

# UC Berkeley

## Research Reports

### Title

Pilot Models for Estimating Bicycle Intersection Volumes

### Permalink

<https://escholarship.org/uc/item/380855q6>

### Authors

Griswold, Julia B.  
Medury, Aditya  
Schneider, Robert J.

### Publication Date

2011

Peer reviewed

## Pilot Models for Estimating Bicycle Intersection Volumes

Julia B. Griswold, SafeTREC,  
Aditya Medury, Institute of Transportation Studies, Berkeley, and  
Robert J. Schneider, SafeTREC

1 **Pilot Models for Estimating Bicycle Intersection Volumes**

2  
3  
4 Julia B. Griswold (Corresponding Author)  
5 Safe Transportation Research and Education Center  
6 University of California, Berkeley  
7 2614 Dwight Way #7374  
8 Berkeley, CA 94720, U.S.A  
9 Telephone: (510) 643-0566  
10 Fax: (510) 643-9922  
11 Email: juliagris@berkeley.edu

12  
13 Aditya Medury  
14 Department of Civil and Environmental Engineering  
15 Institute of Transportation Studies  
16 University of California, Berkeley  
17 Email: amedury@berkeley.edu

18  
19 Robert J. Schneider  
20 Safe Transportation Research and Education Center  
21 University of California, Berkeley  
22 2614 Dwight Way #7374  
23 Berkeley, CA 94720  
24 Telephone: (510) 643-0566  
25 Fax: (510) 643-9922  
26 Email: rschneider@berkeley.edu

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48 Word Count:  $5,030 + (4 \text{ tables} + 1 \text{ figure}) \times 250 = 6,280$  words

49  
50 Submitted to the Transportation Research Board 90<sup>th</sup> Annual Meeting

## 1 **Pilot Models for Estimating Bicycle Intersection Volumes**

### 3 **ABSTRACT**

4 Bicycle volume data are useful to practitioners and researchers to understand safety, travel behavior, and  
5 development impacts. This paper describes the methodology used to develop several simple models of  
6 bicycle intersection volumes in Alameda County, California. The models are based on two-hour bicycle  
7 counts performed at a sample of 81 intersections in the Spring of 2008 and 2009. Study sites represented  
8 areas with a wide range of population density, employment density, proximity to commercial property,  
9 neighborhood income, and street network characteristics. The explanatory variables considered for the  
10 models included intersection site, land use, transportation system, and socioeconomic characteristics of  
11 the areas surrounding each intersection. Four alternative models are presented with adjusted R-square  
12 values ranging from 0.39 to 0.60. The models showed that bicycle volumes tended to be higher at  
13 intersections surrounded by more commercial retail properties within 1/10 mile, closer to a major  
14 university, with a marked bicycle facility on at least one leg of the intersection, surrounded by less hilly  
15 terrain within 1/2 mile, and surrounded by a more connected roadway network. The models also showed  
16 several important differences between weekday and weekend intersection volumes. The positive  
17 association between bicycle volume and proximity to retail or a large university was greater on weekdays  
18 than weekends, while bicycle facilities had a stronger positive association and hilly terrain had a weaker  
19 negative association with bicycle volume on weekends than weekdays. Further testing and refinement is  
20 necessary before accurate count predictions can be made in Alameda County or other communities.

1   Bicycling is increasing in many United States communities. Nationwide, the proportion of commuters  
2   traveling to work regularly by bicycle grew from 0.45% to 0.55% between 2006 and 2008 (1,2). Cities  
3   such as Portland, New York, Seattle, and San Francisco documented increases in bicycle counts over this  
4   same time period (3,4,5,6). Bicycling is a convenient and economical mode for commuting, small  
5   shopping trips, transit access, and recreation. Estimates of bicycle activity are valuable to planners,  
6   engineers, designers, public health professions, and others. However, there are very few tools available to  
7   estimate the number of bicyclists that pass by specific locations or use particular intersections within an  
8   urban area. A predictive model of bicycle volumes can be used to:

- 9       • Quantify bicycle exposure in safety analyses, providing the denominator in the calculation of  
10      crash risk;
- 11      • Identify priority locations for bicycle facility improvements or safety measures; and
- 12      • Estimate changes in bicycle volumes that will occur with new developments, roadway changes,  
13      or new transit projects.

## 14   **STUDY PURPOSE**

15   The purpose of this study is to present several preliminary models that can be used to estimate bicycle  
16   intersection volumes. A simple model structure, loglinear regression, was chosen so that the models  
17   would be easy to apply using geographic information systems (GIS) and simple spreadsheet software.  
18   Since the analysis was conducted in one urban area (Alameda County, California), more research is  
19   needed to refine the model equations and determine the applicability of the results for other communities.  
20   These models are designed for planning purposes to help identify built environment characteristics  
21   associated with higher and lower levels of bicycling and show general differences in bicycle volumes at  
22   intersections throughout a city or region. Applications that require precise bicycle volume data for  
23   individual intersections should use actual bicycle counts from each location.

## 24   **PREVIOUS STUDIES**

25   Several techniques have been developed to analyze bicycle activity in different locations. Overlay  
26   mapping techniques, or sketch-plan methods, are useful for planning and prioritization (7,8,9). However,  
27   they are not typically calibrated to actual bicycle counts. Most previous efforts to model bicycle activity  
28   have used zonal areas or the individual as the unit of analysis. Some regional travel demand models have  
29   been adjusted to estimate bicycle mode share within a certain area (10), but these models are usually  
30   estimated at the traffic analysis zone (TAZ) level and are not fine-grained enough to capture intersection-  
31   level bicycle activity. Models that examine behavior at the level of the individual are valuable for  
32   identifying factors that determine mode and route choice, but they are difficult to use for volume  
33   estimates. Most of these models are based on detailed household travel survey data, which can be  
34   expensive to collect.

