Guide for Travel Model Transfer

Peer Reviews

Data

$f(x) = ln \sum_n$

Tools

Webinars

Analysis

TMIP
Better Methods. Better Outcomes.

TMIP

Better Methods. Better Outcomes.

1. Report No. FHWA-HEP-	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Guide for Travel Model Transfer		**5. Report Date** October 2014
		6. Performing Organization Code
7. Authors Thomas F. Rossi, Cambridge Systematics, Inc. Chandra R, Bhat, University of Texas, Austin		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Cambridge Systematics, Inc. 100 CambridgePark Drive, Suite 400 Cambridge, MA 02140		**10. Work Unit No. (TRAIS)**
		11. Contract or Grant No. **DTFH61-13-P-00162**
12. Sponsoring Agency Name and Address United States Department of Transportation Federal Highway Administration 1200 New Jersey Ave. SE Washington, DC 20590		**13. Type of Report and Period Covered** 8/29/2013 to 10/31/2014
		14. Sponsoring Agency Code HEPP-30

15. Supplementary Notes

The project was managed by Sarah Sun, COR for Federal Highway Administration

The report was reviewed by Raw, Jeremy (FHWA); Pihl, Eric (FHWA); Yoder, Supin (FHWA); Murakami, Elaine (FHWA); Gardner, Brian (FHWA) and Cervenka, Kenneth (FTA)

16. Abstract

Travel demand models are important tools used in transportation planning to forecast the amount of travel expected to occur under various scenarios and to estimate the effects of changes in transportation supply or policies. The mathematical relationships in these models depend on parameters that reflect the sensitivity of travelers' choices to the model's input variables. The process of model transfer from one region to another has become commonplace as a way to obtain model parameter estimated when sufficient local data for model estimation. This report provides guidance to transportation planners on alternative approaches to model transfer based on the amount of available resources. The guidance, which is based on research that has been conducted on model transferability, sis supplemented with simple examples of the various model transfer methods.

17. Key Words Travel demand model transferability, model estimation, model validation	**18. Distribution Statement** No restrictions. This document is available to the public at http://www.fhwa.dot.gov/planning/tmip/.

19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 76	**22. Price** N/A

Guide for Travel Model Transfer

Original: October 2014
Final: October 2014

Prepared by:
Cambridge Systematics, Inc.
with Chandra R. Bhat, Ph.D.

Thomas F. Rossi
Cambridge Systematics, Inc.
100 CambridgePark Drive, Suite 400
Cambridge, MA 02140

Prepared for:

 Federal Highway Administration

Table of Contents

List of Tables

1.0 Introduction

Travel demand models are important tools used in transportation planning to forecast the amount of travel expected to occur under various scenarios and to estimate the effects of changes in transportation supply or policies. Travel models are most often applied for metropolitan areas but may also be applied for smaller areas (subareas, corridors, cities) and for larger areas (states, parts of states, and multistate corridors). In these models, the outputs are values for useful variables such as highway volumes, transit ridership at the route level, and the amount of travel from one part of the modeled area to another. These outputs are estimated using a set of mathematical relationships where the inputs are a set of known or readily estimated variables such as the features of the transportation network, the costs of travel, and the amount and locations of activity in the region, as measured by the population characteristics and measures of commercial activity, often employment by type.

The mathematical relationships in travel models depend on **parameters** that reflect the sensitivity of travelers' choices to the model's input variables. These parameters vary based on the differences in the reason for travel (activity or trip purpose), the characteristics of travelers (income level, vehicle availability, household composition), and the environment in which travel is made. Planners do not know the "true" values of model parameters; they are generally estimated from reported travel behavior data.

Parameters are used throughout the travel modeling process. Some examples include:

- Trip generation rates;
- Gravity model friction factors and K-factors;
- Mode choice utility coefficients;
- Time of day factors;
- Volume-delay function parameters; and
- Parameters of utility functions in components of choice models used in activity based models.

Typically, travel model parameters are estimated using observed local travel data, usually from surveys. The most common type of survey used for this purpose is the household activity/travel survey, which collects information on all trips made by all persons in the respondent households on a given travel day. This practice is consistent with the manner in which travel models are applied, where travel is estimated for all households or persons living in the model region. In some cases, other surveys may be used to provide supplemental information for travelers, such as transit riders, who may not be well represented in household surveys.

Travel surveys can be difficult and expensive to conduct in a manner that leads to sufficient sample sizes for model parameter estimation, especially in smaller planning regions where resources may be quite limited. This means that, in many areas, surveys are not conducted regularly, and thus there are insufficient recent survey data for model parameter estimation.

