NC 73

Transportation / Land Use Corridor Plan

Table of Contents

Technical Appendix:

Technical Appendix

Demographic and Economic Forecasts for the Route 73 Corridor

Prepared for: HNTB Corporation
By: Thomas R. Hammer, Ph.D.
December 3, 2003

CONTENTS

APPENDICES

I. Summary & Assessment of Corridor Findings

North Carolina Route 73 is a highway extending somewhat over thirty miles from I-85 in Cabarrus County to Lincolnton in Lincoln County. For the most part it is a low-capacity road traversing sparsely developed terrain, but its alignment crosses a portion of the Charlotte region with very high recent growth and still higher growth prospects. Due to concern about the rising usage and future adequacy of Route 73, local public officials and other concerned parties have come together to address the issue, and after long effort have secured funding for a consultant study. The purpose of the study is to promote an upgrading of the roadway, in some appropriate time frame, in a fashion consistent with all relevant community objectives. To some extent this study has become a pilot project in proactive, community-based, multi-jurisdictional transportation planning.

The steering committee for the NC 73 Transportation/Land Use Corridor Study includes representatives of three counties, four cities, three chambers of commerce, four regional planning organizations and the North Carolina DOT. The prime contractor for the study is HNTB Corporation (Charlotte). The four subcontractors are: Tom Sawyer Company, for public involvement; S/K Transportation Consultants, Inc., for access management; the UNCC Center for Transportation Studies, for implementation; and the present economist for socioeconomic forecasting.

The present document describes the derivation of demographic and economic forecasts to support the conceptual design of future Route 73 improvements. The key elements of the proposed roadway design were determined at an engineering session on September 9, 2003, using forecasts delivered to HNTB on September 4. The forecasts came from a modified version of a regional modeling framework that had been developed for other purposes. This framework utilized “top-down” allocation procedures, which had various strengths and weaknesses as discussed later. The economist’s work plan called for incorporating “bottom-up” perspectives if feasible, but this was ruled out by the timing of the Route 73 project relative to other activities. Hence the present text will focus upon top-down forecasting and its outputs (which were modified in only one area).

The corridor designated by HNTB for analysis of Route 73 transportation demand is an oblong-shaped area extending seven to ten miles north-south and about thirty-five miles east-west. It covers the northwestern corner of Cabarrus County, the northern extreme of Mecklenburg County, and roughly the eastern half of Lincoln County. The Mecklenburg portion accounts for less than a third of the corridor’s land area but nearly half of its current population. Development of the corridor has been strongly influenced by the presence of I-77 and frontage on Lake Norman, which divides the portions of the corridor in Mecklenburg and Lincoln counties.

The corridor has been partitioned into seven component areas, numbered from east to west. (This text uses the original numbering specified by a map that HNTB sent to all project participants in the summer of 2003. A less detailed map appears here in Figure 10 of Section VI.) Area 1 is a relatively small slice of northern Cabarrus County that covers parts of Kannapolis and Concord. Area 2 is a large district comprised about equally of land in Cabarrus and Mecklenburg counties. Area 3 contains I-77 and most of Mecklenburg’s frontage on Lake Norman. Area 4 is a sliver of land extending south along the lake and the Catawba River. Areas 5 and 6 comprise most of eastern Lincoln County, and Area 7 is a small zone centering upon Lincolnton. The approximate land areas of these corridor segments are listed in the table below.

The present summary and explanation of forecasting results will focus only on population, the descriptor of primary concern for conceptual planning, and will only consider data for years ending in zero. (The baseline year for forecasting was 2002.) Further detail will be provided in Section VI.

Table 1 below gives the population forecasts for the Route 73 corridor through 2030. Due to circumstances explained later in this section, most attention is given to the corridor totals rather than the figures for component areas. The table's last two lines describe, in both absolute and percentage terms, the actual change in total corridor population that occurred during 1990-2000 and the changes forecasted to occur over the next three decades.

The expected gains in corridor population are very large, perhaps disconcertingly large. The numbers say that the population of the corridor as a whole will more than double from 2000 to 2020, from about 105,000 persons to 226,000 persons. (The latter figure is almost exactly double the corridor's 2002 population.) By 2030 the population of the Route 73 corridor is expected to be somewhat above 300,000 persons. The extent to which past trends support this forecast is a matter of perspective. On one hand, the corridor's population increased by 60% between 1990 and 2000, whereas its future gains would range down from 48% to 35% per decade according to the forecast. On the other hand, the forecast calls for steadily rising absolute increases, from under 40,000 persons in 1990-2000 to over 75,000 persons in 2020-2030.

Table 1. Summary of Route 73 Corridor Population Forecasts

	Approx. Size in Sq. Miles	Actual Population 1990	Actual Population 2000	Forecasted Population 2010	Forecasted Population 2020	Forecasted Population 2030
Area 1	23	5,323	7,463	9,649	11,725	15,184
Area 2	65	9,136	14,141	24,971	46,044	72,320
Area 3	43	14,629	36,464	59,610	85,543	111,015
Area 4	14	1,857	3,523	5,954	10,047	14,133
Area 5	39	6,855	10,416	15,701	22,245	28,554
Area 6	106	19,601	24,405	30,198	38,295	48,144
Area 7	8	8,340	8,827	9,786	12,172	15,008
Total Corridor	298	65,741	105,239	155,869	226,071	304,357
10-Yr. Change:						
Number			39,498	50,630	70,202	78,286
Percent			60.1%	48.1%	45.0%	34.6%

The rest of this introduction will be devoted primarily to demonstrating that the numbers in Table 1 are plausible. One way to demonstrate a forecast's plausibility is to discuss and defend the underlying methodology. The main part of the present document is being provided for this purpose. However, most readers cannot be expected to pore through such material at the requisite level of detail, so the discussion here will take another tack. This will involve looking at the forecasted magnitudes from a regional perspective and forming some comparisons with data from other sources. An advantage in this regard is that the figures for the Route 73 corridor are part of a complete set of regional forecasts that can be evaluated at multiple geographic levels.

The Charlotte region addressed by the present forecasting framework includes ten full counties in North Carolina, a small portion of one other North Carolina county, and five counties in South Carolina. (These areas and the sub-county districts that also enter the forecasting process are shown later in Figure 2.) A forecast has been developed for the region as a whole using an economically driven model. It calls for a continuation of rapid population growth, from just under 2 million persons in 2000 to nearly 3.5 million persons in 2030. The first question to be asked is whether this forecast – from which the Route 73 corridor numbers ultimately derive – is realistic.

Table 2 addresses this question by comparing the forecasted regional population levels with the magnitudes that would be obtained simply by extrapolating past ratios of regional population to U.S. population. The first column describes past and future U.S. population (mid-year values in thousands), with the future values consisting of official Census Bureau projections. The table then shows: past regional population; past ratios to U.S. population; future ratios based on straight-line extrapolation of the 1990 and 2000 values; and forecast magnitudes based on the extrapolated ratios. The last two columns then compare these magnitudes with the actual, model-based forecasts for the region.

Table 2 shows that the forecasted population for the region as a whole is just 1.9% higher in 2020, and 4.5% higher in 2030, than the values yielded by extrapolating past relationships between regional and national population. The lower part of Table 2 presents similar computations for the region's ten counties in North Carolina (which contain more than four-fifths of its population). In this case the differences work out at 0% in 2020 and 1% in 2030. Thus the detailed economic forecasting process discussed later has done little to modify the findings obtainable from simple extrapolation.

Other comparisons here make use of population projections from the North Carolina

Table 2. Comparison of Actual and Hypothetical Forecasts for Charlotte Region

	Population, w. Census Projection for U.S.		Ratio, Regional Pop. to U.S. Pop. (000)		Future Regional Pop.		
	U.S. (000)	Region	Actual	Extrap.	Based on Ratio	Actual Forecast	% Difference
Entire Region							
1990	249,623	1,581,866	6.3370				
2000	282,339	1,986,903	7.0373				
2010	309,163			7.7376	2,392,170	2,383,793	-0.4%
2020	336,032			8.4378	2,835,384	2,890,564	1.9%
2030	363,811			9.1381	3,324,546	3,473,294	4.5%
NC Portion*							
1990	249,623	1,283,480	5.1417				
2000	282,339	1,637,001	5.7980				
2010	309,163			6.4543	1,995,437	1,977,749	-0.9%
2020	336,032			7.1106	2,389,402	2,389,289	0.0%
2030	363,811			7.7670	2,825,706	2,854,012	1.0%

* Excluding 50-sq.-mi. part of region in SE Catawba County.

State Data Center (SDC). The State Data Center projects the future population of all North Carolina counties using familiar cohort-survival projection methods. These involve computing birth, death and net migration rates from past data and assuming that similar rates will hold in the future. The SDC projections provide a useful benchmark for evaluating other forecasts because they build upon historical data in an unambiguous and widely accepted manner. What matters for their comparative use is that SDC estimates and other cohort-survival projections do not take into account three factors: 1) the role of economic drivers as possible sources of divergence from past demographic trends; 2) the possibility of interactions among counties (such as spillover of growth from one county to another); and 3) the constraining influence of land availability, i.e., the fact that each county has a fixed amount of land and hence a diminishing supply of opportunities for further development. Comparisons can thus indicate how much impact is being attributed to these factors by other forecasts.

Table 3 compares the SDC projections with the present population forecasts (summed across counties in both cases) for the entire North Carolina portion of the Charlotte region. In an area of this size, the last two factors mentioned above should not matter much. Accordingly, the two sets of estimates are very similar. The numbers generated by the present forecasting framework are slightly below the SDC projections through 2020 and exceed the SDC figure for 2030 by only 1.5%. This finding confirms again that the results offered here are within a reasonable range at the region level.

Significant differences emerge below the region level due to the abovementioned factors, especially land availability and interactions among counties. The upper portion of Table 4 compares the SDC projection with the present forecast for Mecklenburg County alone. The figures provided include ten-year increments as well as percent differences between the estimates. The two sets of predictions for Mecklenburg are also depicted graphically in the upper-left panel of Figure 1 on page 6.

The SDC projection calls for the Mecklenburg population to rise by steadily increasing amounts, as if land availability would not constrain land development to any greater extent in future decades than in 1990-2000. The result is a 2030 population of nearly 1.32 million persons. In contrast, the present forecast recognizes the inevitability that Mecklenburg's growth will taper off, yielding a 2030 forecast of about 1.16 million persons. These figures differ by more than 12%.

It should be noted that land availability constraints tend to apply progressively, rather than allowing an area to grow with abandon until it hits a wall. The mediating mechanism is land value escalation, which deflects greater and greater shares of potential growth away from an area even while there are still large amounts of undeveloped land. For Mecklenburg County, reasonable forecasting procedures could yield a considerably faster

Table 3. Population of Charlotte Region, NC Portion

	SDC Projection	Present Forecast	Percent Difference
1990	1,283,480	1,283,480	
2000	1,637,001	1,637,001	
2010	2,012,275	1,977,749	-1.7%
2020	2,404,790	2,389,289	-0.6%
2030	2,811,388	2,854,012	1.5%

slowdown of population gains than that specified by the present forecast. (A study in 2000 by the present investigator predicted a much more abrupt leveling off of the county's population, though this resulted in part from the release of erroneous population estimates by the Census Bureau before the 2000 census results became available.) In any case, it is clear that the present forecast of Mecklenburg population is more realistic than the SDC projection and may even be on the high side.

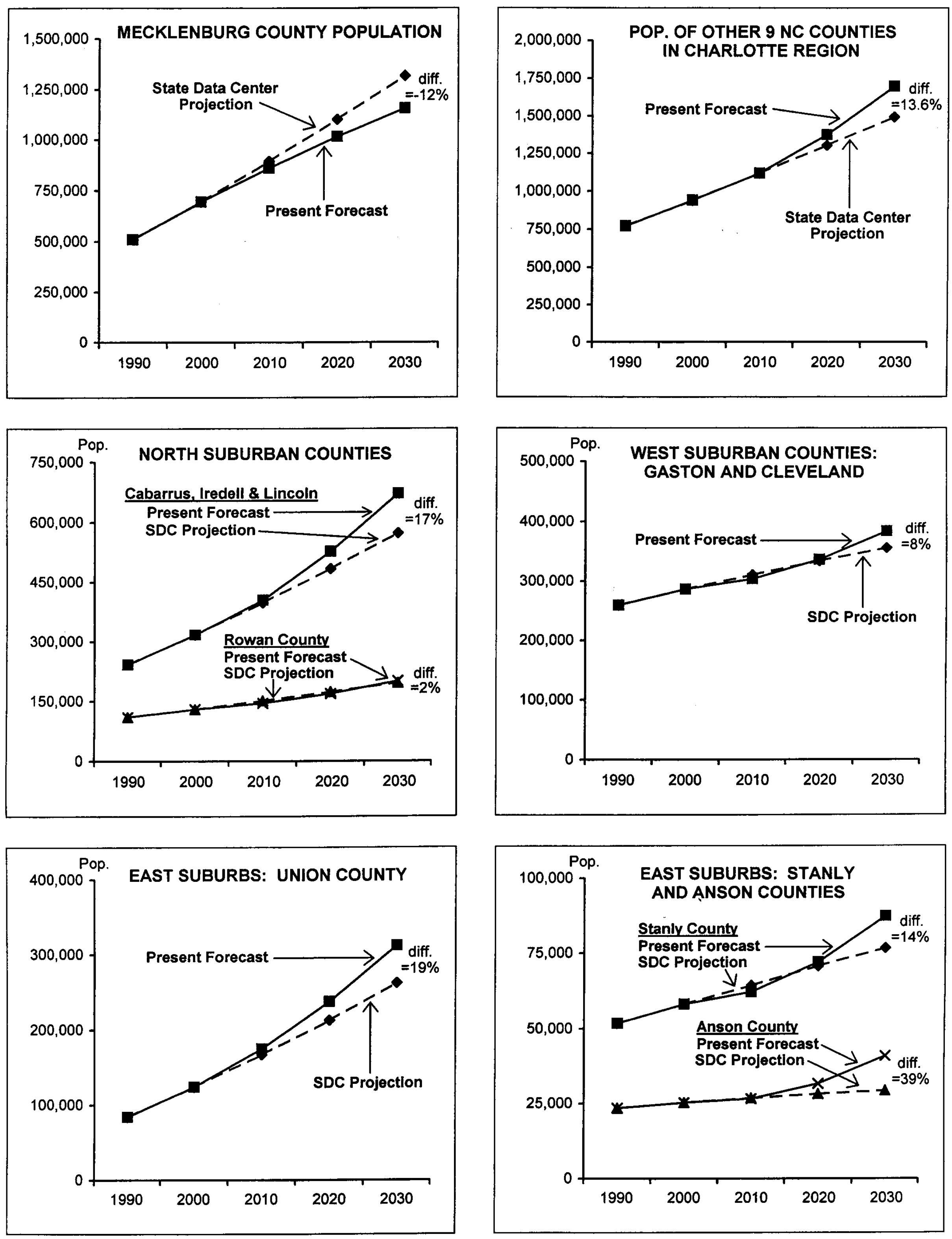
The people that Mecklenburg County can't accommodate will settle elsewhere in the region. This situation is depicted in the lower portion of Table 4 and the upper-right panel of Figure 1. According to the forecasts offered here, the population of the region's other North Carolina counties will rise by progressively greater absolute amounts in future decades (although these counties' annual percentage growth rate will be slightly lower for 2000-2030 than for 1990-2000). The resulting 2030 population is 13.6% above the sum of SDC projections for the given counties.

