Estimating Bicycling Demand

Gary Barnes and Kevin Krizek

Simple and reliable tools for estimating and predicting the amount of bicycling in an area would be useful for a variety of investment and policy decisions. Previous efforts to develop such tools have typically tried to develop demand estimates from basic descriptors of the population, land use, and bicycling facilities of an area. This paper takes an alternative approach by using the idea of deriving estimates of the likely range of total bicycling demand in an area on the basis of census commute-to-work data. The paper makes three contributions. The first is a general discussion of the total amount of bicycling in the United States and how it varies across places, on the basis of a number of surveys and some original data analysis. The second is the development of an argument that predictive models based on land use and transportation factors are unlikely to ever be accurate or useful because of a number of intractable problems. Third, a simple model is developed for estimation of a range of current levels of bicycling in a given geographic area with reasonable and known accuracy and by use of easily available data. While this model stops short of predicting bicycling levels or demand on specific facilities, it is an important first step in reaching these objectives. There is such a high degree of local variation in bicycling rates in the United States that attempts to predict bicycling levels directly without accounting for current levels are unlikely to be consistently successful.

Transportation investment decisions often require estimates or predictions of the amount of bicycling in a given area, as well as how this amount depends on facilities and other conditions. Despite a variety of publications describing efforts to model bicycle demand, no modeling technique, sets of parameter values, or even rules of thumb have emerged as definitive.

A first step in thinking about how bicycling demand can be modeled is to understand the types of questions that the model might be used to answer. Porter et al. list three major questions, which are paraphrased here (*1*):

1. How many people will use a new facility?
2. How much will total demand increase given an improved facility or network?
3. How does bicycling affect public objectives such as congestion and air quality?

The last of these could be expanded to include the benefits to cyclists themselves, such as improved health and recreational opportunities. The answer to this question could be useful politically, in justifying public spending on bicycle-related projects. The answers to the first two questions are likely to be more useful for technical analyses in prioritizing projects, given limited resources.

Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Avenue South, Minneapolis, MN 55455.

Transportation Research Record: Journal of the Transportation Research Board, No. 1939, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 45–51.

Another way of approaching the problem is to note that there are three different demand prediction objectives:

1. Predicting the total amount of bicycling in an area or on a facility;
2. Predicting the marginal amount that total demand will change given a change in facilities or policy; and
3. Identifying areas where inadequate facilities appear to be holding the level of bicycling below its potential, as in the latent demand approach (*2*).

In principle, a model that explains the total amount of bicycling as a function of basic factors, including demographic, policy, and facility variables, would answer all of these questions at the same time. Most past work has taken this kind of approach. FHWA (*3*) and the Texas Transportation Institute (*4*) completed major surveys of nonmotorized modeling techniques in the late 1990s; the majority of the efforts that they describe focused on predicting either commute shares or total bicycle travel by reference to these types of basic factors. More recent work, such as that by Dill and Carr (*5*), has also used this methodology.

The results of these efforts have been mixed. While certain demographic and geographic variables routinely emerge as important, evidence linking bicycle facilities and policies to levels of cycling has proven hard to come by; Dill and Carr note that there is somewhat of a consensus that such evidence has not been established (*5*). In general, it has been hard to find strong relationships because the differences in levels of bicycling across different areas can be very large in relative terms, much larger than can reasonably be explained by differences in the bicycling environments. Unmeasured factors, perhaps cultural or historical, appear to play an extremely large role in determining the level of cycling in an area.

A second, less common type of demand prediction method uses census commute-to-work shares, often combined with other data, to provide an area-specific baseline of bicycle usage; this can help to neutralize or perhaps serve as a proxy for some of the unmeasured factors that can have such a large impact on demand. Epperson, in Miami, Florida, used census data combined with demographic factors to estimate bicycling demand generally (*6*). Goldsmith, in Seattle, Washington, used census data combined with local information to predict likely changes in bicycle commuting due to facility improvements (*7*).

This paper approaches the demand prediction problem more from this second philosophical perspective; that is, to use known information about commuter bicycling to develop estimates of total bicycling levels in an area. These estimates would provide an area-specific baseline that could then be supplemented with other information to predict how the number might change under various conditions. There are three major steps in developing a tool on the basis of this approach.

