In summary, it can be stated that the nonresponse bias for low response rates not only is substantially greater but also affects the trip structure (frequency, choice, purpose, and destination) more than is the case in a more-exhaustive survey sample. As a consequence, the nonresponse error certainly cannot be compensated for by a correction of the share of mobile versus nonmobile persons.

The survey procedure (which includes the main survey and the nonresponse survey) was obviously quite cost-intensive, mainly because of the various follow-up phases. The question arises how these costs can be reduced while essentially the same data quality is maintained. The insights into the response behavior provided by this research might provide the prerequisite for meeting such a goal. It was demonstrated that a systematic bias arises due to the underrepresentation of nonmobile persons. It is therefore essential to take steps that increase the willingness of the nonmobile persons to respond to such surveys and that generate more cost-effective methods to accomplish this objective. It is still necessary to aim for as large a response rate as possible, since the systematic nonresponse bias cannot be compensated for by sociodemographic weighting. A reduction in the follow-up reminders cannot be recommended. At the moment, cost savings might be suggested (assuming that the results of this research are transferrable) by means of correcting the portion of mobile persons on the basis of the research results presented in this paper prior to the sociodemographic weighting of results. Another procedure would be to determine the ratio of mobile to nonmobile persons on the basis of a subsample of nonrespondents. This approach would be justifiable on the basis of this research, since the trip structure is practically unaltered by the nonrespondents.

ACKNOWLEDGMENT

We gratefully acknowledge the support of the Ministry of Transport, Federal Republic of Germany, and the city of West Berlin. We also owe special thanks to Otto G. Förg, Bernhard Schwertner, and Anke-Jutta Bergob for their extensive efforts in the data preparation that supports this research.

REFERENCES

1. P.R. Stopher and A.H. Meyburg. Survey Sampling and Multivariate Analysis for Social Scientists and Engineers. Heath, Lexington, MA, 1979.
2. W. Brög. Continuous Travel Survey (KONTIV). Ministry of Transport, Federal Republic of Germany, Munich, 1975-76.

Assessment of Land-Use and Socioeconomic Forecasts in the Baltimore Region

ANTTI TALVITIE, MICHAEL MORRIS, AND MARK ANDERSON

Accuracy of forecasts for population, labor force, employment, and car ownership from 1962 to 1975 in the Baltimore area are examined. Comparisons are made at three levels of zonal aggregation—city and suburbs, traffic districts, and traffic zones. The lack of information about household size and household income made inferences from the results incomplete. The results show that regionwide forecasts were accurate for all the variables except population. However, allocation of these forecasts between city and suburbs, to traffic districts, and to traffic zones was quite inaccurate. The correlation coefficient between predicted and actual changes varied from 0.93 to 0.17 for the city zones and from 0.28 to 0.02 for the suburban zones. The corresponding ranges at the traffic-district level were from 0.86 to 0.61 and from 0.36 to 0.30, respectively. The results in the paper point toward large errors and uncertainties in the independent variables of traditional travel-demand models.

The importance of socioeconomic and land-use variables to travel forecasts requires no elaboration. Forecasts of population, labor force, employment, car ownership, income, and other such variables are routinely made for 15-20 years into the future.

In spite of the popularity of hindsight, the accuracy of forecasts of land-use and socioeconomic variables is rarely examined. In fact, we know of no other study that has reported on the matter.

In this paper, forecasts of Baltimore-area population, labor force, employment, and car ownership by traffic zone made in 1962 for 1980 are interpolated for 1975 and compared with the actual 1975 figures as given by the Baltimore Regional Planning Council.

The comparison is made at three levels of zonal aggregation--city and suburbs, traffic districts (68), and traffic zones (484). These levels of aggregation were chosen to pinpoint the location of inaccuracy in forecasts. It is noted that 14 zones or 2 districts were eliminated from the analysis because of lack of 1962 data. These areas were on the very outskirts of Baltimore.