Table 4. Population Comparisons for Major Regional Components

	SDC Projection		Present Forecast		% Difference
	Total	Increment	Total	Increment	(in Totals)
Mecklenburg Co.					
1990	511,433		511,433		
2000	695,454	184,021	695,454	184,021	
2010	894,288	198,834	859,864	164,410	-3.8%
2020	1,102,003	207,715	1,015,638	155,774	-7.8%
2030	1,317,738	215,735	1,157,311	141,674	-12.2%
Rest of Region (NC Portion)					
1990	772,047		772,047		
2000	941,547	169,500	941,547	169,500	
2010	1,117,987	176,440	1,117,885	176,338	0.0%
2020	1,302,787	184,800	1,373,652	255,767	5.4%
2030	1,493,650	190,863	1,696,701	323,049	13.6%

The central and lower panels of Figure 1 examine whether the present forecasts allocate reasonable amounts of growth to suburban areas, given the expected magnitude of population spillover from Mecklenburg. For the northern area consisting of Cabarrus, Iredell and Lincoln counties, the present study predicts a 2030 population 17.2% above the corresponding sum of SDC projections. (Iredell is included in this group because it is subject to largely the same growth forces as the Route 73 corridor even though it lies further north.) The northern area's excess relative to SDC is thus 3.6 percentage points higher than the excess predicted for all North Carolina suburbs combined. This gap is mainly due to the relatively low forecasts that have been obtained here for the western area – Gaston and Cleveland counties – on the basis of economic considerations. For the region's eastern suburbs (Union, Stanley and Anson), the present forecasts exceed the corresponding SDC projections by nearly 20% in aggregate, yielding a gap of around 6 percentage points relative to the suburban average.

Figure 1. POPULATION FORECASTS FOR COMPONENTS OF THE CHARLOTTE REGION



It is noted parenthetically that Figure 1 illustrates the predictive advantages of taking economic influences and spatial interactions into account. The present forecasts for Gaston and Cleveland counties acknowledge the impacts of recent manufacturing losses and likely future losses by falling below the SDC population projections for many years. But areas cannot keep on losing the same factories forever, and economies that are surrounded by growth will sooner or later restructure themselves, so the present forecasts eventually turn upward in a way that cohort-survival projections cannot. Similarly, the forecasts for Stanly and Anson counties follow upwardly concave patterns due to the expectation of rising population spillover from Mecklenburg and Union, a phenomenon that cannot be acknowledged when counties are addressed in isolation.

Figure 2 returns to the question of whether relatively high forecasts are reasonable for the region's northern and eastern suburbs. This figure is a map showing the region's counties plus the county subdivisions – referenced as "districts" – that have served as forecasting units in the modeling framework. (These are indicated crudely with dashed lines.) There are 46 forecasting districts including four outlying districts that consist of whole counties.

The shaded areas in Figure 2 are the districts that are highest in growth potential on the basis of current indicators. The criteria used to select them are listed in the upper right corner of the figure, and the relevant numbers are presented and explained on page A1 of this document's appendix section. The chosen districts are those favorable for growth in all four of the following respects: ample land still available for development; high recent population growth; high recent employment growth; and upper-income households accounting for a disproportionate share of total households. (Income is an important growth indicator because upper-income households confer status on residential neighborhoods and because a region's wealthier persons tend to be the ones who decide where business establishments will locate.) The choice among districts is relatively clear-cut. There are nine that comply with all four of the selection criteria, while only two of the others comply with more than two criteria.

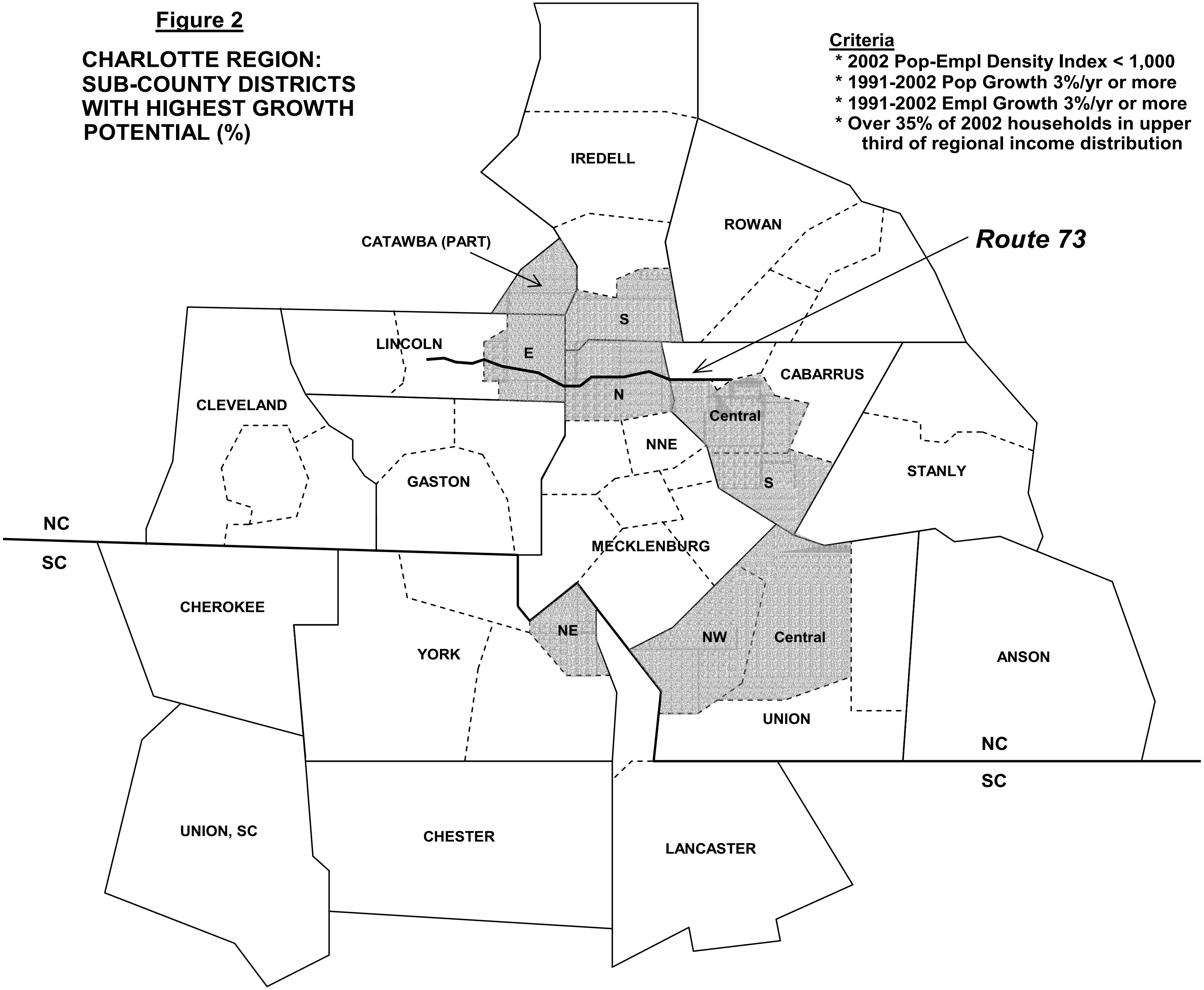
Six of the region's nine districts that are highest in growth potential form a continuous band of territory north of Charlotte that overlaps substantially with the Route 73 corridor. The alignment of Route 73 crosses or touches three of these six zones. As for the remaining districts with high-growth designations, two are located in Union County and one in York County.

A bellwether for directions of growth in the region has been Mecklenburg County's north-northeast district (labeled "NNE" in Figure 2), which contains Harris Boulevard and its environs. From 1991 to 2002 this area developed at a phenomenal rate, achieving double-digit percentage gains in population and nearly double-digit gains in employment *per year*. (See page A1.) By some measures this district, not downtown Charlotte, now represents the epicenter of regional growth. But the retarding effects of limited land availability are in store for this district, which prevented its selection as an area of high future growth potential. Hence much of the growth pressure now focused there promises to move outward into northwestern Cabarrus and northern Mecklenburg – which is to say, into the Route 73 corridor.

Figure 2

CHARLOTTE REGION: SUB-COUNTY DISTRICTS WITH HIGHEST GROWTH POTENTIAL (%)

Criteria
* 2002 Pop-Empl Density Index < 1,000
* 1991-2002 Pop Growth 3%/yr or more
* 1991-2002 Empl Growth 3%/yr or more
* Over 35% of 2002 households in upper third of regional income distribution



All present evidence says that the region has two dominant growth vectors, one pointing north-by-northeast out of downtown Charlotte and one pointing southeast into Union County. (A third vector of somewhat less strength points down I-77 into York County.) Notwithstanding the very large gains forecasted here for Union County, the prospects for the northern vector are arguably more robust due to the presence of two interstate highways and Lake Norman, plus the area's stronger existing complement of high-growth economic activities.

But still, one might ask, what must be assumed to get forecasts for the Route 73 corridor as high as those in Table 1? The long answer consists of the methodological discussion that weighs down the present document. It says, in effect, that the present forecasts have been obtained from a statistically calibrated model that takes into account the location of every activity relative to everything else in the region, measured in both absolute and incremental terms, while factoring in available land and past change in the given activity. But the short answer is that the corridor results offered here can be replicated using only one variable.

Suppose we go back to the SDC county projections and simply allocate their 2000-2030 population changes to smaller areas on the basis of how much land was available for development in 2002. Such computations are presented in a table that appears on appendix page A2. The counties in question are Cabarrus, Mecklenburg and Lincoln, and the allocation units are the portions of each county within the Route 73 corridor (by "area") plus the remainder of the county. Available land is estimated on the basis of total land area and the levels of population and employment already present. The formula that does the estimating is explained on page A2. It incorporates a parameter for which the most appropriate values vary among circumstances, but the crudeness of the formula is not pivotal to the demonstration as shown momentarily. What must be kept in mind about this exercise is that except for the available-land formula, the computations and results have no connection with any aspect of the present forecasting methodology.

The outcome is that, when SDC-projected population growth is allocated to areas inside and outside the Route 73 corridor strictly on the basis of available land, the total corridor population in 2030 works out almost the same as the forecast in Table 1, namely 300,583 persons versus 304,357 persons. The outcomes obtainable with realistic values of the available-land parameter range from 286,688 persons (6% below the present forecast) to 315,500 persons (4% above the forecast). This level of agreement can be interpreted as follows. The preconditions for rapid population gains in the vicinity of Route 73 already exist, embedded in the historical record and hence in the SDC projections. All that the Route 73 corridor needs to provide is land.

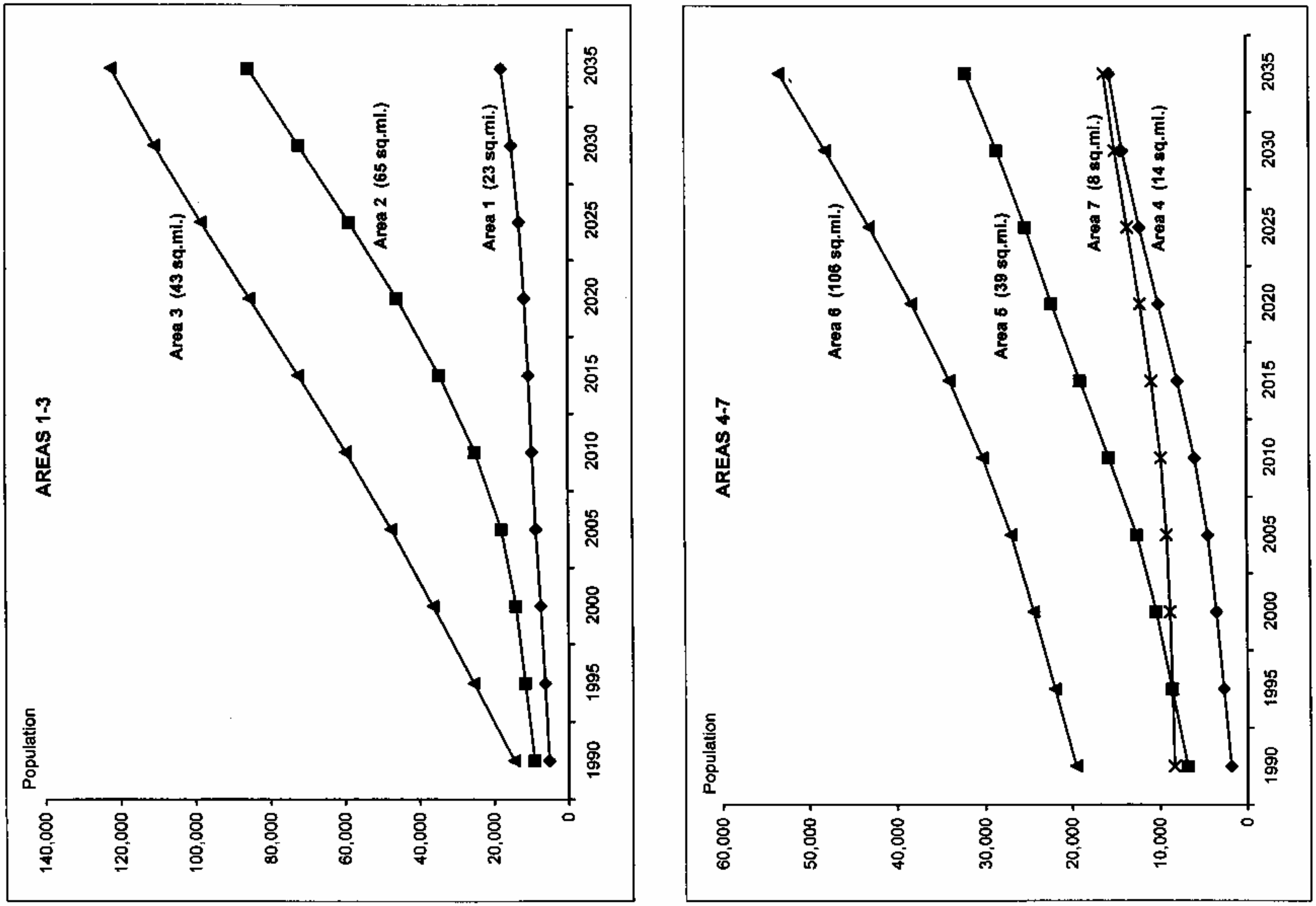
The particulars of this exercise can certainly be debated. For example, one can point out that the distance of northern Mecklenburg from downtown Charlotte may cause the land in this area to develop less intensively and/or less rapidly than Mecklenburg's available land at more central locations. On the other hand, one can note that the above exercise fails to acknowledge the strong eastward tilt of prospective growth in Lincoln County, hence understates the corridor's likely population gain in east Lincoln. In any case, when one asks the central question – Where are all the people going to go? – it is hard to avoid concluding that a big share of them will be going to the Route 73 corridor.

