Preliminary Draft for Review

Longitudinal Calibration of UrbanSim for Eugene-Springfield

Transportation and Land Use Model Integration Program
Task 3E

Prepared for

Transportation Development Branch
Oregon Department of Transportation
555 13th Street, NE
Salem, Oregon 97310

By

The University of Washington
Box 353055
Seattle, Washington 98195

And

Parsons Brinckerhoff Quade & Douglas, Inc.
ECO Northwest
Hunt Analytics
KJS Associates, Inc.

10 July, 1999

Revised 27 August, 1999

This document describes results of the longitudinal calibration of the UrbanSim model for the Eugene-Springfield metropolitan area for the period 1980 to 1994. The work represents Task 3E in the TLUMIP II workplan, and is intended to inform discussion of the design specifications for the second-generation Oregon model. These results have only recently been compiled, and are therefore not fully reflected in the draft second-generation model design.

The document is organized into four sections. Section 1 describes the plan for the calibration process, and describes the data requirements and the results of the initial data compilation. This work was conducted by a coordinated effort between the staff of ODOT and LCOG, with assistance from Parsons Brinckerhoff and the University of Washington. This section draws on a series of technical memoranda that systematically address the topic. For brevity, the memos are left in their original form and presented in Appendix A.

Section 2 describes the work undertaken by the University of Washington to integrate the data described in Section 1 and prepare it for use in the model calibration. Section 3 then describes the calibration procedures and results. Finally, Section 4 summarizes key findings and provides some considerations for discussion of the second generation Oregon model design.

This work was facilitated by grants from the National Science Foundation, the University of Washington, the National Cooperative Highway Research Program and the Federal Highway Administration. These grants have funded substantial development of the UrbanSim model and software implementation since the completion of the Oregon Prototype Metropolitan Land Use Model in May of 1998. The longitudinal calibration process has made substantial use of the extensions to the UrbanSim model made during the past year, which are described in Appendix A.

1 Phase I: Initial Data Collection and Calibration Plan

The first phase of the longitudinal calibration task was the initial data collection and processing to develop the 1980 base input data to run the calibrations from 1980 to 1994, and to develop intermediate year and end year calibration targets. In addition, the first phase developed a strategy for the longitudinal calibration process. The data requirements were described in the first three technical memoranda (see Appendix B). The next three memoranda are preliminary descriptions of the results of the data compilation by ODOT and LCOG, with some assistance from Parsons Brinckerhoff. The seventh and eighth memoranda then describe the dynamic process of running the model from 1980 to 1994, and the specifics plans for the calibration process.

Since these topics are thoroughly addressed in the memoranda, only a brief summary of key points is included here for quick reference. The major points drawn from the first phase of the project are the following:

1.1 Creation of a 1980 Base

The use of 1980 as the base year for the longitudinal calibration process necessitated reconstructing a parcel database, business establishments, and household data for 1980, almost two decades later. One should not be surprised to learn that this proved to be quite challenging. The LCOG staff was able to provide historical parcel databases and GIS coverages for parcels, and data for business establishments. ODOT obtained these databases and undertook the initial spatial processing to organize and integrate the data into a form suitable for this task. Several difficulties emerged during this process, each of which is discussed below.

1.1.1 Missing Commercial Square Footage

The 1980 parcel file, like its 1994 counterpart, did not contain square footage estimates for nonresidential buildings. Since the model uses square feet of commercial space to allocate businesses, this was a critical piece of the effort to compile a useful database. The only source of these building square footage estimates were from LCOG supplemental inventories done by coding assessor manual records, between 1994 and 1997. This building inventory contained most, though certainly not all, commercial buildings that existed by 1994. Many of these records, though again, not all, had a year built value which could be used to identify buildings that should be present in the 1980 database. This was essentially the strategy we pursued, though clearly it would not be as accurate as it would have been if the assessor files had maintained these data as part of the 1980 database.

1.1.2 Missing Year Built

The second major data development issue was that many of the parcels in 1994 that contained a building did not have a value for the year built. This made the identification of buildings that should have been in existence in 1980 more difficult. The implications of incorrectly assigning buildings to either 1980 or the post 1980 period is that the model calibration would be forced to attempt to replicate invalid trends. Our intent was to create as clean as possible a base year and end year, in order to maximize the utility of the calibration exercise for learning about the model dynamics. The priority of this caused substantial delay in the process of compiling the database.

1.1.3 Matching 1980 and 1994 Parcels

In order to assign buildings that were in the 1994 inventory (and could be identified as existing in 1980) to the appropriate 1980 parcel, the initial approach was to link the two databases on parcel identifier. This turned out to be too inaccurate for several reasons, all obvious in hindsight. First, parcels are generally split when they are subdivided for development (consider a farm that is converted to a housing subdivision), which causes the creation of many new parcel numbers that have no match in the 1980 database. Second, and perhaps less obvious, is that tax assessors may occasionally renumber parcels, rendering a match occasionally invalid.

These difficulties precipitated turning to other approaches to matching 1980 and 1994 parcels. The final approach taken was to adjust the coordinates of the 1980 parcels to the newer coordinate base that the 1994 parcels were in, so that a GIS overlay operation could be performed to match them. LCOG undertook this coordinate adjustment, which made possible the spatial matching of parcels.

These were the major difficulties encountered in the initial phase of data compilation. Further discussion and analysis as part of the second phase of this work is described in Section 2.

1.2 Longitudinal CalibrationStrategy

The longitudinal calibration strategy was developed by Waddell and Hunt, and is documented in the last two of the technical memoranda in the Appendix. Some of the key points are summarized here.

1.2.1 Integration with LCOG Travel Model

The UrbanSim model was designed to interface to an existing travel model such as the LCOG four-step model (or to activity-based models as in Honolulu). The actual process of interfacing existing travel models with the UrbanSim model requires that the output files from the land use model be reformatted for use in the trip generation step of the travel model, and that the results of the travel model, in a binary emme2 database, be made available for access by the land use model. This interaction can be carried out every single year, but practical constraints and common sense suggest a more parsimonious use of the travel models. While the land use model runs one simulation year in less than two minutes, the travel models take -- a bit longer.

The current LCOG travel model consists of a series of steps that make interaction of the land use and travel models relatively time-consuming. A project has been initiated by LCOG, with assistance from Parsons Brinckerhoff and Tim Hier, to automate and streamline more of the travel model system, which will substantially facilitate this interaction. The results of this effort to streamline the travel model are not yet available. To date, then, the calibration and testing of UrbanSim, doe to time constraints, has not been done interactively with the LCOG travel models. The travel utilities from 1980 were used throughout the calibration testing to date.

1.2.2 Calibrating Parameters that Affect Model Dynamics

The basic purpose of the longitudinal calibration is to calibrate a number of parameters in the model system that govern how the dynamics of the model operate over time. These are mostly coefficients that affect the market clearing process, price adjustments, and the reaction of developers to changing market conditions. A review of the model equations led to an identification of two sets of parameters that would be valuable to explore within the calibration process. The major focus of the effort was to be on what we termed ‘heuristic’ parameters, to represent the intention of searching heuristically across a range of values for parameter values that would attempt to maximize the model performance over time. A second set of parameters we termed 'one-time' parameters, to reflect the potential of making adjustments in some parameters such as mobility rates, that we would not attempt to search in a heuristic process.

1.2.3 Specification Testing

A second component of the calibration strategy was to examine the specification of model components that would influence the model dynamic behavior, and to make alterations where the results suggested testing alternative specifications. This was the most tentative aspect of the original calibration plan, since there were concerns that little time would be available for this work.

2 Data Integration and Synthesis

Once the data compilation by LCOG and ODOT was essentially completed in the first Quarter of 1999, the data were transferred to the University of Washington for final integration and synthesis, and to load the model database for the calibration. At this stage of the process, substantial additional diagnostic assessment of the data was conducted.

The work performed in this phase benefited from ongoing research funded by the Federal Highway Administration, through KJS Associates. In that project, Environmental Systems Research Institute and the University of Washington have been subcontracted to develop tools to facilitate data preparation for transportation and land use models (not just UrbanSim). Early development of prototype software tools to support the data preparation process was used to assist the longitudinal calibration process, using it as a test case for the development of these data preparation tools. A description of the current functions performed by the data preparation tools is presented as Appendix B.

The following are the major steps that were taken during this phase of the process.

2.1 Reconciliation of 1980 and 1994 Parcels

As noted earlier, difficulties in matching parcels between 1980 and 1994 necessitated devloping a spatial matching process. Once the coordinates for the 1980 parcel coverage were adjusted by LCOG, the parcel centroid coordinates from the 1994 parcel coverage were overlaid on the 1980 parcel coverage to develop a cross-reference table of 1980 and 1994 parcel identifiers. These relationships were spot-checked for accuracy, and appear to be reasonably robust.

2.2 Land Use Coding

The classification of land use categories for use in the model was performed on 1980 and 1994 parcel databases using a combination of the original codes in these databases. Due to inconsistencies and gaps in the data coding, use of only one field to assign land use classes was deemed inadequate. As a result, a procedure was developed to execute a series of operations to fill missing data and resolve inconsistencies in a systematic way. These operations are described in Appendix C.

2.3 Synthesis of Missing and Inconsistent Data

One of the most common problems with large databases such as tax assessor parcel files is that they contain numerous problems of missing and inconsistent data. It would probably be fair to say that this has been the source of much of the concern about the viability of the UrbanSim model: it makes heavy use of large parcel databases that are subject to these kinds of errors. There are at least two approaches to dealing with the messy data problem. The first would be to use more aggregate data (and a consistent model specification), perhaps coupled with a data synthesis procedure to create consistent disaggregate data from more aggregate sources. The second approach is to attempt to use the large datasets, but to develop a set of procedures that could evolve into an expert system, to use multiple indicators to identify and fill missing or problematic data with synthesized values. Both approaches may be reasonably valid for modeling purposes, but the second remains closer to the original detailed data. We have adopted the second approach for the current application.

The tools that have been developed to date for identifying missing or inconsistent data and synthesizing values for them are operational, and have been used to process the data provided by LCOG and ODOT into a form suitable for use in modeling. The following sections provide a description of these procedures.

2.3.1 Parcel Data:

- Add units – if a residential ALU parcel contains no units, but has improvement value, units are added. The median units per improvement value of the surrounding parcels in the same ALU is used to calculate the number of added units.

- Fix land value – if the parcel land value is blank, we calculate a new value based upon the median land value per acre of the surrounding parcels in the same ALU.

- Fix improvement value – if a parcel contains a building but has no improvement value then the improvement value is set using the median improvement value per sqft of the surrounding parcels.

2.3.2 Business Merging:

- Check Mapping – most of the businesses are assigned with an initial parcel mapping. If this mapped parcel is not a developed parcel then the mapping is invalid and the business is not placed. Businesses that map to residential parcels are placed immediately. Businesses in commercial parcels are subjected to the steps below.

- Check Sqft – if there are no sqft on the mapped parcel and there is improvement then the median sqft per improvement value of the surrounding parcels is used to calculate a sample sqft value. If this sample sqft implies a sqft per employee ratio below a predetermined value then the business is moved, otherwise the calculated number of sqft is created and the business is placed. In the case of no sqft and no improvement the business is moved.

- Moving a parcel – we search the parcels in the surrounding quarter mile, starting with the closest, for an appropriate location. To qualify, a parcel must provide enough space so that the sqft per employee ratio is within a valid range for the parcel's ALU. If a valid parcel is found then the business is moved. If no valid parcel is found then space is created on the original parcel and the business it not moved.

2.3.3 Household Merging:

- Group households and parcels – the household to parcel mapping is done via tract and blockgroup. Every household and parcel is grouped by tract and blockgroup. If a group of households has no corresponding group of parcels then the households are not placed.

- Create units – if the number of households in this tract/blockgroup is greater than the number of units, additional units are created. The only exception occurs when this tract/blockgroup is on the border of the study area. In this case additional units are not created and surplus households go unplaced.

- Place households (preserve ALU) – for each tract/blockgroup pair we iterate through the households and place them in a parcel that is in the correct ALU. Those households without a match are preserved for the next matching.

- Place households (ignore ALU) – the remaining households are placed in the correct tract/blockgroup ignoring the ALU constraint.

2.4 Results of Database Compilation

The final results of the database compilation are shown in Tables 1-5. The results of the data synthesis are evident by comparing Tables 1 and 2 for 1980, and 3 and 4 for 1994. Note that the most substantial changes arising in the data synthesis process appear to be the creation of a significant quantity of nonresidential square footage, based on business locations that were missing buildings.

