A calibrated labeling method for generating bicyclist route choice sets incorporating unbiased attribute variation

Joseph Broach
PhD Student
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
jbroach@pdx.edu
phone: 971-285-7045
fax: 503-725-8480
corresponding author

John Gliebe
Assistant Professor
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
gliebej@pdx.edu
phone: 503-725-4016
fax: 503-725-8770

Jennifer Dill
Associate Professor
Nohad A. Toulan School of Urban Studies and Planning
Portland State University
PO Box 751
Portland, OR 97207-0751
jdill@pdx.edu
phone: 503-725-5173
fax: 503-725-8770

Paper submitted for presentation and publication at the 89th Annual Meeting of the Transportation Research Board January 10-14, 2010, Washington D.C.

Original submission date: August 1, 2009
Revised submission date: November 15, 2009

Word count: 6210 (text) + 6 x 250 (figures and tables) = 7710

ABSTRACT

Discrete choice model estimation requires specifying the alternatives actually considered for each observed choice. In route choice problems based on real world travel observations, generally only the chosen route is observed, and the rest of the choice set remains hidden from the analyst. In dense travel networks thousands of paths may connect a given origin or destination, necessitating methods for generating a reasonable subset of options.

A new method is proposed for generating deterministic (non-random) route choice alternatives. The technique modifies the labeled routes method, in which multiple criteria are optimized individually to generate attractive routes. The proposed method offers two potential improvements. First, instead of one optimum route per criterion label, a set of optimal routes is generated by allowing a sensitivity parameter to vary. Second, a calibration step fits the alternative shortest path deviation distribution to observed behavior. The resulting process is more flexible than traditional labeled routes while maintaining strong links to behavior and reducing the potential for attribute bias.

The proposed calibrated labeling method is applied to bicyclist route choice in a dense, urban network. Results suggest that the proposed technique outperforms existing methods on several key criteria. In addition, explicitly linking choice set generation to observed travel patterns creates a more intuitive behavioral link than many existing strategies. The proposed method should be immediately useful for route choice modeling in similar contexts. Furthermore, the basic framework could be more broadly applicable for choice set generation.

INTRODUCTION

Random utility-based discrete choice models generally require two steps to estimate. First, a choice set must be specified. The choice set may consist of every possible alternative (the universal set) or a subset of available options. Choice sets may be identical across all observations, or they may vary by observation, individual, or group. Once a choice set with at least two options has been specified for each observation, the choice model is then estimated conditional on the choice set specification.

Typically, much less attention is placed on the choice set formation step, even though model estimation depends crucially on the considered alternatives' accuracy. In transportation modeling, route choice problems have spurred the most innovation in choice set formation theory and practice. Route choice problems bring choice set generation front and center. Even a moderately dense travel network defines hundreds if not thousands of alternative routes connecting each origin and destination. In fact, the universal set is usually not even enumerated, instead remaining implicit in the network structure. A so-called master set stands in for the intractable universal and in part or whole constitutes the consideration set for each observation.

The consideration set would ideally contain all possible routes with a non-zero probability of being chosen. Put another way, it should exclude all those routes which are dismissed without their attributes being fully weighed. The non-compensatory process implied is not compatible with a utility framework, which has a basic assumption that attributes are traded off with one another. As an example, a potential route which is dismissed solely because it is too long should not be included in a consideration set. If it were to be included, it would supply false information about the contribution of all other attributes in the choice process.

In practice, choice sets are typically unobservable, and ideal consideration sets are out of reach. Instead, efforts focus on minimizing biases as much as possible by specifying reasonable choice sets consistent with what data are available. Computational costs are also important in practice as some generation algorithms may take days to run even for relatively small networks (1). Finally, it seems desirable to limit the number and importance of choice set generation parameters exogenously fixed by the researcher.

Current route choice set generation techniques tend to focus almost to a fault on the replication of observed routes. It cannot be overemphasized that the goal of choice set generation is to reproduce the actual choices considered as faithfully as possible. For estimation purposes, it is the composition of the entire choice set that matters, not only—and perhaps not even most importantly—that the observed route is "covered." This is the point of departure for the technique developed in this paper.

In the process of developing a bicyclist route choice model from observed behavior over a large, dense urban network, existing techniques were found inadequate for generating reasonable choice sets. A new generation process was developed by adapting and extending the labeled routes technique. The proposed technique seeks to make behavioral and parameter assumptions more explicit. An important contribution is the incorporation of a calibration step that reduces the generation procedure's sensitivity to analyst input, reducing the chance of introduced bias and making the generation step more readily transferable to other contexts. A secondary contribution is the generation of reasonable choice sets for non-recreational urban bicycling trips, which pose special challenges for alternative selection.

The remainder of the paper is organized as follows. Section 2 provides a brief review of the literature on route choice set generation and validation techniques. Section 3 describes a proposed extension of the labeled routes method. Section 4 applies the proposed method and three other techniques to observed bicyclist route choice data in Portland, Oregon. Section 5 concludes.

REVIEW OF ROUTE CHOICE SET GENERATION TECHNIQUES

Bovy (2) provided a recent comprehensive review of route choice set generation theory and techniques. He usefully divided the various methods into three main types: path search, constrained enumeration, and probabilistic. The technique proposed in this paper falls in the path search category. Bovy (2) also divided choice set generation into two distinct steps. In the first step, a master set is generated from the universal set. In this step some subset of all the routes between a given origin and destination pair is discovered by a specified algorithm. In the second step, the master set is reduced by prescribed filters to form a consideration set. The consideration set is presumed to include all routes that the traveler fully considers in the final choice process. Different techniques place different emphasis on the discovery versus the filtering step. The proposed method focuses on discovery.