These demonstrations should be adequate to establish the reasonableness of the present forecast for the Route 73 corridor as a whole. Any rigorous analysis of the region and its dynamics will yield conclusions somewhere in the same ballpark. So the focus now shifts to the seven individual areas within the corridor. The forecasts for these areas have been presented numerically in Table 1 and are depicted graphically in Figure 3 on the next page. The graphical presentations cover five-year rather than ten-year intervals and extend through 2035.

When addressing individual components of the corridor, the present forecasting methods push the envelope of reliability. The forecasts have integrity for what they are, but what they are is an incomplete reflection of the factors that can influence growth. The following remarks will say a word in their defense before describing their weakness.

The forecast for Area 2 – the territory that straddles the Cabarrus-Mecklenburg line – is the case most likely to inspire skepticism. This area's population is predicted to increase during 2000-2030 no less than fivefold: from 14,141 persons to 72,320 persons. Such an increase sounds outlandish until one considers that: A) Area 2 offers 65 square miles of largely virgin territory located in and adjacent to one of the hottest counties in America; B) the populations of two neighboring zones – Area 3 and Mecklenburg's north-northeast district – respectively increased during 1990-2000 by 149% and 228%, which are rates that would yield fivefold increases in just 18 and 14 years; and C) if Area 2 reaches 72,320 inhabitants as predicted, while acquiring a commensurate number of jobs, over two-thirds of its land will still remain undeveloped. (See Figure 4). The idea that a suburban district this size can go from being sparsely developed to one-third developed in a time period longer than a generation is by no means far-fetched.

Figure 3. POPULATION OF ROUTE 73 CORRIDOR BY AREA



The forecast for Area 2 and the less eye-popping numbers for other areas are in fact perfectly reasonable estimates of what the market will try to bring about in the Route 73 corridor. The problem is that growth can reflect determinants other than market demand. There are also “supply-side” factors. These tend to assume greater relative importance as one considers smaller geographic scales, and the subdivisions of the Route 73 corridor are below the size threshold at which they threaten forecast reliability.

Supply-side factors include all the circumstances that determine the amounts of land available at various locations for various kind of development, with given levels of infrastructure support and development suitability. In concept it is possible for a top-down forecasting procedure – i.e., one that allocates predetermined regional totals among smaller areas – to cover supply-side and demand-side influences in a balanced fashion. The present forecasting framework has a demand-side emphasis, however, due to its reliance upon statistical calibration of predictive relationships using large samples of empirical data. This approach rules out detailed consideration of supply-side factors because quantitative expressions of land use policy, infrastructure availability and natural land characteristics cannot feasibly be obtained for large numbers of observations.

The calibrated allocation model takes into account rough estimates of the land in each area that remains available for development at each point in time. Also, because it relies substantially upon extrapolation and its equations are pegged to reproduce 1990-2000 conditions exactly, the framework captures the influence of land use controls and natural land features insofar as they remain unchanged from the recent past. But the framework cannot express the impacts of changes in supply-side factors, such as new land use policies linked to new comprehensive plans, or new infrastructure projects of such scale that they transcend the average rate of infrastructure provision in the past.

The resulting forecasts are objective – which is what they gain from statistical calibration of relationships – but they are objective as demand-side estimates. Changes in supply-side conditions must be factored into them as a follow-on activity (if such changes can be predicted), or else must be accepted as a source of random error.

The corridor’s Area 4, for example, might represent a worst case in terms of expected forecast reliability due to its small size and natural characteristics. Much of this 14-square-mile area lies along the Catawba River and reportedly has environmental constraints. If so, a combination of natural deterrents and environmentally based land use controls could turn away as much as half of the growth that the market would otherwise place in this area. The population of Area 4 would then increase from 3,523 persons in 2000 to only 8,828 persons rather than 14,133 persons in 2030.

Area 2 is nearly an opposite extreme (exceeded in this regard only by Area 6). Given that Area 2 has 65 square miles of largely developable land, draconian measures would be required to turn away half of the growth that market forces intend for this zone. Development diversion of such magnitude is not unachievable. In the Washington-Baltimore region, thirty years of extremely stringent land use controls have blocked well over half of the growth that would otherwise have occurred in certain districts of northern Baltimore County and northwestern Montgomery County, both exceeding 100 square miles. (Keeping such areas pristine has the effect of requiring more and more workers to in-commute from Pennsylvania and West Virginia, but local objectives are served because the extra pollution goes elsewhere.) It is hard to believe that jurisdictions in North Carolina would ever become so restrictive, but big percentage impacts from supply-side measures cannot be ruled out.

Yet still one must ask: Where are all the people going to go? Land development that is diverted from one area will occur somewhere else, and market forces will try to place it as close as possible to the area of diversion. This is why top-down, demand-side forecasts become more reliable as larger areas are considered.

Rather than ending this discussion on a waffling note, there is a felt obligation to find a point at which we can make a stand. This requires moving back up the geographic scale, but not all the way to the corridor as a whole. Given everything known, we feel that definitive statements can be made about the two portions of the Route 73 corridor located east and west of the lake and the river – that is, the Cabarrus-plus-Mecklenburg portion and the Lincoln portion. These composite areas both measure about 150 square miles and are referenced respectively as the East Corridor and the West Corridor.

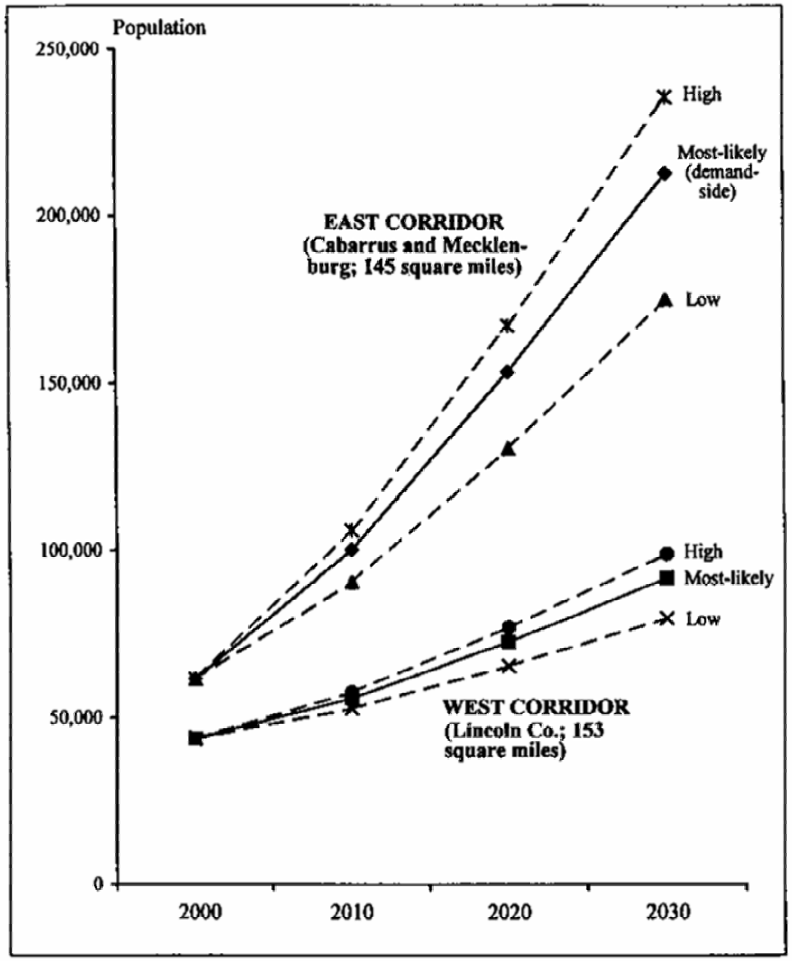
The given assertions are presented in Figure 4 on the next page, where they appear numerically on the left-hand side and graphically on the right. They consist of three sets of numbers for the East Corridor and three for the West Corridor: a most-likely forecast, a low forecast and a high forecast. In each case the first is a sum of the numbers already presented in Table 1. It is called a “most-likely” forecast because demand-side estimates can in fact be considered most likely for districts of this size in a region lacking commitment to large-scale growth management. The low and high forecasts have been computed from the most-likely numbers by subtracting and adding percentages of post-2000 population change. As shown at the bottom of Figure 4, the percentages allow for demand-side forecasting error, for the omission of supply-side factors, and for the strong probability that the foregoing errors won’t both be extreme in the same direction. The percentages work out to –25% for the low forecast and +15% for the high forecast. The downside margin is greater than the upside margin because supply-side factors can push growth away from an area more easily than they can attract growth in excess of demand.

On the right-hand side of Figure 4, the most-likely forecasts are plotted with solid lines and the low and high forecasts are plotted using dashed lines. The latter are not meant to convey “probable” ranges, but bet-the-farm ranges. The population of the East Corridor in 2030 *will* be between 174,886 persons and 235,310 persons, while the West Corridor’s 2030 population *will* be between 79,691 and 98,914 persons. (Never mind that your economist lamentably won’t be around in 2030 to see this happen.)

The next section presents an overview of the forecasting methodology. Readers interested mainly in results might touch base with this section (pages 15-16 and 23-24) and then skip to Section VI for the full set of Route 73 corridor forecasts.

Figure 4. COMPOSITE POPULATION FORECASTS WITH RANGES

	Population 2000	2010	2020	2030	2030 Pop. Per Sq.Mi.
MOST-LIKELY					
East Corridor					
Population	61,591	100,183	153,359	212,651	1,467
Ten-Year Change	30,646	38,592	53,175	59,292	(41% de-
Ten-Year % Chg.	99.0%	62.7%	53.1%	38.7%	veloped)
West Corridor					
Population	43,648	55,685	72,712	91,706	599
Ten-Year Change	8,852	12,037	17,027	18,993	(17% de-
Ten-Year % Chg.	25.4%	27.6%	30.6%	26.1%	veloped)
LOW FORECAST					
East Corridor					
Population	61,591	90,535	130,417	174,886	1,206
Ten-Year Change	30,646	28,944	39,882	44,469	(34% de-
Ten-Year % Chg.	99.0%	47.0%	44.1%	34.1%	veloped)
West Corridor					
Population	43,648	52,676	65,446	79,691	521
Ten-Year Change	8,852	9,028	12,770	14,245	(14% de-
Ten-Year % Chg.	25.4%	20.7%	24.2%	21.8%	veloped)
HIGH FORECAST					
East Corridor					
Population	61,591	105,972	167,124	235,310	1,623
Ten-Year Change	30,646	44,381	61,152	68,186	(45% de-
Ten-Year % Chg.	99.0%	72.1%	57.7%	40.8%	veloped)
West Corridor					
Population	43,648	57,491	77,072	98,914	646
Ten-Year Change	8,852	13,843	19,581	21,842	(18% de-
Ten-Year % Chg.	25.4%	31.7%	34.1%	28.3%	veloped)



Determination of Forecast Ranges

Post-2000 growth is changed by a constant % relative to most-likely forecast.

	Low Fore.	*High Fore.*
Allowance for error in demand-side forecast	*-10%*	*+10%*
Allowance for omission of supply-side factors	*-20%*	*+10%*
Adjustment for unlikelihood of joint extremes	*+5%*	*-5%*
Difference in growth from most-likely forecast	*-25%*	*+15%*

Estimation of "Percent Developed"

Population per square mile is divided by an estimated "full development" density of 3,600 persons per sq.mi., obtained by assuming that 75% of developed land is residential and 75% of this is net residential land, yielding 360 net residential acres per square mile. With 4 housing units per net acre and 2.5 persons per unit, this gives 3,600 persons per sq.mi.

II. Overview of Forecasting Methodology

The forecasting approach applied to the Route 73 corridor was developed by the present investigator in a series of forecasting studies over the past three years. These projects addressed: the Metrolina region (2000); metropolitan Washington (2001); metro Atlanta (2001); the Asheville region (2001-02); the Charlotte region (2002); Henderson and Haywood counties, NC (2003); and the Washington-Baltimore region (2003). Each addressed demographic and economic variables on an integral basis, and each employed a top-down sequence in which forecasts were generated successively for the nation, the target region and one or more sets of component areas within the region. The most demanding technical task, which underwent various refinements from study to study, was the development of a model to accomplish the allocation of growth increments from the region to smaller areas.

The purpose of this section is to summarize the origin of the Route 73 forecasts for anyone who does not require a full methodological discussion. Also, this section is used to make some general observations about forecasting philosophy and interpretation of results (which will not be repeated later). Readers lacking methodological interest but wanting to see detailed results for the Route 73 corridor should read the next page and the last two pages of this section (23 and 24), then proceed to Section VI.

As applied in the present study, the top-down approach has involved five different levels of geography as shown below. The national and regional levels were linked by region-specific predictive relationships rather than a full allocation process. All other progressions from higher to lower levels of geography were allocation steps. The forecasting framework was originally designed so that final forecasts would be obtained by allocating variables directly from the region level to the district level, bypassing counties, but county-level forecasting was reinstated for the reasons noted later.