DATA AND METHOD

Three things need to be said about the data and method. First, the data pertain to the Baltimore area. In the 1962 study, this area was divided into 796 traffic zones. Some time later, the traffic zones were redefined, which resulted in 498 traffic zones. Equivalency between the old and the new traffic zones is achieved by means of a zone-equivalency table that assigns certain percentages of the old zones to new zones. This introduces a source of error. Percentage allocations of old zones to new zones cannot be done in a faultless manner. This problem will be examined briefly later in the paper.

Second, the 1980 forecasts were interpolated for 1975 by using both linear and logarithmic mathematical forms. The former provided better agreement for areawide figures for population, labor force, and employment (jobs). The latter provided a better match for car ownership [Table 1 (1)]. Thus, the linearly interpolated figures are chosen as the basis of comparison for population, labor force, and

employment, and the logarithmic interpolation was used for car ownership.

Third, forecasts will be evaluated in terms of both absolute numbers and change from 1962 to 1975.

As a comment to Table 1, it is noted that it is unfortunate that predicted figures for household size are not available. Thus, it is not clear whether population projections are off because fewer households moved to the area than were predicted or because household size declined. If fewer households have moved to the area, labor-force participation rate and household car ownership have increased from the projections made in 1962. On the other hand, it is possible that family size has gone down and that household car ownership and labor-participation rates have been predicted correctly. It is not known which of these two sources of error is more important.

ANALYSES AND RESULTS

Allocation of Activities Between City and Suburbs

To begin, the predictions for city and suburban areas are compared with the situation that existed in 1975. The data are shown in Table 2 (1). Again, linear-interpolation figures are used for all the variables except cars, which is interpolated by using the logarithmic form.

It is seen from Table 2 that the total (normalized) population is allocated reasonably well between the city and the suburbs. Car ownership is overpredicted in the city and underpredicted in the suburbs. The same applies more strongly to labor force and jobs. In fact, the location or relocation of jobs into the suburbs is substantially underpredicted.

The data in Table 2 are brought into a sharper focus when the forecasts are viewed as changes from 1962 to 1975. These changes are shown in Table 3 (1).

It can be seen from Table 3 that changes in the number of cars and in the labor force have a low percentage of error for the suburban areas, whereas other changes have been poorly predicted. In particular, the drop in labor force in the city was much larger than anticipated and the number of jobs created in the city was only half what was anticipated; also, the increase in car ownership in the city was much less than was predicted.

In general, the total change was predicted quite well except for population. This total change was inaccurately divided between the city and suburbs. The misallocation may mask important and interesting demographic changes that were not foreseen in 1962. Key information on income, household size, unemployment, and labor-participation rates would be desirable to make speculations about these unforeseen changes worthwhile.

Tables 2 and 3 are, of course, important from the point of view of travel-demand forecasting. According to these tables, population is the only variable that is substantially mispredicted as a total. However, predictions of allocation of activities between the city and the suburbs resulted in significant mispredictions, not only for population but also for employment and labor force. Because travel demand is directly dependent on these variables, travel forecasts may be critically affected by geographic misallocation of activities. The allocation of activities into geographic areas smaller than the city and the suburbs is examined next.

Allocation of Activities into Traffic Districts

The quality of the forecasts deteriorates rapidly when allocation to geographic areas smaller than the

Table 1. Actual and forecast values for Baltimore in 1975.

		Interpolation of 1980 Forecast			
Variable	Actual 1975	Linear	Error (%)	Logarithmic	Error (%)
Population	1 749 125	2 000 592	+14	2 079 342	+19
Cars	693 627	643 974	-7	693 508	0
Labor force	773 522	777 496	+1	816 295	+6
Employment	776 765	763 464	-2	816 690	+5

Table 2. Actual and forecast values for city and suburbs of Baltimore, 1975.

