but in terms of "transitions" from one space to another. The approach requires coding of a detailed network, which is then treated as a "graph." Topological methods are used to characterize the properties of the network (graph) through such measures as connectivity (number of other nodes that connect to each node), depth (average number of steps between nodes), and integration (ease of access from other nodes). Integration is the key variable, whose formula compares an ideally connected graph with the one in question to determine a measure of accessibility for each node in the network. The quantified measures of accessibility and connectivity are then used to generate movement "potentials," which are then correlated with counts. The correlations are then used to predict volumes on a street-by-street basis for the defined study area.

Illustrative tests of Space Syntax in the United States have occurred in the City of Oakland, CA, for pedestrian planning (Raford and Ragland, 2003) and in relation to bicycle travel in Cambridge, MA (McCahill & Garrick, 2008). In the McCahill & Garrick example, the correlation of Space Syntax measures and observed bicycle volumes in the Cambridge, MA, bicycle network was tested. The "choice" segment indicator was used as the means of predicting relative cyclist volumes on facilities, using road centerline maps in place of the traditional "axial maps," and ArcGIS to compile information on segments from spatial analysis and census statistics. A linear regression was developed to reveal the best correlation between existing bike volume counts at 16 intersections, census population, and employment data to serve as productions and attractions, plus various Space Syntax measures. The researchers determined that the method was useful in predicting bike volumes in a network and could be useful in designing more efficient networks.

In the City of Oakland, Raford and Ragland used Space Syntax to forecast pedestrian volumes for safety analysis in the City's pedestrian master plan. Space Syntax was used to leverage existing count data from a sample of 42 intersections into forecasts of pedestrian volumes at 670 intersections citywide. However, because Space Syntax assumes an even population distribution, the researchers supplemented the model by using Census population and employment data to allow for distortions caused by major generators. Discrepancies in forecasting accuracy (remaining after the adjustments) included a tendency to underestimate volumes on high-volume streets and on streets connecting to three Bay Area Rapid Transit (BART) stations. However, the researchers believed that additional enhancements (e.g., including auto volumes and speeds and using more specific land use characteristics) could help improve accuracy.

Because of the lack of clarity in how Space Syntax works and that it is proprietary, it has not been possible to fully evaluate Space Syntax's capabilities, so it is not included in the best-practice recommendations. However, users can investigate further if the features of the tool seem interesting or useful.

Direct Demand Models

Direct demand models have been the accepted practice for estimating pedestrian or bicycle facility demand for some time. The NCHRP Project 08-78 background review recorded use of these methods back in the 1970s (Benham & Patel, 1977). Their structure is to explain observed levels of bicycle or pedestrian activity on facilities (links) or at intersection (points) as recorded through counts, using a range of factors that describe local context. This is usually done using regression modeling techniques, with the calibrated models then applied back on all or a subset of the sampled system of intersections or links to assess their accuracy in replicating choices.

Variables often used to represent context in these types of models include the following:

- Population or employment densities, sometimes differentiated by type (e.g., populations differentiated by age, gender or income, or employment categorized as office or retail).
- Population or employment activity levels within a nominal buffer distance of ¼ or ½ mile from the intersection.
- Land use mix, measured either through an index (e.g., entropy) or implicitly through corresponding buffered activity levels.
- Characteristics of the facility, including type of bike path and sidewalk existence and sufficiency.
- Interaction with vehicle traffic (e.g., adjacent speeds or volumes, intersection approaches with crosswalks, sidewalk widths, on-road versus off-road bike facilities).
- Transit availability (e.g., transit frequency and stop density).
- Major generators (e.g., proximity to universities, schools, recreation, neighborhood shopping, major transit centers, and civic centers).

Numerous examples of models in this genre are cited in Table 4-2 and documented in Appendix 7 of the Contractor's Final Report under the Aggregate Demand Methods discussion. Because each is unique, it is difficult to name one or two that are exemplary; however, among those that have undergone the most development and had access to the best data resources are the Seamless Travel pedestrian and bicycle models developed by Alta Planning & Design in San Diego (Jones, et al., 2010) and the Santa Monica pedestrian and bicycle demand models (Fehr & Peers, 2010).

Seamless Travel Models

In the Seamless Travel study, pedestrian and bicycle models were developed to predict approach volumes at intersections during the 7 to 9 A.M. period on weekdays. Manual counts from a sample of 80 intersections supported the analysis. Counts were supplemented with traveler intercept surveys at 25 locations to obtain additional data, although the surveys did not identify the type of trip in progress.

The Seamless pedestrian model is of the following form:

$$P_{AM} = 1.555 + 0.723\ ED + 0.526\ PD - 1.09\ R (R^2 = 0.516)$$

where
P_{AM} = Morning peak pedestrian count
ED = Employment density within 0.5 mile
PD = Population density within 0.25 mile
R = Presence of retail within 0.5 mile

So the model predicts that A.M. peak-period walk trips will increase in proportion to adjacent employment and population density and decrease in the presence of retail activity. Even though these are probably work-related trips, given the time of day, it is not immediately clear why retail activity would have a negative effect on walk trip levels. Employment density carries a higher coefficient than population density, again presumably related to these being primarily work trips, although the buffer radii are different for population and employment and elasticities were not provided.