The first part of the paper describes the results of several surveys and other measurements of general bicycling demand completed over roughly the last decade. The aim is to bring together the results of these many different measurements to show that the statistics are all roughly consistent when their different time frames are considered and place general bounds on the sizes of numbers that are likely to be observed.

The second part of the paper argues that, for a variety of reasons, the common demand modeling objective for the development of relationships between facilities and use by comparing different geographic areas is not likely to provide models that are consistently successful. The reasons are derived in large part from some problematic findings from the authors' attempt to develop a demand model for the Twin Cities (Minneapolis–St. Paul, Minnesota) area.

The third part of the paper discusses a simple model that relates current total bicycling rates to census commute-to-work shares. Estimates of this relationship are described across several geographic scales. This method is advantageous because it is simple to estimate, understand, and explain to policy makers and has a known range of accuracy.

AMOUNT OF BICYCLING IN THE UNITED STATES

This section describes the results of several surveys that have measured general bicycling demand and that have been completed over roughly the last decade. The primary objective here is to bring together the results of many different measurements, to show that they are roughly consistent when their different time frames are considered, and to place general bounds on the sizes of the numbers that are generally likely to be observed. A secondary objective is to demonstrate how a conceptual framework in which there is a distribution of bicycle riding frequencies over the population can reconcile the various measures of bicycling demand.

MEASUREMENTS OF BICYCLING FREQUENCY

Most of the available information about the amount of bicycling addresses the number of cyclists, as opposed to the number of trips or the number of miles of cycling. Because of the amount of information that is available about riding frequency, this is the measure of bicycling demand used here. The end of this section briefly addresses some other measures.

The surveys and other sources that address the frequency of bicycling produce a wide variety of results. Each source asks about a different time frame; the number of people who cycle in a week will be larger than the number of people who ride in a day. A key distinction to keep in mind is that (empirically) adults are considerably less likely to ride a bike than children, regardless of the time frame being considered. These two groups must therefore be studied independently to avoid confusion or ambiguity. This is generally not an issue with most bicycling surveys, which tend to focus on adults, but it is a factor in deriving numbers from general travel data collection surveys. In the ensuing discussion and tables, the data refer to adults ages 18 years and older.

Measures of the number of people who ride a bicycle in a given day were derived from two sources. The Twin Cities Travel Behavior Inventory (TBI) from 2000 and 2001 was a daily diary survey of about 5,000 households in the Minneapolis–St. Paul metropolitan statistical area (MSA) (*8*). This was done primarily during the spring and summer. The National Household Travel Survey (NHTS) of 2001 was a similar survey done over the entire United States (*9*); roughly 25,000 households were sampled in the general survey examined for this study. This survey was done over an entire year, which makes it possible to measure seasonal variations. Both of these surveys involved households that kept travel diaries on a randomly assigned day; these days were spread throughout the week and throughout the year for each geographic area.

NHTS also identified households in about 20 MSAs and 34 states, which allowed the calculation of averages for these areas. It should be noted that the sample sizes for many of these were fairly small, so the number for a specific area could be well off the true value. However, this probably gives a reasonable estimate of the range of values that might be observed over areas with large populations. NHTS also asks whether the individual completed bicycle trips during the last week; again, it is possible to calculate this at the level of specific MSAs and states.

Several national bicycling-specific surveys address time periods longer than a week. Rodale Press has reported on U.S. surveys done in 1992 and 1995 (*10*). It reported on the percentage of adults who bicycled in the last year, and it is possible to calculate the percentage of adults who bicycled in the last month. The Bureau of Transportation Statistics conducted a U.S. survey that asked about riding done during the summer of 2002, defined as a 3-month period (*11*). A more general Minnesota Department of Transportation survey from 2003 asked whether respondents bicycle for exercise but did not ask about frequency (*12*). The 2002 National Sporting Goods Association survey asked about participation in a variety of recreational activities; here, the standard was riding a bike at least six times in the year (*13*).

Finally, the U.S. Bureau of the Census asks detailed questions, including mode choice, about the commute to work of about 10% of the residents of the United States. These are summarized for use by transportation planners in the *Census Transportation Planning Package* (CTPP) (*14*). These data have the advantage of being by far the largest and most geographically comprehensive bicycle-related data sample available. The disadvantage is that they capture only commute-to-work trips, which are a small minority of all bicycling trips (*11*). Table 1 summarizes the results from the sources described here and in the preceding paragraphs.