1) Nation

2) Region (territory centering upon Charlotte, larger than present MSA)

3) Counties (15 counties plus a small portion of another)

4) Districts (46 forecasting units including 42 sub-county areas and four whole counties)

5) Sub-districts (15 forecasting units defined to allow description of the 7 component areas in the Route 73 corridor)

The counties comprising the Charlotte region as presently defined are shown at the top of the next page. This territory substantially exceeds what is now the official Charlotte Metropolitan Statistical Area (MSA), for the reason that it is intended to cover the full extent of metropolitan development thirty years hence. The generous definition has turned out to be strategically important because the allocation procedures have assigned significant amounts of the region's growth after 2020 to the outlying counties. The region includes a 52-square-mile portion of southeastern Catawba County, even though Catawba is part of the Hickory-Morganton-Lenoir MSA, because this served the transportation-related purposes for which the framework was originally developed.

Counties Comprising the Charlotte Region

Charlotte MSA		Counties Outside MSA	
North Carolina	**South Carolina**	**North Carolina**	**South Carolina**
Cabarrus	York	Anson	Cherokee
Gaston		Cleveland	Chester
Lincoln		Iredell	Lancaster
Mecklenburg		Stanly	Union
Rowan		Catawba (part)	
Union			

The 46 districts that constitute the next level of geography have already been shown graphically in Figure 2 (on page 8). The number of districts in each county depends largely on its current level of urban development. Anson County and three of the South Carolina counties are undivided. Two counties are divided into two districts; three have three districts; five have four districts; and Mecklenburg County is divided into eight districts.

The Route 73 corridor project has required the addition of another level of geography because the corridor's seven component areas are all smaller than the 46 districts originally targeted by the forecasting framework. The corridor extends across five of the districts: two in Cabarrus County, one in Mecklenburg, and two in Lincoln. Only the Mecklenburg district is located entirely within the corridor. Furthermore, three of the corridor's component areas overlap district boundaries rather than falling within a single district. This situation has required the creation of 15 forecasting units that could be assembled into the seven corridor areas. These sub-districts are listed below under the names of their parent districts.

The baseline year for the forecasting process is 2002. This choice has helped to capture impacts of the recent economic slump, but required extensive data estimation.

Cabarrus NW	Cabarrus Central	Mecklenburg N	Lincoln East	Lincoln Central
Area 1 (part)	Area 1 (part)	Area 2 (part)	Area 5	Area 6 (part)
Area 2 (part)	Area 2 (part)	Area 3	Area 6 (part)	Area 7
Remainder	Remainder	Area 4	Remainder	Remainder

Demographic variables have been estimated by using the Census Bureau's 2002 intercensal population estimates for counties to guide the extrapolation of detailed data from the 1990 and 2000 censuses. Values of industry-specific employment for 2002 have been obtained by using 2002 data from InfoUSA, a proprietary source, to update 2001 figures from the federal data system. Forecasts have been obtained for 11-year intervals from 2002 to 2035. Values of variables for all intervening years ending in 5 or 0 have then been interpolated by fitting third-degree polynomial equations to the 2002-35 data.

Forecasting Philosophy

In all applications, the forecasting approach has had two pivotal features: 1) treatment of the metropolitan region as a single unit for forecasting purposes, with the forecasted values of regional variables held fixed in all subsequent forecasting steps; and 2) reliance upon statistically calibrated equations to allocate regional totals (increments) among smaller areas. These features plus data availability issues determine most aspects of the methodology.

The focus upon the region as a unit follows largely from an assumption that long-term demographic trends are economically driven. That is, population and household changes are ultimately determined by what happens to employment. Though it may seem obvious to assume that people will follow jobs, this is not always true outside the U.S. For example, there are parts of Europe where people live in the same places for centuries and governments feel obliged to arrange jobs for them. In such circumstances it might be possible to forecast local population using simple projection methods and then to estimate employment on a derivative basis. But in America, jobs come first. Since a metropolitan economy in the U.S. is functionally integrated, and since many Americans compensate for economic determinism by living a long way from their workplaces, this means that the component areas of a metropolitan region are highly interactive. Hence no part of a metropolis can be forecasted reliably in isolation. Theoretically it might be possible to forecast all regional magnitudes and their spatial distributions simultaneously, but in practice the complexities of the situation make this impossible to do without relying heavily upon subjective judgment. So the best solution is to split the problem into two parts and address them sequentially. This means first treating the region as a unit and then worrying about where its gains of activity will go.

In the subsequent task of allocating regional magnitudes among smaller areas, the use of a statistically calibrated “model” brings advantages of objectivity, rigor and reproducibility of results. Objectivity is no small matter. Without the discipline imposed by formal quantitative methods, the forecasting process tends to become political, in a broad sense if not a narrow sense. People trying to imagine the unimaginable – i.e., what the world will be like decades in the future – can easily be drawn into focusing upon what should occur rather than what is most likely to occur. Urban planners and others with a professional or personal stake in shaping the future are particularly susceptible. (The strong preference of many planners for bottom-up forecasting comes from the flattering notion that they, through the design of land use controls and mass-transit facilities, will be telling future development where to go.) Forecasts can verge into being prescriptive rather than predictive, and while prescriptive forecasts may have their value, the present investigator is not in that business. So the approach described here mandates the use of allocation relationships established through formal analysis of empirical data. Statistical calibration confers advantages of realism as well as objectivity, because the interactions of urban activities over space are so complex and multifaceted that it is very hard to specify the existence, much less the magnitude, of relationships without recourse to historical evidence.

The question then is what geographic areas will serve as observation units for the calibration of predictive relationships. The ideal situation from an aesthetic standpoint is a study design in which all observation units are contained within the target region. A recent Washington-Baltimore project was able to utilize such a design due to the large size of that region. Among the 27 counties and independent cities comprising the Washington-Baltimore CMSA, the more populous counties were subdivided to yield a total of 78 districts. These districts formed the statistical sample for analysis of past data as well as the forecasting units for describing the future. But the intuitive appeal of this study design was purchased at the cost of two major drawbacks. First, going below the county level to obtain an adequate statistical sample vastly increased the effort required for data assembly. While demographic descriptors were not problematic, obtaining industry employment figures for multiple years required the use of elaborate and error-prone estimation methods because the federal data collection system has never provided much information for areas smaller than counties. And second, even the 78 districts comprised a less-than-optimal statistical sample, though they well exceeded the number required by theoretical considerations. Given the complexity of linkages and the levels of random error encountered in urban modeling, only a triple-digit sample size can assure the development of balanced equations that spread predictive responsibility across an appropriate number of significant variables. In most regions this criterion cannot be met by subdividing counties more finely, because smaller districts become increasingly subject to influence by unique, unpredictable events, so the divisions add more noise than explanatory power.

These considerations have caused all other applications of the present forecasting approach to use statistical samples containing geographic areas outside as well as inside the target region. Using only whole counties (and independent cities) for this purpose has then made it feasible in terms of data collection to analyze samples of more than 150 observations. The allocation model involved in addressing the Route 73 corridor was calibrated to data for 227 counties in 29 separate metropolitan areas. These were the metro areas in the eastern U.S. that most resembled the Charlotte region in terms of present and future population.

Reliance upon external data for model calibration requires an assumption that the dynamics of urban expansion operate in largely the same fashion from place to place. This is not a bold assumption in the United States, where urban commonality has been noted and lamented for decades. (As the folk song said: "I've seen yer towns / They're all the same.") The uniformity assumption does not require that all metropolitan areas exhibit the same spatial patterns at a given point in time, but only that the relationships governing their development on the margin be essentially the same. It need not be true, for example, that most metro areas contain a network of outlying industrial cities like those in the Charlotte region. The statistical sample must only contain enough similar cases to yield equations that, when applied to a context of strong growth emanating from the core, can replicate Charlotte's potential for converting nearby towns into satellite cities. The Washington-Baltimore study provided a valuable demonstration that the relationships found repeatedly in cross-metropolitan analyses, such as the tendency of upper-income households to attract other kinds of growth, are also observed when the analytical focus is restricted to a single region.

From these aspects of the study design – i.e., the downward allocation of predetermined regional forecasts using equations calibrated to large samples of county-level data – follow the main issues discussed in the remainder of this section, such as the allocation model's reflection of demand-side versus supply-side factors and the question of how far down the geographic scale the allocation should extend.

National and Regional Forecasting

The given forecasting approach starts with the estimation of future national employment by industry. Simply adopting a national forecast from an external source might have advantages, but the federal government no longer engages in multi-decade employment forecasting and there are problems with using proprietary forecasts.

All of the national forecasts prepared to date have extended through 2030. (Regional forecasts for 2035 have been obtained by extrapolation.) The national forecasting task draws upon two sets of data: a ten-year forecast of employment by industry from the Bureau of Labor Statistics (BLS), which is updated every two years, and the Census Bureau's long-term projections of the U.S. population by age and sex. The key assumption is that the national employment three decades from now will be constrained by the number of working-age persons available to staff the economy. The factual basis for this assumption is the rapid aging of the nation's population and the massive retirement of "baby boomers" scheduled to occur after 2010. The premise that 2030 employment will be demographically limited allows a total figure to be established for that year by applying extrapolated employment participation rates to the population projections for age-sex groups. The industry specific employment figures in the BLS forecast are then extrapolated forward, with modifications, and reconciled with the 2030 total to yield a profile of industry-specific national employment through that year.

Table 5 below summarizes the national forecast that underlies the present results for the Route 73 corridor. Gains in the nation's total population are expected to taper off substantially from the 13% rate achieved during the 1990s. The biggest influence on employment, however, will be the fact that well over half of all population growth after 2010 will be supplied by persons aged 65 and above. The number of persons aged 16 through 64 will then be rising by less than 3.5% per decade. So even though employment should increase relative to the population of prime working age, it will decline after 2010 relative to the population as a whole (as shown respectively by the last and next-to-last columns of the table). Based on present estimates, the result will be a drop in the ten-year rate of employment increase from 12.6% in 2000-10 to only 6.7% in 2010 20 and 5.3% in 2020-30.

Forecasts of regional employment and demographics are obtained by forming straightforward linkages between the regional economy and the national economy, under the abovementioned assumption that long-term regional growth is economically driven. The process starts with descriptions of the national economy and regional economy using a 42-industry classification scheme. (The industry descriptors utilized in these tasks and in many later allocation steps actually consist of worker earnings rather than numbers of employees, but only employment is referenced here to avoid confusion.) For each year from 1969 – the start of data availability – through the baseline year, the

Table 5. SUMMARY OF NATIONAL FORECAST

	Population (Midyear)		Total Employment			
	Number (000)	Percent Change	Number (000)	Percent Change	Per Capita	Per Person Aged 16-64
2000	282,339	---	135,208	---	0.479	0.739
2010	309,163	9.5%	152,218	12.6%	0.492	0.751
2020	336,032	8.7%	162,462	6.7%	0.483	0.775
2030	363,811	8.3%	171,003	5.3%	0.47	0.791

employment in each regional industry is split into "basic" and "population-serving" components through the application of an input-output table. Basic employment in each regional industry is then expressed as a ratio to total employment in the corresponding national industry. Simple linear regression is used to establish time trends in these ratios. The time trends are then extrapolated forward through 2030, and the resulting ratio values are applied to future national employment to forecast regional basic employment. Lastly, future population-serving employment is derived from basic employment using the input-output table, and the two industry components are combined for each industry and year to yield overall profiles of the future regional economy.

Demographic forecasts are obtained by finding a regional population profile for each future year that yields a labor force consistent with the expected employment level. This is done via cohort-survival projection methods, which start with the derivation of historical birth, death and net migration rates for the region. Using projected values of these rates, the cohort-survival tableau simulates the transition of the regional population across each future decade. Labor force participation rates are applied to the results, and the net migration rates in the tableau are scaled so that the projected number of employed persons in each year – after allowing for unemployment and net commuting – is equal to the forecast of total employment already established.

The use of input-output analysis to partition the regional economy renders the regional forecasting process somewhat complicated in execution. (There are actually many different input-output tables for different years, and their use involves forward and backward applications of matrix inverses.) But in substance the process is mechanical and does no more than implement an assumption that the past long-term relationships between regional economic drivers and national industries will continue to hold.

The previous section has looked briefly at the relevant regional forecasts from the Charlotte regional forecasting framework. As noted there, the 2030 population of the region as a whole works out 4% higher than the figure that one would obtain simply by projecting forward the past ratios of regional population to national population. For the region's ten counties in North Carolina, the difference is just 1%, and a similar margin holds for the present forecast versus the sum of cohort-survival projections from the NC State Data Center.

County and District Forecasting

The allocation model responsible for the Route 73 corridor forecasts consisted of 35 equations: three to predict demographic variables and 32 to address employment in different industry groups. The three demographic variables in question were lower-income, middle-income and upper-income households (defined relative to the regional income distribution). As in other applications of the given forecasting approach, households became the leading demographic variables below the region level, while other demographic descriptors such as population were forecasted outside the allocation model using supplementary relationships.

Each equation was calibrated by using multiple regression analysis to "explain" the 1990-2000 changes in the given variable that were observed in the 227 counties comprising the study sample. The explanatory variables tested in each analysis dealt with: changes in all economic and demographic magnitudes during the prior decade; conditions prevailing at the start of the 1990-2000 interval; and contemporaneous changes in selected variables besides the one under analysis. Most of the independent variables were complex expressions that described conditions outside as well as inside the area to which a value pertained. These were "proximity" measures in which a given change or initial condition would be weighted by an inverse function of distance to the subject area and summed across all areas in the region. An additional complication was that all types of explanatory variables but one were weighted by a function of available land in the subject area (which contained a parameter that was estimated as part of the calibration process). Composite predictors of this nature were needed to express the manner in which the urban development process balances the attractive force of existing activity against the dispersive force of land scarcity.

The dependent variables were arranged in four groups and analyzed in that order, with the same ordering followed when their equations were applied for predictive purposes. The groups were important because variables in a given group were eligible to serve as predictors of contemporaneous changes in the variables addressed later in the sequence. The variables in the first group pertained to economic sectors having a high degree of locational independence within a metro area, such as manufacturing, whereas the last group covered economic sectors with strong local-serving propensities such as retail trade. The household variables were placed in third position and thus could be predicted by some industries and serve as predictors for others.

A best version of each equation was selected for inclusion in the allocation model on the basis of statistical significance and other criteria. The last step in the calibration process consisted of applying the selected equations to "predict" the 1990-2000 changes in variables for all component areas of the study region. The discrepancies between actual and predicted values were then inserted into the predictive framework as local adjustment factors.

Forecasts were obtained by assembling 2002 values of all the relevant variables and applying the model recursively to three time intervals: 2002-13, 2013-24 and 2024-35. In each case the changes predicted for one time interval served as inputs to the next round of forecasting. All of the variables, relationships and procedures were set up so that the model accomplished an exact regional allocation, i.e., so that the final versions of forecasted quantities always summed to the pre-established regional totals.