Repeated Shortest Path Search Methods

The most common route choice set generation techniques involve repeated least cost path calculations. In successive iterations, a vector of link costs is modified such that multiple least cost paths may be found. Cost variable values, the variables themselves, or cost function parameters may be altered between searches. The changes applied may be systematic or stochastic. When systematic changes are applied, the resulting choice set is said to be deterministic in the sense that the alternatives generated are fixed. When stochastic changes are applied, the choice set produced will vary from run to run and will therefore be more sensitive to the number of iterations specified. In general, advantages of path search techniques include the potential to use common network software, reasonable computation cost, and relatively straightforward links to behavior.

Among deterministic path search techniques, *K* shortest paths and labeled routes are most common. The *K* shortest paths approach may be executed by finding the exact *K* least cost paths according to some criterion; however, the exact approach is seldom used as it tends to generate many minor variations on the shortest path (3). Instead, the *K* shortest paths problem is usually solved by either eliminating (“link elimination”) or penalizing (“link penalty”) links in previously selected paths. The link penalty approach is usually preferred, since it maintains network continuity (3). Although distance and travel time are the most common criteria, other network variables have been used (see, for example, 4). From a behavioral perspective, the idea that choice sets are anchored to a least-cost path seems plausible for most trips. Potential bias is introduced by the analyst’s choice of a stop point. In a dense network, it will be unreasonable to search exhaustively, and so *K* must be set to some arbitrary upper bound. In the link penalty approach, the analyst must also specify the incremental penalty, which in practice has been somewhat arbitrary. Perhaps more troubling, no rationale is given in this method as to why the *K*th path might be attractive and unique enough for the traveler to consider it a reasonable option. Finally, the method leaves no room for traveler heterogeneity either in tastes or network knowledge. The size and composition of each choice set is fully determined by the choice of *K* and the network composition. In theory, *K* could vary across travelers or choice situations, but in practice it has not. Ramming (3) did provide an example of varying the link penalty spatially.

One way of addressing the lack of route heterogeneity that arises with the *K*-shortest path approach is to introduce random disturbances to link costs (3,5). A single criterion is used, but for each draw the vector of link costs is simulated by drawing disturbances from a specified distribution. The analyst selects the distribution, distribution parameters, and the number of draws. Ramming (3) suggested that a behavioral rationale for the randomized link costs may lie in the misperception of link attributes by the traveler. Imperfect measurement or actual variability (e.g. of travel time due to congestion or incidents) could be additional rationales. To date, it is unclear whether and how any of these phenomena map to a probability distribution. Existing studies have chosen normal distributions out of convenience with standard deviation parameters based on observed travel time perception errors (3,5). An inherent feature of the simulation technique is that the probability of drawing a particular route or set of routes is unknown (6). Ideally, the random utility model estimated in the second modeling step would be adjusted to account for the probability of the generated choice set, but there is no obvious way to do so. In terms of replicating observed paths, Bekhor et al. (5) reported that simulated paths perform about equally to the deterministic link penalty approach. Computational costs were reported to be magnitudes lower for the simulation, but the application performed for this paper fails to corroborate that finding (5). Finally, there remains the question of whether the generated alternatives are realistic and whether they are sufficiently distinct from one another to be recognized.

Representing a different approach, labeled routes is a deterministic path search technique that addresses some of the behavioral criticisms of the *K* different paths approach. The method of labeled paths was first developed in Ben-Akiva et al. (7). Rather than multiple iterations using a single criterion, in labeling the criterion itself changes. The number of generated alternatives is equal to the number of criteria, or labels where each labeled path optimizes a criterion. The labels are based on network features believed important to travelers, including attributes such as distance, delays, road hierarchy, and even scenery.

The behavioral rationale for labeled routes is appealing. Faced with a large number of possible routes and costly information search, travelers initially screen routes along single dimensions. From the set of "best" candidate routes (e.g. shortest, quickest, straightest, fewest delays, etc.), travelers make more detailed comparisons among a more manageable consideration set. A specification problem immediately arises. First, the truly optimal route for a given criterion may involve an unreasonable detour. For this reason, labels other than distance or travel time are actually specified as parameterized cost functions including distance or time. Ben-Akiva et al. (7) explicitly set the parameters of these cost functions to maximize the reproduction of observed routes, although they acknowledge that solving for an exact solution is not feasible. In other examples, the cost function parameters are specified without further explanation (3,5). A second problem is the relatively limited number of alternatives generated. Limited data and limited knowledge of traveler tastes combine to restrict the number of labels. Existing applications have specified from four to sixteen label functions; however, in many cases labels will generate identical or very similar routes that must be discarded (1,3,5,7). For some routes no alternative to the observed route will be discovered. Finally, although in theory choice situation or traveler heterogeneity could be incorporated by varying labels or label function parameters, in practice this has not been done. The technique proposed in the present paper draws heavily from the labeling approach while attempting to improve on the limitations just mentioned.

Other techniques

Constrained enumeration techniques address concerns about route feasibility, attractiveness, and distinctness (1). In this method, all paths between a given origin and destination are discovered subject to a series of constraints. Prato and Bekhor (1), for example, use a branch and bound algorithm that includes parameterized constraints on direction, travel time, detours, similarity to other paths, and left turns.

Probabilistic route generation perhaps comes closest to the theoretical ideal and is most integrated with subsequent choice models. Few applications to real choice problems have been carried out, and it remains to be seen whether they are practical for such applications. Examples include Implicit Availability Perception (IAP) models (8), and the recently proposed importance sampling for route choice technique (6).

Combining Methods

Alternative paths generated by different methods can and have been combined to form a more diverse choice set. Bekhor et al. (5) combined labeled routes with simulated shortest paths. Prato and Bekhor (1) merged labeled routes, link elimination, link penalty, and simulated shortest paths. The improved coverage and variety of the combined sets may be offset by murkier behavioral foundations.

Validation Techniques

There has been relatively little discussion of how to judge choice set quality or how to compare sets generated with different parameters or methods. Really, there are few compelling metrics to use. Occasionally, elicited consideration sets may be gathered to help verify generation techniques, but this is burdensome, costly and therefore rare in route choice data (but see 1,8).