35   Two recent studies in Southern California have developed predictive models of bicycle volume,  
36   using linear regression with bicycle counts as the dependent variable and demographic and land use  
37   measures for the independent variables. Four land use and transportation system explanatory variables  
38   (afternoon bus frequency, land use mix, density of residents under age 18, and proximity to the bicycle  
39   network) were used to predict weekday afternoon peak hour bicycle volumes at intersections in the City  
40   of Santa Monica (11). However, the ordinary least squares regression equation used for the model could  
41   produce negative predictions. This modeling issue was addressed in San Diego County by using the  
42   natural logarithm of the dependent variable, but the San Diego model was limited to two explanatory  
43   variables (employment density and length of nearby multi-use trails) for predicting weekday 7 a.m. to 9  
44   a.m. bicycle volumes (12).

45   An additional bicycle modeling study explored applying Space Syntax measures in a bicycle  
46   model. A Space Syntax measure representing direct paths in the street network was combined with land  
47   use variables to predict morning peak hour bicycle volumes at intersections and street segments in  
48   Cambridge, MA (13). The model performed fairly well, but the sample size was small (n=16) and the  
49   50

1 land use variables (population and employment density) explained most of the variation in the bicycle  
 2 volumes. In addition, the Space Syntax measure required special software to calculate.

3 Table 1 summarizes the factors that have been associated with bicycle activity in these modeling  
 4 efforts and other bicycle studies.

5

6 Table 1. Factors Associated with Bicycling in Previous Research

Variable	Relationship with Bicycle Activity	Source
<b>Environmental and Land Use Variables</b>		
Nearby population density	+	McCahill and Garrick (2008) (13)
Nearby employment density	+	McCahill and Garrick (2008) (13), Jones et al. (2010) (12)
Land use mix	+	Haynes and Andrzejewski (2010) (12)
Proximity to downtown	+	Dill and Voros (2007) (14)
Slope	-	Aultman-Hall, Hall, and Baetz (1997) (15), Dill and Gliebe (2008) (16)
<b>Transportation System Variables</b>		
Nearby street connectivity	+	Dill and Voros (2007) (14)
Proximity to a freeway	+	Dill and Voros (2007) (14)
Amount of bicycle lanes nearby	+	Dill and Carr (2003) (17)
Amount of multi-use trails nearby	+	Jones et al. (2010) (12)
Presence of bicycle facility	+	Dill and Gliebe (2008) (16), Stinson and Bhat (2003) (18)
Proximity to bicycle network	+	Haynes and Andrzejewski (2010) (12)
Afternoon peak hour bus frequency	+	Haynes and Andrzejewski (2010) (12)
Motor vehicle volume	-	Stinson and Bhat (2003) (18)
Major arterial	-	Stinson and Bhat (2003) (18)
Parallel parking permitted	-	Stinson and Bhat (2003) (18)
Smooth pavement	+	Stinson and Bhat (2003) (18)
<b>Socioeconomic Variables</b>		
Age 18-55	+	Dill and Voros (2007) (14)
Age	-	Xing, Handy, and Buehler (2008) (19)
% of nearby population under age 18	-	Haynes and Andrzejewski (2010) (12)
Education level	+	Xing, Handy, and Buehler (2008) (19)
Income	+	Dill and Voros (2007) (14)
Household automobile availability	-	Dill and Voros (2007) (14), Dill and Carr (2003) (17)