	City Zones			Suburban Zones			Actual Totals	
Variable	Actual	Forecast	Error (%)	Actual	Forecast	Error (%)	1975	1962
Population	845 035	942 813	+12	904 090	1 057 779	+17	1 749 125	1 624 138
Cars	227 165	238 461	+5	466 462	455 047	-2	693 627	438 564
Labor force	357 420	379 748	+6	416 102	397 748	-4	773 522	616 659
Employment	417 015	462 970	+11	359 750	300 494	-16	776 765	542 692
Population (normalized)	845 035	824 304	-2	904 090	924 821	+2	–	–

Table 3. Actual and forecast changes from 1962 to 1975 for city and suburbs of Baltimore.

	City-Zone Change			Suburban-Zone Change			Total Change		
Variable	Actual	Forecast	Error (%)	Actual	Forecast	Error (%)	Actual	Forecast	Error (%)
Population	-106 641	-8 863	-92	231 628	375 637	+62	124 987	366 774	+193
Cars	22 748	34 044	+50	232 315	220 902	-5	255 063	254 946	0
Labor force	-23 760	-1 432	+94	180 623	162 269	-10	156 863	160 837	+3
Employment	41 955	87 910	+110	192 118	132 862	-31	234 073	220 772	-6
Population (normalized)	-106 641	-3 020		231 628	128 073		124 987	124 987	

city and the suburbs is required.

Table 4 (1) shows the average absolute error, the correlation coefficient (ρ), and the Theil U coefficient between the actual and forecast values for population, cars, labor force, and employment at the traffic-district level.

It may be seen from Table 4 that total predictions for the city districts are quite good. Forecasts for the suburban districts are still fairly good, but their overall accuracy is about one-half that of the city districts. It may be noted that allocation of jobs, especially to suburban districts, has been predicted poorly; on average, they are 50 percent off.

Again, when the allocation of changes from 1962 to 1975 is considered, the quality of the forecasts drops. The average absolute error remains the same, but correlations between the predicted and actual changes are about one-half of those between the totals. Interestingly (and unlike the prediction of the totals), the prediction of changes is only slightly (if at all) better for the city districts than for the suburban districts.

Of course, the use of average-error figures can be misleading. More accurately, many districts are reasonably well predicted and few districts have been predicted very poorly. For example, one suburban district had a population of about 5000 in 1975; in 1962 it had been predicted to have a population of more than 23 000. Another suburban district had been predicted to have about 2500 jobs in 1975; in reality it had nearly 21 000 jobs. Based on visual observation, 5-10 percent of the city districts (one to two districts) was predicted poorly, whereas 15-20 percent of the suburban districts (seven to nine districts) was predicted quite poorly.

It is not surprising that allocation of activities to the city districts is predicted better than allocation of activities to the suburban districts. The city districts had already been built at the time of the forecast. Knowledge existed about population and employment, and trends of change may also have been known. The situation is different in the case of the suburban districts. Often a fair amount of suitable vacant land exists for development to take place. It is always difficult to predict which tracts will develop, since this depends on choices of many individuals and firms. It is left for further studies to show how sensitive travel forecasts are to errors in input data. Nonetheless, a guess is made that prediction of changes in travel demand is subject to substantial uncertainty.

Allocation of Activities into Traffic Zones

The same pattern of accuracy observed at the district level holds for zonal-level predictions except that, relatively speaking, at the zonal level the errors are much larger than at the traffic-district level. Table 5 (1) gives the same statistics as Table 4 for the zonal level.

The prediction of totals for the city zones is made with half the precision of the prediction of totals for the suburban zones. The exception that confirms this rule is employment, which is done equally poorly for both the city and the suburban zones.

When only the allocation of forecast changes is considered, these forecasts are wholly inaccurate for both city and suburban zones. Theil's U coefficient is nearly equal to 1, except for the labor force in city zones, for which the value is 0.35. This observation was confirmed by plotting actual versus forecast changes on graph paper. Such plots showed that if a dozen well-predicted zones were removed--zones that gave direction to the plots--the plots formed a circle. This shape indicates a completely random pattern of predictions. The plot for labor force also suggested that the reason for the good correlation coefficient and low Theil U value was due to a single well-predicted zone. Without that extreme value, the plot was effectively a circle.