The strengths of this model-based allocation approach include its objectivity, as already noted, and its ability to capture a wide variety of relationships and spatial interactions. Its weaknesses derive from the severe limits on types of variables that can feasibly be collected for large-sample model calibration. Because whole classes of variables must be omitted, the factors driving the model (other than regional totals) are limited to earlier values of the target variables themselves – i.e., to demographic and economic descriptors – plus functions of distance, land area and density. The most important omissions are factors that typically must be measured at a fine-grained level of detail (and often are hard to quantify in a relevant fashion) such as land use controls, natural land characteristics and availability of infrastructure. Since these factors mostly affect the supply of land suitable for development, and since the factors that allocation models do cover are mostly predictors of development demand, the limitations of such constructs can be summarized by calling them demand-side models.

Calibrated allocation models are nevertheless able to capture some supply-side influences due to two circumstances. First, such models can express the general role of land availability using crude measures that consider total land area (minus large-scale deductions like military bases, wetlands and parks) and existing development density. Second, because the model equations operate partly by extrapolation and are pegged to replicate past conditions in the subject areas, they implicitly cover all supply-side factors to the extent that the future impacts of these factors equal past impacts. For example, the allocation process in the recent Washington-Baltimore study yielded reasonable forecasts even for districts strongly affected by restrictive growth management policies, because these policies had been in place throughout the years consulted for model calibration. But what models of the given type cannot do is capture the influence of future changes in supply-side factors, such as exceptionally large infrastructure projects or shifts toward more or less stringent development controls. They basically assume that the tendency of public actions to restrict or encourage growth will resemble the conditions prevailing in the calibration period (at present meaning the 1990s).

Demand-side factors ordinarily determine the large-scale pattern of development in a region. Widely differing policy environments in Washington-Baltimore have managed to shift the long-term balance of growth between Maryland and Virginia, but supply-side effects on this scale are unknown further south. In North Carolina, county-level forecasts from a calibrated allocation model should ordinarily be reliable, to the extent that any forecast is reliable, with little or no adjustment for omitted supply-side influences. But supply-side factors gain in potential importance at progressively smaller geographic scales, so the question is how far below the county level a model application should extend.

Once an allocation model exists, obvious benefits can be gained by leveraging the rigor of this approach as far as possible. There are no theoretical reasons why equations calibrated to county-level data cannot be applied to smaller areas. (The calibration sample for the Charlotte model included independent cities and some other geographic units that were only a fraction as large as most counties, and the model's heavy reliance upon "proximity" variables as predictors served to assure that the estimated relationships would not presuppose observation units of any particular size.) Furthermore, demand-side forecasts from an allocation model can serve useful purposes below the geographic scale at which their supply-side omissions become serious, so long as users acknowledge the nature of the forecasts and their prospective need for adjustment.

Another relevant consideration, however, is that reducing an area's size increases the likelihood that its historical data will be dominated by individual events and spurious influences. Such data aberrations can yield forecasts that are unrealistic even as demand-side estimates. So given this consideration along with the supply-side problem, recent studies have adopted a rule that any sub-county area used as a forecasting unit should either exceed 50 square miles in spatial extent or have a current population above 25,000 persons. The 46 districts ultimately selected for use in the Charlotte allocation model all exceeded 52 square miles or 33,000 persons.

County-Level Versus District-Level Allocation

The plan was to generate final forecasts by allocating regional growth directly to the district level (although a county-level application of the Charlotte model was used in late 2002 to obtain preliminary outputs). County descriptors would then emerge simply as sums of district forecasts. Results obtained from direct region-to-district allocation had passed muster in prior studies and proved favorable again in the 2003 Washington-Baltimore project. However, the Charlotte-area forecasts obtained in the summer of 2003 from direct region-to-district allocation were found unacceptable. The aspect that seemed unrealistic was an excessive northward tilt in the region's long-term growth pattern, both demographic and economic.

The emergence of a region-to-district allocation problem in the Charlotte case probably had to do with the magnitude of regional growth plus the asymmetrical pattern of recent development. The asymmetry is illustrated below in Table 6, which divides the close-in areas around the central Mecklenburg district into three zones and presents values of the four growth indicators used earlier in Figure 2 (on page 8). The North zone leads the West zone and the South & East zone in terms of all four criteria. The margins are small for density (which is inversely related to growth potential) and upper-income households, but very large for the two measures of recent growth. Given this impetus, the model went somewhat overboard in allocating future gains to the North.

The solution was to develop a hybrid forecasting tableau. This approach was based on region-to-county allocation, but it reflected intra-county growth patterns by using sums of district-level variables as inputs to the county forecasts. Most of these variables were land-weighted "proximity" measures, and their nonlinearity caused the substitution to make a substantial difference. (The allocation equations per se were the same in all model applications.) The county level forecasts obtained from this tableau served as controls upon the district forecasts. That is, the final district results consisted of county forecasts allocated to districts in proportion to the independently derived values of district variables. This strategy of embedding a district-level model within a county-level model seemed to impose an appropriate degree of restraint upon the forecasted regional development pattern.

Sub-District Allocation

Given their small sizes, the component areas of the Route 73 corridor were known to be pushing the envelope of reliability for top-down, demand-side forecasting. The economist's consulting agreement stated that bottom-up inputs would be incorporated in the corridor forecasts if permitted by the timing of the project relative to other activities, but that otherwise only top-down numbers would be provided. When the time came for delivery of forecasts to support conceptual design of Route 73 alternatives, the economist had

Table 6. DESCRIPTORS FOR MAJOR ZONES AROUND CENTRAL MECKLENBURG

	West	South & East	North
Land Area (Sq.Mi.)	468	662	441
Growth Indicators:			
Density Index	817	994	804
1991-2002 Pop. Growth	22%	43%	76%
1991-2002 Empl. Growth	24%	29%	87%
Upper-Income HH Share	32%	44%	45%

Area Definitions:
West: Gast E & SW; Meck NW & SW; York N
South & East: Cab S; Meck ENE, E & S; Union NW & Central; York NE
North: Cab C & NW, Ire S; Linc E; Meck N & NNE

seen only a few bottom-up numbers that were months away from reconciliation with top-down forecasts. Hence the results offered here are strictly products of the top-down forecasting sequence, with the one exception noted momentarily.

The above text (page 16) has described the 15 sub-districts required to translate outputs for the region's five relevant districts into forecasts for the seven component areas of the Route 73 corridor. This translation was accomplished by turning the hybrid forecasting tableau just discussed into a three-level framework, with a sub-district model embedded in a district model embedded in a county model. Values of predictors based on the model equations were generated at all three levels, but unlike the relationship between the district and county predictors, there was no feedback from the sub-district level to the district level. The only linkage was a top-down allocation of district results to sub-districts in proportion to the model-based predictions for the latter. The numbers were then summed as appropriate to yield forecasts for the seven corridor areas.

One after-the-fact adjustment of the numbers has occurred in response to inputs received during the review process. The failure of the top-down forecasting procedure to acknowledge the impacts of special infrastructure projects has been judged a critical weakness in eastern Lincoln County, where the upgrading of Route 16 to a freeway will clearly yield "extra" growth increments. To correct for this situation, the population predicted by the top-down model for area 5 in 2035 has been advanced to 2025, and the forecast for Area 6 has been advanced from 2029 to 2025. The forecasts for other years and demographic variables have been raised accordingly. These adjustments – involving 8.794 to 10,898 persons in 2025-2035 – have been compensated by deductions from the forecasts for Union County, NC, which were previously believed to be on the high side. (The deductions were allocated among the four Union districts in proportion to post-2002 growth.) These changes are incorporated in all the figures presented in this report.

Overview of Regional Growth
The Charlotte region as presently defined consists of fifteen counties plus a fraction of one other county. These areas have already been listed in Section II.

Table 6 describes the region's population and gives three measures of 1990-2000 population change: absolute population gain, share of the region's total gain, and compound annual rate of change. Rather than addressing all counties individually, this table presents data only for Mecklenburg County and the region's "inner ring" and "outer ring." (The inner ring is defined as Cabarrus, Gaston, Union and York counties, while the outer ring covers the rest of the region.) The lower portion of the table offers some comparison data for the Atlanta metropolitan area.

Table 6. SUMMARY OF REGIONAL POPULATION GROWTH

	Population		Population Change		Compound Annual Rate of Change
	1990	2000	Number	Share	
Charlotte Region:					
Mecklenburg Co.	511,433	695,454	184,021	45%	3.1%
Inner Ring	489,736	609,719	119,983	30%	2.2%
Outer Ring	580,697	681,730	101,033	25%	1.6%
Total	1,581,866	1,986,903	405,037	100%	2.3%
Atlanta Metro Area, 1970-2000:					
Fulton Co.				9%	1.0%
Inner Ring*				59%	3.4%
Outer Ring**				32%	4.0%
Total				100%	2.9%

* Defined as Clayton, Cobb, DeKalb, Douglas and Gwynnett counties.
** Equals the 14 other counties in the Atlanta MSA as defined since 1980.

As applied in the present study, the top-down approach has involved five different levels of geography as shown below. The national and regional levels were linked by region-specific predictive relationships rather than a full allocation process. All other progressions from higher to lower levels of geography were allocation steps. The forecasting framework was originally designed so that final forecasts would be obtained by allocating variables directly from the region level to the district level, bypassing counties, but county-level forecasting was reinstated for the reasons noted later.

The counties comprising the Charlotte region as presently defined are shown at the top of the next page. This territory substantially exceeds what is now the official Charlotte Metropolitan Statistical Area (MSA), for the reason that it is intended to cover the full extent of metropolitan development thirty years hence. The generous definition has turned out to be strategically important because the allocation procedures have assigned significant amounts of the region's growth after 2020 to the outlying counties. The region includes a 52-square-mile portion of southeastern Catawba County, even though Catawba is part of the

An unusual feature of the Charlotte region is that its central area – referring to both Mecklenburg County and Charlotte per se (which contains over three-fourths of the Mecklenburg population) – has been growing much faster than its suburbs. During the 1990s Mecklenburg County captured 45% of the regional population gain, as compared with 30% and 25% shares for the inner and outer rings, and outpaced the rings in terms of growth rates by 0.9% to 1.5% per year. This was a nearly unique situation in the eastern U.S., equaled only by Wake County in the Raleigh-Durham-Chapel Hill area.

Metropolitan Atlanta – used here as a convenient basis of comparison – has represented an extreme in terms of overall growth, but has exhibited the usual geographic pattern of expansion. This pattern is that a metro area's inner ring captures the largest absolute population gains, while its outer ring achieves the highest percentage growth. Meanwhile its central county lags far behind in both respects. Metro Atlanta followed this pattern throughout the last three decades (described collectively in Table 6). The fact that Fulton County gained population at all was largely due to the county's highly elongated shape, which preserved the extremities for recent development.

The Charlotte region has differed from the norm in having not only a fast-growing central county but also a relatively slow-growing outer ring. During the 1990s, while the region as a whole was exceeding the national population growth rate by over a percentage point per year, the region's outer ring was less than 0.4% above the U.S. rate. In part this finding reflects the inclusion of some presently nonmetropolitan counties in the Charlotte region, but it still represents a significant difference between the study area and many other fast-growing districts such as metro Atlanta.

The archetypical pattern of U.S. urban growth is outward expansion into a thinly populated hinterland, driven by centrifugal forces involving relative land value and land availability. The Charlotte pattern tends instead to involve a coalescence of formerly disjoint communities under the influence of growth forces emanating from the core. Mecklenburg County is surrounded by small towns and cities that have always had their own hinterlands and their own sources of economic support. Urban expansion is now linking these areas to Charlotte and each other, in addition to creating new communities from scratch.

The fourteen counties comprising what we have called metro Atlanta's outer ring had an average of fewer than 25,000 inhabitants in 1970, when the Atlanta area stood roughly where greater Charlotte stands today. In contrast, the Charlotte region's ten outer-ring counties (excluding Catawba) have an average population exceeding 50,000 persons. This difference reflects the presence of more substantial urban centers. Much of greater Charlotte's land development involves accretion around these traditional centers rather than amorphous sprawl of the Atlanta variety. The fact that metro Charlotte is building upon an existing urban network explains the relatively slow growth of its outer ring in two ways. First, the relatively high initial populations of its outlying counties moderate the percentage changes produced by spillover growth from other areas; and second, the overall employment gains achieved by these counties are limited by their traditional dependence on slow-growing industries. The latter fact refers in particular to dependence on the textile industry, which was the largest source of manufacturing jobs for all but two of the Charlotte region's counties in 1990, and all but three in 2000.

The expansion of the Mecklenburg County economy during the 1990s was highlighted and to a major extent catalyzed by the explosive growth of banking activity in the county, which involved the evolution of NationsBank (now Bank of America) and First Union (now Wachovia) into financial institutions of global stature. On the other hand, the banking sector per se accounted for only a little over 10 percent of the county's 1990-2000 gains in employment and earned income, and 17.5 percent of its gain in "final demand" as defined below. The business services industry created 2.6 times as many new jobs as the banking sector. Banking was also well exceeded in job creation by three composite industry groups (namely trade, finance-insurance-and-real-estate other than banking, and services besides business services). The implication is that Mecklenburg's rapid economic growth promises to resume after the present slump even though the 1990s experience in banking can hardly be repeated. To some extent this continued growth will involve a filling-in of the new economic role pioneered by the bankers – i.e., a further acquisition of accountants, lawyers, consultants and other professionals linked to big-time finance – but it will also reflect myriad other sources of momentum that still operate throughout the county economy.

The two important points here are that: A) the Charlotte region promises to keep growing rapidly; and B) Mecklenburg County will only be able to accommodate so much of this growth. The 1990s situation in which Mecklenburg absorbed the lion's share of regional population and employment gains cannot continue, simply because Mecklenburg has a fixed supply of land. Mecklenburg captured over 45% of the region's additional population during the 1990s, but according to the forecasts described later will only be capturing 33% a dozen years from now and less than 24% after 2025. The corresponding figures for employment change are 69%, 48% and less than 45%. What this means is that suburban development – the integration process mentioned above – will proceed faster in many areas than observers tend to expect from past trends.

Variables Used in Forecasting

An outline of the overall forecasting sequence has been given in Section II. The present introductory discussion is limited to some comments on data inputs.