Some people ride almost every day; others may ride only once a year. The longer that the time frame being considered is, the more people there are who will have ridden a bicycle at least once. It is possible to divide the population into different frequencies of riding in a manner consistent with the numbers, presented above, derived from different time frames. If each member of a group of people has a probability p of riding a bicycle on a given day, then the expected fraction n of that group that will ride at least once in a span of x days is given by the formula:

$$n = 1 - p^x \tag{1}$$

Groups with different riding probabilities p will generate different expected numbers of riders over a given time frame, and the numbers from each group can then be summed to arrive at a population total. Table 2 shows an example of how the population can be allocated to groups with different probabilities of riding on a given day to match known overall population bicycling rates over different time frames. These riding probabilities and population frequencies are mathematically consistent with about 1% of adults riding a bicycle on a given

TABLE 1 Measures of Adult Bicycling Frequencies

Source and Area	Measure	Average	Range
TBI, Twin Cities MSA	% per day	1.4%	—
NHTS, U.S. total		0.9%	.56% winter .88% spring–fall 1.1% summer
NHTS, U.S. MSAs		—	0.2%–2.4%
NHTS, U.S. states		—	0.0%–2.2%
NHTS, U.S. total	% per week	6.7%	—
NHTS, U.S. MSAs		—	4.5%–12.7%
NHTS, U.S. states		—	3.5%–12.4%
Rodale	% per month	—	16.6%–21.2%
BTS	% per summer	27%	
Rodale	% per year		37%–46%
NSGA	% 6 times per year	10.7%	—
Mn/DOT	% that ever ride	50%	—
U.S. Census	Commute to work %	0.4%	
U.S. Census, MSAs			0.1–1.4%
U.S. Census, states			0.1–1.1%

BTS = Bureau of Transportation Statistics; NSGA = National Sporting Goods Association; Mn/DOT = Minnesota Department of Transportation

day, 5.3% riding a bicycle on a given week, 16% riding a bicycle on a given month, 29% riding a bicycle in the summer, and 40% riding a bicycle in a year, with 50% sometimes riding a bicycle, although not necessarily in a given year.

The numbers derived from the population frequencies do not exactly correspond to the national averages over the medium time frames. This is probably because the national averages may be slightly overestimated in these cases. Intermediate time frames such as "this week" or "the last month" contain some room for personal interpretation; a person who rode 10 days ago might consider that to be close enough to count as part of the last week. Evidence that this is happening can be seen in the fact that the fraction of adults in NHTS who report that they rode a bicycle in the last week is more than seven times the number that rode a bicycle on the survey day. Given that survey days covered all days of the week and that every day will not be a completely new set of people, this result should be logically impossible.

TABLE 2 Possible Population Distribution of Bicycling Frequencies

Frequency of Cycling	% of Adults
3 of every 4 days	0.1
1 of every 2 days	0.2
1 of every 4 days	0.5
1 of every 10 days	1.2
1 of every 20 days	3
1 of every 50 days	10.0
1 of every 100 days	15
1 of every 200 days	20
Never	50

If the frequency table in Table 2 is roughly right, there are some interesting implications. The top four lines describe the people who ride a bicycle at least once every 10 days. They are 2% of the adult population, or 5% of the adults who cycle in a given year. However, they constitute 42% of the riders on any given day. That is, the 5% most active cyclists generate about half the riding days; the other 95% generate the other half. Because so many of the trips are generated by such a small number of people, a relatively small part of the population can have a big impact on the total amount of cycling. If 4% of the public were in the "frequent" category rather than the 2% that probably are now, that could conceivably lead to a 40% increase in the total amount of biking. Something like this may be what is happening in areas that generate very high levels of bicycling.

Evidence from TBI and NHTS, although it is not exactly consistent, shows that on the average day when an adult rides a bicycle, he or she rides for about 40 min. NHTS also reports distances; however, these seem to be extremely unreliable. By considering the total daily ride durations in these data, assuming plausible average speeds, and assuming that those people who ride longer times will also go faster, a likely daily average distance is perhaps 7 to 10 mi. Those people who ride more than 60 min in a day, while they are only a quarter to a third of all cyclists in a given day, ride about two-thirds of the total miles.