The numbers of both origins and destinations of trips are dependent on variables shown in Table 5; some variables, such as income and household size, are still missing. Because of such direct dependency and because of substantial uncertainty in allocating activity changes to the traffic-zone level, prediction of changes in travel demand must be subject to large errors, since changes from the base line are bound to occur even if the region is not experiencing growth or decline.

Comparison of District- and Zone-Level Forecasts

Traffic zones have traditionally been used in transport planning for pinpointing origins and destinations of trips and thus for defining trips. Most of the summary information relevant to

Table 4. Statistics for total forecasts and forecast changes at traffic-district level.

	City Districts (N = 26)					Suburban Districts (N = 42)				
Variable	Absolute Error	ρ_1	ρ_2	U_1	U_2	Absolute Error	ρ_1	ρ_2	U_1	U_2
Population	5341	0.98	0.61	0.13	0.73	6928	0.85	0.36	0.24	0.62
Cars	1070	0.94	0.40	0.12	0.65	3743	0.79	0.30	0.27	0.52
Labor force	2172	0.94	0.86	0.15	0.47	3140	0.85	0.31	0.24	0.54
Employment	3708	0.98	0.64	0.13	0.58	3787	0.74	0.36	0.41	0.72

Notes: The Theil U coefficient is equal to 0 for perfect predictions and has an upper bound of 1. The subscript 1 refers to the total forecast and subscript 2 to the forecast changes.

Table 5. Statistics for total forecasts and forecast changes at traffic-zone level.

	City Zones (N = 205)					Suburban Zones (N = 279)				
Variable	Absolute Error	ρ_1	ρ_2	U_1	U_2	Absolute Error	ρ_1	ρ_2	U_1	U_2
Population	1217	0.87	0.41	0.22	0.81	2106	0.36	0.28	0.46	0.78
Cars	331	0.89	0.40	0.23	0.75	982	0.24	0.02	0.51	0.79
Labor force	538	0.80	0.93	0.27	0.35	855	0.39	0.03	0.46	0.80
Employment	1151	0.77	0.17	0.36	0.91	948	0.71	0.17	0.49	0.88

transport decision making is provided at the traffic-district level. The information for the traffic-district level is obtained by adding the zonal figures that make up the traffic district. For this reason, it is of interest whether traffic districts could be used directly for predicting travel demands and especially whether this is warranted on the basis of accuracy of predictions of the socioeconomic and land-use forecasts.

Table 6 summarizes the relevant statistics of forecasting accuracy at the zonal and district levels for total forecasts and for changes from 1962 to 1975. These include the statistics given in previous tables and the actual and predicted means and the root-mean-square error (RMSE).

The data in Table 6 suggest that the prediction of totals at the district level is more accurate than it is at the zonal level. Theil's U coefficient for districts is about one-half that for the zones, and the RMSE, as a percentage of the mean, is also twice as much for the zones as for the districts. The activity levels in city zones or districts are predicted better than in suburban zones and districts; actually, the accuracy of predictions for city zones is quite comparable to the accuracy of predictions for suburban districts.

The data on changes also show that changes from 1962 to 1975 are allocated better at the district level than they are at the zonal level. The advantage that district-level allocations of changes have over the zonal-level allocations is, however, less pronounced than is the allocation of totals. This is in part due to the general inaccuracy in allocating changes even at the most aggregate level of city versus suburbs.

At any rate, Table 6 suggests a general conclusion that the district-level allocations are superior to the zonal-level allocations in the suburban areas and the allocation of employment in the city is substantially better accomplished at the district rather than at the zonal level. Because of the importance to travel demand of the location of jobs, the problem whether traffic zones or districts should be used from the point of view of accuracy of travel forecasts merits serious consideration and study.