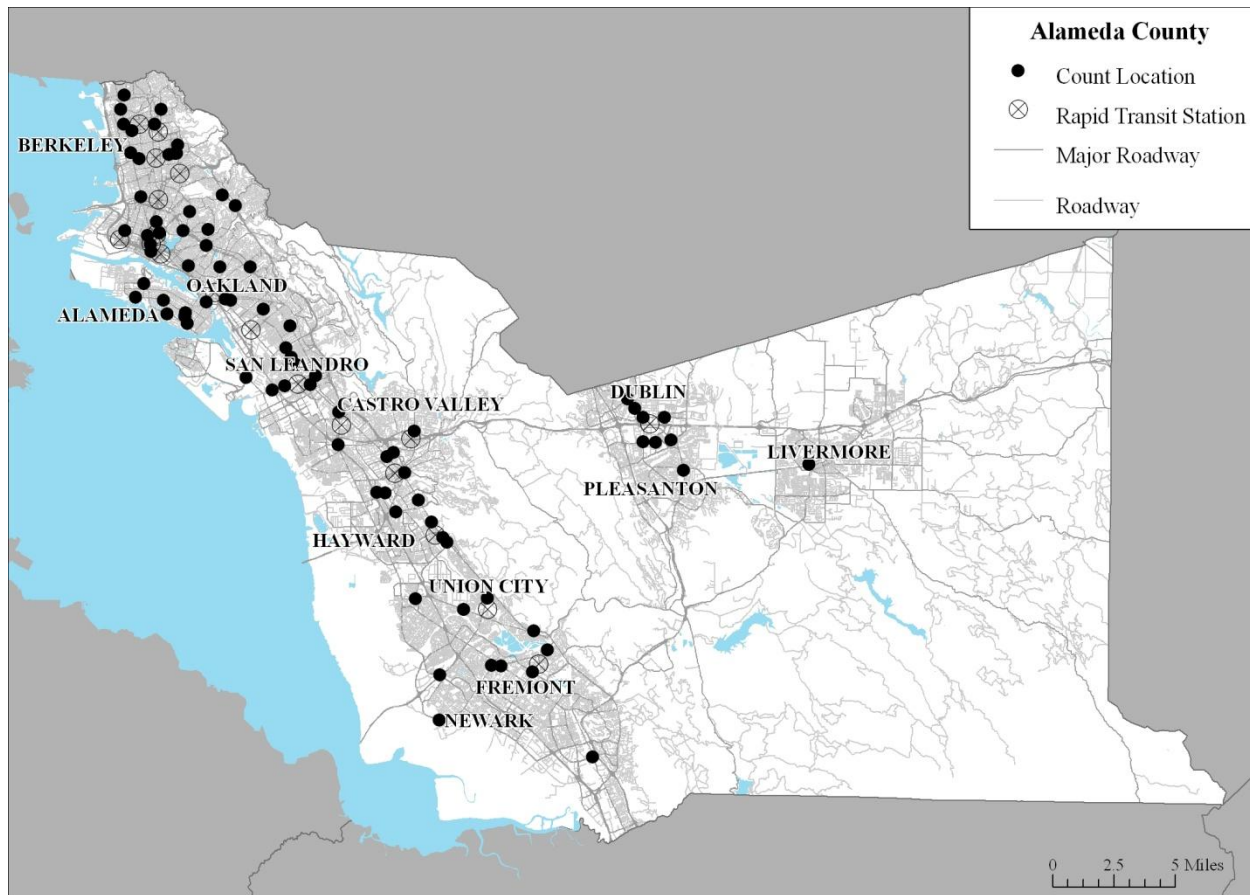
7

8

## 9 METHODOLOGY

10 This section describes the study area, the bicycle count collection, and the development of the explanatory  
 11 variables. The model is based on bicycle counts taken at 81 intersections along arterial and collector  
 12 streets in Alameda County, California (Figure 1).

13



1  
2 Figure 1. Map of Manual Bicycle Count Locations  
3

#### 4 Study Area

5 Alameda County is located on the eastern side of the San Francisco Bay across the Bay Bridge from San  
6 Francisco, and it stretches east and south towards the San Joaquin Valley. Its population in 2008 was  
7 estimated to be 1.47 million (2). The western portion of the county was largely developed in the early to  
8 mid-twentieth century and contains several small downtown commercial areas and many streetcar suburbs  
9 with grid street patterns. The hill areas and the eastern portion of the county were mostly developed in  
10 the second half of the twentieth century; these areas are more suburban in nature, with curvilinear street  
11 patterns. The largest city in the county is Oakland, with an estimated population of approximately  
12 366,000 (2). In 2008, estimates are that 48 percent of the county population was white, 25 percent Asian,  
13 22 percent Hispanic or Latino, and 13 percent black (2).  
14

#### 15 Bicycle Counts

16 The bicycle counts for this study were collected in the spring of 2008 and 2009 in conjunction with a  
17 study of pedestrian exposure and crash risk. While the pedestrian study determined that the counts would  
18 be at intersections, the count locations were also appropriate for understanding bicycle exposure because  
19 approximately half of bicycle crashes occur at intersections (20). Counts were taken at 50 intersections in  
20 2008 and 31 in 2009. All intersections were along arterial or collector roadways. A strategic process was  
21 used to select intersections, ensuring that the set of study sites represented areas with a wide range of  
22 population density, employment density, proximity to commercial properties, neighborhood incomes, and  
23 other socioeconomic characteristics. In addition, the study intersections had a variety of roadway  
24 characteristics. Of the 81 intersections, 16 (20%) had a marked bicycle facility (e.g., bicycle lane, bicycle

1 boulevard, or shared lane marking) on at least one approach. The choice of intersection locations is  
 2 described in detail in Schneider, Arnold, and Ragland (21) and Schneider et al. (22).

3 Each bicycle was logged according to the movement that was made at the intersection—straight,  
 4 right turn, or left turn from one of the four intersection legs—which made a total of 12 possible  
 5 movements. Bicycles being ridden either on the street or on the sidewalk were included, but bicycles  
 6 being walked were logged as pedestrians. The counts for each movement were summed together to find  
 7 the total bicycle volume for the intersection. Counts were conducted for two 2-hour periods at each  
 8 intersection. One was taken on a Tuesday, Wednesday, or Thursday, and one was taken on a Saturday.  
 9 The time of day for the counts varied depending on data collector scheduling. Counts were taken during  
 10 12 p.m.-2 p.m., 2 p.m.-4 p.m., 3 p.m.-5 p.m., or 4 p.m.-6 p.m. on weekdays and 9 a.m.-11 a.m., 12 p.m.-2  
 11 p.m., 3 p.m.-5 p.m., or 4 p.m.-6 p.m. on weekends. Additional information about the count methodology  
 12 is provided in Schneider, Arnold, and Ragland (21) and Schneider et al. (22). Table 2 includes summary  
 13 statistics for the bicycle count data.