Replication of observed routes is by far the most common measure of choice set generation performance. Results are usually given as the percentage of observed routes “covered” by generated routes given overlap thresholds of 70 to 100 percent (1,3,5,7,9). Bekhor and Prato (10) generalized the overlap measure with their efficiency index. The efficiency index takes into account the number of routes generated in excess of the number of relevant routes. This at least penalizes inefficient search algorithms, but the number of relevant routes must still be defined.

While reproducing observed routes is a necessary condition for reproducing a consideration set, it is not a sufficient one. After all, it is the composition of the full choice set that matters, not only that the observed route is included. For example, a complete enumeration scheme would always maximize observed route overlap, but the resulting choice sets would be of poor quality in most cases. Choice set size, and the distribution of sizes, should also be considered. Smaller is not necessarily always better, of course. A choice set generation scheme consisting of the observed route plus one variant would maximize overlap and minimize choice set size, but this probably would not provide the attribute variation needed for estimating an informative model. Thus, sufficient and preferably unbiased attribute variation is another important quality. Finally, though perhaps most difficult to quantify, generated routes should be reasonable. One definition of reasonableness might be that sufficiently improbable routes are excluded. For example, routes inferior to others on all attributes, or paths in the far tail of the chosen route distribution on any specific attribute might be considered unreasonable. The choice set generation technique proposed in the present paper attempts to consider the full range of choice set performance criteria detailed above.

Other possible criteria include computational cost, model estimation performance, and prediction accuracy. Computational cost has great practical importance, but comparisons are difficult. In general, and this paper is no different, modelers will have spent much more time optimizing the routine they are presenting and much less optimizing the comparison routines, which may be straw men. Also, different algorithms will scale differently with increasing network density and sample size, and some will require significant post-processing. The need to create custom software could also be considered a computational cost. Still, it is worth reporting program run times to at least provide an upper bound on expected performance.

Model estimation performance is of course the overarching goal, but one must be careful in using fit statistics to gauge choice set quality. It is easy enough to show that if one is going to add a non-chosen alternative, adding a very inferior alternative will improve fit statistics relative to adding a more reasonable one. Furthermore, adding inferior routes may vault other meaningless variables into statistical significance, but this should not be confused with an improved consideration set. Spurious routes not fully considered only add noise to the estimation; the goal is to reproduce the choice set as accurately as possible given available information. That said, the reasonableness of parameter estimates in sign and relative magnitude may provide valuable information about the appropriateness of the generated choice set.

Prediction accuracy has been used as a measure of choice set quality (9). The ability to predict of course depends strongly, and inversely, on choice set size. If prediction is the end goal, then it may be worth considering, but if attribute sensitivities are more important, predictive ability may be inappropriate.

DESCRIPTION OF PROPOSED CHOICE SET GENERATION TECHNIQUE

A modification of the route labeling choice set generation technique is proposed with two key differences. First, instead of generating a single “optimum” route for each attribute label, the proposed method generates multiple optimal paths. This result is achieved by altering the label cost function parameter. Second, the range over which the parameter varies is calibrated using the observed distribution of shortest path deviations. The resulting method puts labeling on a more even footing with more complex simulation and constrained enumeration techniques.

Behavioral Model and Assumptions

First, a behavioral rationale and associated assumptions underlying the technique need to be established. Consider a traveler facing a new trip from A to B over a dense, urban street grid. The universal choice set defined by the network may easily number in the thousands. Typically, the traveler is considered to start with this overwhelming universal set and proceed to "rule out all those alternatives that prove not to be sufficiently satisfactory or useful when considered by aspect" (2, p. 48). Instead of elimination by aspects, perhaps a more realistic way to think about the traveler's behavior is *addition* by aspects. When one looks at a map, it is to immediately find attractive routes, not to find all the unattractive ones and see what is left.

As with most methods, the assumption here is that the traveler chooses a least cost time or distance path as an anchor and searches out from it to discover paths with certain advantages over the shortest path, which of course has an obvious attraction. Advantages are assumed to be captured by specific network attributes such as stops, traffic signals, turns, and, in the case of non-motorized travel, pedestrian and bicycle facilities, slope, traffic volume, and so forth. It is assumed that as the traveler's search (whether using prior knowledge, a map, exploration, or some other means) extends out from the shortest path, candidate routes will require an increasing advantage on some other aspect to be considered relevant. Eventually, the traveler reaches a threshold where even a maximal difference on another attribute cannot offset the added distance or time, and she stops her search.

The resulting choice set includes the shortest path plus all of those paths which "distinguished" themselves on some aspect as the search progressed out. The collected routes represent a set of optimal mixes of distance/time and single attributes falling off toward some deviation constraint. In practical terms, say, for a bicyclist, a set of routes might include the shortest path with no bike facility, a longer path half on a bike facility, and a path longer still but entirely on a bike facility. None of the paths are clearly inferior to the others along the two considered dimensions. If the shortest path in the example included a bike facility for half its length, then the middle choice would no longer be relevant.

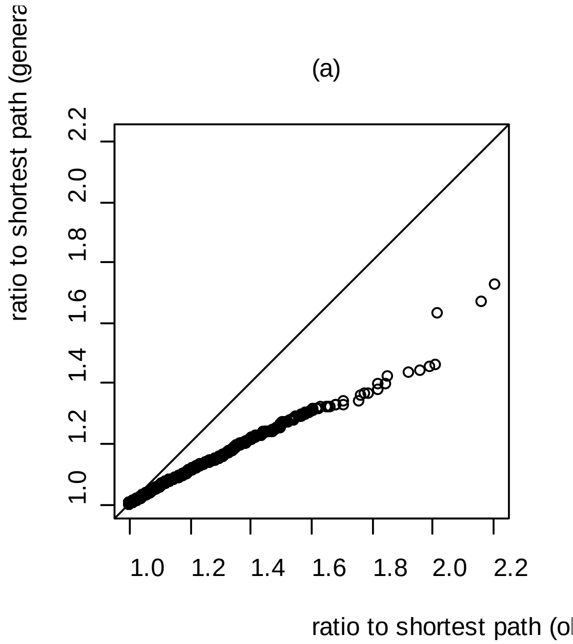
At some cost band the traveler stops searching out from the shortest path. The stop point probably varies by traveler and situation, and so it makes sense to think of it not as a constant but as a distribution of values. One obvious candidate for this distribution is the distribution of observed shortest path deviations. Certainly, it seems that a path with a commonly observed deviation is more likely to be considered than one with a deviation so great it is rarely or never observed. The assumption, and it is probably the strongest of the proposed method, is that the expected distribution of shortest path deviations in choice sets is directly related to the distribution of shortest path deviations in observed routes. Actually, it seems likely that the variance in choice set shortest path deviation would be somewhat greater than the variance in chosen routes, but an assumption that they are identical may not be far off.