MODELING BICYCLING DEMAND

Traditional approaches to modeling bicycling demand are derived from the standard methods used to forecast automobile travel. That is, they start from basic information about the people and the transportation environment in an area and use this in some way to predict an amount of bicycle travel, either directly or as the solution to a mode choice problem in a larger travel model.

This section discusses some problems with the use of this approach to the modeling of bicycling demand, some of which appear intractable. The arguments are based in part on some of the facts about bicycling discussed in the previous section and in part on some preliminary findings from the authors' attempt to estimate a demand model for the Twin Cities area. While this model is not described here, in part because of the lack of useful results, it is used to illustrate some of bicycle demand modeling more generally.

There are several reasons why a bicycling demand model derived from basic information about land use, demographics, and the transportation system is likely to be of limited utility. These can be illustrated in part by the present attempt to develop a demand model, in which a statistically significant result that off-road paths were associated with lower per-person levels of bicycling for nearby residents was found. This result makes no sense intuitively; at worst, residents should ignore the paths. Empirically, Davis and Wicklatz found that off-road facilities in the Twin Cities were in fact much more intensively used in all parts of the city than other options, such as streets and on-street bike lanes (*15*). That result was not due to an obviously underspecified model; a wide variety of demographic and land use variables were included in the regressions. There seem to be four major possible reasons for this problematic outcome.

One is a possible shortcoming in the analysis, in that it is possible that the way in which the facilities were defined did not correspond to how people perceive them. For example, many of the suburban off-road facilities run next to busy highways, with all the associated crossing of driveways and roads. They are off-road in the sense that there is a barrier separating them from the road, but they are not off-road in the sense of eliminating potential conflicts or of being appealing to ride on. However, the development of a more general measure of the bicycling environment, going beyond the simple number of miles of facilities, is a difficult problem for many reasons.

A major reason is that a large fraction of bicycle riding is recreational. Intuitively, the sorts of land use and transportation facilities that would be ideal for utilitarian riding (dense development, a grid network, etc.) seem very different from what would be ideal for recreational riding (infrequent intersections, density of little importance). That is, the value of a facility may depend on the use to which it is being put. As a related point, the skill level of the rider likely also influences perceptions of the riding environment. These are significant conceptual difficulties, since it would seem that there is no single land use–transportation type that is ideal for all bicycling activities or people and, hence, no unambiguous way of defining the quality of the environment.

The second problem with this sort of model is that there are large and seemingly random differences from one place to another. In one area analyzed in Minneapolis, 16% of the adults made bike trips on the day that they were surveyed, while the rate in many other areas was 0%. Even across entire metropolitan areas or states, differences of a factor of 10 can be seen, as shown in the previous section.

Some well-documented population and land use characteristics are associated with higher levels of bicycling. For example, people with college educations are more likely to bicycle, but the difference is on the order of a factor of 2 compared with that for less educated people. Similar differences exist for factors such as income, development density, and gender. None of the known factors, alone or together, can come close to explaining why people in some places are 10 or more times as likely to ride bikes as people in other places. Other attitudinal and possibly historical factors seem to dwarf the effects of the factors that planners and policy makers can control.

Because the impacts of the unobservable variables are so big relative to those of the variables of interest, it seems highly likely that what is being observed, both in the authors' model and in other models, is the effect of attitudinal variables acting on policy variables through spurious correlations. The authors' model seems to have been driven by geographic spikes in riding that happened to be positively correlated with some facility measures and negatively correlated with others but that in a causal sense had little or nothing to do with any of them. It seems possible that these types of spurious correlations might also be driving the results of other work of this type in the literature, given the typically low explanatory power of these models.

A third problem is that low levels of bicycling cause the range of sampling error to be many times larger than the sample mean for any realistic sample size. The effect is that the regression is trying to match measured variable values that could be off by a factor of 5 or more from their true values. A sample of 1,000 people would yield nine cyclists on a given day at the national average level; the 95% confidence interval for this sample ranges from 3 to 15 cyclists. This is a large difference in relative terms, and the observed extremes could easily just be sampling aberrations. Yet, these inaccurate measurements could strongly influence the estimated parameter values.