REDEFINITION OF ZONES AS SOURCE OF ERROR

Between 1962 and 1975 the traffic-zone structure was changed in the Baltimore area. In 1962 there were about 800 zones; these were consolidated into approximately 500 zones in 1975 by percentage allocation of old zones to new zones. From 1962 to 1975, 58 city zones and 80 suburban zones remained unchanged. It is therefore of interest whether the redefinition of zones alone introduces a substantial error.

Table 7 (1) lists some summary statistics for all zones and for the zones unaffected by the redefinition of zone boundaries. It can be seen from Table 7 that the allocation of activities to the zones unaffected by zone redefinition is done more accurately than it is to all zones. The exception is allocation of jobs to suburban zones, in which the unaffected zones fare less well.

So the results in Table 7 give a new twist to the results obtained earlier. At least some of the allocation error by zone must be attributed to the redefinition of zones. On the other hand, the lack of redefinition of these zones may imply that they are well-defined and well-established areas and, as such, easier to make predictions for than other zones. There may also be other reasons for forecasting success for these few zones. To pursue detailed analysis of such causes would require a good knowledge of the area and its historical development for such analysis to be of value. Due to lack of such knowledge, the matter was not researched further.

CONCLUSIONS

The conclusions of this paper are tentative and quickly stated. First, regionwide forecasts for cars, labor force, and employment were made with

Table 6. Comparison of relevant statistics of forecasting accuracy at zonal and district levels.

	Total Forecasts				Changes in Forecast			
Variable	City Zones	City Districts	Suburban Zones	Suburban Districts	City Zones	City Districts	Suburban Zones	Suburban Districts
Population								
Mean								
Actual	4122	32 501	3240	24 955	-520	-4102	830	5515
Predicted	4599	36 262	3757	21 526	-43	-341	1346	8944
RMSE	1619	7420	2772	9452	1621	7452	2773	9439
Correlation coefficient	0.87	0.98	0.36	0.85	0.41	0.61	0.28	0.36
Theil U	0.22	0.13	0.46	0.24	0.81	0.73	0.78	0.62
Cars								
Mean								
Actual	1108	8737	1672	11 106	111	875	833	5531
Predicted	1163	9172	1631	10 835	166	1309	792	5260
RMSE	480	1691	1482	5067	479	1692	1482	5071
Correlation coefficient	0.89	0.94	0.24	0.79	0.40	0.39	0.02	0.30
Theil U	0.23	0.12	0.51	0.27	0.75	0.65	0.79	0.52
Labor force								
Mean								
Actual	1747	13 747	1491	9907	-116	-914	647	4300
Predicted	1852	14 607	1426	9470	-7	-55	582	6918
RMSE	859	3444	1165	3985	859	3453	1165	3982
Correlation coefficient	0.80	0.94	0.39	0.85	0.93	0.86	0.03	0.31
Theil U	0.27	0.15	0.46	0.24	0.35	0.47	0.80	0.54
Employment								
Mean								
Actual	2034	16 039	1289	8565	205	1614	689	4574
Predicted	2258	17 807	1077	7155	429	3381	476	3163
RMSE	1652	5181	1825	6891	1654	5194	1824	6889
Correlation coefficient	0.77	0.98	0.71	0.74	0.17	0.64	0.17	0.36
Theil U	0.36	0.13	0.49	0.41	0.91	0.58	0.88	0.72

Table 7. Summary statistics for all zones and zones unaffected by redefinition of boundaries.