The variables entering the forecasting process consist of demographic measures from the census and economic variables from those few sources that release data for counties and smaller areas. The latter variables are essentially limited to employment and earnings by industry. Beyond employment and earnings, nearly all statistics available at the county level or below either lack comprehensiveness (as is the case for data from the five-year economic censuses), or address limited subjects (e.g., building permits). The only other variables that come into play are land area and distance measures used in computing proximity variables for sub-regional allocation purposes. Parenthetically, the shortage of small-area economic data is one of the reasons why the process of allocating regional forecasts to smaller areas relies exclusively upon single-equation modeling. The variables like savings, investment, output and financial flows that drive national models are totally missing at the county level. In their absence, the burdens and limitations of simultaneous-equation modeling are unjustifiable.

The key demographic descriptors in the forecasting process are population by age/sex group and households by income. Regional forecasting focuses mainly on population, which is linked to the regional economy via labor force participation rates. Households become the leading demographic variables in the sub-regional allocation process, with household distributions by income assuming particular importance. The final products of the forecasting sequence include various other quantities that are estimated outside the allocation model using supplementary relationships. Among these are households broken down by numbers of persons and autos per household; housing units classified by occupancy and tenure; and median household income and housing value.

On the economic side, employment and earnings are classified using the SIC industry code, even though all of the relevant federal data sources have now switched to NAICS, because every component of the forecasting framework relies upon historical information as well as data for recent years. Several detailed matrices have been developed to implement NAICS-to-SIC conversions, and various aspects of the methodology have been shaped to minimize errors from that source.

The forecasting sequence uses different numbers of industries at different levels, namely: 62 for national forecasting; 49 for regional forecasting (with a compression to 40 for input-output computations); and 32 industries for sub-regional allocation. The 49-industry classification is a slightly condensed version of the two-digit SIC code and is shown on appendix page A16. The 32-industry grouping differs from it primarily in failing to cover individual two-digit manufacturing industries. (Even though they are not covered by separate model equations, two-digit manufacturing industries are carried through the regional allocation process using industry-mix calculations.) Lastly, the final employment forecasts for sub-county districts are tabulated using an eight-category system relevant for transportation planning.

The definition of employment used throughout the forecasting process and the present report counts each worker only once, at his or her primary job. An area's total employment – i.e., number of at-place jobs – under this definition equals the number of workers living in the area, after any required adjustment for net commuting. This is the concept of employment used by the Bureau of Labor Statistics (BLS) when describing labor force, employment and unemployment magnitudes. Importantly, the one-job-per-worker definition yields employment magnitudes as much as 20% below the employment figures reported by the Bureau of Economic Analysis (BEA) in its Regional Economic Information System, which is the most comprehensive and hence most often-used source of small-area data. The difference comes from the fact that BEA statistics cover part-time as well as full-time employment. Any activity that yields self-employment income or a payroll tax deduction, no matter how small or short-term, is counted by the BEA as a job. Hence, definitional issues must be kept in mind when comparing the present results with employment magnitudes seen elsewhere.

Though employment is the ultimate concern, most forecasting steps in the present approach deal with economic activity denominated in terms of employee earnings (which include wages, salaries, tips and some fringe benefits such as employer contributions to health and retirement plans). The reason is that data from the BEA regional information system are essential for both regional analysis and sub-regional allocation modeling, and the BEA files describe earnings in much more industrial detail than employment. The process thus involves many points at which earnings are converted to employment or vice versa – not all of which will be mentioned – and the conversions relate BEA earnings to BLS employment, not BEA employment.

A further complication is that, despite the reliance upon BLS definitions, most of the employment figures used as input to the forecasting process are actually obtained from other sources (since the BLS does not routinely release county-level information). These sources consist primarily of: 1) the BEA regional information system; 2) County Business Patterns; and 3) InfoUSA. BEA numbers are used mainly for the manufacturing and government sectors, where the conversion factors required to obtain BLS-definition employment are usually close to unity. County Business Patterns (CBP) is a data series offered by the Census Bureau. It covers part-time employment but not self-employed persons, which means that CBP-to-BLS conversion factors can be above or below unity. The great advantage of CBP is its inclusion of establishment-size distributions that are invaluable in getting around disclosure regulations. (Federal law prohibits the release of data that would disclose – or even give hints about – the characteristics of individual establishments. As a result, employment figures for small areas and small industries are very often suppressed. Usable estimates can be obtained, however, from algorithms that squeeze information out of CBP size distributions.) InfoUSA is a proprietary source that reports employment for individual establishments. It has many liabilities, including a capacity for wild errors and a systematic tendency to under-report public employment. InfoUSA statistics are used because they can describe employment for any geographic areas, no matter how small, but the results are always pegged at the county level to numbers from the federal data system. Lastly, occasional use is made of statistics from state Employment Security offices, which can provide very up-to-date descriptions but have the problem of being sample-based.

The most serious data assembly problems involve employment data for sub-county areas in past years. The only historical statistics available from the federal system (not counting the fragmentary descriptions from five-year economic censuses) consist of County Business Patterns data for zip codes. At the zip-code level, CBP only provides total employment plus establishment-size distributions for individual industries, so almost all numbers must be estimated from the size distributions. (This process is not quite as shaky as it sounds, at least for areas with hundreds of establishments, due to the law of large numbers and the frequent ability to pin down the sizes of large establishments through recourse to county-level data.) Ideally, zip-code statistics are used only to estimate percent changes in industry employment over time, which can be applied to recent employment levels based on other sources. The historical descriptions for small areas back-calculated in this fashion are then summed and reconciled with more reliable county profiles. This procedure has been followed in the present investigation.

National Forecast

The process of forecasting national employment by industry has already been described in Section II. A total employment figure was obtained by assuming that the nation's long-term employment growth would be demographically constrained. Then an industry breakdown was derived by extrapolating a ten-year BLS forecast across the rest of the forecast period and reconciling the results with the total already established. Due to the acknowledgement of a demographic constraint and the use of a one-job-per-worker definition of employment, the resulting national employment totals for future years are lower than most projections prepared elsewhere. (See Table 5 on page 19.)

Notwithstanding the expected slowdown in employment growth after 2010, the national employment forecast is optimistic in its expression of faith that the American economy will retain an ability to employ the potential workers available. The numbers used later in the forecasting sequence have incorporated a pessimistic assumption, however, about the effects of the post-2000 economic slump. Economists often treat a recession as a temporary deviation from the long-term growth trend, meaning that it can be ignored in long-term forecasting so long as a recession year or recovery year is not used as the baseline for projections. But the present investigation has credited the recent slump with a permanent loss of growth relative to the national trend passing through 2000. After conversion of the national forecast from employment to earnings, the earnings magnitudes have been adjusted downward to subtract one year's growth from the earnings gain in each industry that would otherwise be expected to occur during 2000-2010.

Partitioning of the Regional Economy

In the approach applied here, regional forecasts are obtained by: 1) quantitatively linking the regional economy to the national economy; 2) projecting the regional-national linkages into the future; 3) applying them to the national forecast; and 4) translating the resulting regional magnitudes into full economic and demographic descriptions. The regional-national linkages are limited to economic variables, except in one area, and do not cover the whole regional economy. The approach basically consists of taking the regional economy apart, estimating future trends in the sectors considered to be its drivers, and re-assembling it to obtain aggregate descriptions. Much of the effort involves the use of input-output analysis to isolate the economic drivers, which are not whole industries as conventionally defined, and to establish their relationships with the rest of the economy.

Input-output models are basically expanded versions of the familiar economic base multiplier model, which says (when applied on the margin) that any independent economic stimulus in an area will have "ripple effects" yielding an overall growth increment larger than the original stimulus. Input-output analysis expresses multiplier effects on an industry-specific basis by using a table of purchase coefficients to trace the individual transactions required to support an industry expansion. In static terms, input-output modeling attributes all economic activity to a set of industry components that are collectively called "final demand." These are generally not whole industries but the estimated shares of industries that bring in revenue from the outside world. The shares assigned to final demand are typically large for manufacturing and other goods-producing activities and small to moderate for most population-serving functions (although such differences are fading in the post-industrial era).

An input-output table for the Charlotte region was obtained in 2000 from BEA, and the same resource has been utilized in the present study. The table is denominated in terms of earnings (which operate as surrogates for the output amounts that would directly describe inter-industry transactions) and has been extensively modified to enforce consistency with information from other sources. The modifications have yielded multiple versions of the table to describe the region at different points in time. This is essential in long-term applications of input-output because many input relationships have changed in the past and can be expected to change further in the future. (As an extreme example, the I-O coefficients expressing industry demands for business services have increased more than sixfold on average between 1969 and 2002, and will rise another 50% by 2035.) The framework for input-output analysis thus includes ten different versions of the table for years between 1969 and 2002, plus a systematic method for generating future versions. Each version is a square 40-sector matrix covering most but not all of the industries in the 49-group classification referenced above.

Partitioning a regional economy into final demand and other activity can be accomplished by working "backward" through an input-output table (when the table takes the form of a matrix inverse rather than an array of input coefficients). The normal use of such a table is multiplication by a final-demand vector to determine total industry outputs. The table can also be used to infer final demands from outputs, however, by employing iterative procedures to home in on the unique final-demand vector that exactly generates the given outputs. Using a spreadsheet designed for this purpose, the present study has obtained final-demand vectors for all historical years covered by versions of the input-output table, then has estimated final demands for other years by interpolating industry shares. (In this and all subsequent steps, both final demand and output have been denominated in terms of earnings.) After final demands for future years have been estimated via linkages to national industries, conventional "forward" applications of input-output are used to obtain overall descriptions of the future regional economy.

By convention, input-output models treat all government activity as final demand, but in fact most local government functions and some state and federal functions play driven rather than driving roles, to no less an extent than sectors like retail trade. Hence the present study has divided the three levels of government into "endogenous" and "exogenous" components on the basis of various factors. Endogenous government is linked by an empirical relationship to conditions within the region, while exogenous government is treated like a final-demand sector.

In these and all other forecasting steps involving monetary amounts, the amounts are expressed in constant 1999 dollars. Using 1999 as the reference year for inflation adjustment is a convention based on the fact that household income distributions from the population census are expressed in 1999 dollars (because census respondents were asked in April of 2000 to describe the incomes they received in the previous calendar year).

Forecasting of Final Demand

Regional forecasts are obtained by linking final-demand earnings in each regional industry to total U.S. earnings in the same industry, so that future levels of final demand can be estimated from the national economic forecast. By developing estimates of earnings and final demand through 2002, the present study has obtained a thirty-four-year historical record for use in establishing the regional-national relationships, given the 1969 starting date of the BEA data source.

The steps involved in forecasting final demand are as follows. First, the value of final demand in each regional industry and each historical year is expressed as a ratio to total U.S. earnings in the same industry and year. Second, a simple linear time trend is fitted for each industry across the 34-year period of record (or a shorter period; see below). Third, these time trends are extrapolated into the future. Fourth, the ratios indicated by the extrapolated time trends are multiplied times national earnings in the respective industries for the future years covered by the national forecast. And fifth, the resulting estimates of regional final demand are adjusted for consistency with the baseline values in a manner to be described.

The graphs on pages A3 through A15 in the appendix to this document show the historical values of final demand in the Charlotte region and the time-trend relationships linking these quantities to national earnings. The figures offer 37 sets of graphs pertaining to final demand (including exogenous government) plus one set addressing endogenous government and a final set covering total regional earnings. In each case the left-hand graph describes regional final demand (or other earnings) and the right-hand graph plots the ratios of these figures to the corresponding national earnings. The graphs cover only 37 industries rather than the 40 sectors in the input-output table because three pairs of input-output industries – agriculture and mining plus two pairs of manufacturing industries – have been combined.

The straight lines in the right-hand graphs on pages A3 through A15 are the statistically estimated time trends. The long-term strength of the Charlotte region's economy is readily apparent from the fact that the 34-year time trends for all but one of the 37 industries – textile products manufacturing – are upward-sloping, denoting gains in the region relative to the nation. It must be remembered, however, that the data points describe the portions of industries identified as final demand, not total earnings in the given industries. The explosive uptrends seen for a number of financial and service industries express the rates at which these sectors have become sources of basic support for the regional economy, not their overall rates of growth.

A preliminary forecasting exercise was conducted in late 2002 using versions of these graphs that incorporated less current data. It addressed all industries using time trends for the full 34-year historical period, even though there were cases in which quite different trends would have been obtained by fitting the lines to shorter periods of record. The rationale was a desire to avoid subjective judgments and keep the forecasting process as mechanical as possible, along with a belief that any unreasonable results would cancel out across industries. This strategy was later modified, however, based on a judgment that the resulting forecasts systematically overstated growth prospects in manufacturing. The region's manufacturing sector had shown an overall tendency to gain less rapidly relative to the nation in the late 1980s and the 1990s than in earlier years; and the later patterns were clearly more reflective of what could be expected in the future.

The final forecasts have therefore been based on the use of 20-year time trends – that is, relationships fitted to data for 1983-2002 rather than 1969-2002 – for all but three manufacturing industries. The 20-year trends are the shorter lines that appear in most of the right-hand graphs on pages A3 through A8. For three industries (paper products, printing & publishing, and rubber & plastic products) the 20-year trend lines are virtually indistinguishable from the 34-year trends. In all nine other cases the 20-year trends possess less upward slope than the 34-year trends, or else slope downward rather than upward. The three manufacturing industries for which 34-year trends have been used are textile products, electrical equipment and transportation equipment. For textile products a 20-year trend line would have hit zero when extrapolated to 2030, whereas using 20-year trends in the other two cases would have yielded unrealistically high rates of future growth.

Final demand forecasts for regional industries have been obtained by extrapolating trend lines across the forecast period and applying the resultant regional-national ratios to industry earnings in the national forecast. This process has included an adjustment step wherein "forecasts" have been obtained from the trend-line equations for 2000 through 2002 as well as later years. The adjustment consisted of raising or lowering the constant term (A-coefficient) in each equation to make the average forecast for 2000 through 2002 equal the actual average value of final demand for those years. Pegging the equations in this fashion took care of situations wherein the trend lines failed to explain the most recent data points in the historical sample. The forecasts for later years were then obtained by using the pegged equations rather than the original equations.

Development of Overall Forecasts

The remaining calculations needed to obtain overall economic and demographic forecasts for the region are simple in concept but not in practice. The basic steps involve: 1) using conventional "forward" applications of input-output to translate the forecasted values of final demand for each future year into descriptions of total earnings by industry; 2) forecasting endogenous government earnings outside the input-output framework (which does not deal with the local-serving aspects of government); 3) converting the resultant economic profiles from earnings to employment using projected values of earnings per employee; and 4) finding the future regional population levels that yield the required total number of workers, given assumptions about net commuting into the region (which is expected to remain very minor).