*The **correct way** to transfer model parameters is to transfer the **entire** model, while examining whether the set of parameters needs to be scaled to the application context.*

This has led to the practice of **model transfer** in many areas. Model parameters that have been estimated in one context (the "estimation context") are used to create parameters for use in the "application context."

The process of model transfer has become commonplace in transportation planning as resources for planning continue to be constrained. Since the late 1970's, the National Cooperative Highway Research Program (NCHRP) has

published guidebooks that provide transferable parameters, the latest of which, *NCHRP Report 716: Travel Demand Forecasting: Parameters and Techniques*, was released in 2012. The parameters provided in these reports reflect national averages, based in part on data from the National Household Travel Survey (NHTS). However, while these publications have provided some guidance to planners transferring parameters, the discussion of the transfer process is limited, and alternatives to using the national parameters are not discussed in great detail. (A partial literature review on model transferability is provided as an appendix to NCHRP Report 716.)

*Transferring just a subset of parameters to the new model is risky as parameter estimates can be affected by the presence of other variables in the model, and transferring parameters related to some variables without considering the effects of other variables can result in **incorrect** model sensitivities.*

It should be noted that the correct way to transfer model parameters is to transfer the entire model, while examining whether the set of parameters needs to be scaled to the application context. In practice, a simplified method has been commonly used where just a subset of parameters is transferred to the new model. This partial transfer is risky, as parameter estimates can be affected by the presence of other variables in the model, and transferring parameters related to some variables without considering the effects of other variables can result in incorrect model sensitivities.

This document discusses suitable methods and good practices for transferring model parameters. Planners in many cases may have superior alternatives to the use of national average parameters, especially when some local data are available. The research performed for the development of NCHRP Report 716 included a synthesis of transferability studies that indicated mixed results for **model** transferability and suggested ways of improving transferability results, including use of small amounts of local data in an updating process when transferring **models**. This document seeks to present that knowledge to provide planners along with practical guidance in determining the best transfer methods for their unique situations. The guidance includes variable specification, model updating procedures using some local data, procedures to use when local data are unavailable, and use of data from sources such as NCHRP Report 716.

It should be noted that this document reflects the best information that can be provided at the time of its publication. Additional research into model transferability may provide further information to practitioners over the next several years.

1.1 Concepts of Model Transfer

This report focuses on spatial transfer of **models**, that is, a behavioral process estimated in one region (the estimation context) being used to estimate travel in another region (the application context) at about the same period in time. Before discussing spatial transfer, it is important to note that other forms of transferability are also implicitly assumed in travel models (Hansen, 1981). First, it is assumed that broad behavioral postulates, such as utility maximization, can be

It is important to include, as input to the model, anything where the analyst might be able to provide insight into future conditions.

used to explain travel behavior in a given context. Analysts also assume that model forms (such as linear regression, multinomial logit, or nested logit) are transferable, even if the model parameters are estimated directly using local data. One of the most critical assumptions, though, of travel models is that of temporal transferability of behavior (and therefore model parameters) over time within the same region.

An axiom of travel modeling is that the underlying behavioral process driving observed travel patterns remains stable (i.e., is transferable) over time: that what people do today is a reasonable gauge of what they might do next year or twenty years from now. In practice, that axiom asserts that the parameters that are estimated and calibrated for a base year condition will stay the same for the future scenarios for which the model produces forecasts. For example, the sensitivity of mode choice with respect to travel time is assumed to be the same in the future as it is today.

Model coefficients from a particular estimation context may not be appropriate for an application context for several reasons:

- There may be **true differences in behavioral responses** between the application and estimation contexts. Because model parameters apply to large segments of the traveling population, even in disaggregately applied models, the parameters reflect "average" behavior for the segments to which they apply. Thus, if the average person within a segment in one region behaves differently than the average person in another, the parameters from the estimation context might not be valid in the application context. Simply put, without analyzing data from both regions, we do not know how the sensitivity of travelers in Region A to factors affecting travel differs from that of travelers in Region B.

- There may be **limitations of model specification**. For example, in a mode choice model, the sensitivity to cost may be a function of income; if this is not considered, and the income ranges are different between the estimation and application contexts, the model will not have good transferability.

- There may be **differences in contextual factors** such as general attitudes, social mores, and geography between regions that are difficult to explicitly capture using model variables. For example, Region A where the central city has a wage tax may show different commuting patterns than Region B that does not. That would not mean that people in Region B have a different sensitivity to the tax (which would be a difference in behavioral response, the first bullet above), only that the residents of Region B do not have to pay the tax.