Operationalizing the Behavioral Model

The proposed method was implemented as follows:

- Identify shortest paths for each origin-destination pair.
- Define a set of attribute labels based on survey responses regarding important factors, previous research results, and data availability.
- For each label, specify a label cost function of form $L_i = \beta * c_i + (1-\beta) * x_i$, where L_i is the label value of link i, β is a weighting parameter between 0 and 1, c_i is the base cost (distance or time) of link i, and x_i is an attribute cast as a disutility. For example, x_i might be the distance link i traverses *without* a bike facility.
- Set an initial minimum value for β and a step size that specifies how much β will decrease with each iteration. At the limits, β=0 returns the path that minimizes attribute x regardless of distance, and β=1 returns the shortest path.
- Starting with β = 1 – step size, minimize the label cost function, and with each iteration decrease β by the step size until β<min(β). The combination of step size and min(β) determine the maximum number of unique routes each label generates. For example, with min(β)=0.1=step size, a maximum of nine routes would be returned.
- Calibrate min(β) by fitting the generated shortest path deviation distribution to the observed distribution. Optimizing a fit statistic such as minimizing the Kolmogorov-Smirnov (K-S) test statistic could be implemented; however, fitting by eye using quantile-quantile (Q-Q) plots as shown in Figure 1 accomplishes what is needed (for the K-S test, see 11). It may be convenient to first fit the observed route deviations to a known distribution to minimize the effect of outliers and different sample sizes.
- Repeat the process for each label. Combine generated alternatives, observed routes, and, if desired, shortest paths. Duplicate and, if desired, highly overlapping routes should be filtered.

If heterogeneity is suspected in the distance/time constraint over market segments or even individuals, it is a fairly straightforward extension to incorporate it in the choice set generation process. The distribution of generated shortest path deviations are simply calibrated to the appropriate subset of observed deviations based on the segmenting scheme. Segments may be hypothesized from prior knowledge and then the segments' deviation distributions compared. Alternatively, and intriguingly, the segments could instead be suggested from the choice model estimation itself. For example, perhaps commuters are found in the choice model to be more time sensitive than non-commuters. This result could be used to segment the choice set generation, potentially improving the choice sets for a second round of estimation. Another method of introducing heterogeneity in choice set generation is to withhold certain labels for given trip types or individual characteristics. For example, perhaps labels related to traffic volume should be excluded from off peak trips, since traffic volume variation may be greatly reduced outside of peak periods. Introducing heterogeneity in these ways has not been thoroughly tested but is included here in hopes of spurring further research.


(a)
ratio to shortest path (genera
1.0
1.2
1.4
1.6
1.8
2.0
2.2
ratio to shortest path (ol

available in Dill and Gliebe (12). In addition to the proposed method, three additional choice set generation methods are included for comparison: K-shortest paths (link penalty), simulated link costs, and labeled routes. Comparisons are made based on observed path replication, choice set sizes, shortest path deviations, and choice model estimation. Constrained enumeration techniques are not represented because the basic path search algorithm could not be readily adapted to such techniques. Comparison with branch and bound search would be a natural extension for future research.

The bicycle network used is particularly challenging due to its size and density. The network consists of over 88,000 undirected links connecting almost 67,000 nodes. Within the Portland city boundary, where more than 80 percent of travel took place, network density is about 178 undirected links/km^2 (460 undirected links/mi^2). Much of the network has a dense grid layout, which makes parallel routes quite common and often only 61m (200 ft) apart. From a route choice perspective, the result is many competing routes with little variation in distance.

Applying the Proposed Calibrated Labeling Choice Set Generation Method

Label attributes were chosen based on survey questions administered to participants, existing stated preference models (13,14), and data availability. In the end, eleven labels were selected, and they are described in Table 1. Distance was chosen as the base cost variable. Casting the attributes as labels requires care to ensure that the resulting label cost function is both reasonable and easily interpretable. Discrete link attributes such as presence of a bike facility are cast as dummy variables such that x_i is equal to 0 if the attribute is present and equal to c_i else. With this formulation and $\beta = 0.5$, it can be shown that if the labeled attribute is entirely absent along the shortest path, a maximum deviation constraint of two times the shortest path is implied. Continuous link attributes such as elevation gain or traffic volume are cast as proportions of an upper percentile from the attribute's distribution over the network multiplied by c_i. For example, traffic volume is cast as the proportion of the 95th percentile of link Annual Average Daily Traffic (AADT). With $\beta = 0.5$ and 95th percentile AADT = 20,000, a candidate route twice as long as the shortest path would be included only if the longer route had average AADT equal to that of the shortest path less 20,000. Casting the attributes as labels may appear tedious, but the calibration step provides feedback; specifically, unrealistic labels are quickly spotted as they result in either very large or very small shortest path deviations in generated routes.

The label cost functions were iteratively minimized by finding shortest paths using Dijkstra's algorithm (15). Many commercial software packages solve the shortest path problem and can be scripted to iterate through the set of β parameter values. For this exercise, however, the shortest path algorithm was implemented in the Python programming language (16). This allowed all of the methods used to share the same core optimization algorithm, making computational comparisons more meaningful.

