated population growth, if any remains, is then accumulated for later re-allocation to those counties with remaining developable DLUs.

The potential for spill-over development is one of the most interesting parts of the model. Spill-over occurs when there is insufficient developable land in a city or county to accommodate that city/county's forecast population growth. In such cases, the unallocated increment of population growth is accumulated for potential re-allocation (or spill-over) into a neighboring municipality, unincorporated area, or county. This is not to suggest that it will always be possible to accommodate spill-over growth. Depending on the types of local policies being simulated, it may not be possible for the unallocated population growth from one city or area to be re-allocated to another city or area.

The DLU Profitability Potential Sub-Model

The growth allocation process (Steps 3 through 5, above) is itself fairly mechanistic. The most interesting part of the Spatial Allocation model, and its conceptual heart, is the process for "scoring " each DLU according to its potential economic profitability if developed (Step 1, above). This scoring process attempts to capture the key decision variables and processes used by residential developers to screen and rank potential sites for development. The scoring process calculates a *Profitability Potential Index* for each DLU based on the following maximization function:

DLU Profit Potential per acre (d) =

$$\text{MAX} < [\text{GP\_Den}(d)*\{(\text{HsePrice}(i) - \text{ConCost}(i) - \text{SiteCost}(i,d) - \text{LandCost}(i,d)\}], [\text{Spec\_Den}(d)*\{(\text{HsePrice}(i) - \text{ConCost}(i) - \text{SiteCost}(i,d) - \text{LandCost}(i,d) - \text{Rezone}(i)\}] >$$

where: i indicates each city or policy sphere
d indicates each DLU

GP_Den is the maximum residential density per acre proscribed under the current zoning ordinance or general plan.

Spec_Den is the maximum residential density per acre if the site were "up-zoned" to the city-wide average density.

HsePrice is the median price of new homes in each city. It is a measure of expected revenue per home constructed.

ConCost is the cost of constructing the typical new home in each city. It does not include costs related to the quality of the site, but does include costs such as fees and generally required subdivision improvements.

SiteCost is a measure of the extra cost of building the typical new home depending on site characteristics such as slope or soil quality.

LandCost is the cost of raw land per home. It varies by city, by site quality, and by current zoning.

Rezone is the per-home cost of rezoning the site to the higher speculative density.

For a given DLU, this function compares the profit potential of developing a typical new home under the current zoned density with the profit potential of developing the same home at a higher, speculative density. On the one hand, if the per-unit cost of rezoning a particular site (*Rezone*) is less then the expected increase in profit associated with rezoning the site, then the calculated DLU profit potential (per acre) includes the extra profit potential associated with the rezoning. On the other hand, if the cost of rezoning exceeds the likely profit increase associated with a higher density, then the DLU profit potential per acre is based on current zoned densities.

Both parts of the maximization function are themselves profit functions. *HsePrice* is the current median price of new housing in the community, and, as such, is a measure of the revenue associated with building the typical new home. All else being equal, developers would prefer to develop those sites upon which the finished home will earn the highest price. This variable is assembled city-by-city from local property transaction or census data, and is currently exogenous to the BASS II model.

ConCost is the cost of constructing the typical new home in each community. *ConCost* includes the hard costs of constructing the home itself, as well as the cost of standard subdivision improvements and local fees; it does not include construction costs which vary with site quality or site location. All else being equal, one would assume that developers would prefer to build in communities in which construction costs (including fees and required subdivision improvements) are low. *ConCost* varies by city, and has been assembled based on surveys of major residential developers and municipal building departments.

SiteCost is the cost of preparing the site so that it could accommodate the typical new home. For example, *SiteCost* cost may include the additional cost of building on a steep slope, or the cost of building on poor soils, or the cost of building in areas lacking basic utility services. All else being equal, one would presume housing developers would prefer DLUs in which the site costs of constructing homes are low. SiteCost varies by DLU and by community.

LandCost is the estimated cost of raw land per home. It varies by city and is generated using a simple hedonic-price model of recent housing transactions.

Both profit functions are computed on a per-home basis. They are each then converted to profit per acre using one of two densities. *GP_Den* is the density of residential development as permitted by current zoning ordinances and/or the city's General Plan. *GP_Den* varies by DLU. *Spec_Den* is the maximum permitted DLU density, assuming a developer were successfully able to upzone that DLU to equal city-wide or county-wide densities. The difference between *GP_Den* and *Spec_Den* varies by DLU, and is a measure of the speculative potential of that DLU. All else being equal, developers able to secure sites zoned for low-density development, and then get those sites rezoned to higher densities, will realize much greater revenues. In fact, this is precisely how many of the more successful land developers operate. They secure lower-priced sites that are "under-developed" or "under-zoned" relative to the market, and then work through the development entitlement process to raise permitted densities.