The first point cannot be addressed directly without application context data. The transferability problems arising from limitations of the model specification (the second point) can be alleviated to some extent by using and market segmentation schemes and carefully defining variables. The problems arising from differences in "difficult-to-explicitly-capture" contextual factors can be partially addressed by using updating procedures; however, such procedures require use of a small sample of data from the application context. In fact, several studies have shown that the validity of model transferability is substantially increased by the use of such updating procedures. However, it must be noted that simply assuming that behavioral reactions in one region are the same as those in another region has no basis in research. Regardless of the methods used, the success of a transferred model depends on being able to defend the plausibility and suitability of the model structures and parameters that have been transferred.

1.2 Model Components and Parameters That May Be Transferred

As noted above, the correct way to transfer model parameters is to transfer the entire model, while examining whether the set of parameters needs to be scaled to the application context. The "entire model" in this case refers to a major model component that is typically estimated on

its own. In the context of four-step travel models, the major components may include the following (not all models have every one of these components, of course):

- Trip production
- Trip attraction
- Trip distribution
- Mode choice
- Time of day
- Vehicle availability
- Highway assignment
- Transit assignment

Within the same complete model, one might estimate trip production models from household activity/travel survey data, trip attraction models from establishment survey data, and mode choice models from combined household and transit survey data. Highway and transit assignment parameters are not estimated at all and may be initially specified from experience from the literature, from previous models, or models in other regions.

Transferring the "entire model" refers to any of these types of components listed above that are estimated separately. (The set of components for an activity based model, of course, would be much different.) In a particular model setting, analysts may choose to estimate some model components and transfer others, depending on the sufficiency of local data with which to estimate models and the availability of appropriate models to transfer.

Only some types of model components are amenable to being transferred. Some types of components are not likely to have a true estimation context; for example, highway assignment model parameters are not estimated from local data (though they are often calibrated using local data). Other model components have parameters that are often estimated but may be too area-dependent to be transferred to another region. The following list summarizes the types of four-step model components that fall into each of these categories:

Amenable to model transfer:

- Trip production
- Trip attraction
- Mode choice
- Time of day
- Vehicle availability

Not likely to have a true estimation context:

- Highway assignment
- Transit assignment

Models with parameters that may be too area-dependent to be transferred:

- Trip distribution

For activity based models, the categorization may be less clear, and more research is needed. Assignment models will still fall into the second category, and components related to location choice (tour and trip level destination choice, workplace and school location) are likely to fall into the third category.

1.3 Transferability Considerations

Rigorously demonstrating that transferability of a model is correct involves using a model from the estimation context to predict travel behavior in the application context and comparing the results to those from a model estimated directly from data in the application context. In other words, one would estimate a model using survey data from Region A and apply the model to input data for Region B, and then compare the results to those from a model estimated using survey data from Region B. If the differences in results (including sensitivity to changes in input parameters) are statistically insignificant, transferability has been demonstrated.

Naturally, such a transferability study is unlikely to be undertaken by a planning agency since it would require more effort than simply estimating models from local data, and there is no need for model transfer when sufficient local estimation data exist. Therefore such studies have been performed primarily in the research environment. A summary of relevant research is presented in Appendix A, with more details provided in Appendices C and D. This summary shows mixed results with respect to the validity of transferability of a variety of typical travel model components.

It would be beneficial if additional research was conducted to understand the conditions under which model transfer is likely to succeed. For example, there is some evidence that model transfer may be more successful if the estimation and application contexts are "similar" but there has been no systematic effort to quantify what constitutes "similarity". It would also be beneficial to understand which model components are more (or less) transferable, so agencies might target data collection in topic areas where local data will significantly improve representations of travel behavior.

Given the gaps in research performed to date, planners must do the best they can with the available resources. If there is insufficient local data to estimate the parameters of a particular model component that is needed for the planning analyses an agency is required to perform, there is no choice but to transfer a model. In this case, planners must be able to communicate the assumptions underlying models and parameters they choose to transfer, and must have candid discussions with decision makers about the types of assumptions that those decision makers are prepared to make, and the risks entailed if the assumptions prove to be wrong. This document can facilitate such discussions by presenting different techniques to improve the transfer process as well as providing information about their benefits and risks.

If there is some data for the application context, but not enough to perform full model estimation, there is evidence that using the local data to update the models from the estimation context can substantially improve the model transfer. This potential improvement in the model transfer process is discussed further below.

When transferring a model, practitioners should, to the extent that they can, confirm that the estimation context model has been estimated using sufficient and appropriate data and has been validated completely. This should be done using complete documentation of the estimation and validation processes for the estimation context model.

1.4 Methods for Model Transfer

Model transfer may be enhanced when some data are available in the application context. Updating procedures that make use of the application context data can greatly improve model transferability. Five broad types of updating procedures may be used when transferring models when some data are available in the application context. These are the following:

- **Simple (Naïve) Transfer Method**. In this approach, the model from the estimation context is simply used "as is," and parameters may be revised during model calibration.