TABLE 1 Attribute Labels

Label		min( )
Maximize all bike facilities	link length*(1-bike facility dummy)	0.2
Maximize on-street bike lanes	link length*(1-bike lane dummy)	0.3
Maximize improved and unimproved bike routes	link length*(1-bike route dummy)	0.2
Maximize improved bike route	link length*(1-improved bike route dummy)	0.2
Maximize off-street bike paths	link length*(1-bike path dummy)	0.3
Minimize upslope	(upslope/90th percentile observed travel upslope)*length	0.1
Minimize traffic volume	(AADT/95th percentile AADT)*length	0.1
Minimize stop signs and traffic signals	(stop dummy + signal dummy)*length	0.1
Minimize turns*	left turn dummy*100m + right turn dummy*50m	0.5
Minimize adjacent employment density†	(emp. density/99th percentile emp. density)*length	0.1
Minimize adjacent commercial land use†	commercial share*length	0.1

*turn penalties calculated from observed travel time estimation, converted to distance
†thought to proxy for on-street parking turnover and driveway access frequency

Calibration and filtering were implemented as follows:

- Perform initial label run with min(β)=0.5 and step size=0.1.
- Calibrate to fitted observed distribution of shortest path deviations. The calibration does not need to be overly precise. Using Q-Q plots like those in Figure 1, each label could be fitted in one to four iterations by changing min(β) in 0.1 increments. Table 1 includes calibrated min(β) values.
- Step size was left at the initial value of 0.1, which seemed to strike an acceptable balance among choice set size, route variation, and computation time.
- Generated labels were combined, shortest paths and observed paths were added if not generated, and routes overlapping more than 90 percent with others were removed.

Applying the Other Choice Set Generation Techniques

Brief descriptions of the other techniques used are provided here. Readers should refer to the sources cited for more detailed information about applying each method. In each case, the shortest path was added if not found by the algorithm, and alternatives were filtered to remove routes that overlapped others more than 90 percent. The same Python program used for the proposed calibrated labeling method was modified to execute each of the other methods.

For the *K*-shortest and simulated shortest path methods, the number of iterations was set to 88, the maximum number of potential alternatives generated by the proposed calibrated labeling method. The labeled routes method cannot generate additional alternatives without additional labels, one of the shortcomings of that method.

K-shortest Paths Link Penalty

A De La Barra link penalty approach was followed (3,4). Distance was used as the cost criterion. After each iteration, the length of each link along the shortest path was multiplied by a penalty factor of 1.04, which was the middle value used by Ramming (3) and Bekhor et al. (5) for an urban network.

Simulated Shortest Paths

The simulated least-cost path approach described by Ramming (3) and Bekhor et al. (5) was adapted. For each iteration, link lengths were drawn from a lognormal distribution with mean equal to the natural logarithm of the actual link length and standard deviation (of the natural logarithm) equal to 1. The lognormal distribution avoids the problem of negative length values that can occur with the normal distribution used elsewhere (3,5). Turn penalties based on estimated observed travel time effects (about 100m/330ft for left turns and 50m/165ft for right turns) were added after initial runs showed paths weaving through the grid in unrealistic patterns.

Labeled Routes

The set of labels from the proposed method were used, but the β parameter was fixed at 0.5, judged to be a reasonable mix of attribute and distance sensitivity. It implies that most alternative routes will deviate no more than 100 percent from the shortest path, consistent with observed behavior. Each label generates only a single optimal route as is the case in other applications (3,5,7).

Comparing the Generated Choice Sets

Table 2 provides descriptive statistics. In some cases, no alternative to the observed route was generated. This is not necessarily a shortcoming, since for those mostly short trips a traveler could be captive to the observed route. The proposed calibrated labeling method replicated the observed route more frequently at all coverage thresholds. Coverage rates are generally lower than those found in auto route choice studies. This likely reflects the density of the network, varied trip purposes, and the inability to use travel time variation to distinguish among parallel routes. Both *K*-shortest path and simulation found at least one alternative for each trip, while the proposed method and labeled route method left 15 and 54 trips, respectively, as captive to the chosen route.

Also provided are efficiency index scores derived from formulas in Bekhor and Prato (10). The efficiency index gives equal weight to replication of the observed route and the number of alternatives generated excess to the expected number of relevant routes. Setting the number of relevant routes is somewhat arbitrary at this point. Here it is fixed to 2 to make results comparable to the original work (10). Based on this criterion, the labeled routes technique was most efficient, followed by the proposed method. Since the efficiency index is sensitive to the number of generated routes, fewer iterations were tested for the non-labeled algorithms, but scores did not improve.

TABLE 2 Generated Choice Set Statistics

Method	Unique Alternatives	% routes covered at overlap threshold 100%	90%	80%	70%	Captives*	Runtime†
Proposed	29,090	22.5	29.4	42.3	54.6	15	5h 10m
Efficiency Index ‡		*23.18*	*25.63*	*28.79*	*31.87*		
K-shortest paths (iter=88)	111,221	20.0	26.8	39.5	52.2	0	3h 11m
Efficiency Index		*18.25*	*20.69*	*23.86*	*27.02*		
Simulation (iter=88)	114,110	21.8	27.5	38.8	51.0	0	4h 10m
Efficiency Index		*18.32*	*20.49*	*23.38*	*26.42*		
Labeled routes	11,965	20.4	24.6	35.4	47.1	54	0h 18m
Efficiency Index		*24.91*	*26.82*	*29.59*	*32.51*		

*cases with no alternative (N=1,464)
†runtime on 2.4GHz Intel Core 2 Duo; includes calibration runs for proposed method
‡based on Bekhor & Prato (2009) with relevant routes = 2

The proposed calibrated labeling method took somewhat longer to compute than the other methods. This result is somewhat misleading, however, since calibration runs accounted for about half the time. Excluding calibration, only 2 hours and 17 minutes hours were required. Although not integral to either method, the link penalty and simulation approaches both required multiple runs to arrive at reasonable parameter values.