14  
 15 **Table 2. Summary Statistics for all Two-Hour Bicycle Counts**

Statistic	All Counts	Weekday	Weekend
Number of Counts	162	81	81
Minimum	0	0	1
Maximum	343	343	171
Median	23.5	22	24
Mean	35.8	38.6	33.0
Standard Deviation	41.4	50.3	30.1

16  
 17 **ANALYSIS**

18 The data analysis process for the model estimation included three steps: (a) explanatory variables were  
 19 screened to eliminate those weakly correlated to the dependent variable; (b) the remaining variables were  
 20 screened for collinearity to avoid including strongly correlated variables in the same model; and (c) four  
 21 alternative model structures were chosen for strong goodness-of-fit and significant coefficients.

22  
 23 **Regression Modeling**

24 Loglinear ordinary least squares (OLS) regression was used to estimate a model of bicycle intersection  
 25 volume. In loglinear regression, the dependent variable is transformed using the natural logarithm. This  
 26 is an appropriate method for modeling count data because when the natural log predictions are  
 27 transformed back to counts using the exponential, there are no negative values. A negative binomial form  
 28 of the regression model was also tested during the analysis process, and the model coefficients were  
 29 similar. The loglinear form was selected because it is easy to apply using spreadsheet software and  
 30 because it is easier to interpret the model goodness-of-fit and independent variable coefficients. One  
 31 intersection had a count of zero bicyclists. This count was changed to 0.1 to allow the log to be  
 32 computed. The model formulation is:

33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 (1)

where:

$Y_i$  = bicycle counts at intersection  $i$

$X_{ji}$  = value of explanatory variable  $j$  at intersection  $i$

$\beta_j$  = model coefficient for variable  $j$

41 The modeling process examined the relationship between the bicycle intersection volume and the  
 42 built environment surrounding the intersection, such as nearby land use, transportation system, and site  
 43 characteristics. Some of the measures were evaluated at different scales, including buffer radii of 0.1 mi



1 (161 m), 0.25 mi (402 m), and 0.5 mi (805 m). This study took advantage of explanatory variables  
 2 previously developed for the 81 intersections for pedestrian volume and pedestrian crash models (21,22).  
 3 Several additional factors used in previous studies were also included, bringing the total to more than 70  
 4 variables considered in the analysis. Network distance to UC Berkeley Campus and Oakland City Center  
 5 were included to account for two of the major trip attractors in the county. Connected node ratio was  
 6 calculated using GIS. This measure represented the ratio of three- and four-way intersections to dead-end  
 7 streets in the area around each study intersection. Areas with lower connected node ratios had higher  
 8 proportions of dead ends. The new variables and any existing variables that were used in the final  
 9 bicycle volume models are described in Table 3. Descriptions of the other variables considered for the  
 10 analysis can be found in Schneider, Arnold, and Ragland (21) and Schneider et al. (22).

11 It was necessary to narrow down the number of independent variables to avoid overfitting the  
 12 model. Models with nearly as many independent variables as the number of observations they are based  
 13 on predict volumes poorly. First, correlation coefficients were estimated between each explanatory  
 14 variable and the bicycle counts. Those variables with weak correlation ( $\rho < 0.2$ ) were eliminated from the  
 15 analysis. Next, the remaining independent variables were screened for collinearity; pairs of variables with  
 16 moderate to strong correlation ( $\rho > 0.3$ ) were not included in the same model.  
 17

18 Table 3. Description of Selected Explanatory Variables Considered in the Modeling Process

Variable Name	Variable Description <sup>1</sup>
<b>Intersection Site Characteristics</b>	
SlopeT	Average slope (degrees) of terrain within 1/10 mile
SlopeQ	Average slope (degrees) of terrain within 1/4 mile
SlopeH	Average slope (degrees) of terrain within 1/2 mile
BikeSym	Presence of bicycle markings on any approach (dummy) <sup>2</sup>
<b>Surrounding Land Use And Transportation System Characteristics</b>	
NComPropT	Number of commercial properties within 1/10 mile
UCBDist	Network distance to UC Berkeley Campus edge (mi.)
lnUCBDist	Natural log of network distance to UC Berkeley Campus edge (mi.)
TransDist	Network distance to nearest BART station (mi.)
CCDist	Network distance to Oakland City Center (mi.)
<b>Surrounding Transportation System Characteristics</b>	
IDensT	Intersection density within 1/10 mile
IDensQ	Intersection density within 1/4 mile
IDensH	Intersection density within 1/2 mile
CNRT	Connected node ratio within 1/10 mile
CNRQ	Connected node ratio within 1/4 mile
CNRH	Connected node ratio within 1/2 mile
<b>Other Variables</b>	
Count09	Indicates counts conducted in 2009 (dummy)

19 1. 1 mile = 1.61 km

20 2. Bicycle markings include on-street bicycle facilities: bicycle lanes, bicycle boulevard markings, or shared lane  
 21 markings. They do not include bicycle route signs on approaches without other street markings. This variable does  
 22 not represent the association between multi-use trails and bicycle volume. Multi-use trail proximity was tested in a  
 23 different variable that was not significant in the predictive models.  
 24

### 25 **Factors Associated with Intersection Bicycle Volumes**

26 Several options were considered during the model specification process. The models discussed are shown  
 27 in Table 4. While the adjusted R-square values varied among the models, the F-statistics showed that all  
 28 models were significant. Model A was the best fit model specification based on all counts ( $N = 162$ ,  $F$   
 29  $= 33.9$ ). Number of retail commercial properties within 1/10 mile and presence of on-street bicycle  
 30 facilities both had positive coefficients in this model, indicating that retail activity and bicycle facilities

1 were associated with higher levels of bicycle activity at the 81 study intersections. The average slope  
 2 within 1/2 mile of each study intersection had a negative coefficient, indicating that cyclists tended to  
 3 avoid hilly terrain. These results are consistent with previous research.