Comparison of 1994 to 1980, shown in Table 5, reveals plausible trends in the development of the metropolitan area. While clearly not perfect data, these appear to be adequate to use on the calibration of the longitudinal dynamics of the model for Eugene-Springfield.

3 Calibration Results

The calibration process is described in detail in the last memorandum in Appendix B. This section focuses on changes since the memorandum was written, and on the actual results of the calibration to date.
The calibration process was facilitated by research done under grants from the National Science Foundation and the University of Washington to implement a reusable software architecture for urban and environmental simulation (see Appendix A). The software implementation used for this longitudinal calibration is operational under the new architecture, and this application is its first testing.

3.1 Model Calibrator

One particular extension to the model architecture that was directly informed by the needs of the calibration process was the development of a model calibrator component to streamline the calibration process. The main purpose of the calibrator is to automate the process of submitting multiple simulation runs for purposes of calibration, so that multiple runs can be executed unattended for overnight operation.

3.1.1 Client-Server Configuration

The calibrator is designed as a client-server arrangement, allowing multiple networked computers to be used simultaneously to run calibration runs. The server coordinates the activities of the clients, feeding parameter sets called 'points' in the application to the clients, and compiling their results into a single database.

3.1.2 Manual Points

There are two modes of operation. The first is the processing of 'manual points', which are sets of parameter values manually supplied in a file for use by the calibrator. The user can enter as many combinations of parameters as desired, stepping systematically through the parameter space, or testing specific combinations. This approach was used to generate almost one hundred initial simulation runs, taking the initial estimate for each parameter, and mutating it to user-specified minimum and maximum values.

3.1.3 Simulated Annealing

The second mode of operation for the calibrator is the application of a simulated annealing operation (see Press et al., 1992, p. 444-455). This is a simulation method for optimization problems of large scale, where there is potential for simpler algorithms to become trapped within local minima or maxima. This procedure randomly mutates the parameter values, using a 'slow-cooling' analogy from thermodynamics. At high energy levels, atoms move freely with respect to each other, but as the temperature cools the movement of atoms becomes much more constrained, so that at freezing temperature very solid crystalline forms are produced.

This procedure is also now operational, and has been used to augment the manual points procedure.

3.2 Goodness of Fit

The calibration procedure is based on the maximization of an objective function. In this case the objective function is a goodness of fit measure for the entire model system. The challenge was to produce summary goodness of fit measures that would be useful for calibration purposes (a reasonable objective function), but also to obtain detailed assessment of the various components of the model system. To this end we have implemented several goodness of fit measures, and applied them to specific subsets of the results in addition to generating overall measures that weight the components together.

The following goodness of fit measures have been implemented and are now part of the software application:

1. A scaled least squares sum of the errors between the simulated and observed 1994 values by Traffic Analysis Zone, proposed by Hunt and described in Technical Memorandum 8.
2. A mean error estimate between the simulated and observed 1994 zonal values.
3. An R-squared estimate generated by regressing the 1994 simulated values by zone on the observed zonal values for 1994. This measure can be interpreted as the proportion of the variance in the observed 1994 zonal values that are explained by the model predictions, and has a range of 0 to 1.
4. A correlation coefficient between the simulated and observed 1994 data, which has a range of –1 to 1.
5. A scaled least squares sum of the errors between the simulated and observed change from 1980 to 1994 by zone.
6. A mean error between the simulated and observed changes from 1980 to 1994 by zone.
7. An R-squared estimate generated by regressing the observed change from 1980 to 1994 on the observed change.
8. A correlation coefficient between the simulated and observed change from 1980 to 1994.

Goodness of fit measures were computed for each of the following variables simulated by the model for each zone:

- Total Population
- Total Households
- Households by Income
- Households by Size
- Total Employment
- Employment by Sector
- Acres by Land Use Category

- Housing Units by Residential Land Use Category
- Sqft of Buildings by Nonresidential Land Use Category
- Land Value per Acre by Land Use Category
- Improvement Value per Unit or Sqft by Land Use Category
- Total Value per Unit of Sqft by Land Use Category

The last three of these types of variables are attributes of development by category within a zone. In zones which had no development of a particular type, for example in the office land use category, the land value, improvement value and total value of office land use were considered missing rather than zero, and therefore omitted from the computation of the goodness of fit. These three measures were also weighted by the amount of development of a land use in the zone, to reduce the bias due to the use of zones as the units of measurement. Zones with relatively little development of a particular land use otherwise would weigh as much in the computation of goodness of fit as zones with substantial development, potentially biasing the calibration process towards the problems of small zones.

For the purpose of weighting together for an overall measure, the individual goodness of fit measures for the individual components listed above were grouped into the following categories, with each category given an equal weight in the overall goodness of fit measure:

- Population
- Employment
- Land
- Buildings
- Values

This represents a somewhat arbitrary combination of the component measures, and may be particularly problematic where the underlying measures are not scaled to the same range, as with the mean squared error. For this reason, we examine some of the more critical components in more detail in the results discussed below.

3.3 Calibration Results

3.3.1 Results from Simulated Annealing

The calibration results discussed here are based on two sets of simulations. One is a reasonably large set of simulation runs using the simulated annealing algorithm to perform a global parameter search on all the parameters included in the calibration. For each of these 525 simulation runs, all 8 goodness of fit measures were computed for 77 individual components. The results are saved in a series of 8 files, one per goodness of fit measure, with the full set of calibration parameters used to run the simulation, and the component and overall goodness of fit measure. These results are selectively described in this section, and the full results are available in electronic form on the ODOT ftp site.

A second set of simulations were generated as a set of sensitivity tests on selected aspects of the model inputs that were not subject to the calibration parameter search. These are briefly described as well.

These results provide some useful insights into the behavior of the model over the historical period of the calibration. The simulated 1994 values of key variables of population and employment achieve fairly high correlation measures of .916 and .838, respectively, while the R-squared measures for them are .839 and .703. On the other hand, the results show considerably lower capacity to completely reproduce the observed changes from 1980 to 1994. The correlation between the simulated 1980 to 1994 change in population and employment were .373 and .109, and the R squared values were .139 and .012.

Table 1
Summary of Goodness of Fit Results
Maximum Values Selected From Simulated Annealing

	Correlation 1994	R-Squared 1994	Correlation 1980-94	R-Squared 1980-94
Combined GOF	0.721	0.578	0.193	0.057
Total Population	0.916	0.839	0.373	0.139
Total Households	0.919	0.844	0.388	0.151
Total Employment	0.838	0.703	0.109	0.012
Basic Employment	0.657	0.432	0.269	0.072
Retail Employment	0.846	0.714	0.252	0.063
Service Employment	0.866	0.749	0.247	0.061
Gov/Ed Employment	0.981	0.961	0.147	0.021
Single family Units	0.925	0.854	0.404	0.164
Residential 2-4 Units	0.959	0.920	0.203	0.041
Multi-family Units	0.862	0.743	0.368	0.135
Industrial Sqft	0.789	0.622	0.312	0.097
Warehouse Sqft	0.847	0.718	0.227	0.051
Retail Sqft	0.714	0.509	0.326	0.106
Office Sqft	0.678	0.459	0.457	0.209
Single Family LandValue/Acre	0.829	0.688	0.101	0.106
Residential 2-4 Unit LandValue/Acre	0.863	0.745	0.308	0.095
Multi-family LandValue/Acre	0.916	0.840	0.381	0.149
Industrial LandValue/Acre	0.481	0.231	0.567	0.321
Warehouse LandValue/Acre	0.878	0.770	0.451	0.203
Retail LandValue/Acre	0.840	0.705	0.305	0.093
Office LandValue/Acre	0.857	0.733	0.356	0.127

Note: the results in Table 1 represent a composite produced by selecting the maximum value found for each individual measure listed from the set of simulation tests. They are not from a single simulation run.

The full set of results from the simulated annealing parameter search are provided in electronic format for viewing within a spreadsheet application, since they are too large to format for a printed report. There are eight files, one corresponding to each of the goodness of fit measures, each containing the results of 525 simulations generated by the simulated annealing process. The parameter values used in the run are provided first, followed by the goodness of fit measures (which start in column CU when loaded into Excel).

3.3.2 Results from Sensitivity Tests

A series of sensitivity tests were also conducted, with specific changes made to an initial base simulation run, to gauge the effect of particular input assumptions on the model fit. The sensitivity tests run included the following:

- Run 1 Base run with initial values from cross sectional calibration and initial assumptions for remaining dynamic coefficients
- Run 2 Relax constraints on developing land outside UGB, and in wetlands, floodplains, high slope areas.
- Run 3 Same as 2 but with low household and business mobility rates
- Run 4 Same as run 2 but with all mobility rates set to zero
- Run 5 Same as run 4 but with an extreme penalty on redevelopment (100 times higher cost)
- Run 6 Same as run 2 but with Low mobility rates and low redevelopment penalty
- Run 7 Same as run 2 but with low mobility rates and extreme development penalty
- Run 8 Original mobility rates and no redevelopment penalty.

Results of selected sensitivity runs are shown in Tables 2 and 3 below. These results are from single runs, with only the changes identified above made from the base run. For brevity, only the correlation coefficients on 1994 and on the change from 1980 to 1994 are shown.

Table 2
Correlation Coefficients for 1994
Selected Sensitivity Runs

	Base	Run 2	Run 5	Run 8
Combined GOF	0.691	0.692	0.696	0.683
Total Population	0.912	0.912	0.912	0.910
Total Households	0.912	0.912	0.911	0.905
Total Employment	0.812	0.814	0.839	0.827
Basic Employment	0.641	0.616	0.646	0.632
Retail Employment	0.741	0.758	0.808	0.805
Service Employment	0.805	0.808	0.863	0.823
Gov/Ed Employment	0.980	0.980	0.980	0.981
Single family Units	0.915	0.915	0.915	0.918
Residential 2-4 Units	0.932	0.936	0.908	0.784
Multi-family Units	0.853	0.851	0.852	0.853
Industrial Sqft	0.753	0.760	0.783	0.794
Warehouse Sqft	0.782	0.780	0.795	0.801
Retail Sqft	0.689	0.692	0.687	0.688
Office Sqft	0.619	0.635	0.615	0.604
Single Family LandValue/Acre	0.759	0.764	0.755	0.690
Residential 2-4 Unit LandValue/Acre	0.788	0.791	0.777	0.699
Multi-family LandValue/Acre	0.804	0.809	0.812	0.749
Industrial LandValue/Acre	0.361	0.325	0.348	0.301
Warehouse LandValue/Acre	0.723	0.752	0.742	0.758
Retail LandValue/Acre	0.658	0.653	0.627	0.626
Office LandValue/Acre	0.719	0.727	0.745	0.749

The sensitivity test runs presented here contain what appear to show reasonable correlation between the simulated and observed 1994 values, with the highest correlations for population and housing elements. However, the correlation coefficients in Table 3 for the changes from 1980 to 1994 are generally low, and the coefficients for land values per acre are actually negative. This suggests that in the set of sensitivity tests shown here, the simulated changes from 1980 to 1994 are not mirroring at a zone level the changes observed in the calibration targets. As seen in Table 1, however, there were numerous runs within the larger set of simulations run as part of the simulated annealing that produced correlation coefficients substantially above these sensitivity test runs, with correlation coefficients on the changes in land values ranging from 0.1 to close to 0.6.

Table 3
Correlation Coefficients for 1980 to 1994 Change
Selected Sensitivity Runs

	Base	Run 2	Run 5	Run 8
Combined GOF	0.115	0.129	0.105	0.121
Total Population	0.324	0.330	0.320	0.364
Total Households	0.345	0.349	0.333	0.323
Total Employment	0.040	0.045	0.118	0.072
Basic Employment	0.116	0.032	0.034	0.006
Retail Employment	-0.025	-0.039	-0.006	0.045
Service Employment	0.136	0.154	0.281	0.141
Gov/Ed Employment	0.120	0.109	0.129	0.140
Single family Units	0.363	0.359	0.354	0.335
Residential 2-4 Units	0.068	0.093	0.047	0.063
Multi-family Units	0.128	0.103	0.125	0.209
Industrial Sqft	0.131	0.145	0.211	0.310
Warehouse Sqft	0.107	0.140	0.012	0.068
Retail Sqft	0.046	0.084	0.040	0.052
Office Sqft	0.211	0.268	0.205	0.187
Single Family LandValue/Acre	-0.249	-0.233	-0.252	-0.374
Residential 2-4 Unit LandValue/Acre	-0.091	-0.076	-0.057	-0.265
Multi-family LandValue/Acre	-0.058	-0.049	-0.066	-0.338
Industrial LandValue/Acre	-0.194	-0.179	-0.083	-0.161
Warehouse LandValue/Acre	-0.023	0.040	-0.036	-0.007
Retail LandValue/Acre	-0.139	-0.119	-0.137	-0.166
Office LandValue/Acre	-0.119	-0.078	-0.053	-0.134

One way to assess the behavior of the model over time is through an inspection of the aggregate dynamics that the model produces. For this purpose, one sensitivity run is selected (run 5) and presented in graphical form in Figures x – y below. The overall dynamic interaction between changes in the aggregate population and employment supplied as exogenous control totals, and the model response through demand and supply of real estate, appear to be operating in a reasonable way. As population or employment grow in the aggregate, vacancy rates in affected real estate sectors decline, eventually triggering a supply response which raises vacancy rates again. Lack of perfect equilibrium due to imperfect information and continued external shocks, such as continued change in aggregate population and employment, produce overshooting of real estate construction, and a consistent elevation of vacancy rates, which then triggers a self-regulating drop in construction activity.