The Seamless bicycle model has the following form:

$$B_{AM} = -4.279 + 0.718\ C + 0.438\ ED\ (R^2 = 0.439)$$

where
B_{AM} = Morning peak bike trips
C = Footage of Class I bicycle path within 0.25 mile
ED = Employment density within 0.25 mile

This bicycle model predicts an increase in bike trips based on higher employment density and greater presence of Class 1 bikeways within ¼-mile of the count site.

Santa Monica Models

The pedestrian and bicycle models developed by Fehr & Peers for Santa Monica predict volumes for the 5 to 6 PM peak hour. The pedestrian model has the following form:

$$P_{PM} = 222.18 + 0.00321\ ED + 3.675\ BF_{PM} + 82.695\ SDP - 0.00685\ DO - 5.699\ SL\ (R^2 = 0.584)$$

where
P_{PM} = Evening peak pedestrian volume
ED = Employment density within ⅓ mile
BF_{PM} = PM bus frequency
SDP = Intersection is within shopping district
DO = Distance from ocean
SL = Average speed limit on approaches

This equation predicts that PM peak-period walk trips will increase in proportion to adjacent employment, with higher rates of PM bus service, and if the intersection lies within a shopping district. This equation predicts that PM peak-period walk trips will decline with increased distance from the ocean and with higher adjacent auto speeds. In contrast to the Seamless Travel pedestrian model, this model sees a positive effect from retail proximity, which may be due to a higher proportion of non-work trips occurring during the PM peak.

The Santa Monica bicycle model has the following form:

$$B_{PM} = 1.317 + 0.120\ Ln\ ED + 1.632\ MXD + 0.431\ BN + 0.523\ INT\text{-}4\ (R^2 = 0.401)$$

where
B_{PM} = Evening peak hour bike trips
Ln ED = Log of employment density within ⅓ mile
MXD = Land use mix within ⅓ mile
BN = Proximity to bike routes (intersection is along a bike route or at the intersection of two bike routes, with higher weighting going to better classes of bike facilities)
INT-4 = Four-legged intersection

This equation predicts an increase in bike trips based on higher employment density, mixed land use, proximity to bike routes, and if the intersection is four-way.

The appeal of these models lies in their simplicity and custom quality. Although not easy to construct, they do not require advanced transportation modeling skills and are fairly easy to understand and apply. Aside from the activity counts, most of the data used to construct the context variables are generally available, and model builders are often resourceful in designing the models to use the data that they have.

The caveat with these models is that they trade directness and simplicity for behavioral structure. In effect, they try to explain/predict an aggregate quantity—activity counts in a particular time period—with factors descriptive of the surrounding environment. What results are relationships that may display strong correlations with the activity variable, but cannot be readily shown to "cause" the behavior represented in the counts (which is itself an amalgam of travel activity).

What the NCHRP Project 08-78 research has shown is that accessibility is the most significant determinant of choice, particularly for non-motorized travel, and representing accessibility requires a deliberate effort to simultaneously account for both the opportunities presented through the land use and the ease and efficiency with which the modal networks connect the traveler with these opportunities. It is difficult to apply this relationship in count-based models given that the modeled intersection or link is neither a trip production nor attraction.

Therefore, this guidebook suggests that use of these models should be judicious in how they are developed and when they are used. The following guidelines are suggested:

1. None of these models should be construed as transferrable. Their coefficients are unique to how the models have been

specified (variables included) and the specific location for which they were developed. If an existing model presents an appealing structure, the user is advised to re-estimate the model(s) using identical data for the new study area.

2. The user needs to be aware of the uncertainties associated with modeling "count" data. In almost all cases, the models are blind to the travel behavior represented by the counts (e.g., the purpose of the trip, the sociodemographic characteristics of the traveler, the origin-destination of the trip, and the existence of alternatives). Focusing the counts and models on a particular time period (e.g., A.M. weekday peak for work or mid-day weekend for recreation) can narrow the uncertainty as to the types of trips being observed, but, for other time periods, the mix of trips being modeled may be difficult to surmise.
3. Once the models are calibrated, the user should test their reliability in predicting activity at individual locations and overall for the study area. Although most of the models reviewed have R^2 values of 0.5 or better, they may not be particularly accurate at the level of the individual intersection or link. The Seamless Travel study experimented with methods to adjust the base estimates to account for unusual circumstances (that cannot be directly included in the model), and it may prove worthwhile to review and consider emulating these methods (see http://www.altaplanning.com/caltrans+seamless+study.aspx).
4. Be judicious in the types of applications or decisions to be supported by the models. For example, if measures of network connectivity are not included in the model structure, it would be misleading to estimate demand for a new or improved facility without recognizing that some portion of the new demand predicted may simply be a diversion from some other facility. At the same time, a network improvement that contributes to overall network connectivity may well induce new travel on other portions of the network.

Given the above, it is recommended that the direct demand tools be reserved for either quick estimates or screening in advance of more comprehensive analysis, or for incremental extrapolations from an existing situation. Regardless, the forecast effort should be within the bounds of the explanatory variables in the model and not be used for forecasting new demand or changes within a network. For these types of applications, the user is advised to apply one of the earlier choice-based tools (e.g., the GIS-Accessibility, MoPeD, PedContext, or even the Portland Pedestrian model approach).