4 The natural log of the distance to the UC Berkeley Campus had a negative coefficient, indicating  
 5 that the further the intersection was located from the campus, the less likely it was to have a high bicycle  
 6 volume. The high numbers of bicycles counted at intersections near campus were not surprising since UC  
 7 Berkeley is the largest employer in Alameda County, the largest university in the Bay Area, and has  
 8 limited space for on-campus automobile parking. The natural log form of the variable was used because  
 9 beyond a certain distance the likelihood of bicycling to campus drops dramatically. This variable may  
 10 also be a proxy for general urban, bicycle-friendliness. Berkeley has a well-connected bicycle boulevard  
 11 system, and Oakland has more than 80 miles of bicycle lanes and routes (23,24). Both cities are located  
 12 in the relatively dense, urban, northwestern corner of Alameda County. When tested, the distance to  
 13 Oakland City Center variable was also correlated with the bicycle counts, but it was strongly correlated  
 14 with the UC Berkeley Campus variable. Since the distance to UC Berkeley variable was a better  
 15 predictor of bicycle volumes, it was used instead of distance to Oakland City Center.

16 A dummy variable indicating counts performed in 2009 was also included in the model because it  
 17 was highly significant. This variable is of less interest in this study because it is probably unrelated to the  
 18 built environment; it could reflect changes in behavior due to gas price shock of Summer 2008, the  
 19 recession of 2009, or differences in the sample of intersections where counts were conducted in 2008 and  
 20 2009.

21  
 22 Table 4. Alternative Bicycle Model Specifications

Model Variable	Model A: All Counts		Model B: Weekday		Model C: Weekend		Model D: Weekday Alt.	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Dependent Variable = 2-hr Intersection Bicycle Count								
Constant	3.776	0.185***	3.899	0.262***	3.652	0.255***	-1.127	0.855
NComPropT	0.024	0.007***	0.030	0.010***	0.017	0.010*		
BikeSym	0.477	0.163***	0.437	0.230*	0.517	0.225***	0.459	0.269*
lnUCBDist	-0.458	0.059***	-0.546	0.083***	-0.369	0.081***		
SlopeH	-0.517	0.073***	-0.659	0.103***	-0.375	0.100***	-0.470	0.117***
CNRH							4.634	0.989***
Count09	0.811	0.127***	1.002	0.180***	0.620	0.176***	1.036	0.211***
Overall Model								
Sample size (N)	162		81		81		81	
Adjusted R <sup>2</sup>	0.505		0.600		0.386		0.450	
F-test	33.87***		24***		11.08***		17.38***	

23 NOTE: Coeff. = coefficient and Std. Err. = standard error. \*\*\* = significant at 99% (p < .01); \*\* = significant at 95% (p < .05);  
 24 \* = significant at 90% (p < .10). Model variables are defined in Table 3.

### 25 26 Weekday and Weekend Bicycle Volume Models

27 Although all of its coefficients have intuitive signs and are significant at the 99 percent confidence level,  
 28 Model A is problematic because the sample includes two counts from each intersection (one from a  
 29 weekday and one from a weekend time period). Datasets contain unobserved heterogeneity when  
 30 multiple observations are considered for the same location, and as a result, the standard errors are  
 31 evaluated incorrectly, which could lead to some variables being seen as more significant than they are. A  
 32 first approach to avoid this problem was to estimate Model A with a dummy variable representing  
 33 weekday versus weekend counts, but it was not statistically significant. Therefore, a second approach  
 34 was used. Separate models were estimated for the weekday (Model B) and weekend (Model C) counts.

1 As Table 4 shows, the weekday model (Model B) has the highest adjusted R-square (0.6) of the models  
 2 presented. The coefficients of Model B are similar to Model A, but the coefficient for the bicycle  
 3 markings is only significant at 90 percent. The weekend model (Model C), on the other hand, has a lower  
 4 adjusted R-square than Models A or B. While presence of bicycle markings is highly significant with a  
 5 larger coefficient, number of commercial properties is less significant with a smaller coefficient than in  
 6 the previous models. This result may have occurred because recreational bicycle trips are more common  
 7 on weekends, so the proximity of retail establishments may be less important for attracting weekend  
 8 bicycle activity. Bicycle markings could have a greater influence on weekends because bicyclists may  
 9 have more flexibility to seek routes with bicycle facilities or because less experienced riders who tend to  
 10 prefer marked bicycle facilities are more likely to ride on weekends. Model C also has smaller  
 11 coefficients for the UC Berkeley variable and the slope variable. This is consistent with the expectation  
 12 that fewer people would see the campus as a destination on the weekend and recreational riders might  
 13 seek out hilly terrain.

14 Finally, in consideration that Models A through C include a variable that is particular to this  
 15 Alameda County study area (*lnUCBDist*), a fourth model specification is presented. Model D is an  
 16 alternative weekday model that excludes the UC Berkeley variable. Model D also adds connected node  
 17 ratio within 1/2 mile and excludes the commercial property variable. As expected, connected node ratio  
 18 has a positive coefficient, meaning that areas with fewer dead end streets and cul-de-sacs have more  
 19 bicycle activity. Population density and employment density were included in a variety of model  
 20 specifications, but they did not perform well. Although these variables have been used in previous  
 21 studies, they were not included in the final set of bicycle volume models.