Inherent in the calculation of the DLU Profit Potential Index is the assumption that the various measures of cost (*ConCost*, *SiteCost*, *LandCost*, and *Rezone*) are basically independent of each other, and that new home prices are determined in the marketplace (on the basis of supply and demand) and not on a cost-plus-profit-margin basis. Likewise, it assumes that housing and land markets are competitive, and that residential developers are "price-takers."

Operationalizing the Spatial Allocation Model

Running the Spatial Allocation Model involves, first, the creation of a single large database combining city-based information (such as median home prices, fees, and forecast population growth) with DLU-based information (such as slope, agricultural use, wetland status, and current zoning); second, the calculation of a profit potential measure for each DLU; and third, allocating forecast population growth to appropriate DLUs. This somewhat complicated procedure is accomplished in six separate steps (Figure 5):

1. Tabular files containing the DLU attributes are exported from ARC/INFO (as implemented on the Sun SparcStation), converted to dBase III+ format, and imported into dBase III+ on an IBM-compatible microcomputer.

2. Using dBase III+, a second database containing city information is relationally joined to the DLU attribute database imported in Step 1, above.

3. Using dBase III+, a Profit Potential Index value is calculated for each DLU. DLUs that are inappropriate for development are eliminated.

4. Using dBase III+, the DLUs within each city and its sphere-of-influence are sorted from highest profit potential to lowest profit potential.

5. Using dBase III+, forecast population growth is allocated to DLUs in order of calculated profit potential.

6. Using dBase III+, unallocated spill-over population growth is allocated to unincorporated areas and other counties.

7. The updated DLU database (now including allocated population growth) is imported back into ARC/INFO, where it can be displayed.

VII. COMPLETING THE LOOP: THE ANNEXATION/INCORPORATION MODEL

BASS II runs "five-years" at a time. By this we mean that population growth is projected and then allocated at five intervals. Thus, a 20-year "run" of BASS II is really four sequential five-year "runs." There are three reasons why we project in five-year increments. First, because the Bottom-Up Population Growth Model was estimated on the basis of five-year increments, to use the model for forecasting on anything other than a five-year basis would be inappropriate.

Second, running the model in five-year increments mitigates the possibility that small cities will experience runaway growth. Experience shows that many small cities, particularly those at the urban edge, grow quickly —but only for brief periods. Such cities may experience several years of slow growth before the outward wave of population growth reaches them; followed by several

Figure 5: Spatial Allocation Model Operation

1. Export DLU attribute data from ARC/INFO

2. Using dBase III+, "join" DLU attribute data with corresonding city and census tract data.

3. Using dbase III, generate a Profit Potential Index value for each DLU; eliminate inappropriate DLUs.

4. Within dBase III, sort DLUs within each city and sphere-of-influence from highest profit potential to lowest.

5. Using dBase III+, allocate forecast population growth to DLUs in order of calculated profitability, above.

6. Using dBase III, allocate "spill-over" growth (if any) to unincorporated areas and/or other counties

7. Import updated database (including newly-developed DLUs) back into ARC/INFO for display.

years of extraordinarily high growth rates, as they become the growth wave-front; followed again by several years of moderate growth (as easily developable sites are used up). This cycle of slow growth, followed by rapid growth, followed by moderate growth, may take as many as 50 years or as little as 10 years. Looking only at long-term growth trends tends to even out the cycle or obscure it. As a result, models of small-city growth based on 10- or 20-year growth trends tend either to under-estimate or over-estimate population growth, depending upon where the city is in its growth cycle. As the extraordinary statistical fits obtained in Equation 1 would indicate, models of city population growth based on five-year trends tend to do quite well at capturing the small-city growth cycle.

Third, from a policy perspective, many local land use policies are somewhat transitory in nature. Few policy initiatives are consistently applied over a 20-year period. Cities are constantly altering allowable densities and land uses in response to current issues and citizen concerns. Some policies —such as significantly down-zoning developable sites at the urban edge in the face of continued growth pressure — may produce immediate market responses. Other policies, such as extending mass transit service, may take a generation to affect the spatial pattern of development. Running BASS II in five-year increments facilitates simulating policies that have short-term effects as well as long-term effects.

Running BASS II over several sequential five-year periods requires incorporating feed-back or "up-dating" loops in the model. This means using the results of the first five-year forecast/simulation as initial conditions for the second set of five-year forecasts, and so on. It also means updating the Spatial Database to incorporate the specific results of the Spatial Allocation Model.