Finally, there is always the problem, noted by Dill and Carr (*5*) and others, that even a positive correlation between riding and facilities could be causation in the other direction, that is, that the large number of cyclists creates the political climate to build the facilities rather than that the facilities encourage more riding. For example, bike lanes in some cases may be a response to existing situations such as bikes interfering with traffic. In these cases retrofitted lanes will often be associated with high levels of cycling after they are built. By contrast, lanes in some newer cities in California do not seem to have high riding levels (*5*), possibly because they were designed into new roads; that is, they were built in anticipation of riding rather than in response to it.

Seemingly, the only way around these problems would be to study the same geographic area over a period of time as facilities change. The relevant question for policy is not comparing people living at Location A with different people living at Location B but, rather, comparing the people at Location A with themselves as the provision of facilities changes over time. This would be an expensive prospect by the use of surveys; the development of a low-cost method of counting bikes over a large number of different streets and bike facilities, such as that outlined by Davis and Wicklatz (*15*), would be of great value for this purpose.

MODEL OF TOTAL BICYCLING DEMAND

This section outlines a simple sketch planning method for estimating the number of daily bicyclists in an area by using easily available data from CTPP (*14*). An estimate of the current number of bicyclists in an area could be used for general political purposes to justify expenditures by reference to the number of bicyclists and the benefits that they receive from cycling. However, the more interesting problems for planners are predicting how the number of cyclists will change as a result of a facility or other improvement and knowing how many cyclists use or will use a specific facility.

While this model does not directly address these questions, the authors believe that it is still useful because the answers to these questions will, in general, need to be conditioned on the basis of the number of current bicyclists. This is not to say that the number of cyclists in an area cannot grow; the examples of many cities with high proportions of bicyclers show what is possible. In general, however, any growth will probably be gradual rather than abrupt and will likely depend on continued improvement of the cycling environment. Thus, the rules of thumb developed here are not intended to represent permanent bounds on possible cycling levels but only to provide a range of likely short-term changes.

Similarly, while the use of a given facility will probably depend on a host of site-specific factors, in most cases it will also be limited in the short term by existing bicycling habits among the surrounding population. A thousand daily users may be realistic in an area where 2,000 people a day currently ride bikes; it is probably not realistic in an area where 200 do. This is not to say that facilities are justifiable only in areas that already have a lot of cyclists or that a facility cannot increase the number of cyclists. The point, again, is only to provide an empirical basis for the development of realistic expectations regarding short-term results.

The basic assumptions motivating this analysis are that a large fraction of total bicycling is done by a small fraction of cyclists who ride frequently, as discussed earlier, and that many of these frequent riders are bicycle commuters observed in the census commute-to-work data. The hypothesis tested in this section is that the basic riding frequency table described in the previous section will hold more or less across different areas. Thus, an area with many commuter cyclists will also have more total cycling, and an area with few commuter cyclists will have little total riding. In other words, commuting by bicycle, while it is a small fraction of the total bicycling in a given area, can still be used as a leading indicator of what might be happening with other types of cycling.

Three different geographic divisions are examined to study this issue. The first is a set of 15 MSAs for which CTPP commute-to-work shares could be matched with NHTS (*9*) daily bicyclist counts. The next one is states; there are 34 for which both census and NHTS data were available. The last one comprises 66 zones of the Minneapolis–St. Paul MSA obtained by using data from TBI (*8*), which shows that the basic principle still works at this very different geographic scale.

TBI and NHTS, like most travel diary data, are limited by small sample sizes for specific geographic areas. Because of this and the low level of cycling, the expected number of cyclists in the sample for a given area could vary by a factor of 10 or more from the low to the high end of the range. Ordinary measures of goodness of fit have little meaning in this sort of environment; the focus instead is on more heuristic measures, such as the number of observations that fit within the predicted confidence interval.

Metropolitan Statistical Areas

Combining census data with the NHTS analysis produced 15 MSAs for which both commute-to-work shares by bike and the total percentage of adults biking on the survey day were available. The commute shares ranged from 0.1% (Cincinnati, Ohio, and Dallas, Texas) to 1.4% (Sacramento, California). The daily adult biking shares ranged from 0.18% (Houston, Texas; although this is probably a sampling problem, as 4.2% rode during the previous week) to 2.45% (Portland, Oregon, with Sacramento close behind at 2.25%). Parameter values were estimated as shown in Equation 2; the R^2 value for this equation was about .7.

$$A = 0.3\% + 1.5 * C \tag{2}$$

where A is the percentage of the adult population who bicycle in a day, and C is the bicycle commute share (in percent).