## 22 **PRACTICAL APPLICATION**

23 Practitioners wishing to apply an intersection bicycle volume model should select among Models B, C,  
 24 and D based on the type of counts being predicted (weekday or weekend) and their understanding of the  
 25 built environment in the study area. To extract counts from the models above the following formulation  
 26 should be used:  
 27

28  
 29 (2)

30 where:

31  $Y_i$  = predicted two-hour bicycle count at intersection  $i$   
 32  $X_{ji}$  = value of explanatory variable  $j$  at intersection  $i$   
 33  $\beta_j$  = model coefficient for variable  $j$   
 34

35 When working with the loglinear form of a regression model, it is important to understand how to  
 36 interpret the coefficients correctly. They can be interpreted as the fractional increase in the bicycle  
 37 volume when the explanatory variable increases by 1 unit. Unlike OLS regression, the relationship  
 38 between coefficients, explanatory variables, and the predictions is nonlinear. For example, consider the  
 39 application of Model D to the median variable values for the set of 81 sample intersections. This model  
 40 produces a count prediction of 15 bicycles within a two-hour period on a typical weekday afternoon. The  
 41 median value of the connected node ratio variable (*CNRH*) in the sample is 0.85. Increasing the  
 42 coefficient of this variable from 4.634 to 5.634 brings the prediction up to 35.1 bicycles, an increase of  
 43 134 percent. Another unit increase in the coefficient raises the prediction to 82.4, more than five times  
 44 the original estimate. Along the same lines, examining the median values in Model B, a 20 percent  
 45 increase or decrease in the median value for SlopeH changes the predicted counts by -3 percent and 4  
 46 percent, respectively. Identical modifications to the median value for *NComPropT* change the predictions  
 47 by 36 percent and -26 percent, respectively. This sensitivity of the coefficients should be considered when  
 48 interpreting models.  
 49

## 1 **CONSIDERATIONS FOR FUTURE RESEARCH**

2 The models presented in this paper are preliminary and are not intended to be a final product. They  
3 represent a first attempt at modeling bicycle volumes in one county in Northern California. Bicycle  
4 activity in other communities is not expected to have an identical relationship to the built environment.  
5 Additional models in other states and countries are needed to provide a more complete picture of how  
6 surrounding land use, transportation system, and socioeconomic characteristics are related to bicycle  
7 volumes at specific locations. One of the next steps for refining the Alameda County models is to collect  
8 counts at additional locations and conduct model validation.

9 Ideally, the models would have been estimated using count data that was collected during the  
10 same time-of-day at each location. In this case, however, counts were taken during different times of the  
11 day on weekdays and Saturdays. In theory, the number of bicyclists passing through an intersection  
12 between 12 p.m. and 2 p.m. on a given day is likely to be different than the number of bicyclists passing  
13 through between 4 p.m. and 6 p.m. To account for these time-of-day differences, dummy variables  
14 indicating the count time period were included in several versions of the models. However, the  
15 coefficients of these variables were highly insignificant, indicating that controlling for time of day  
16 provided little additional value for predicting bicycle volumes at the sample of 81 intersections.  
17 Therefore, the time-of-day variables were not included in the models. Prior research on pedestrian  
18 volumes in Alameda County dealt with this issue by using several months of automated pedestrian  
19 counter data from several locations to extrapolate two-hour counts to weekly counts (21). Unfortunately,  
20 since automated bicycle counters must be installed in a roadway, it is more difficult to collect counts from  
21 a variety of locations or to capture both directions of a roadway. The available automated bicycle count  
22 data from two locations in Alameda County could not be used for extrapolation without making gross  
23 assumptions.

24 Weather effects were not incorporated into the final bicycle volume models. Dummy variables  
25 indicating extreme temperatures and cloud cover proved highly insignificant in all model specifications.  
26 This is likely due to the lack of variation in weather during the bicycle count periods. The count data  
27 were collected in the spring months, and counts were rescheduled if it was raining. Alameda County has  
28 a temperate climate, and there were only two counts on days where the temperature was greater than 90°  
29 F (32° C) or less than 50° F (10° C). Weather is expected to show a greater effect on bicycle activity in  
30 areas where the climate varies more.

31 Previous studies have found a significant relationship between the socioeconomic characteristics  
32 of individuals and neighborhoods and likelihood of bicycling (14,17,25). In Alameda County, there was a  
33 moderate correlation ( $0.3 < |\rho| < 0.5$ ) between the bicycle counts and socioeconomic variables such as  
34 surrounding neighborhood income (positive), percentage of rental housing (positive), and percentage of  
35 residents under age 18 (negative). However, these variables were less significant than the built  
36 environment variables in the final models, so they were not included.

37 In Alameda County, most retail and employment attractors are in the flatter areas, although the  
38 hills are popular with recreational riders. A different set of study intersections could have included more  
39 counts along popular recreational routes in the hills. This would likely have produced different model  
40 results, particularly for the slope and proximity to retail variables. However, the 81 intersections were  
41 selected to ensure representation from areas with different surrounding land uses, transportation  
42 infrastructure, and neighborhood socioeconomic characteristics. Intersections were not chosen with the  
43 specific intention to capture high bicycle volumes or popular bicycle routes.

44 The bicycle count dataset included only one zero-count occurrence. As mentioned above, this  
45 value was changed to 0.1 so that the natural log could be computed. Since the motivation behind these  
46 models is to provide reasonable estimates for planning purposes, rather than absolute accuracy, it is  
47 acceptable to assume that a prediction of 0.1 represents no bicycle activity. This approach was expected  
48 to have little effect on the models because it only applied to one count in the sample. Study areas with  
49 less bicycle activity and multiple zero-count intersections would require a different model structure.