If all city boundaries were fixed into perpetuity, the updating process would be straightforward. The first set of outputs of the Spatial Allocation Model, a list of newly developed DLUs, would be used to update the Spatial Database, while the second set of outputs— a city-by-city summary of allocated population growth— would serve as inputs for the next iteration of the Bottom-Up Population Growth Forecasting Model.

In reality, of course, city boundaries are rarely fixed. City boundaries change and do not change over time for several reasons. The traditional practice is for cities to extend their boundaries (almost always by annexing unincorporated county lands) to provide a higher level of public services to growing areas, in order to increase their tax base, and to better integrate newly-developing areas with already developed neighborhoods. More recently, many cities in California have been extending their boundaries outward (and thus their control over land uses) as a way of preventing or reducing development. Still other cities have chosen not to extend their boundaries in the face of citizen pressures as a way of retaining their existing community character.

Annexation is not the only way California cities can expand. Previously unincorporated neighborhoods can, though incorporation, become cities. Neighborhoods incorporate for the same reasons that cities expand: to capture a tax base, to facilitate orderly development, to obtain a higher quality of local public services, and, on occasion, to prevent new development.

Because city boundaries (and spheres of influence) are so essential a part of BASS II, the updating process must necessarily include a procedure for determining which newly developed DLUs are to be annexed to existing cities and which are to be part of newly incorporating cities. Making such determinations is the purpose of the Annexation/Incorporation Model.

The Annexation/Incorporation Model

At this stage of its development, the Annexation/Incorporation Model consists of a simple regression model comparing ten-year annexation activity by city (dependent variable) with separate independent variables describing initial city population, density, location, and growth policy. The sample upon which the model is estimated excludes "land-locked" cities (i.e., those unable to expand their boundaries) but includes cities that did could have annexed but did not. The model takes the following form:

Annex(t, t-10) =

f [Population Change (t, t-10), Density(t-10),
Cnty-Control-DV(t-10), Local-Control-DV (t-10)
Median Household Income (t-10)]

where: t indicates the current time period;
t-10 indicates ten years prior;

Annex is the number of acres annexed by each city during the prior ten years.

Population Change is the change in population in each city during the previous ten years.

Density is the initial density of the city.

Cnty-Control-DV is a dummy variable indicating whether county land use policies make annexation difficult.

Local-Control-DV is a dummy variable indicating whether or not a city has formal growth control program in place.

Median Household Income in each city.

Coefficient estimates and goodness of fit measures for the Annexation Model are shown in Table 4. Subsequent efforts will be made to refine this model and to develop a companion model for projecting incorporation activity.

VIII. SIMULATING DEVELOPMENT POLICY CHANGES

The effects of new regulatory and investment policies upon the location, amount, and intensity of urban development can be simulated in BASS II through three different mechanisms:

1. ***Adding New Spatial Features or Map Layers:*** Adding new features or map layers changes the geometry and characteristics of the set of DLUs— the supply side of BASS II. For example, to simulate the likely impacts of a proposed greenbelt, one would first generate a new map layer showing the precise location of the greenbelt. This new layer would then be merged with the existing set of Developable Land Units (DLUs). The updated DLU list would then indicate which particular DLUs were inside or outside the greenbelt. Such information would be used within the Spatial Allocation Model either to prohibit development within greenbelt DLUs, or alternatively to reduce the densities of development in greenbelt DLUs. To the extent that the greenbelt DLUs would have otherwise been allocated more development, that development would then be re-allocated elsewhere.

2. ***Changing Environmental or Infrastructure Policies that Facilitate or Prohibit Development, or Change the Cost of Development:*** Changing environmental and infrastructure policies can affect the allocation of growth to individual DLUs in three ways.

   First, such policies can affect the calculation of the DLU Profitability Potential Index. For example, the decision not to expand a municipal water district to service a growing city would tend to make development in that city more expensive, thereby reducing the attractiveness of that city to private housing developers. Raising local development fees in certain cities would have a similar effect.

   Second, changing policies can affect which DLUs are precluded from development. For example, the adoption of a county-wide policy to prohibit development on steep hillsides would eliminate steeply-sloped DLUs from development consideration regardless of their private development profit potential.

   Third, changing environmental or infrastructure policies can affect the densities at which new development is allocated. For example, rather than totally prohibiting development from prime agricultural lands, a county government might reduce the maximum development densities allowed on such parcels. Such a change would create a density ceiling for such DLUs, as well as reduce the profitability of developing them.

3. ***Changing Local Zoning and/or Land Use Regulations:*** City and county governments frequently up-zone and down-zone areas, as well as change allowable uses. Such policy shifts can be simulated in two ways. First, previously undevelopable DLUs can become developable (and vice versa). This would be the case when land parcels previously reserved for commercial development are opened up to residential development.