Figure 2 reveals considerable differences in choice set size distribution among the different methods. The large number of alternatives in some choice sets, up to 50 or more in rare cases, may at first seem unreasonable. Given that large choice sets mostly occur on long routes, however, they may represent the total options over a series of decision points. For example, a traveler planning a long route from A to D may reasonably consider 4 options from A to B, 4 from B to C, and 4 from C to D, resulting in $4^3 = 64$ possible route combinations.

The proposed calibrated labeling method and the labeled routes method resulted in considerably more choice set size heterogeneity. The distributions from K-shortest paths and simulation, on the other hand, imply that a single choice set size tends to dominate regardless of individual, trip, and network characteristics. While we argue that the varied choice set sizes seem more sensible, we are not aware of any existing research to confirm either view.

Figure 3 shows the deviation from shortest path distributions for our sample of non-recreational cycling trips. The proposed method is calibrated based on this distribution. K-shortest paths and labeled routes, as specified, produced a higher frequency of low-deviation routes than expected from observed behavior. The simulation choice set produced lower than expected frequencies of both low and high-deviation paths.

Statistics cannot tell the whole story regarding the generated choice sets. Even more convincing were the perceived quality and parsimony of the alternatives, which is only hinted at by the statistical comparison. Alternative routes generated by the *K*-shortest paths and simulation methods were often neither cohesive (e.g. leaving a street only to return a block later) nor distinctive (e.g. parallel routes of similar length with identical attributes). Both labeled route methods produced paths that appeared much better on these measures, although of course this is difficult to quantify.

As a final comparison, results from route choice models estimated with the different alternative sets are presented in Table 3. The path-size logit (PSL) formulation was used to adjust the multinomial logit (MNL) for expected correlation due to route overlap (17). Fit statistics cannot be directly compared, since the choice sets vary across models. Prediction rates would not be meaningful either, since they depend on choice set size and content. Estimations were performed with the Biogeme software package (18,19).

The most striking differences occur in the bike facility parameter estimates. Results using the *K*-shortest path and simulation choice sets show less heterogeneity among facility types and also considerably higher marginal utility effects relative to distance. This likely reflects a lack of variation among routes using bike facilities, since these algorithms choose such facilities only by chance. We argue that the estimated parameters are less reasonable for the *K*-shortest and simulation choice sets. For example, the estimations imply that a cyclist would be willing to travel about 60 percent farther to use an unimproved, signed bike route for an entire trip. Based on the proposed method choice set, the same attribute would be worth a more reasonable 14 percent increase in distance. The same comparison for improved bike routes shows a willingness to ride 123-131 percent farther versus a more plausible 44 percent using the proposed method. Even the relative ranking of facility types differs. While we cannot directly test the appropriateness of these parameters, the differences highlight the importance of choice set generation to model results. Differences in magnitude like those described could have major policy implications.

Estimation using the labeled route choice sets did not perform well. The positive sign on the distance parameter and negative sign on the bike route parameter are counterintuitive. The unexpected distance parameter probably reflects a lack of intermediate distance routes between the shortest path and labeled routes.