50 The independent variables included in the final models show statistically significant associations  
51 between specific land use and transportation infrastructure characteristics and bicycle volumes. However,

1 they do not necessarily imply a causal relationship between any of the independent variables and levels of  
2 bicycling. For example, adding a bicycle facility to an intersection approach may make the roadway a  
3 more attractive place to ride, which could increase the bicycle volume at the intersection. However,  
4 communities often add bicycle facilities to roadways that already have high bicycle volumes in order to  
5 make conditions more comfortable for existing bicyclists. In this case, a high intersection bicycle volume  
6 would precede the bicycle facility. Accordingly, these models are not intended to predict the change in  
7 bicycle volume after installation of a bicycle facility. Instrumental variable methods could improve the  
8 models by accounting for the potential endogeneity of the bicycle facility variable.  
9

## 10 CONCLUSION

11 The preliminary models presented in this paper are simple tools that can be used to estimate bicycle  
12 intersection counts during specific time periods. The analyses performed here contribute to the body of  
13 research on relationships between the built environment and bicycle activity. In particular, the models  
14 showed that bicycle volumes tended to be higher at intersections:

- 15 • Surrounded by more commercial retail properties within 1/10 mile of the intersection
- 16 • Closer to a major university
- 17 • With a marked bicycle facility on at least one leg of the intersection
- 18 • Surrounded by flatter terrain within 1/2 mile of the intersection
- 19 • Surrounded by a more connected roadway network

20 In addition, the models showed several important differences between weekday and weekend intersection  
21 bicycle volumes. The models showed that:

- 22 • The positive influence of commercial retail establishments on bicycle volumes tended to be  
23 greater on weekdays than on weekends
- 24 • The positive influence of proximity to UC Berkeley tended to be greater on weekdays than on  
25 weekends
- 26 • The positive influence of bicycle facilities on bicycle volumes tended to be greater on weekends  
27 than on weekdays
- 28 • The negative influence of hilly terrain tended to be smaller on weekends than on weekdays

29 Further refinement and testing in other study areas is necessary to improve count predictions in  
30 Alameda County or elsewhere. Bicycle volume counts are an important tool to help understand bicycle  
31 crash risk and the impact of the built environment on travel. It is crucial to understand the level of bicycle  
32 activity for bicycles to have a greater impact on funding and policy decisions.  
33

## 34 ACKNOWLEDGEMENTS

35 The authors would like to thank Lindsay Arnold and David Ragland of UC Berkeley Safe Transportation  
36 Research and Education Center (SafeTREC) for their assistance with the project. Special thanks to  
37 Population Research Systems for collecting bicycle counts in Spring 2008 and to all of the volunteers  
38 who collected bicycle counts in Spring 2009. This work was partially funded by the California  
39 Department of Transportation and the Alameda County Transportation Improvement Authority.  
40

## 41 REFERENCES

- 42
- 43 1. United States Census Bureau. *2006 American Community Survey 1-year Estimates*. Available online:  
44 <http://factfinder.census.gov>, 2007.
- 45
- 46 2. United States Census Bureau. *2008 American Community Survey 1-year Estimates*. Available online:  
47 <http://factfinder.census.gov>, 2009.
- 48
- 49 3. City of Portland, OR. *Portland Bicycle Counts 2008*. Available online:  
50 <http://www.portlandonline.com/shared/cfm/image.cfm?id=217489>, October 2008.