Figure 1

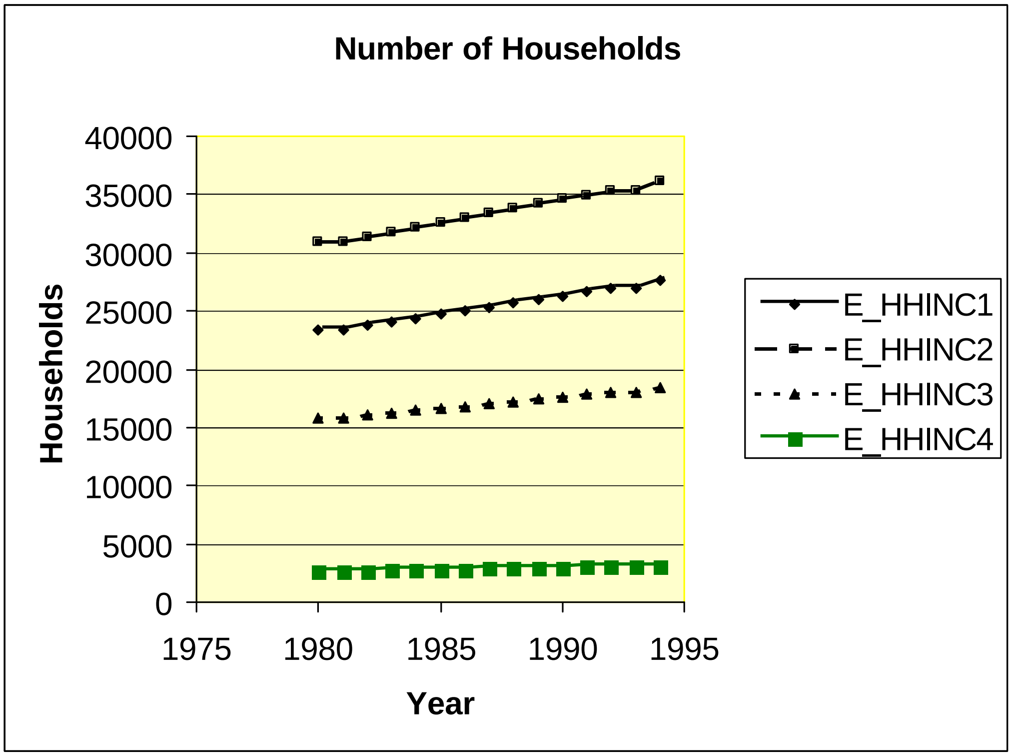


Figure 2

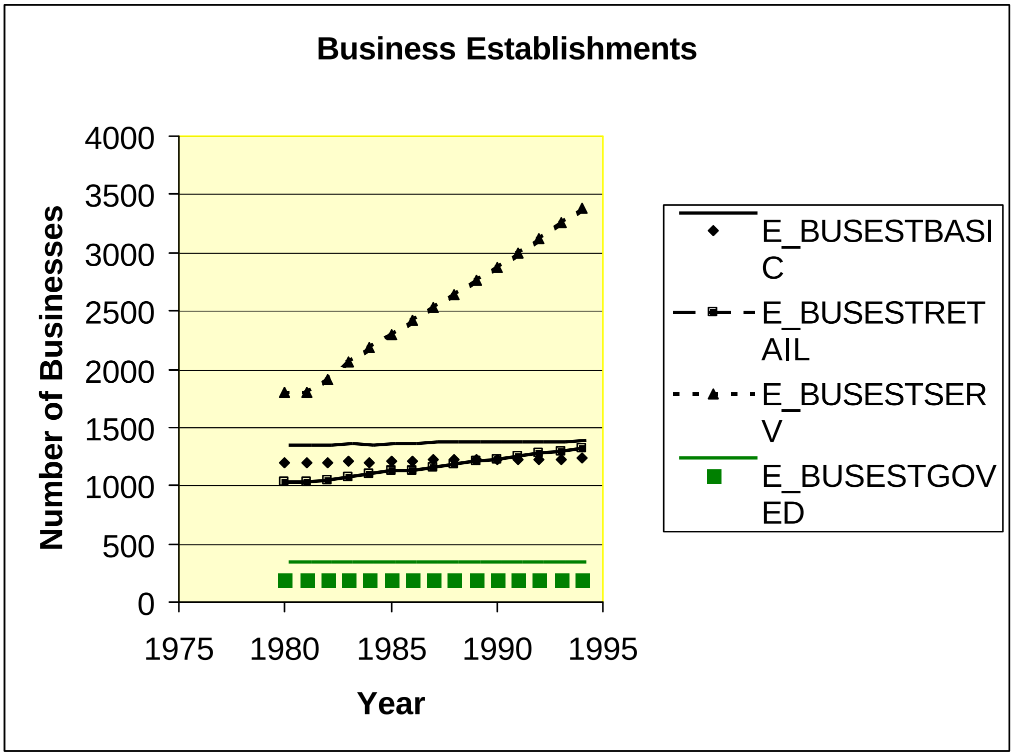


Figure 3

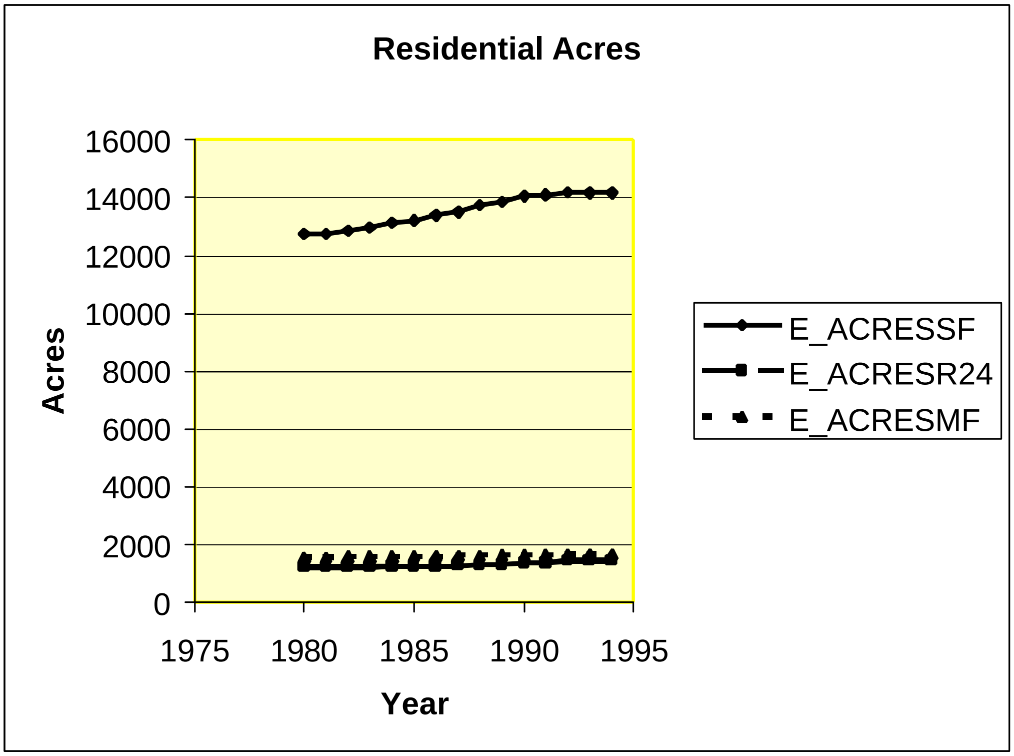


Figure 4

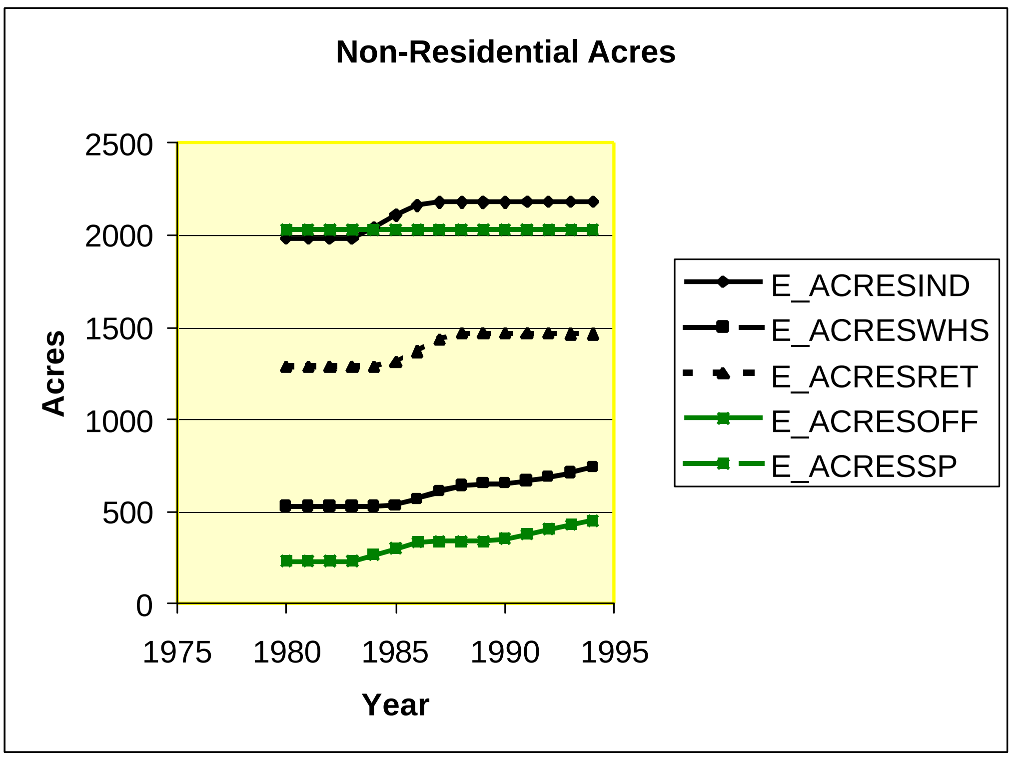


Figure 5

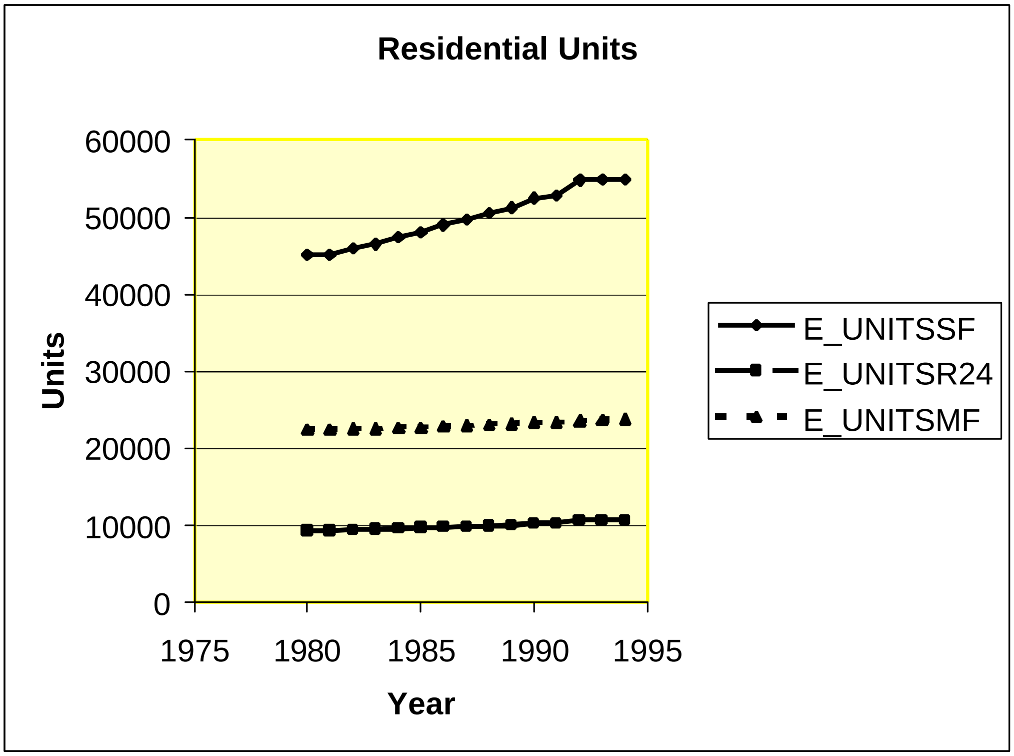


Figure 6

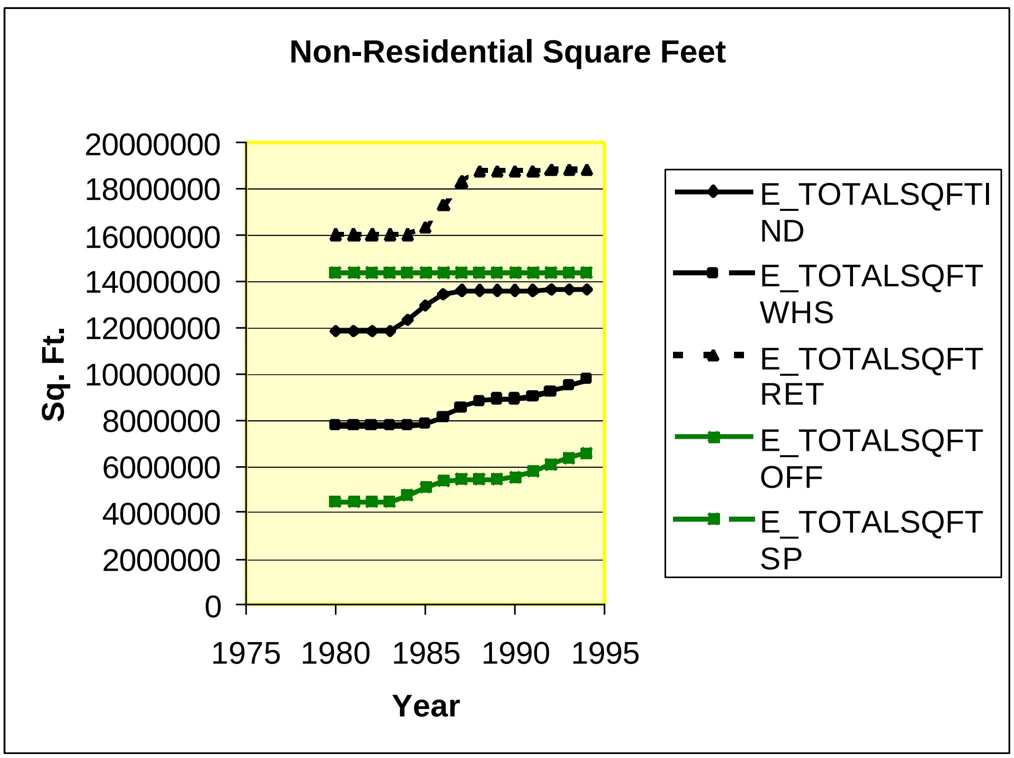


Figure 7

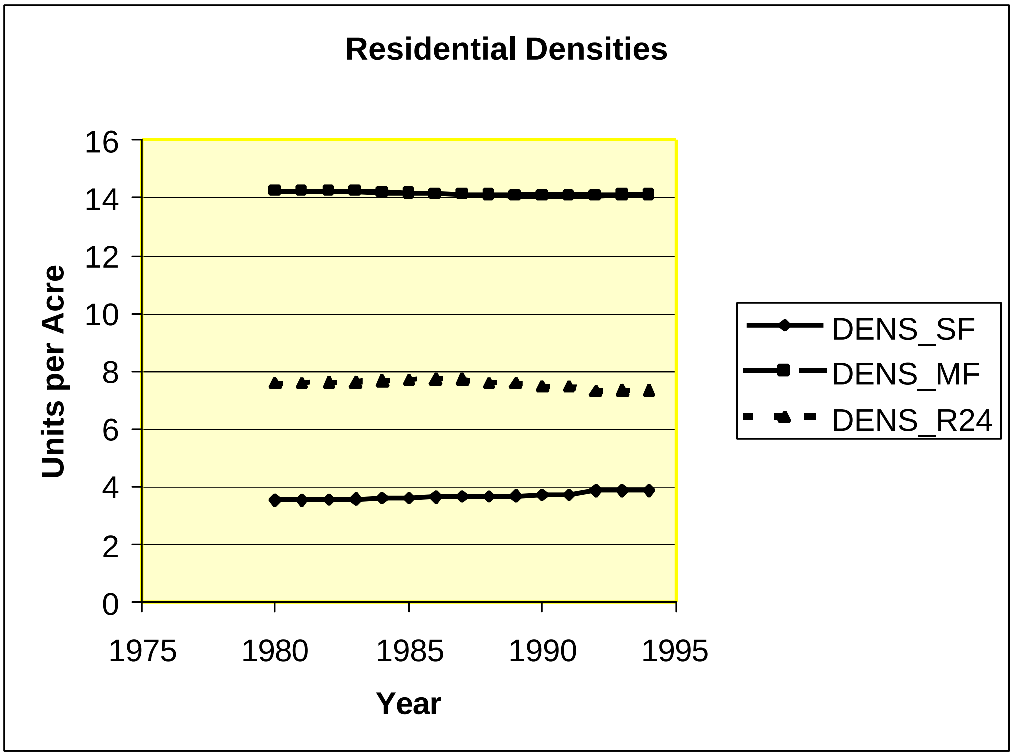


Figure 8

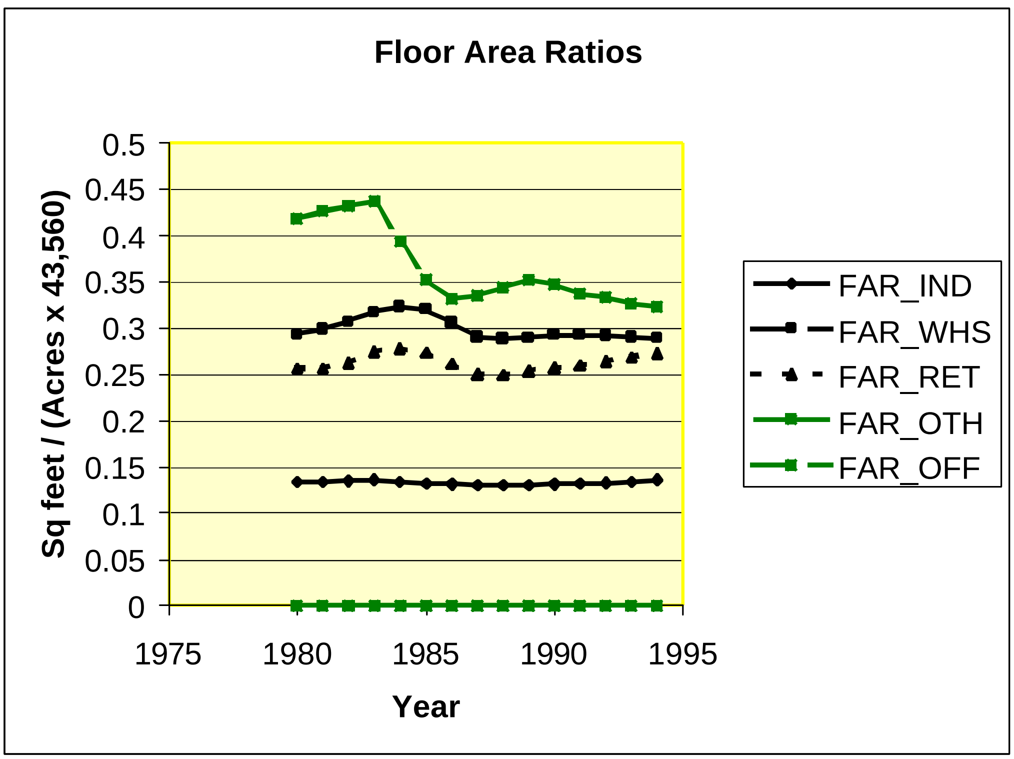


Figure 9

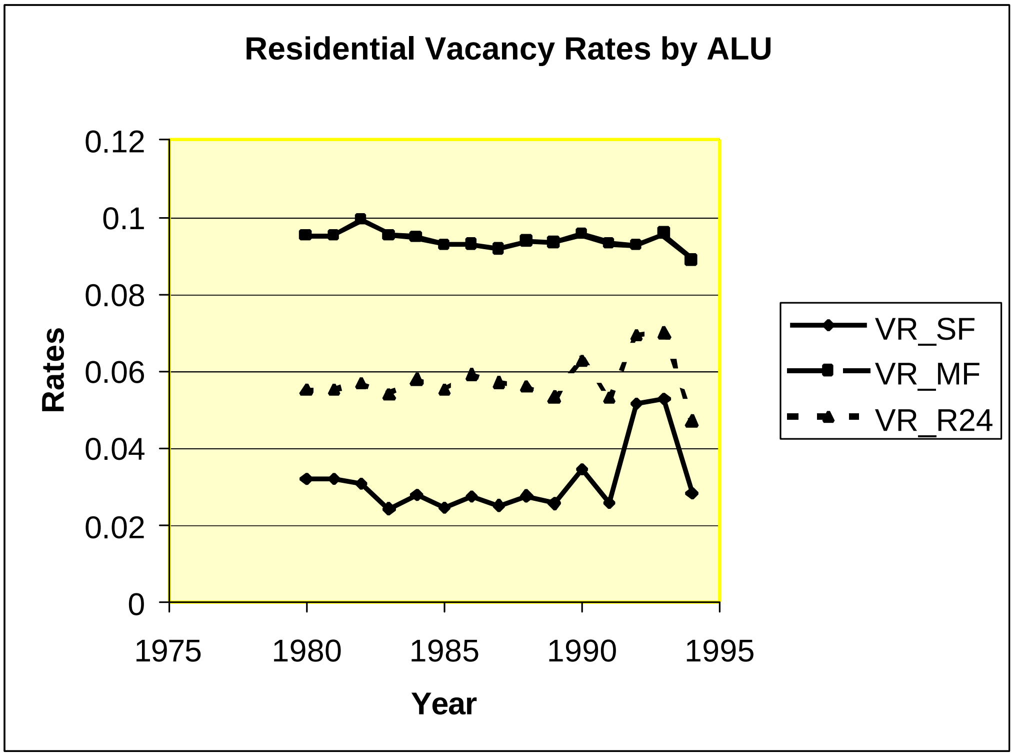


Figure 10

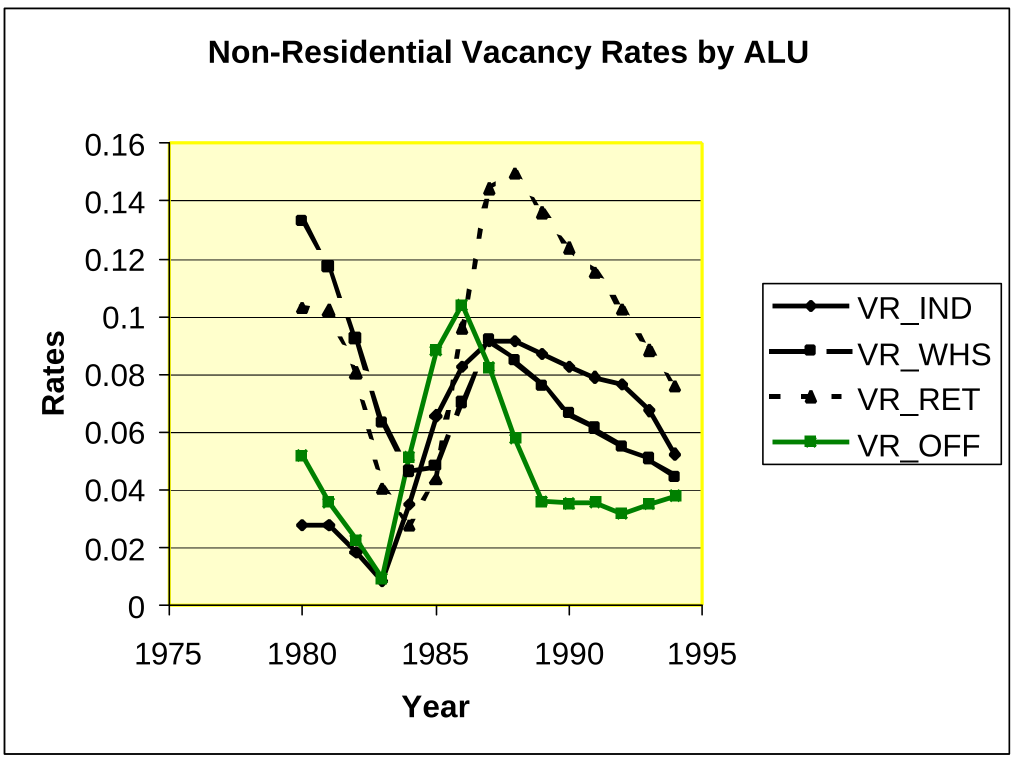


Figure 11

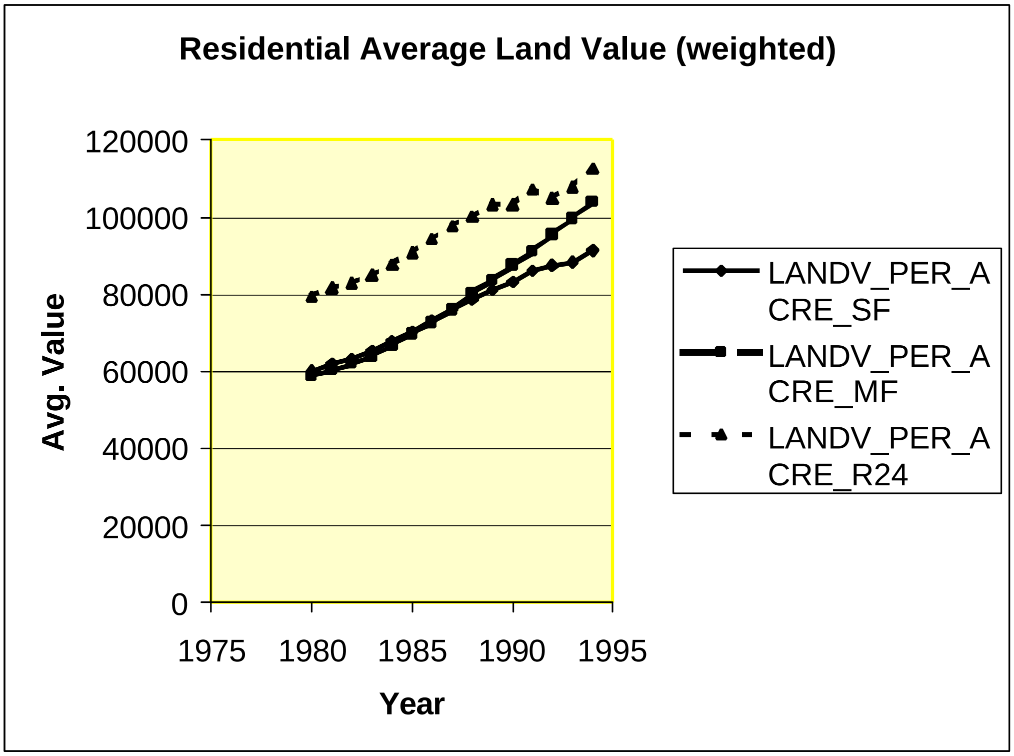


Figure 12

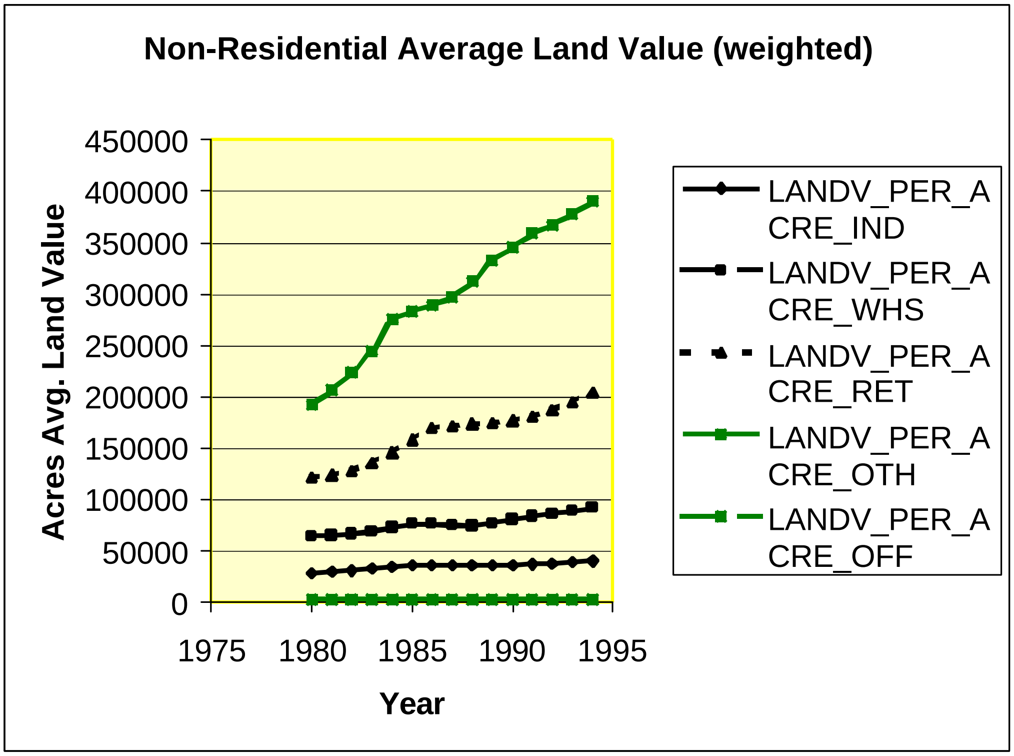


Figure 13

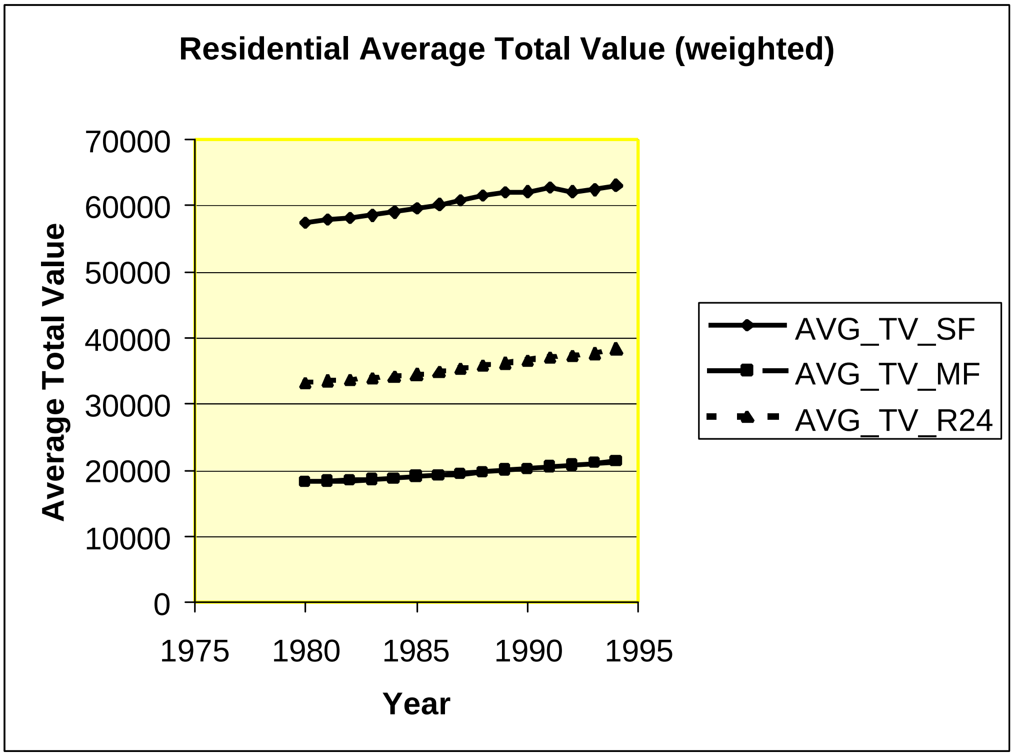
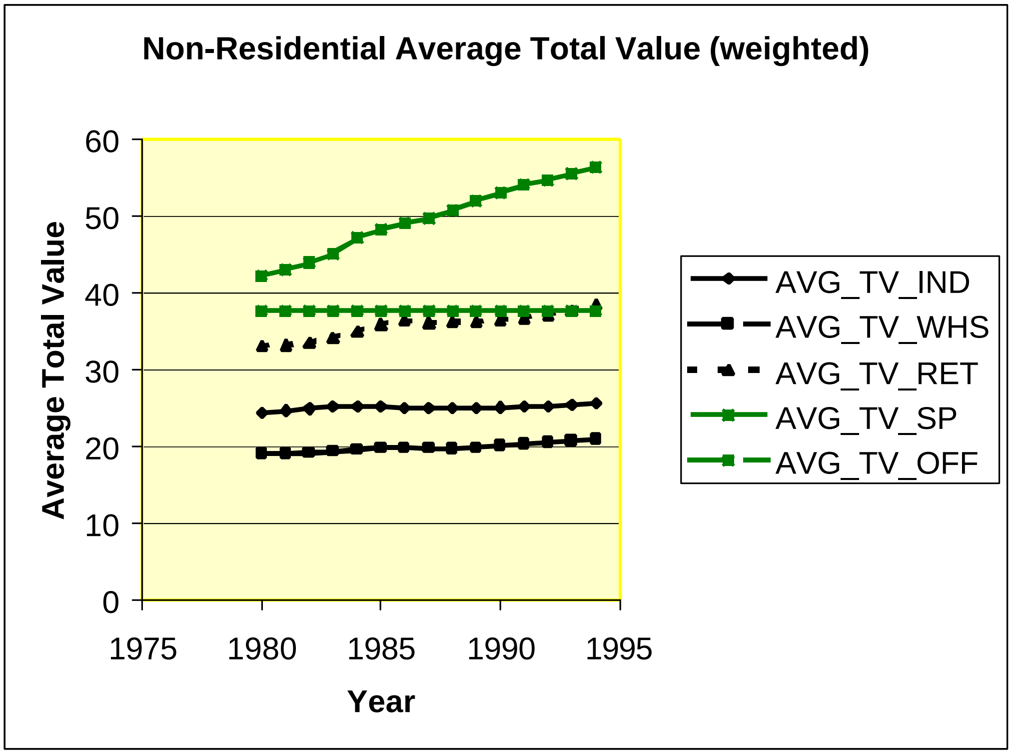


Figure 14



3.3.3 Measurement Error in the Input Data and Calibration Targets