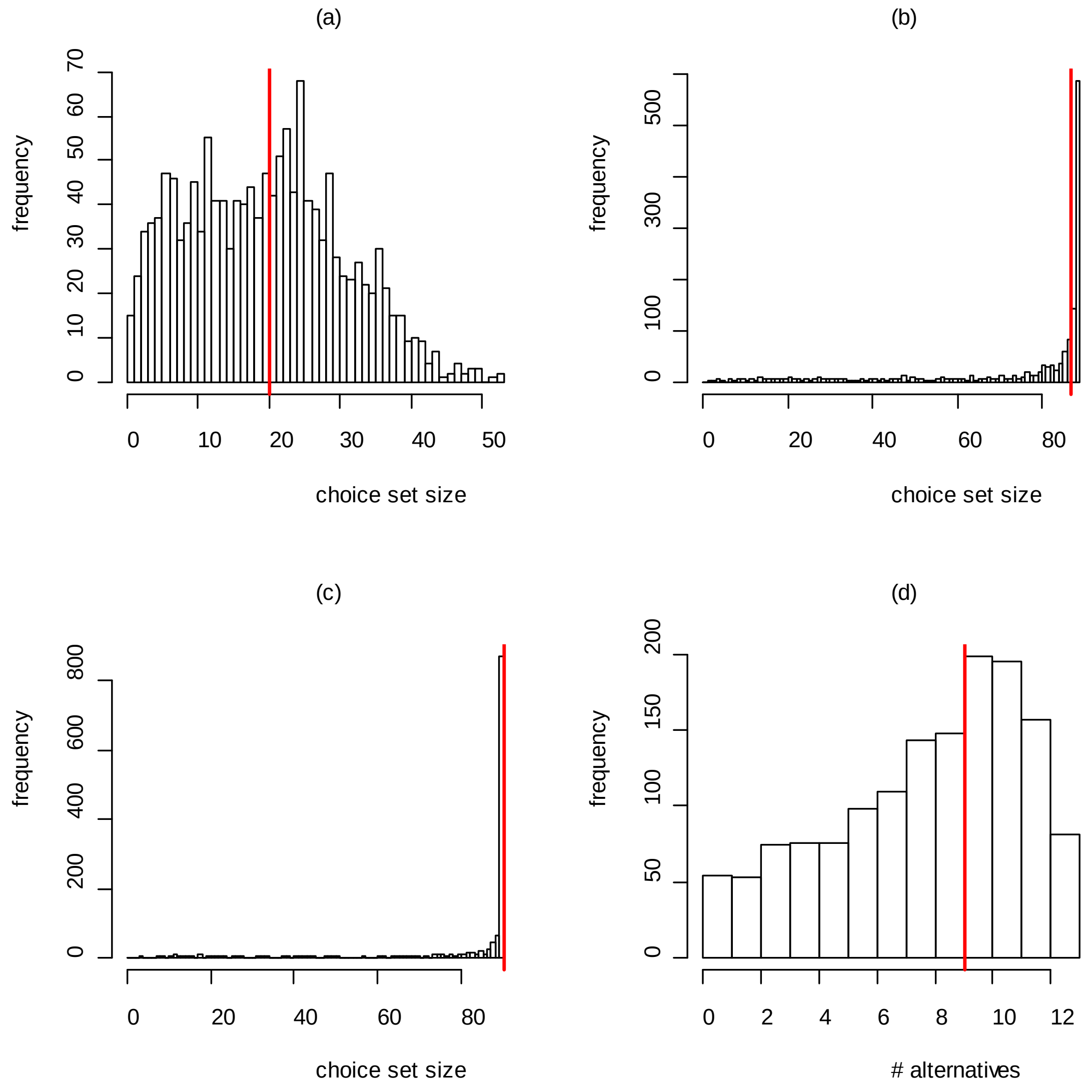


FIGURE 2 Distribution of choice set sizes including chosen route with medians marked by extended lines for (a) proposed method, (b) K-shortest paths, (c) simulation, and (d) labeled routes.

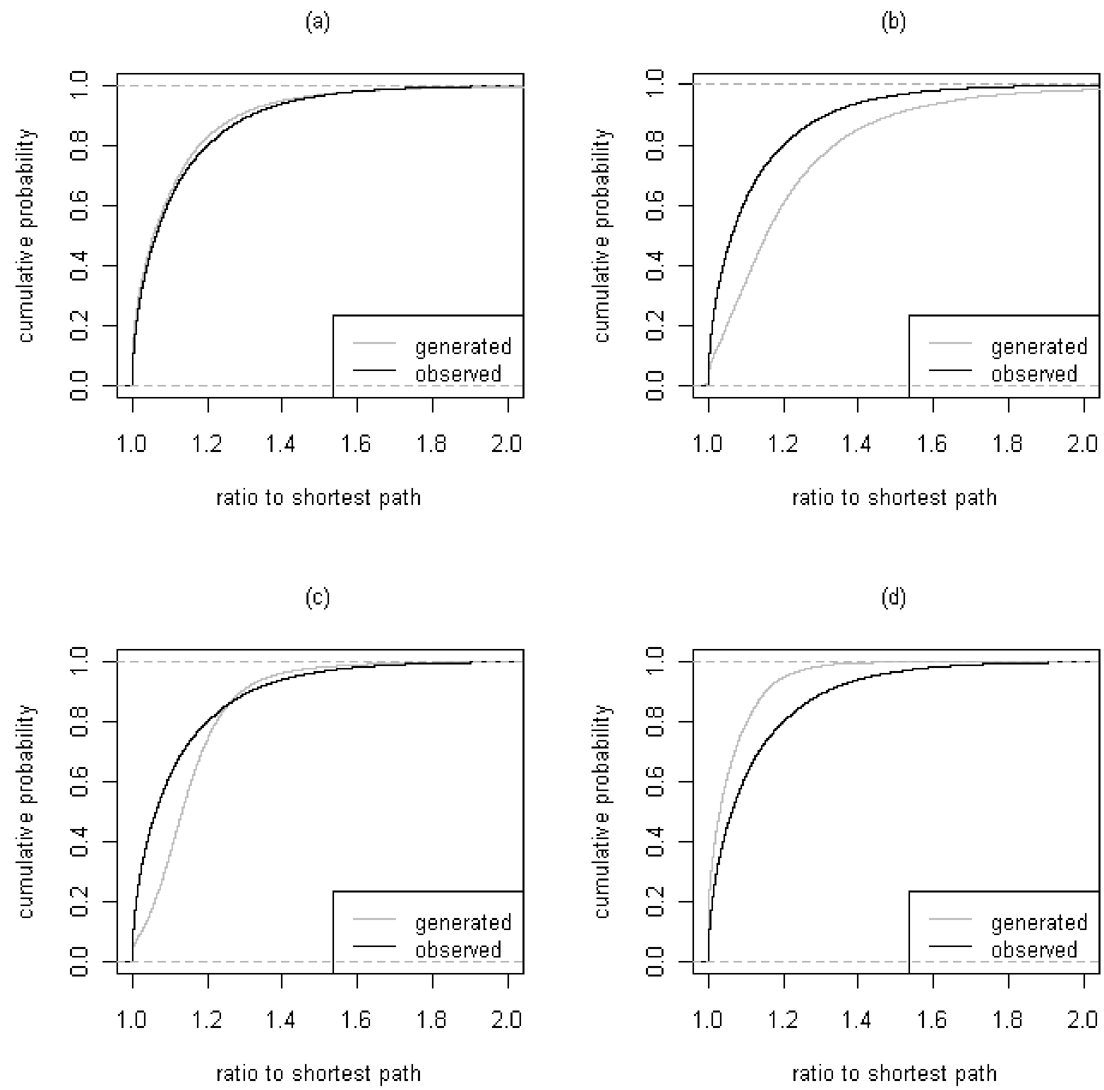


FIGURE 3 Cumulative probability of shortest path deviations for (a) proposed method, (b) K-shortest paths, (c) simulation, and (d) labeled routes; plots truncated at 2.0 for legibility.

TABLE 3 Estimation Results for Path-Size Logit (PSL) Model

	Choice set generation method:							
	Proposed		K-shortest		Simulation		Labeled	
Variable	Param.	t-stat	Param.	t-stat	Param.	t-stat	Param.	t-stat
ln(Distance)	-4.820	-13.29	-8.260	-20.05	-8.900	-16.72	5.420	10.51
Bike route*	0.639	3.14	4.010	14.29	4.170	13.66	-0.709	-3.28
Bike route, improved*†	1.760	9.46	6.640	22.77	7.450	21.59	0.313	1.56
Bike lane, on-street*	1.680	9.48	5.540	21.48	5.510	19.33	0.536	2.82
Bike path, off-street*	2.970	9.01	5.700	11.75	6.160	11.84	1.370	3.57
Traffic volume‡	-0.137	-12.34	-0.228	-15.56	-0.238	-14.94	-0.118	-9.65
Upslope (m/100m)	-0.949	-6.93	-1.910	-11.15	-1.910	-10.00	-1.160	-7.01
Stop signs	-0.045	-7.18	-0.081	-8.81	-0.121	-11.58	-0.041	-5.43
Left turns	-0.227	-10.57	-0.370	-17.01	-0.315	-12.40	-0.185	-7.86
Right turns	-0.041	-1.93	-0.204	-9.60	-0.375	-15.06	-0.032	-1.38
ln(Path-size)	1.900	23.52	2.520	22.19	1.200	15.71	0.640	5.35
N	1,449		1,464		1,464		1,410	
LL(0)	-4058.6		-6209.5		-6224.5		-2873.9	
LL	-3258.1		-3408.8		-3517.0		-2434.3	
Rho-square	0.197		0.451		0.435		0.153	