This equation can be used to generate a predicted total riding share for each city. Given this predicted share and the NHTS sample size, a 95% confidence interval of the expected number of adult bicyclists in the sample can be calculated by assuming a binomial function. For 14 of the 15 cities, the actual number of bicyclists fell within this confidence interval. The one exception was Chicago, Illinois, which generated 19 actual cyclists, compared with a predicted number of 9.

The performance of this model at predicting the observed number of cyclists for the cities with the biggest samples (and, presumably, the most reliable numbers) is quite good, again, with the exception of Chicago: New York City had 20 predicted cyclists and 23 actual cyclists; Los Angeles, California, had 23 predicted cyclists and 22 actual cyclists; San Francisco, California, had 21 predicted cyclists and 19 actual cyclists; and Boston, Massachusetts, had nine predicted cyclists and seven actual cyclists. At the low and the high ends of the commuter cyclist ranges, Cincinnati had one predicted cyclist and one observed cyclist, Dallas had three predicted cyclists and three observed cyclists, Portland had six predicted cyclists and 10 observed cyclists, and Sacramento had nine predicted cyclists and eight observed cyclists. Portland was among the cities with the worst predicted number of cyclists, but the number was still within the 95% confidence interval. Overall, as Figure 1 shows, the hypothesis that overall bicycling rates will correlate with bicycle commuting rates seems to be supported; indeed, the correlation seems quite strong at this geographic level.

The equation is also exactly consistent with the findings for the United States as a whole (0.4% commute share, 0.9% total daily cyclists) and with a division into larger and smaller cities, for which the same data were observed.

States

There were 34 states with data from both the census and NHTS. Alabama had the lowest bicycle commute share at 0.07%; Oregon the highest at 1.07%. Arkansas had the lowest rate of total bicycling at 0% (this is again a sampling problem, as 3.4% of the people in Arkansas surveyed rode a bicycle during the preceding week), and Florida had the highest at 2.21%. Equation 3 shows the estimated parameter values, which are slightly different from those observed at the MSA level. The R^2 value of this model was about .3.

$$A = 0.4\% + 1.1 * C \tag{3}$$

By using either these parameter values or those derived from the MSA level, the same predictive results emerge. Of the 34 states, 30 have actual counts within the 95% confidence interval of their

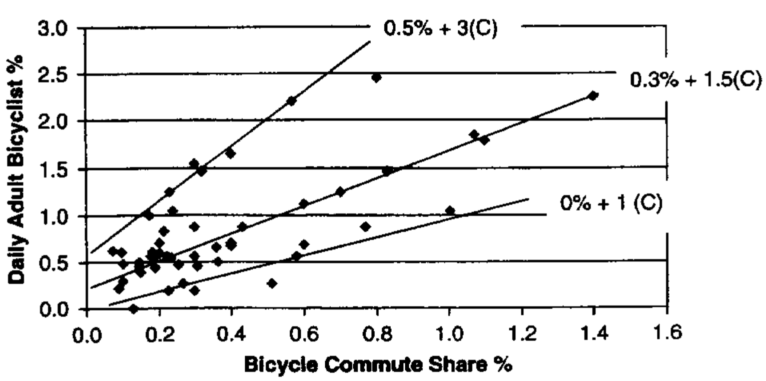

FIGURE 1 Daily bicyclists and commute share, combined MSAs and states.

predicted values; the exceptions are all underpredicted. Of the states with good sample sizes (more than 1,000), the values for about half were predicted with good accuracy (less than 1 standard deviation); the values for other half were farther off the mark, with predictions both too high and too low.

Twin Cities Zones

The final level of analysis considered variations within the Minneapolis–St. Paul MSA by using 66 zones that had been defined for a different project. These were largely based on political boundaries, with the two central cities divided into a number of zones on the basis of natural and artificial divisions and neighborhood characteristics. The populations of the zones ranged from about 10,000 to 30,000. While this analysis was based on TBI, which was a large local survey, there were still only 139 adults in the survey who made bike trips (and whose home location could be mapped to a zone in this area), and a third of these were in four zones in Minneapolis. Thus, the estimated bicycling rates for most of the zones are extremely unreliable. The results of this regression are shown in Equation 4:

$$A = 0.6\% + 2.5 * C \qquad (4)$$

The relatively high slope and intercepts of this equation are likely a reflection of the outlier nature of the Twin Cities compared with the data for the areas on which the previous two regressions were based. That is, depending on the measurement, the Twin Cities have an overall adult bicycling rate of 1.6% to 2%, which is quite high compared with their bike commute share of 0.4%. Thus, overall bicycling there is about twice as high as would be predicted by the earlier regressions; given this, it is perhaps logical that the estimated parameter values with data drawn from this region would be about twice as high as well.