It should be noted that the goodness of fit measures are all computed by comparing against calibration targets that are based on the database assembled as input for the model to begin in 1994. The target database includes the data synthesis steps described for the 1980 database, for consistency. Unfortunately, the data synthesis steps may introduce further measurement error in the calibration targets. A systematic comparison of the model input data used for 1980 and 1994 to the LCOG data available for households and employment reveals, counter-intuitively, that the errors between the datasets is greater in 1994 than in 1980. So while the data appear reasonable for use in the kind of calibration and testing done within this project, they do contain a significant amount of noise that the calibration process is forced to attempt to simulate as actual change. At best, the measurement errors reduce the efficiency of the calibration exercise.

To begin to assess the potential measurement error, a series of comparisons are made between the data used as input to the model or a calibration targets, and the data generated by LCOG for input to their travel models. These comparisons are limited to households and population, since the LCOG data did not contain other variables such as land values, housing units, or commercial square footage. Figures 15 and 16 compare the household count by zone in 1980 and 1994, respectively, between the caliobration data and the LCOG estimates by zone. These two graphs reveal a substantial agreement between the two estimates, but also show that there are several significant outliers where there is considerable disagreement. Figure 17 compares the 1980 to 1994 household change observed in the LCOG data to that observed in the model input data, and shows that while the 1980 and 1994 comparisons look quite good, the comparison of the changes appears to compound the errors. A similar pattern is observed in the employment data in Figures 18-20.