*mutually exclusive, measured as proportion of route

†Improved bike routes or "bike boulevards" include traffic calming, traffic diversion, and intersection priority in addition to signage.

‡measured as average estimated AADT along route

CONCLUSIONS

Generating a reasonable choice set is a critical but often undervalued step in discrete choice model estimation. For modeling route choice decisions, in which the universal set of alternatives may number many thousands, alternative generation is particularly challenging. The proposed calibrated labeling method modifies the labeled routes technique to improve flexibility and promote consistency of the entire choice set with observed behavior.

An application of the proposed calibrated labeling method and three other common choice set generation techniques was performed on a revealed preference bicyclist route choice problem. The proposed technique outperformed the other methods in terms of replicating observed routes efficiently. It also generated heterogeneous choice set sizes, which we argued were more plausible, and produced more reasonable choice model parameter estimates. Perhaps most importantly, the generated routes have a clear behavioral explanation. The proposed method did, however, take somewhat more computation time and required the specification of a set of attribute labels.

The proposed calibrated labeling method demonstrated considerable potential even in its most basic form. It should be immediately useful as an alternative choice set generation technique for route choice modeling in similar contexts. The only significant hurdle would be the attribute label specifications. Future research might explore route choice model sensitivity to changes in label set composition. The basic framework could also prove useful in any discrete choice context in which choice set formation can be hypothesized as a bounded, attribute-based search. Possible improvements and extensions include formalizing the calibration method and introducing heterogeneity to choice set generation parameters across classes.

REFERENCES

1. Prato, C.G., and S. Bekhor. Applying Branch-and-Bound Technique to Route Choice Set Generation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1985*, TRB, National Research Council, Washington, D.C., 2006, pp. 19-28.
2. Bovy, P.H.L. On Modelling Route Choice Sets in Transportation Networks: A Synthesis. *Transport Reviews*, Vol. 29:1, 2009, pp. 43-68.
3. Ramming, S. *Network Knowledge and Route Choice*. PhD thesis. Massachusetts Institute of Technology, Cambridge, 2001.
4. Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta. A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks. In *Proc., 13th International Symposium on Transportation and Traffic Theory*, Pergamon, Lyon, Frnace, 1996, pp. 697-711.
5. Bekhor, S., M.E. Ben-Akiva, and M. Scott Ramming. Evaluation of Choice Set Generation Algorithms for Route Choice Models. *Annals of Operations Research*, Vol. 144:1, 2006, pp. 235-247.
6. Frejinger, E., M. Bierlaire, and M. Ben-Akiva. Sampling of Alternatives for Route Choice Modeling. *Transportation Research Part B*, in press, Retrieved July 1, 2009, doi:10.1016/j.trb.2009.03.001.
7. Ben-Akiva, M.E., M.J. Bergman, A.J. Daly, and R. Ramaswamy. Modelling Inter Urban Route Choice Behavior. In *Proc., Ninth International Symposium on Transportation and Traffic Theory*, VNU Science Press, Delft, The Netherlands, 1984, pp. 299-330.
8. Cascetta, E., F. Russo, F.A. Viola, and A. Vitetta. Random Utility Models with Implicit Availability/Perception of Choice Alternatives for the Simulation of Travel Demand. *Transport Research Part C*, Vol. 9:4, 2002, pp. 249-263.
9. Prato, C.G., and S. Bekhor. Modeling Route Choice Behavior: How Relevant Is the Composition of the Choice Set? In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003*, TRB, National Research Council, Washington, D.C., 2007, pp. 64-73.
10. Bekhor, S., and and C.G. Prato. Methodological transferability in route choice modeling. *Transportation Research Part B*, Vol. 43, pp. 422-437.
11. Chakravarti, I.M., R. G. Laha, and J. Roy. *Handbook of Methods of Applied Statistics, Volume I*, John Wiley and Sons, New York, 1967, pp. 392-394.
12. Dill, J., and J. Gliebe. Understanding and Measuring Bicycling Behavior: A Focus on Travel Time and Route Choice. Oregon Transportation Research and Education Consortium, Portland State University, Final Report, December 2008, Retrieved July 23 http://www.otrec.us/main/document.php?doc_id=966.
13. Stinson, M., and C. Bhat. Commuter Bicyclist Route Choice: Analysis Using a Stated Preference Survey. In *Transportation Research Record: Journal of the Transportation Research Board,No. 1828*, TRB, National Research Council, Washington, D.C., 2003, pp. 107-115.

14. Hunt, J. D., and J. E. Abraham. Influences on Bicycle Use. *Transportation*, Vol. 34:4, 2007, pp. 453-470.
15. Daskin, M.S. *Network and Discrete Location: models, Algorithms, and Applications*, John Wiley and Sons, New York, 1995, pp. 43-47.
16. Python Programming Language. http://www.python.org/
17. Ben-Akiva, M.E., and M. Bierlaire. Discrete Choice Methods and Their Applications to Short Term Travel Decisions. In *Handbook of Transportation Science*, Kluwer, Dordrecht, The Netherlands, 1999, pp. 5-33.
18. Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models , *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.
19. Bierlaire, M. (2008). An introduction to BIOGEME Version 1.7, http://biogeme.epfl.ch.