Three circumstances complicate this process. First, the cohort-survival tableau that yields population projections given an overall level of net migration cannot be solved for employment. That is, it cannot be structured so that simply entering the number of workers required by the economy yields a description of future population by age and sex. Second, the treatment of endogenous government creates a feedback loop from the population forecast to the economic forecast. Thus the economic side cannot be finalized before dealing with demographics. And third, the derivation of earnings forecasts per se is complicated by a need to adjust the input output matrix so that the region's overall economic "multiplier" is held constant. All of these circumstances create situations in which solutions must be found by iteration. In each case a unique solution exists, but it cannot be found by solving the relevant equations analytically. Using iterative methods instead is not problematic since all the individual systems are linear (essentially meaning that the exact solution can be found in three tries), but the process as a whole is protracted because it involves an iteration within an iteration within an iteration.

Starting on the demographic side, cohort-survival modeling is a means of deriving population projections by looking at the transition of each age-sex cohort over time. A "cohort" refers to persons of one sex who occupy a given age bracket at one point in time (e.g., females aged 25 to 29 in 1990) and an age bracket advanced by "t" years at a time "t" years later (e.g., females aged 35 to 39 in 2000). Cohort-survival modeling rests on the truism that the number of persons occupying a cohort at the end of a time interval equals the number at the beginning of the interval plus three components of change: 1) births; 2) deaths (entering negatively); and 3) net migration of cohort members into or out of the geographic area under study. Births only affect the first cohort or cohorts (e.g., persons under age 10 when "t" equals 10), and deaths are mainly relevant for the oldest age groups, leaving net migration as the principal component of change for persons in most age brackets.

The first step in constructing a cohort-survival tableau for a study area is quantifying the components of change for all age-sex cohorts across some historical interval. The common procedure is to draw upon the census for age-sex distributions of the population in the interval's beginning and ending years. Then births and deaths during the interval are estimated at the necessary level of detail by drawing upon all available sources of vital statistics, and net migration is obtained for each age-sex cohort by subtraction. The present study has executed this procedure for the 1990-2000 interval, using national as well as local data for guidance in allocating births and deaths across age categories. The pattern of net migration thus obtained for the Charlotte region is typical for a fast-growing urban area.

A cohort-survival tableau is converted into a "model" by making assumptions about the components of population change that allow them to be estimated for future years. In the present study, the births and deaths that occurred during the historical interval have been expressed as percentage rates, using average cohort populations in the beginning and ending years as divisors, and the operative assumptions have focused on future trends in these rates. (In the case of births, the divisors have pertained to females of child-bearing age, and the variations among rates have reflected differences in child-bearing propensity across age groups.) All birth and death rates have been projected into the future on the premise that rates will change in the same direction as during 1990-2000, but by annual amounts only half as great. This premise yields a further general decline in death rates and a continuation of the trend toward child-bearing at older ages.

In the case of net migration, the 1990-2000 values for age-sex cohorts have been expressed not as rates but as a percent distribution of total net migration (covering both males and females). The operative assumption is that the same percent distribution will hold in all future intervals. Given a cohort-survival tableau for a future interval that contains the appropriate birth and death rates and an age-sex breakdown of initial population, this treatment of migration means that a total net migration figure is the only input required to obtain a full population projection for the interval's ending year (which will serve as the initial year of the next interval). The only catch is that the solution must be found iteratively, because the linkage of births and deaths to average population across the interval makes them dependent upon quantities yet to be determined.

Endogenous government creates another complication because this sector is demographically driven and hence cannot be forecasted in advance. As already noted, input-output covers all government activity as a source of demand for other industry outputs, but none as an internally determined part of the economy. This arrangement is unrealistic because government – particularly local government – contributes importantly to every area's complement of nonbasic activity. The solution in the present study has been to designate a portion of government as endogenous and link it to the region's total population. As in other cases, the predictive relationship has been estimated on the basis of ratio values. It is shown graphically in the middle section of page A15. This arrangement creates an iterative step in the forecasting process because the endogenous government earnings associated with any given population scenario may, and generally will, yield a level of total employment that is inconsistent with the economic forecast on which the population scenario is based.

Last comes the need to control the regional economic multiplier that is imbedded in the input-output table. The multiplier is defined here (in static terms) as the ratio of total earnings to the sum of final demand and exogenous government earnings. It reflects the general level of interactivity within an area economy, and it can only be determined after-the-fact by generating and examining forecasts. The multiplier has a critical bearing on forecast magnitudes since it determines the overall growth yielded by any given gains in final demand, which means in the present scheme that it can shape the relationship between the regional economy and the national economy. Hence there is a need to avoid imparting any overall bias to the forecasts when modifying the input-output table to obtain versions for different years.

The resolution in the present study was to constrain the matrix modification process so that the implicit multiplier was held constant throughout both the historical period and the forecast period. The constant value (2.5816) was obtained as an average across the historical period in an initial phase of the economic partitioning process. This constant was enforced for individual years by scaling all of the off-diagonal elements of the input-output matrix (and the diagonal elements minus appropriate constants) by parameters that were specific to each industry and year. When creating versions of the table for future years, these parameters were extrapolated over time and imbedded in a routine that reduced the matrix adjustment process to the selection of a single parameter. The catch, again, was that solution values of this parameter had to be found by iterative methods because the input-output system with its adjustment step could not be solved analytically.

The outcome for each forecast year was a process of choosing initial values and then solving successively for population, endogenous government earnings, total earnings and total employment. This yielded another population profile, and the process was iterated until convergence. The loop from earnings back to employment involved the application of earnings-per-employee figures to the provisional earnings forecast and employment participation rates to the provisional population projection, with iteration to reconcile the resultant totals. The employment participation rates used in this process were assumed to change over time in parallel with national rates. The assumptions in both cases were that: the rates for both males and females aged 16 through 20 would hold constant; the rate for males aged 21-64 would decline half as fast as during the 1990-2000 decade; the rate for males aged 65-74 would rise by 10% per decade; and the gaps between male and female rates for both the 21-64 and 65-74 age groups would close at the same pace as during 1990-2000 (causing gains in the female 21-64 rate to exceed declines in the male 21-64 rate).

The final economic forecasting step consisted of expanding the 40 sectors covered by the input-output table back to the 49 industries in the main classification. This was accomplished by analyzing historical data to obtain predictive relationships for industry shares, then extrapolating the relationships forward and applying them to the forecasts for combined industries.

Overview of Regional Results
Table 7 below gives the forecasts of total population and employment thereby obtained for the Charlotte region. A table describing future employment by detailed industry is presented on page A16 of the appendix.

The region is expected to achieve strong gains in population and employment throughout the forecast interval, although it is never expected to equal the growth rates of 2.31% per year for population and 2.32% per year for employment that occurred during the 1990s. Total population is expected to approach 3.5 million persons by 2030 and reach about 3.78 million persons in 2035, up from just below 2 million persons in 2000. Total employment will rise from about 1.08 million jobs in 2000 (and under 1.07 million in 2002) to approximately 1.89 million jobs in 2030 and 2.02 million in 2035.

The average annual rate of employment growth for the current five-year period will be relatively low due to the post-2000 economic slump, but thereafter the region's rates of employment change should stay above 2% per year until late in the next decade. The pace of employment growth will then decline markedly as population aging – which will take hold somewhat later than in the nation as a whole – reduces the share of persons in the age groups with high employment participation. Population change will occur more evenly. The compound rates of population change will vary only between 1.81% and 1.93% per year through 2030, although the region's absolute population gains will steadily increase.

Table 7. FORECASTS OF REGIONAL POPULATION AND EMPLOYMENT

	Population			Employment		
	Forecasted Value	5-Year Change	Annual % Rate	Forecasted Value	5-Year Change	Annual % Rate
2000	1,986,903			1,081,764		
2005	2,179,103	192,200	1.86%	1,157,798	76,034	1.37%
2010	2,385,288	206,185	1.82%	1,289,746	131,948	2.18%
2015	2,624,430	239,141	1.93%	1,440,057	150,311	2.23%
2020	2,889,969	265,540	1.95%	1,593,245	153,188	2.04%
2025	3,175,350	285,380	1.90%	1,743,995	150,750	1.82%
2030	3,474,012	298,662	1.81%	1,886,992	142,997	1.59%
2035	3,779,397	305,386	1.70%	2,016,921	129,929	1.34%

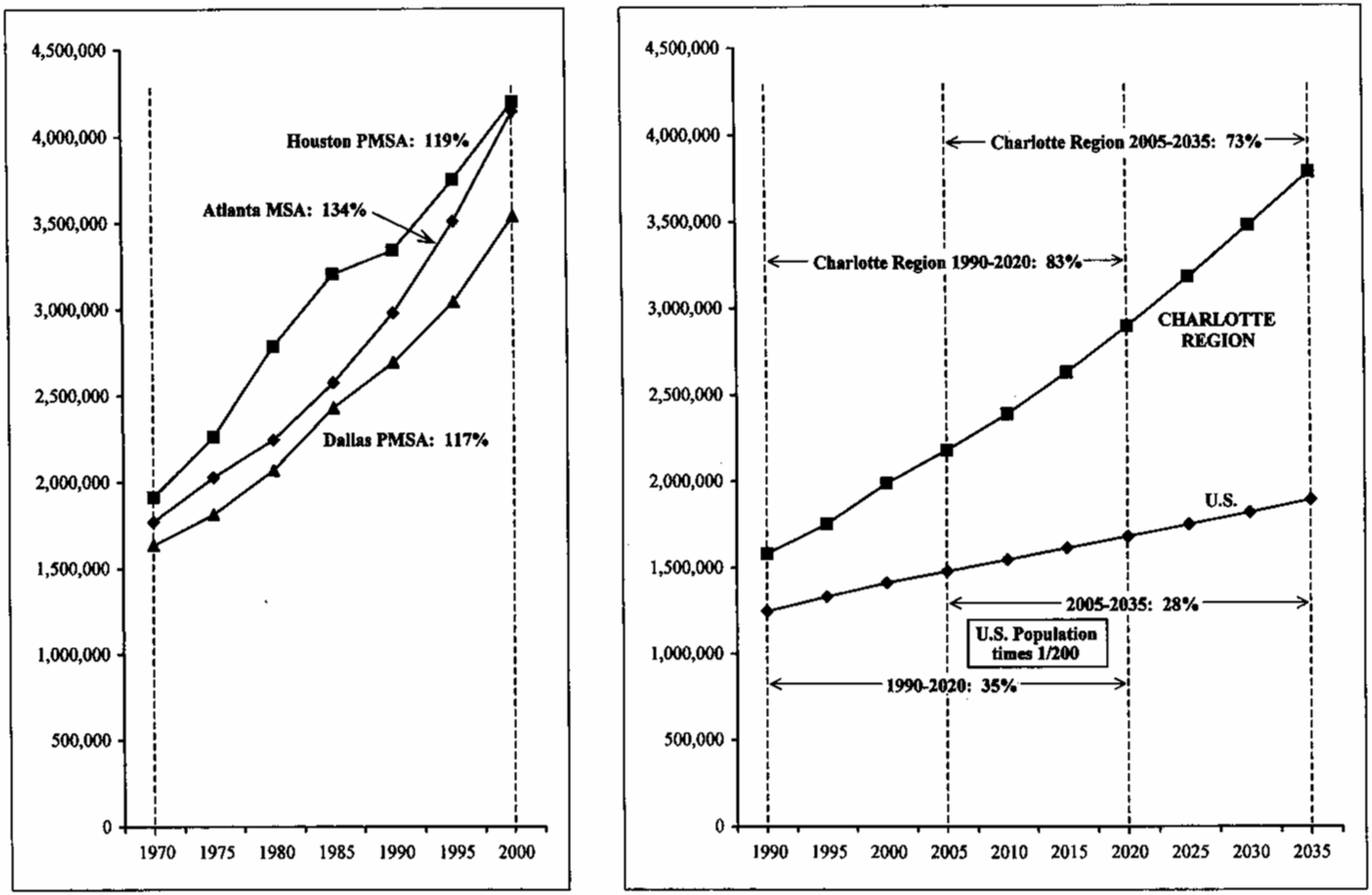
In case the growth forecasted for the Charlotte region seems frightening, Figure 5 on the next page may help to place it in perspective. Greater Charlotte has been and will continue to be a boomtown, but it will never be in a league with the growth leaders of the Sunbelt. This is shown by comparing the Charlotte region's growth trajectory with the population gains achieved by the three most exuberantly expansive metro areas in the southern U.S., namely Atlanta, Dallas and Houston. The left-hand panel of Figure 5 describes the population trends in these three comparison areas from 1970 through 2000, while the right-hand panel plots the Charlotte region's population from 1990 through 2035.

The latter graph also includes the past and projected population of the nation as a whole (divided by 200). Both the horizontal and vertical axes of the two graphs are plotted at the same scales, although the right-hand graph covers more years.

The populations of metropolitan Atlanta, Dallas and Houston all more than doubled between 1970 and 2000. Dallas brought up the rear with a 30-year gain of 117%, while Houston recorded a 119% increase despite the oil crash of the mid-1980s. And comfortably – or uncomfortably – out in front was metro Atlanta with a 30-year population gain of no less than 134%.

Given the present forecast for the Charlotte region and its performance since 1990, the region's highest 30-year percent change in population will be an 83% gain for the period from 1990 to 2020. The 30-year percent changes for the region will then trend downward to 73% for the 2005-2035 interval. Thus, Charlotte will not come within thirty percentage points of the increases posted by the three monsters of the south. In fact, the Charlotte region's peak gain of 83% during 1990-2020 will only be midway between the national growth rate of 33% for that period and Atlanta's 30-year record of 134% for 1070-2000. So the future expansion of the Charlotte region will be robust but by no means unprecedented.

Figure 5. LONG-TERM POPULATION GAINS IN THE CHARLOTTE REGION, AS COMPARED WITH THE UNITED STATES AND THREE OTHER SOUTHERN METRO AREAS



IV. County & District Forecasting

Data Sample and Model Structure
Sections II and III have already described the geographic areas targeted by this investigation, the nature of the top-down forecasting sequence, the data inputs utilized, the development of regional forecasts, and the issues involved in allocating forecasts below the region level. (The strengths and weaknesses of the approach have also gotten some attention in Section I.) The present section addresses the specific methods used to obtain county-level and district-level forecasts and the results obtained in the course of implementing these methods. As noted previously, the principal task has consisted of statistically calibrating a regional allocation model using data for metropolitan counties located outside as well as inside the Charlotte region.