Figure 15

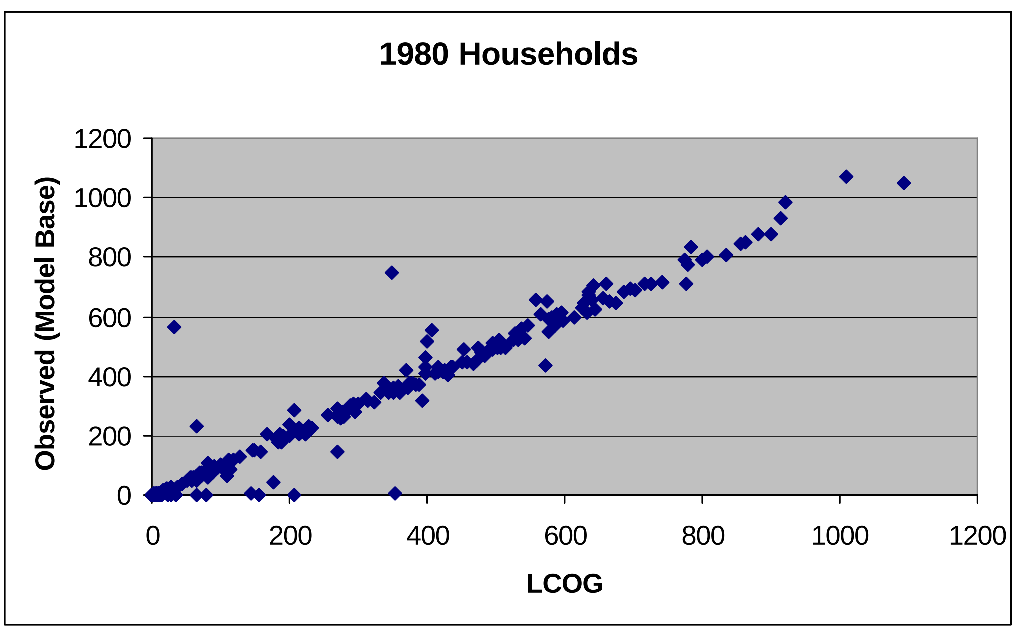


Figure 16

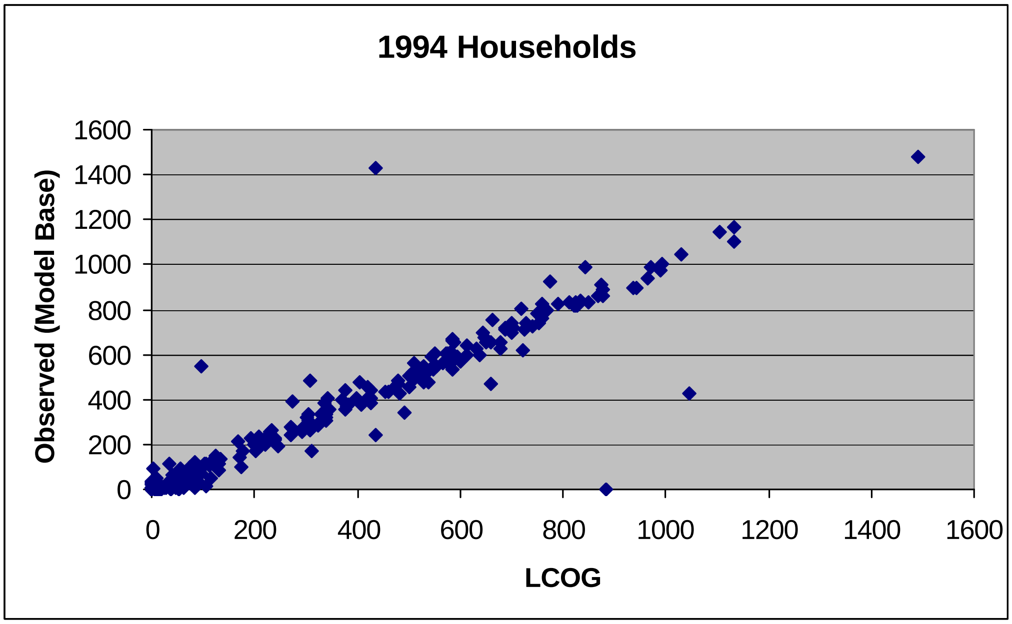


Figure 17

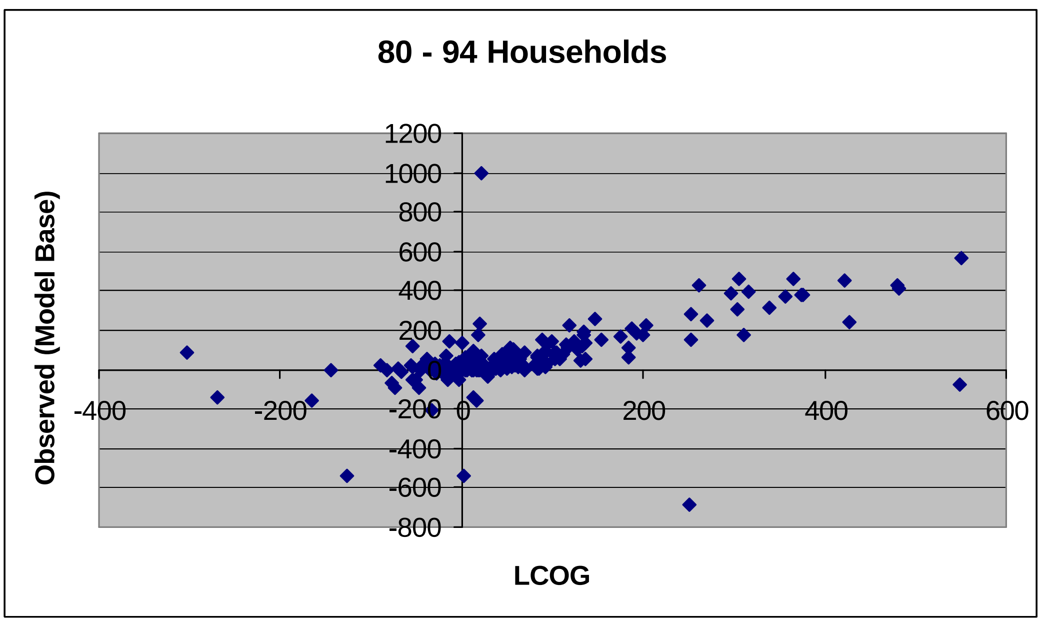


Figure 18

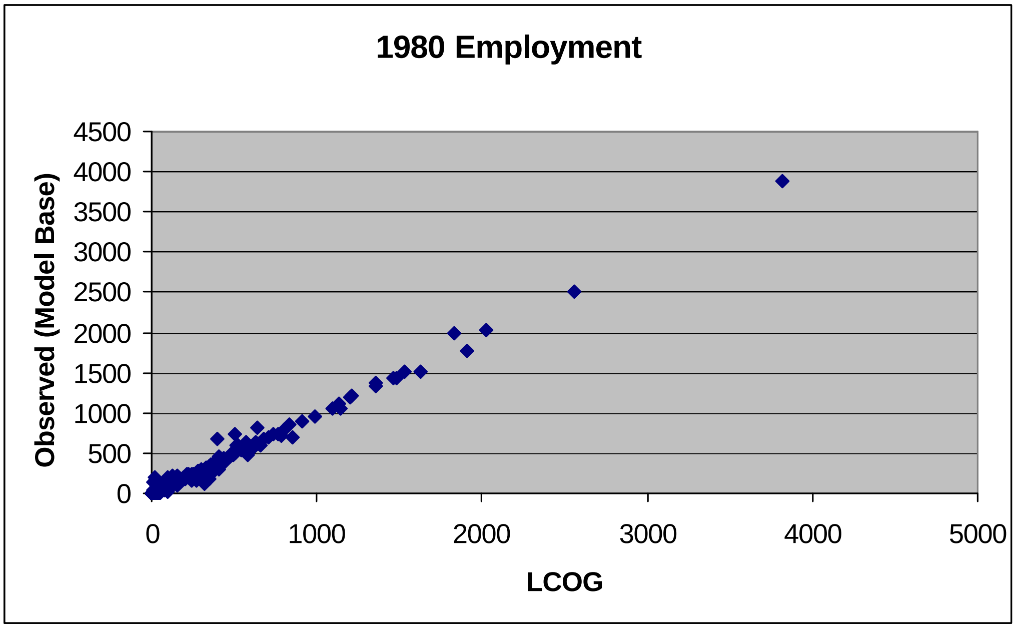


Figure 19

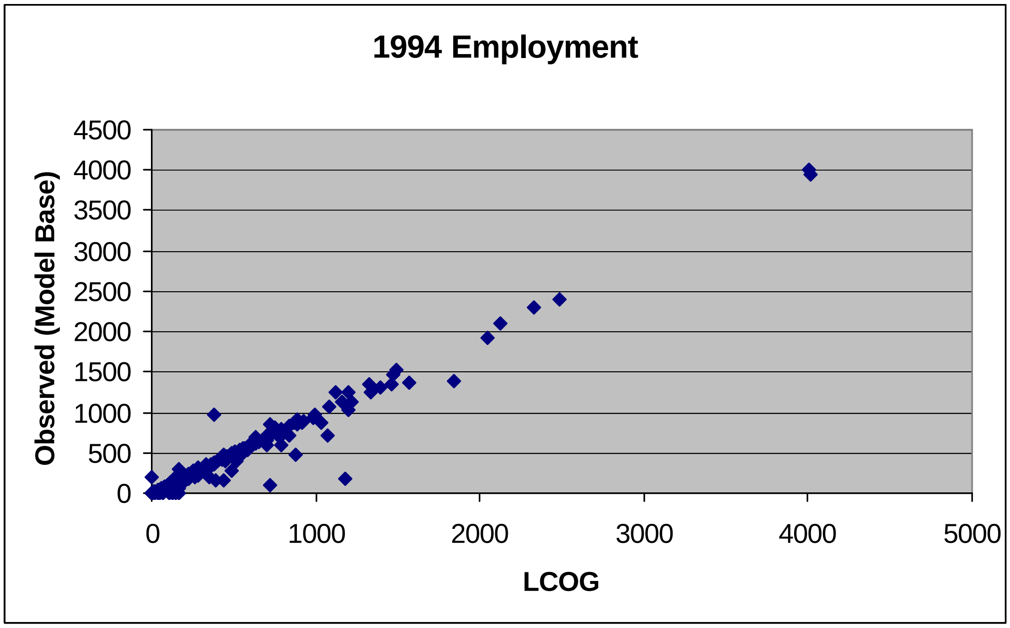
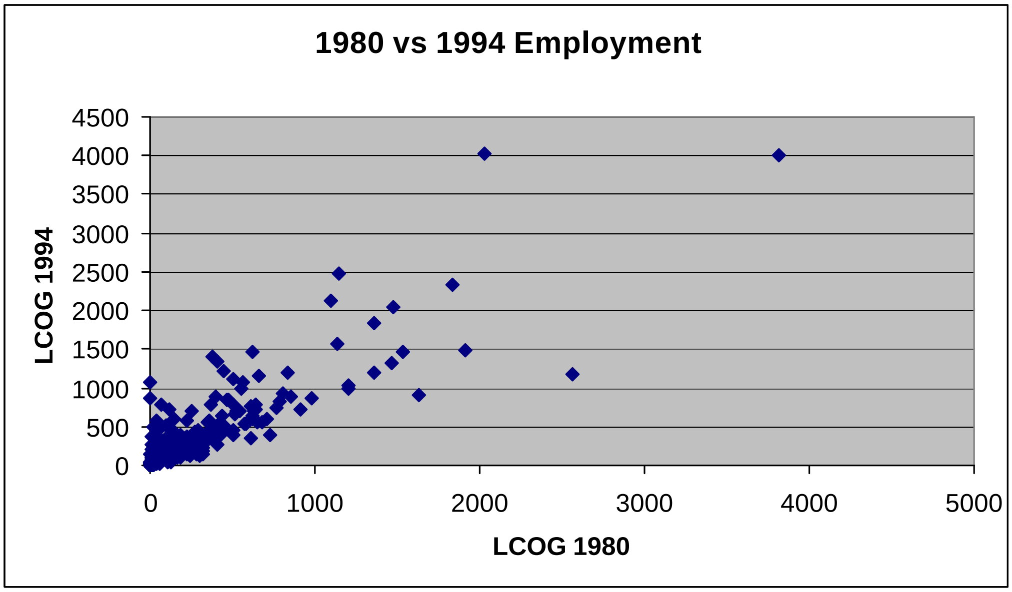


Figure 20



These tests of the agreement between the LCOG zonal estimates and the data used as input to the calibration and as calibration targets for 1994 are fairly limited, in that there is no real external benchmark to compare against, and only two variables are compared. Nonetheless, they provide some assurance that the basic data used in the calibration is reasonably accurate and consistent, but that changes from 1980 to 1994 will be particularly prone to measurement errors. This will undoubtedly limit the effectiveness of the longitudinal calibration to some extent, but placing additional effort into cleaning the data to reduce these errors is not likely to be feasible or warranted.