In most of the zones the sample size was too small to present an interesting prediction problem; that is, for these zones both the prediction and the actual count were either one or zero. For those 15 zones where the predicted number was two or more, the predictions for 12 were within the 95% confidence interval, while three zones had actual values in excess of the predicted range. Although the predictions for the zones with high rates of bicycling were not accurate in absolute terms, the general relationship between commuting and total bicycling held in general. The six zones where commuting by bike exceeded 2% generated six of the seven highest rates of overall daily bicycling.

CONCLUSION

On any given day, roughly 1% of the adults in the United States ride a bicycle. Over large geographic areas such as metropolitan areas or states, this number could range roughly from about 0.3% to 2.5%. Over smaller areas, such as specific parts of metropolitan areas, the range could go as high as 15%. These variations are far larger than can reasonably be explained by differences in formal policies and facilities. It seems that local or even subcultural attitudes and perhaps history play a substantial role in the perception of bicycling as an appealing or even normal thing for an adult to do, although without further study it is difficult to imagine how these factors might exert their influence.

When the actual percentage of cyclists in an area is not known, it can be estimated with reasonable accuracy by considering the bicycle commute-to-work share. A most likely value would be 0.3% plus 1.5 times the commute share; this was the best fit at the MSA level and also describes the United States as a whole. Figure 1 shows lines representing rough boundaries on the observed values for daily cyclists, as they relate to bicycle commute shares at the MSA and the state levels. In fact, these lines appear to represent three distinct relationships between these two variables that are observed in the data, but at this point this must be considered a sampling coincidence.

The model described here has important practical advantages. It is simple enough to be understandable to those who make funding decisions and provides a known range of possible outcomes derived from a wide variety of locations and different geographic scales. However, it does fall short of the modeling ideal of directly describing a relationship between the provision of bicycling facilities and the amount of bicycling that will take place. The formulas derived here simply describe the amount of bicycling that is currently taking place; they do not relate this amount in a causal way to explanatory factors or explain how it might change. The authors believe that this compromise is necessary because of the findings described in the first two parts of this paper.

By helping the planner to estimate a range of the number of bicyclists currently riding in a given geographic area, the model establishes a baseline that can be used to develop more informed estimates about how this number might change given a change to the facilities or cycling environment. Such a baseline is necessary for any more detailed estimates or predictions because there is such a high degree of variation in bicycling demand levels in different locations. This model represents a first step in such a methodology; the question of how to get from a general estimate of current bicycling levels to predictions about general or facility-specific future levels is left to later research.

More qualitative research that can be used to obtain a better understanding of the outliers could also be useful. Some MSAs and states have very high or very low levels of bicycle commute shares or daily adult bicyclists. Over such large populations, this seems unlikely to be very much due to demographic differences. More detailed case studies of places that generate these very high or very low rates of bicycling could be enlightening, especially if soft factors such as culture and attitudes can be probed in some systematic way.

ACKNOWLEDGMENTS

The authors thank NCHRP and the Minnesota Department of Transportation for providing financial support for this research.

REFERENCES

1. Porter, C., J. Suhrbier, and W. Schwartz. Forecasting Bicycle and Pedestrian Travel: State of the Practice and Research Needs. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1674*, TRB, National Research Council, Washington, D.C., 1999, pp. 94–100.
2. Landis, B. Using the Latent Demand Score Model to Estimate Use. In *Pro Bike/Pro Walk 96 Resource Book.* Presented at the Ninth International Conference on Bicycle and Pedestrian Programs, Portland, Maine, Sept. 1996, pp. 320–325.
3. *Guidebook on Methods to Estimate Non-Motorized Travel.* Publication FHWA-RD-98-166. FHWA, U.S. Department of Transportation, 1999.