- 1  
2 4. New York City Department of Transportation. *NYC Commuter Cycling Indicator*. Available online:  
3 [http://www.nyc.gov/html/dot/downloads/pdf/commuter\\_cycling\\_indicator\\_and\\_data\\_2009.pdf](http://www.nyc.gov/html/dot/downloads/pdf/commuter_cycling_indicator_and_data_2009.pdf), April  
4 2009.  
5
- 6 5. City of Seattle Department of Transportation. 2009 Downtown Bicycle Count Results. Available  
7 online: <http://www.cityofseattle.net/transportation/bikeinfo.htm>, 2009.  
8
- 9 6. San Francisco Municipal Transportation Agency. *2008 San Francisco State of Cycling Report*.  
10 Available online: <http://www.sfmta.com/cms/bnews/documents/2008SFStateofCyclingReport.pdf>, 2008.  
11
- 12 7. Turner, S., G. Shunk, and A. Hottenstein. *Development of a Methodology to Estimate Bicycle*  
13 *and Pedestrian Travel Demand*. Texas Transportation Institute, Research Project Number 0-  
14 1723, Report 1723-S, Texas Department of Transportation, Federal Highway Administration,  
15 September 1998.  
16
- 17 8. Landis, B.W. The Bicycle System Performance Measures: The Intersection Hazard and  
18 Latent Demand Score Models. *ITE Journal*, Vol. 66, No. 2, 1996, pp. 18-26.  
19
- 20 9. City of Alexandria, VA. *Alexandria Pedestrian and Bicycle Mobility Plan*. Available online:  
21 <http://alexandriava.gov/localmotion/info/default.aspx?id=11418>, June 2008.  
22
- 23 10. Purvis, C. *Incorporating Effects of Smart Growth and TOD in San Francisco Bay Area Travel*  
24 *Demand Models: Current and Future Strategies*. Metropolitan Transportation Commission,  
25 Available Online:  
26 [http://www.mtc.ca.gov/maps\\_and\\_data/datamart/research/Incorporating\\_Smart\\_Growth\\_MTC\\_](http://www.mtc.ca.gov/maps_and_data/datamart/research/Incorporating_Smart_Growth_MTC_models.pdf)  
27 [models.pdf](http://www.mtc.ca.gov/maps_and_data/datamart/research/Incorporating_Smart_Growth_MTC_models.pdf), 2003.  
28
- 29 11. Haynes, M. and S. Andrzejewski. GIS Based Bicycle & Pedestrian Demand Forecasting Techniques.  
30 Travel Model Improvement Program Webinar. Available online:  
31 [http://tmip.fhwa.dot.gov/sites/tmip.fhwa.dot.gov/files/fehr\\_and\\_peers.pdf](http://tmip.fhwa.dot.gov/sites/tmip.fhwa.dot.gov/files/fehr_and_peers.pdf), April 29, 2010.  
32
- 33 12. Jones, M., S. Ryan, J. Donlon, L. Ledbetter, D.R. Ragland, and L. Arnold. *Seamless Travel:*  
34 *Measuring Bicycle and Pedestrian Activity in San Diego County and its Relationship to Land Use,*  
35 *Transportation, Safety, and Facility Type*. Caltrans Task Order 6117. California Department of  
36 Transportation, 2010.  
37
- 38 13. McCahill, C. and N.W. Garrick. The Applicability of Space Syntax to Bicycle Facility Planning. In  
39 *Transportation Research Record: Journal of the Transportation Research Board*, No. 2074,  
40 Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 46–51.  
41
- 42 14. Dill, J. and K. Voros. Factors Affecting Bicycling Demand: Initial Survey Findings from the Portland,  
43 Oregon, Region. In *Transportation Research Record: Journal of the Transportation Research Board*,  
44 No. 2031, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 9–17.  
45
- 46 15. Aultman-Hall, L., F.L. Hall, and B.B Baetz. Analysis of Bicycle Commuter Routes Using Geographic  
47 Information Systems: Implications for Bicycle Planning. In *Transportation Research Record: Journal of*  
48 *the Transportation Research Board*, No. 1578, Transportation Research Board of the National  
49 Academies, Washington, D.C., 1997, pp. 102–110.  
50

- 1 16. Dill J. and J. Gliebe. Understanding and Measuring Bicycling Behavior: A Focus on Travel Time and  
2 Route Choice. OTREC-RR-08-03. Oregon Transportation Research and Education Consortium,  
3 2008  
4
- 5 17. Dill J. and T. Carr. Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them,  
6 Commuters Will Use Them. In *Transportation Research Record: Journal of the Transportation Research*  
7 *Board*, No. 1828, Transportation Research Board of the National Academies, Washington, D.C., 2003,  
8 pp. 116–123.  
9
- 10 18. Stinson, M.A. and C.R. Bhat. Commuter Bicyclist Route Choice: Analysis Using a Stated Preference  
11 Survey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1828,  
12 Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 107–115.  
13
- 14 19. Xing, Y., S.L. Handy, and T.J. Buehler. Factors Associated with Bicycle Ownership and Use: A  
15 Study of 6 Small U.S. Cities. Presented at 87th Annual Meeting of the Transportation Research Board,  
16 Washington, D.C., 2008.  
17
- 18 20. W.W. Hunter, L. Thomas and J.C. Stutts. *BIKESAFE : Bicycle Countermeasure Selection System*.  
19 FHWA-SA-05-006. FHWA, U.S. Department of Transportation, 2006.  
20
- 21 21. Schneider R.J., L.S. Arnold, and D.R. Ragland. Pilot Model for Estimating Pedestrian: Intersection  
22 Crossing Volumes. In *Transportation Research Record: Journal of the Transportation Research Board*,  
23 No. 2140, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 13–  
24 26.  
25
- 26 22. Schneider R.J., M.C. Diogenes L.S. Arnold, V. Attaset, J. Griswold, and D.R. Ragland. Association  
27 between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California.  
28 Forthcoming, *Transportation Research Record: Journal of the Transportation Research Board*, 2010.  
29
- 30 23. City of Berkeley Transportation Division. Bicycle Boulevards in Berkeley. Available online:  
31 <http://www.ci.berkeley.ca.us/ContentDisplay.aspx?id=6650>, 2010.  
32
- 33 24. City of Oakland Public Works Agency. City of Oakland Pedestrian and Bicycling Information.  
34 Available online: <http://www.oaklandpw.com/Page122.aspx>, 2010.  
35
- 36 25. Xing, Y., S.L. Handy, and T.J. Buehler. Factors Associated with Bicycle Ownership and Use:  
37 A Study of 6 Small U.S. Cities. University of California, Davis, Available online,  
38 [http://www.des.ucdavis.edu/faculty/handy/TTP\\_seminar/Bike\\_Draft\\_11.14.pdf](http://www.des.ucdavis.edu/faculty/handy/TTP_seminar/Bike_Draft_11.14.pdf), 2008.  
39

## **LIST OF TABLES AND FIGURES**

### **TABLES**

Table 1. Factors Associated with Bicycling in Previous Research

Table 2. Summary Statistics for all Two-Hour Bicycle Counts

Table 3. Description of Selected Explanatory Variables Considered in the Modeling Process

Table 4. Alternative Bicycle Model Specifications

### **FIGURES**

Figure 1. Map of Manual Bicycle Count Locations