3.3.4 Graphical Analysis of Sensitivity Tests

To further analyze the results of selected simulation runs, a set of comparisons of the 1994 and 1980 to 1994 change were made in graphical form. In Figures 21-25, comparisons are made between the model household input data for 1980 and the model input data for 1994 (observed), the base simulation test, and runs 2, 5 and 8. These are followed by a similar comparison for employment in 1980 and 1994 in Figures 26-30. The results shown in these figures confirm two overall findings. First, that there are many zones that are shown to contain significant losses in households or employment according to the 'observed' model input data (Figures 21 and 26), and that these declines are not tracked by the simulation runs. In the employment data, there is a substantial loss of employment in one zone, that on further inspection appears to reflect the closure of a large Weyerhouser plant in zone 31 in Springfield. It is not surprising that the simulations do not mirror these large discrete events.

Figure 21

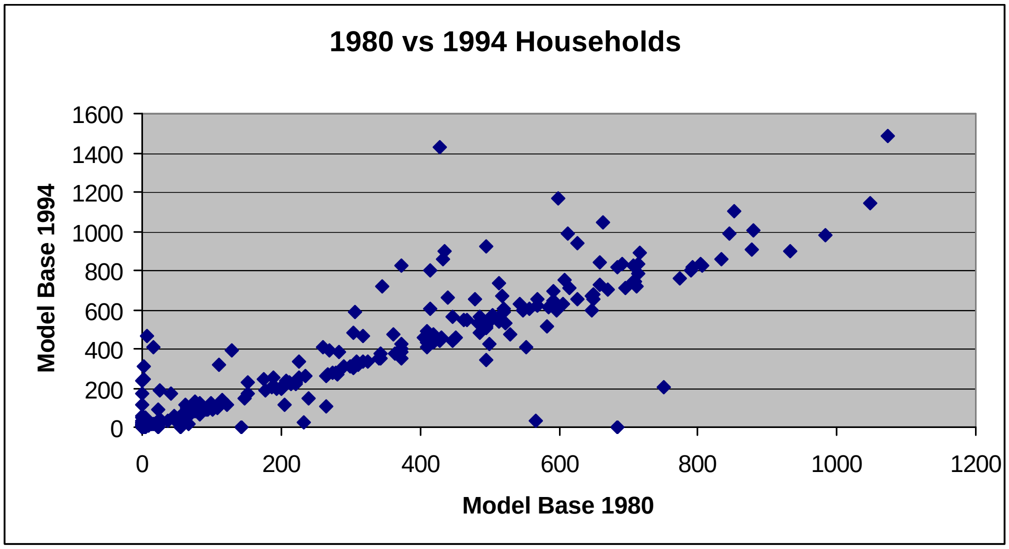


Figure 22

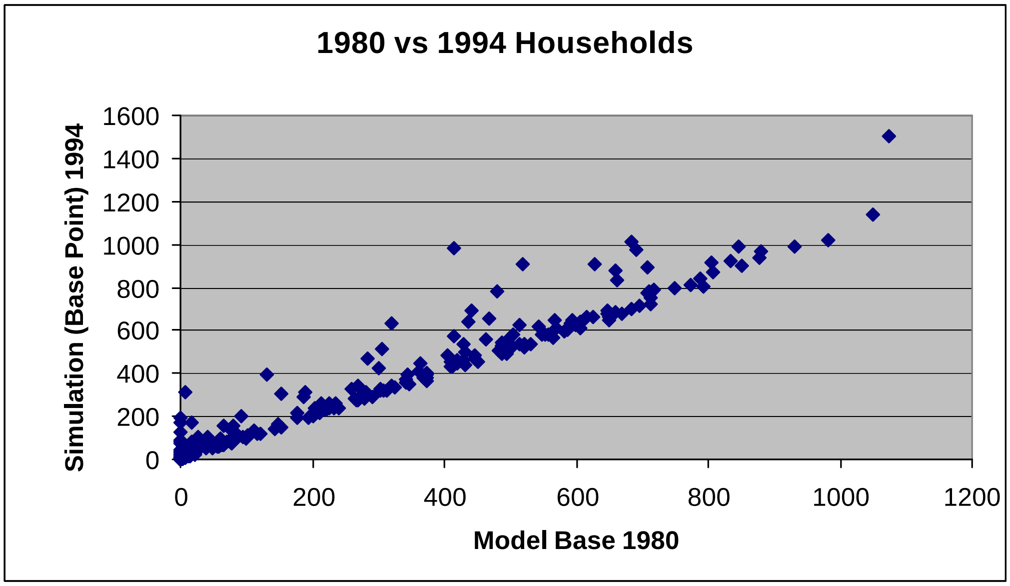


Figure 23

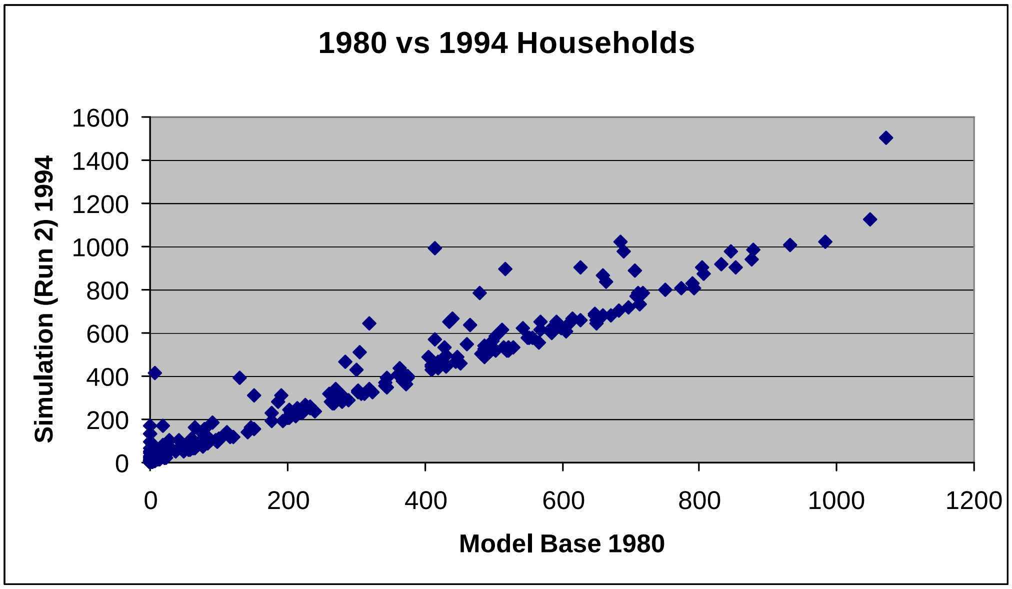


Figure 24

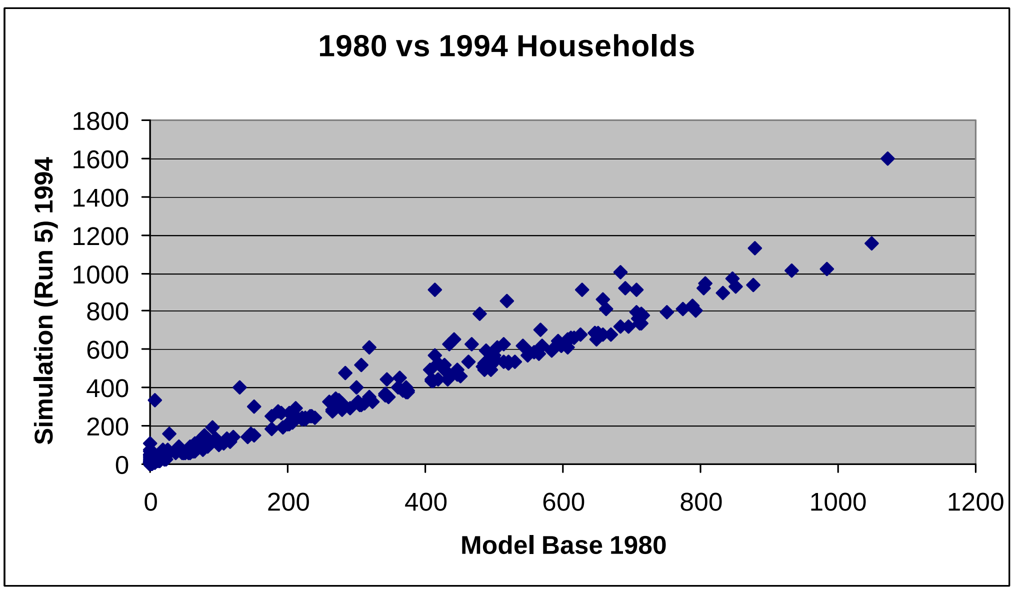


Figure 25

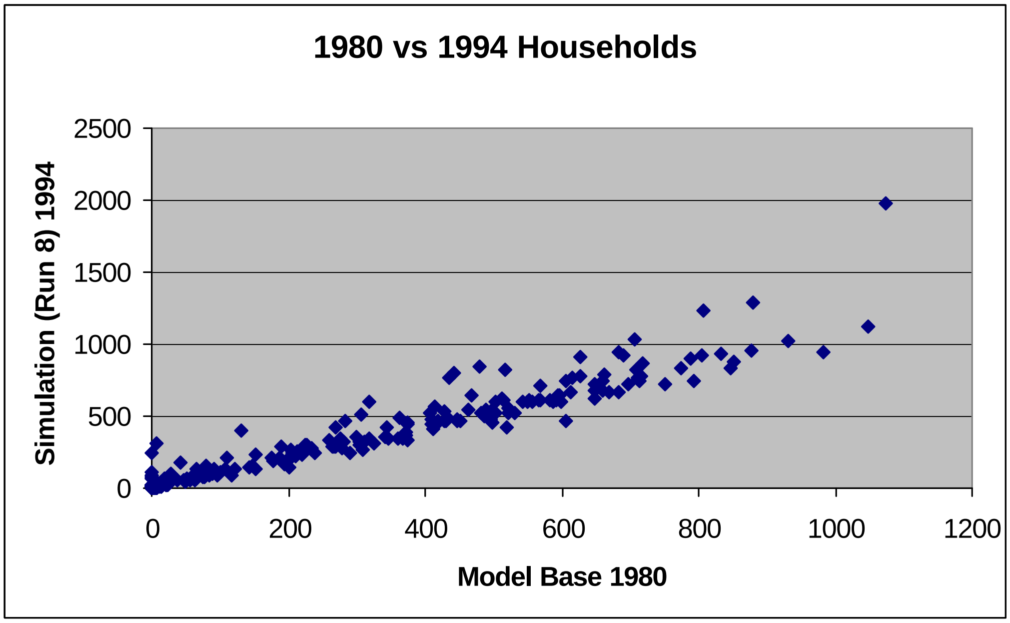


Figure 26

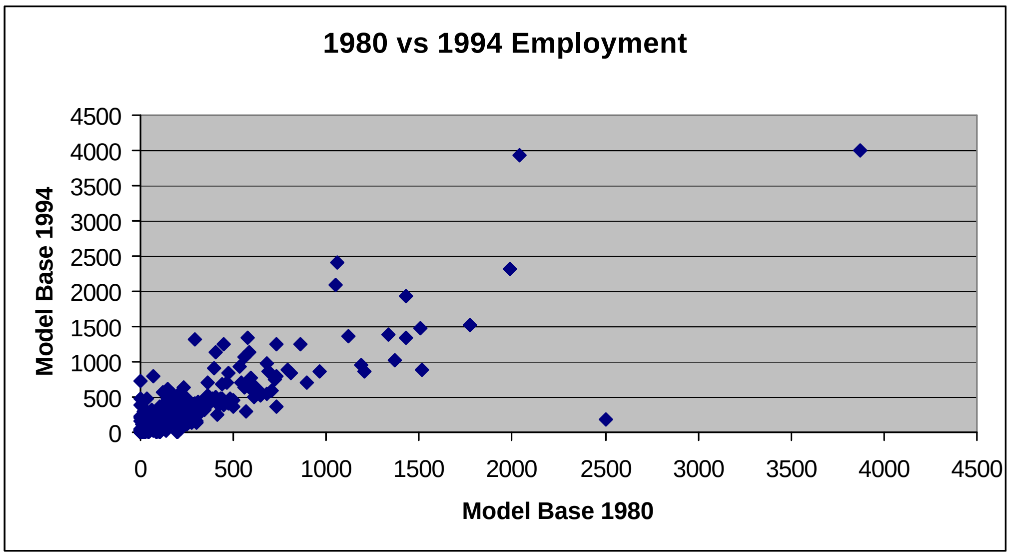


Figure 27

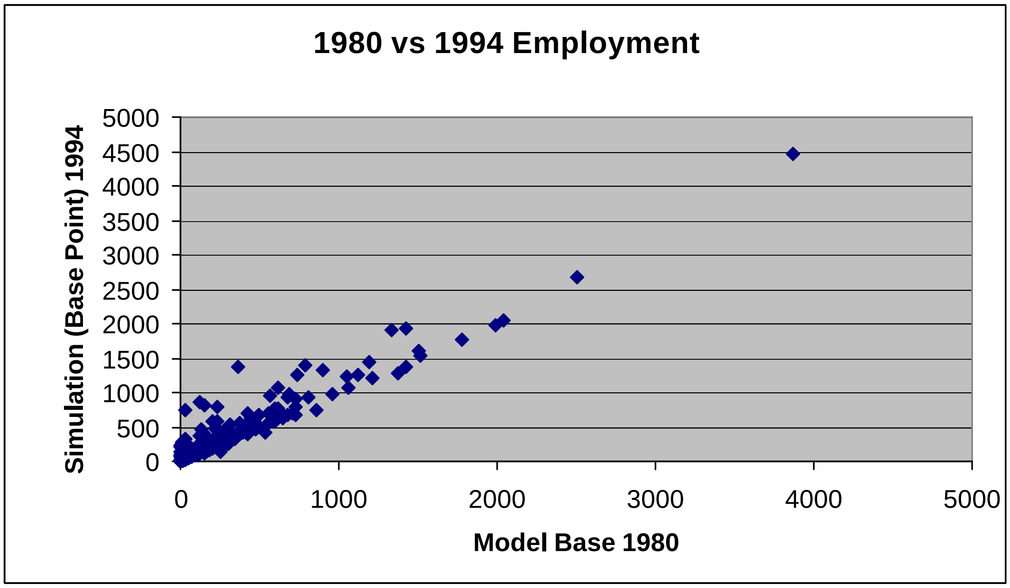


Figure 28

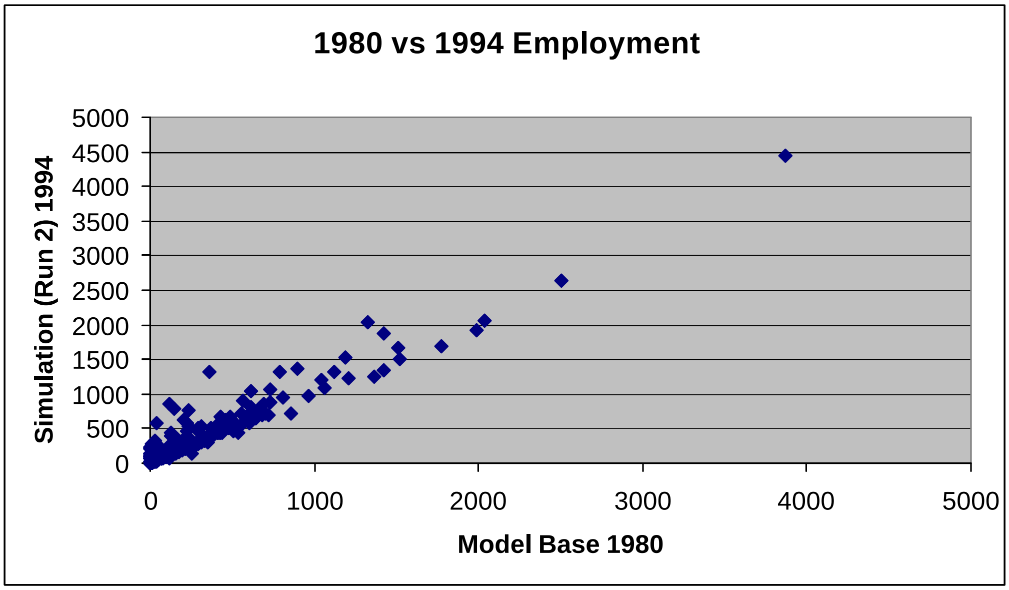


Figure 29

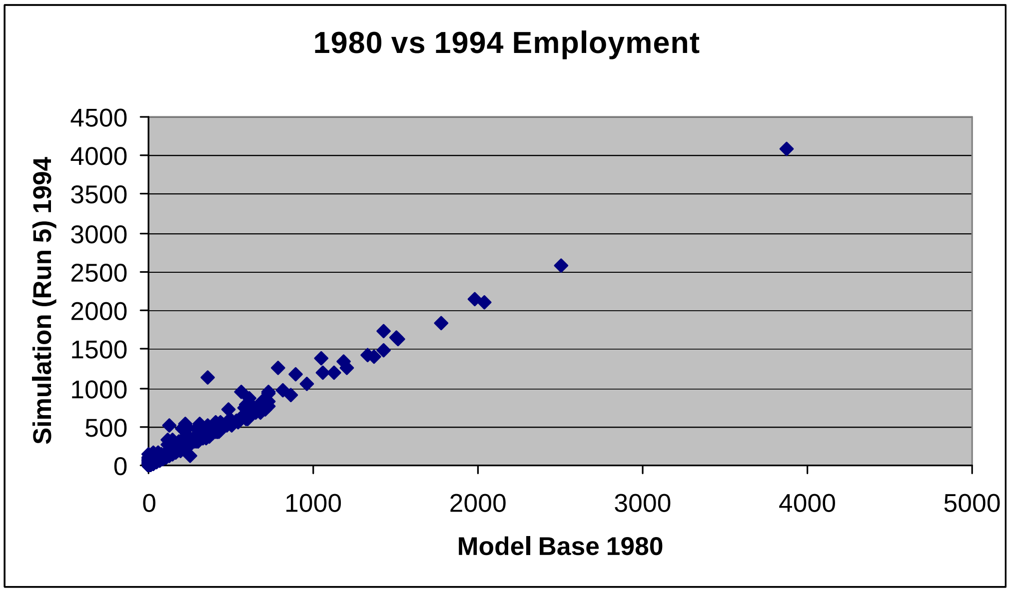
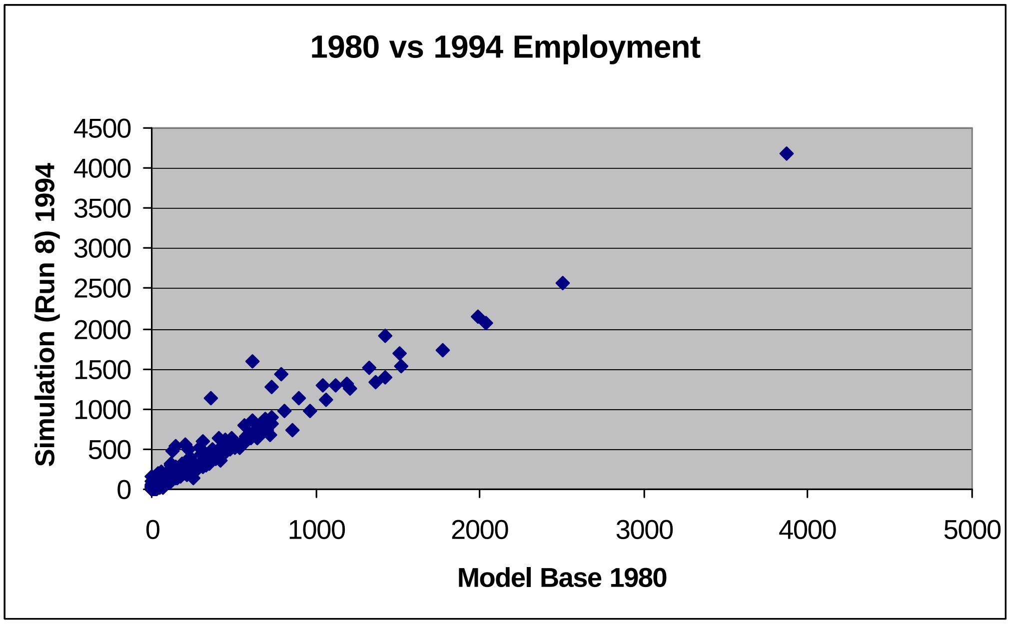


Figure 30



3.3.5 Map Analysis of Sensitivity Tests

A final step in analyzing the sensitivity tests was taken to visualize the simulation results in map form. This provides a useful way to identify spatial patterns in the differences between simulated and observed changes, and complements the analysis presented in the preceding section. These maps are compiled in Appendix D, and include spatial representations of the differences between the LCOG and model input data, as well as between the observed and simulated changes in population, employment, housing units and commercial space by type.

Some particular changes in the employment data from 1980 to 1994 warrant notice. The loss of employment in zone 31 associated with a Weyerhouser plant closure has already been mentioned, and shows up as a significant discrepancy between the simulated and observed changes in these maps. In addition, the opening of the northgate mall and the other retail and office concentrations in several zones adjacent to the north side of the downtown area were not picked up by the simulation results. These are significant development and business events that the model did not anticipate in any of these tests. It is not likely that major business and development events could ever be well simulated, especially with a high degree of spatial detail. This is why a capacity has been developed in the model for user-input of business and development events. This capacity was not used in the longitudinal calibration.

4 Summary of Findings and Recommendations for Gen 2

The calibration results presented in the preceding section through goodness of fit measures and visualization in graph and map forms, provide a first useful assessment of the model system over a known period of history. The results are mixed, at this point. Some of the key findings are:

- The overall model dynamics, as shown in the graphs of the aggregate behavior of the model over time, appear reasonable.
- Goodness of fit measures comparing the simulated to observed 1994 values appear fairly strong, with the best results for population and housing variables, good results for employment, and less satisfactory results for selected building types, and for land values.
- Goodness of fit measures on the changes from 1980 to 1994 appear relatively low, at least in the sensitivity tests, though reasonable results for each were obtained in selected runs from the simulated annealing process. This suggests that further analysis of the simulated annealing results may yield insight into parameter combinations that improve the overall model performance over time.