The observation units for the model calibration process were determined by selecting all metropolitan areas (MSAs and CMSAs) in the eastern the U.S. that had three or more counties and one to five million inhabitants. The size limits were chosen to place the Charlotte region – which has two million residents today and will approach four million by 2035 – in the middle of the observed range. The geographic limit, which excluded metro areas west of Kansas City, was chosen to maximize the general relevance of observation units to greater Charlotte. Another factor was that many metropolitan counties in the west were undesirably large in spatial extent and/or had large amounts of undevelopable mountain land (which would have been hard to acknowledge adequately in measures of land availability). Requiring metro areas to have at least three counties was simply a means of maximizing their statistical value, given that the allocation model would operate entirely by forming inter-county comparisons. These criteria yielded a collection of 29 metropolitan areas in a territory bounded roughly by Hartford, Tampa, San Antonio and Minneapolis.

The individual observation units were counties and political jurisdictions equivalent to counties, where the latter included St. Louis and eight independent cities in Virginia. (Richmond-Petersburg and Norfolk-Virginia Beach-Newport News, the two relevant Virginia metros, had some other independent cities that were combined with adjacent counties because BEA statistics were only available on this basis.) The chosen 29 metros contained a total of 227 observation units, hereafter referenced collectively as counties. These included the eight counties in the Charlotte region (other than Catawba) that are not officially metropolitan at present.

The allocation model was structured in such a way that all variables on both sides of an equation summed to zero for each metro area, as discussed below. In the regression analyses this feature caused a loss of one degree of freedom for each metro. Thus the maximum degrees of freedom equaled 227 - 29 = 198 minus the number of independent variables retained in a regression. This number was a maximum rather than a constant because, for reasons noted in the next section, the observations for one or more metro areas were deleted from the sample in nearly half of all analyses. The outcome was that degrees of freedom ranged from 159 to 196 and averaged 189. Given the characteristics of the data explained later, these numbers were not overly generous.

A common practice in small-area forecasting is to focus on changes in variables over time rather than absolute magnitudes, because explaining the operation of growth forces at the margin is easier than accounting for each area's entire development history. The present allocation modeling approach followed this convention and thus was devoted to the prediction of increments. This meant that the model was calibrated to data for a recent time interval and forecasts were developed recursively by estimating changes over a series of future intervals.

Census data availability necessitated the use of 1990-2000 as the model calibration interval. Since past change was expected to be a strong predictor of current change, the 227-county calibration database included values of all area descriptors for 1980 as well as 1990 and 2000. The study design was then somewhat complicated by the use of a 2002 baseline year and the need to prepare forecasts through 2035. Given the incremental format and the 33-year gap between 2002 and 2035, the choice was to generate forecasts for three successive 11-year intervals: 2002-2013, 2013-2024, and 2024-2035. Hence the overall forecasting process involved time intervals of three different durations, namely ten years for model calibration, eleven years for forecasting, and twelve years (1990-2002) for expressing past-change variables as predictors of 2002-2013 changes.

The forecasting process was recursive in the sense that outputs from one round of forecasting would serve as inputs to the next. For example, the first round consisted of predicting changes in the target variables for the 2002-2013 interval. The results of this exercise were used to compute values of all variables for 2013, which then allowed the prediction of 2013-2024 changes in the second round of forecasting. The 2024 values enabled the third-round estimation of 2024-2035 changes and the computation of county descriptors for 2035.

Formulation of Variables

The functional form used to express the dependent variable in a regression analysis effectively determines the null hypothesis tested by the analysis. For example, if the dependent variable in a county-level economic analysis is employment in the banking industry – not expressed as a change – the null hypothesis is that banking employment equals a constant for all counties, plus some amount of random error. Such a case amounts to a straw-man null hypothesis because it is so easily rejected. The counties in almost any cross-sectional sample will vary a great deal in general scale (ranging in the present sample from Ohio County, Indiana, with 5,623 people to Harris County, Texas, with a population of 3.4 million), and the scale differences will be reflected in essentially all socioeconomic measures. This means that the banking variable just mentioned could be statistically "explained" by any area descriptor ranging from number of household pets to number of corporate scandals. Expressing the dependent variable as a simple absolute quantity stacks the statistical deck in favor of the regression's independent variables, whatever they might be. Avoiding this kind of bias should be the first requirement of cross-sectional analysis, yet the temptation to inflate statistical significance in this manner has conquered generation after generation of regional analysts.

The modeling approach applied here expresses all dependent variables in a functional form that creates a plausible null hypothesis. This hypothesis is that all counties in a metro area change at the same percentage rate. If the metro-area total for some economic or demographic measure increases by P percent over a time interval, the null hypothesis says that for each of the area's component counties, the change in this measure will equal P times the county's initial-year value (plus a normally distributed error term). Independent variables entering the regression equation can only achieve statistical significance by successfully predicting county deviations from metro-average growth. Along with casting independent variables in a critical light, this formulation has the advantage of yielding tenable equations in cases where no independent variables are found significant, which sometimes happens. Future values of variables in this form are predictable by an allocation model because metro-area totals are available for all descriptors from the regional forecasting process.

The formula incorporating these features is shown below. The symbol X denotes an economic or demographic measure to be predicted by an equation, and the underlined version X stands for the metro-area total of the same measure. In each case there is a subscript indicating whether a value pertains to the initial year or the end year of the interval addressed by the equation. Since the calibration interval is 1990-2000, the initial and end years for dependent variables in regression analyses are always 1990 and 2000, respectively. When the equation is used to generate forecasts, the initial and end years are 2002 and 2013 in the first round, 2013 and 2024 in the second round, et cetera.

Dependent variable expressing relative change in measure X = Xend-year – (Xinitial year * Xend-year / Xinitial-year)

When an equation is used in forecasting, a prediction of the above quantity for a future interval is sufficient to determine a value of Xend-year , because Xinitial-year is known from the previous forecasting round and the regional totals have already been established. Variables computed from the above formula have the characteristic – shared by all independent variables as well, as shown below – that they sum to zero across the observations for each metro area. This means among other things that regressions are always run with no constant term (which would have a value of zero if included).

A quantity computed from the above formula is called a "dev-change" variable because its values reflect deviations from the regional rate of change. In the present approach, all independent variables entering a regression analysis are computed using the four formulas shown below. In two cases they are dev change variables and the formulas are the same as that already stated. The symbols X and X again stand for some economic or demographic measure observed at the county level. (These variables are computed similarly for areas smaller than counties when targeted in forecasting applications.)

Current dev-change. Computed in the same fashion as the dependent variable. Can be used as a predictor only if the given economic or demographic sector appears earlier in the forecasting sequence than the one being addressed by the equation. (See Section III.)
Past dev-change. Computed using the dev-change formula already stated except that "initial-year" refers to the start of the prior interval and "end-year" refers to its end (which is the initial year for the current interval).
Dev-share. The difference between a county's initial-year value of a measure and what the value would be if the county resembled its metro area in terms of the given sector's share of a larger whole. (For example, if the measure is number of households in an income category, the larger whole is total households.) Equals the following expression, where Y stands for the larger whole and metro values are underlined.

Dev-share (X) = Xinitial-year – (Yinitial-year * Xinitial-year / Yinitial-year)

Dev-mean. The difference between a county's initial-year value of a measure and the regional mean of the given measure.

equals the following expression, where N is the number of counties in the metro area.

Dev-mean (X) = Xinitial-year – (Xinitial-year / N)

The last two forms are sometimes called "initial" dev-share and dev-change variables because they describe conditions at the beginning of an interval. Two special circumstances apply when an independent variable pertains to the same economic or demographic sector as the one being predicted. First, in economic equations these are the only independent variables that ever refer to detailed industries (as opposed to the three aggregate industry categories mentioned below). And second, these are the only cases in which housing losses and industry-mix effects are deducted from Xinitial-year before further computations occur. The prior subsection has already mentioned that such deductions from Xinitial-year occur in the derivation of certain dependent variables.

Two other complications involve available-land weightings and equation divisors. In all independent variables except the dev-change case just mentioned, X does not actually refer to an economic or demographic measure per se, but to such a measure (or a proximity measure as defined below) times an available-land weighting. The next section describes the purpose of available-land weightings and the manner in which they are derived. These weightings are held constant at initial-year values when multiplying both Xinitial-year and Xend-year. Nevertheless they must be applied before dev-change, dev-share and dev-mean formulas are evaluated because they vary across observations for each metro area. The equation divisors are factors used to divide all variables on both sides of a regression equation in order to reduce heteroscedasticity problems, as is also explained in the next section. They do not affect the computation of independent variables (i.e., can be applied after the above formulas have been evaluted) because they assume a constant value for each metro area in each equation.

Specific Variables

The ultimate outputs of the forecasting process consisted of county or district employment by industry, but as noted in Section III, most intermediate forecasting steps were conducted in terms of worker earnings rather than numbers of employees. The earnings in question included wages, salaries and certain fringe benefits and were always expressed in constant 1999 dollars.

The allocation model addressed 32 industry groups, corresponding to the 49-industry classification used in regional forecasting except that all manufacturing activity were treated as a single industry. Manufacturing was not addressed in detail because obtaining complete county-level manufacturing profiles for all 227 counties would have required inordinate effort and because previous studies have shown that significant allocation equations were hard to obtain for detailed manufacturing industries. Better results were achievable by adjusting manufacturing totals for industry mix when applying the model for allocation purposes.

The 32 industries were grouped into three categories: "industrial" activity, producer services and consumer services. Section III has mentioned the importance of the groupings to the modeling sequence. Except in two cases, independent variables always referenced earnings in whole categories rather than individual industries.

On the demographic side, the allocation model only dealt directly with three variables, namely the numbers of households in three income groups. Past studies had shown that all other relevant demographic variables could be estimated successfully on a derivative basis given their initial levels plus changes in households by income.

Households were grouped on the basis of relative income rather than absolute income. For each year covered by the model calibration database, all households in each metro area were assigned to three groups of equal size – referenced as the lower-income, middle-income and upper-income groups – on the basis of detailed income tabulations from the census. The household measures for individual counties then consisted of numbers of households in the three groups. Among the 227 geographic units covered by the sample, the city of St. Louis had the lowest 2000 income profile with 55% of all households in the lower-income group, 29% in the middle-income group and only 16% in the upper-income group. Hamilton County, Indiana had the highest profile with group shares equaling 17%, 26% and 57%, respectively. Middle-income households were more evenly spread across counties than the two extreme groups, with 2000 shares varying by only sixteen percentage points (from 26% to 42%).

A last introductory point involves proximity measures. These are predictors embodying the dictum that the three important things in real estate are location, location and location. For real estate entities ranging from a single land parcel to a whole county, what matters is relative location – i.e., where the land is located relative to everything else in the built environment. Relative location can only be expressed via composite variables that consider the entire metropolitan distribution of the influence (“attractor”) under consideration and include weightings by distance from the subject area.

In the present study, each proximity variable involved an attractor consisting of households in one of the three income groups or earnings in one of the three industry categories. For a given county, the value of a proximity variable was computed by summing the values of the attractor across all counties in the given metro area, when weighted by an inverse function of distance to the county for which the variable was being measured. The inverse function was the reciprocal of adjusted inter-county distance raised to an exponent of 2 or 2.5. The distances were straight-line distances in miles between county centroids. The formula was as follows.

Proximity measure showing the influence of attractor Z on activities in county j =
Sum across all counties i (including i = j) of:
Zi / (Dij + Qj + T)P

where:

Zi = The value of the given attractor for county i;

Dij = Distance from county i to county j;

Qj = Intra-county impedance for county j (expressed in miles);

T = Terminal impedance (constant); and

P = An exponent equaling 2 or 2.5.

Intra-county impedance referred to distance of travel within a county. It was estimated using a geometrically based function that varied as the square root of county land area and equaled K at 100 square miles. Terminal impedance, T, was a constant for all observations and expressed the cost of travel regardless of distance. It is most easily understandable as terminal time, i.e., the time required to walk to one’s car and so forth, but was expressed as a distance. Prior modeling studies had assigned relatively high values to these parameters – namely K=5 and T=5 – in order to keep each proximity measure from being dominated by the attractor magnitude for county j itself. However, proximity measures with K=3 and T=3 were also tested in the present study and often retained in regression equations. (This change compensated for self-imposed limitations on other variables that are explained in the next section.) These lower values of K and T always accompanied the higher of the two exponents. Thus each proximity measure entering each regression analysis was always offered in three versions, based on P, K and T values of 2,5,5; 2.5,5,5; and 2.5,3,3.

The distance between each pair of counties was computed on the basis of: difference in latitude; difference in longitude; a constant expressing miles per degree of latitude; and a function expressing miles per degree of longitude as a function of latitude. The computations utilized two sets of latitude and longitude measures. One described the geographic center of a county, while the other described the centroid of the county's households (computed from 1990 census tract data). The distances used in the model calibration process were based on weighted averages in which the geographic center was weighted by one-fourth and the household centroid was weighted by three-fourths. These weightings reflected an assumption that most of the new development "attracted" to a county during 1990-2000 would be located near existing households. In applications of the calibrated model to the study region, however, the weightings were progressively shifted toward the county geographic centers to reflect a probable filling-up process.

The proximity measures were used in the same fashion as other descriptors to compute dev-change, dev-share and dev-mean variables. That is, the generic quantity X referenced in the previous subsection would be an proximity measure obtained from the formula on the previous page times an available-land weighting. A dev-change variable would incorporate proximity measures for both the initial and end years of an interval, whereas dev-share and dev-mean variables would only involve initial-year proximity. Given these functional forms and the three different versions of proximity noted above, each regression analysis tested either nine or twelve different proximity variables – depending upon whether or not current dev-change variables were usable – to examine the influence of each attractor. Proximity variables were only allowed to enter regression equations with positive coefficients.

Allocation Modeling Issues

The original design of the allocation model was shaped by a number of issues involving potential uses of the forecasts. Most of these can be treated briefly here due to reductions in their importance or other circumstances.

The first issue was the possible need to avoid negative numbers in the outputs of allocation modeling, i.e., in the household and employment changes forecasted for future intervals. Negative numbers posed no problem in the allocation process itself because the allocation model would express all variables as deviations around expected or average values. About half of all input and output quantities would be negative, and positive and negative values would be treated in a perfectly symmetrical fashion. The problem would arise if there were a further district-to-TAZ allocation, because such a step would be oriented toward allocating positive changes on the basis of positive influences (e.g., allocating residential land development to areas suitable for development). Negative growth would be awkward to accommodate and could yield counterintuitive results.

A plan was adopted that would use different measures to avoid negative household