	City Zones (N^a = 58)				Suburban Zones (N^a = 80)			
Variable	Absolute Error	Absolute Error[a]	ρ	ρ^a	Absolute Error	Absolute Error[a]	ρ	ρ^a
Population	0.30	0.19	0.87	0.96	0.65	0.35	0.36	0.62
Cars	0.30	0.24	0.89	0.94	0.59	0.37	0.24	0.64
Labor force	0.31	0.23	0.80	0.92	0.57	0.34	0.39	0.68
Employment	0.57	0.40	0.77	0.92	0.74	0.73	0.71	0.38

[a]Zones unaffected by redefinition of zone boundaries.

good accuracy; however, population was substantially overpredicted. Second, allocation of these forecasts to traffic districts and zones was inaccurate. Statistical indicators showed that the allocation of activities to districts was more accurate than their allocation to traffic zones. The allocation of predicted changes was especially inaccurate; at the zonal level it was essentially random. This conclusion is tempered because of the redefinition of zones that occurred during the forecasting period. Because of the grave inaccuracy in those zonal projections, research should be undertaken to examine whether traffic districts could be successfully used to predict travel demands without unduly increasing the uncertainty in travel-demand predictions. Third and last, it needs to be mentioned that progress has been made since 1962 in methods for allocating activities to geographic areas. The use of present methods in 1962 might have resulted in better allocations and forecasts. By the same token, the world is more complex and uncertain now than it was in 1962, and we doubt that we are really more knowledgeable now than we were in 1962 of the many causes that affect spatial choices. The uncertainty in forecasts of socioeconomic and land-use variables, whether at the zone or district level, is large and, with certainty, here to stay.

ACKNOWLEDGMENT

This research was supported by a U.S. Department of Transportation contract from the Federal Highway Administration to the State University of New York at Buffalo.

REFERENCE

1. A. Talvitie, M. Morris, and M. Anderson. An Assessment of Land Use and Socioeconomic Forecasts in the Baltimore Region. State Univ. of New York, Buffalo, Working Paper 782-04, June 1979.

Components of Change in Urban Travel

GERALD S. COHEN AND MICHAEL A. KOCIS

Home-interview travel surveys in two upstate New York areas—Buffalo and Rochester—were conducted in the early 1960s and repeated in the early 1970s. An analysis of the changes in travel and household characteristics for both areas shows some surprising patterns as well as many that support the current theories of urban growth. Travel increased 8 percent and 37 percent in Buffalo and Rochester, respectively, over an 11-year period. However, average trip rates and trip lengths remained relatively constant over time, whereas automobile-ownership levels, number of households, and average travel time increased. In general, the increase in person kilometers of travel over time resulted primarily from an increase in the number of households rather than from increasing trip rates or lengths. The theory that travel-time budgets are stable holds for travelers and, to a lesser extent, for households. New highway construction does not appear to have generated large numbers of new trips but has had a greater impact on trip origins and destinations. Analyses of various stratifications of the data showed generally similar results.

How do area and household trip rates and trip lengths change over time? Do area characteristics or system investments cause the changes? What sort of similarities and differences emerge when one looks at two different areas? Trip rates and lengths are inputs to the computer-simulation process, and temporal instability must be adjusted if future forecasts are to have validity. If results are transferable from one area to another, a literature search may reduce the need for a local survey. In an attempt to answer such questions, we describe travel patterns observed in Buffalo and Rochester, New York, over an 11-year period.

Many cities in the United States have conducted comprehensive travel home-interview surveys at two different times. However, the growth in the areas surveyed often makes comparison over time difficult. Atlanta, Georgia, for example, had a comprehensive survey in 1961 that covered an area of 588 km^2 (227 $miles^2$), whereas the area surveyed in 1972 was 6068 km^2 (2117 $miles^2$) (1). The Niagara Frontier region of New York--Buffalo and Niagara Falls--conducted home-interview and cordon-line surveys in 1962 and 1973. Rochester, New York, conducted similar surveys in 1963 and 1974. The type of information obtained was similar for both areas. The survey design for the more-recent surveys permits direct comparison with the earlier surveys, since the area of the first survey is a major subset of the later survey. Thus, the analyst is able to compare travel changes in two cities over time and to note differences between the two areas.

Examination of results in different areas suggests that areas should be treated on an individual basis and that trip rates, in general, are not transferable (1). The study of travel changes in