- Spatial patterns of change in the simulated results appear fairly consistent with the overall observed patterns of change, with the greatest agreement in those variables

with higher goodness of fit measures, such as households and housing. The spatial comparisons show the least similarity for commercial building stock changes, where there appear to be large declines and gains in the observed data that are not matched in the simulations.

- Measurement errors in the input and calibration target data appear fairly significant, particularly for the 1980 to 1994 changes, and may be adversely influencing the calibration results.
- Specific major development events such as the Weyerhouser plant closure or the construction of a major shopping mall, are not well simulated and probably cannot be effectively simulated by any model.

While it may be possible to make further improvement in the model performance by more effectively mining the simulated annealing results, further cleaning input or calibration target data, or adding business and development events for major changes, or making minor specification changes beyond those already undertaken in this analysis, there are probably more productive changes in the model structure that would be more helpful in evolving the model system and supporting the development of the second generation model system. We close this report with a brief consideration of some of these proposed changes.

4.1.1 Land Development Model

One of the key deficiencies in the current model implementation is in the design of the land development and redevelopment component. The principal weaknesses are that the model component is structured as a deterministic process that simulates the construction of development where it is most profitable, rather than considering this a stochastic choice process in which profit expectations are tempered by risk considerations. In addition, the land development model does not incorporate any spatial context in the decision to build a development project, so a vacant parcel in the middle of an industrial park might appear the same as one in a residential subdivision, if it had the same land use plan designation and other attributes.

A high priority for further development of the model will be to reformulate it as a logit model structure, with profit and risk components included in the utility functions. A new specification for this model component is in draft stage, and development of the software infrastructure to support the simulation of the nested choice process is underway. Similarly, a means of measuring the spatial context around a site will need to be developed to address the second limitation in the current implementation.

4.1.2 Grid Implementation

In order to address the need to measure spatial context, a grid spatial infrastructure has been proposed, and is now under development. The proposed specification would allow cross-referencing of parcels and grid cells, allowing the development model to remain at the parcel level where data support this, or be run in a more aggregate mode using only the cell infrastructure. This flexibility will require that the software be well generalized.

4.1.3 Visualization

One of the clear lessons of the calibration process is that integrated visualization tools are not a luxury, but a necessity. While it was possible to generate the graphs and maps for this report using external tools such as Excel and Arcview, the amount of labor it required severely constrained its effectiveness for testing and evaluating input data and simulation results. An infrastructure for producing these kinds of visualization through tools that are directly linked to the model system is now underway.

4.1.4 Integration with Statewide Model

One of the key remaining issues to be addressed is the integration of the metropolitan scale model with a statewide model, to form the second generation Oregon models. This discussion is beyond the scope of this report, and is not treated here. Hopefully, the results of the longitudinal calibration testing that have been compiled here will assist in further refinement of the model design and specification in support of this activity.