No updating of any kind is made to the estimation context parameters when applied in the application context.

- **Transfer Scaling**. This approach entails the estimation of the scale and constants in the application context and assumes that the rest of the model coefficients are transferable from the estimation context. The application context utility function scales and constants are estimated from a small application context sample, and the remaining utility function parameters are assumed to be transferable from the estimation context. This procedure works when the smaller amount of application context data can be sufficient to estimate these two parameters.

  The specific way in which the scaling would occur depends on the model form. For example, in the case of a logit model, let the utility functions from the estimation context be represented by Equation 1-1.

$$U_{Ei} = B_{E0i} + B_{E1i}X_{1i} + B_{E2i}X_{2i} + \ldots + B_{Eni}X_{ni} \tag{1-1}$$

  Where:

  U_{Ei} = Utility of alternative i in the estimation context model

  X_{ki} = kth variable for alternative i in the estimation context model ($k = 1, \ldots, n$)

  B_{Eki} = Estimated coefficient of variable X_{ki} in the estimation context model

  The application context model is estimated with utility functions given by Equation 1-2.

$$U_{Ai} = B_{A0i} + B_{Ai}(B_{E1i}X_{1i} + B_{E2i}X_{2i} + \ldots + B_{Eni}X_{ni}) \tag{1-2}$$

  Where:

  U_{Ai} = Utility of alternative i in the application context model

  B_{A0i} = Estimated constant term for alternative i in the application context model

  B_{Ai} = Estimated scale for the application context model

  This process differs from the simple transfer approach in that B_{A0i} and B_{Ai} are estimated using the application context data. If B_{A0i} is set to equal B_{E0i} and B_{Ai} is assumed to equal 1.0, this is equivalent to the simple transfer approach.

  For more information on transfer scaling, see Karasmaa (2003), Section 3.7.2, pages 49-50..

- **Bayesian Updating.** This involves the use of a classical Bayesian process to update the estimation-context parameters using the application-context data. In this approach, parameter estimates from a small application context sample are combined with the estimation context parameter values estimated from a larger data set, using a classical Bayesian analysis to yield a transferred set of parameters. The idea of Bayesian updating is to optimally combine the coefficients obtained from the application and estimation contexts, accounting for the variances (i.e., precision) of the coefficient estimates in the two contexts. This is done by computing a weighted average of the coefficients from the two contexts, the weights being equal to the inverse of the variance of the coefficient estimates. Bayesian updating is based on a combination approach that attaches more weight to more precise estimates.

In Bayesian updating, the updated trip rate parameter estimates can be obtained through Equation 1-3.

$$\hat{\theta}_{updated} = \frac{\left[\dfrac{\hat{\theta}_{prior}}{\left(\hat{\sigma}^2_{prior}\right)} + \dfrac{\hat{\theta}_{updating}}{\left(\hat{\sigma}^2_{updating}\right)} \right]}{\left[\dfrac{1}{\left(\hat{\sigma}^2_{prior}\right)} + \dfrac{1}{\left(\hat{\sigma}^2_{updating}\right)} \right]}$$

(1-3)

Where:

$\hat{\theta}_{updated}$ = New estimated trip rate parameter for application context

$\hat{\theta}_{prior}$ = Estimated trip rate parameter for estimation context (from survey)

$\hat{\theta}_{updating}$ = Estimated trip rate parameter for application context (from small survey sample)

$\hat{\sigma}^2_{prior}$ = Estimated trip rate variance for estimation context (from survey)

$\hat{\sigma}^2_{updating}$ = Estimated trip rate variance for application context (from small survey sample)

For more information on Bayesian updating, see Ortuzar and Willumsen (2011), Section 4.5.3 (page 145).

- **Combined Transfer Method**. This approach is a generalization of the Bayesian updating process that includes transfer scaling (i.e., it combines the previous two methods). This method relaxes the assumption of the Bayesian approach of identical behavioral model parameters between the estimation and application context, which may be too strong in the case of different urban areas where there may be a real difference in the model parameters in the populations (sometimes labeled as transfer bias). In this case, one may obtain updated parameters as the minimum mean squared error estimate of the original and updated parameters.

  For more information on the combined transfer method, see Karasmaa (2003), Section 3.7.4 (pages 51-52).

- **Joint Context Estimation**, which involves the estimation of a single model using data from both the estimation and application contexts. Therefore, this approach requires not only the model parameters from the estimation context, but the original estimation data set (e.g. the original survey data) as well. In such estimations, a variety of specifications can be used and tested. The basis of joint context estimation is that one or more variable coefficients may be the same across the estimation and application areas, and so there is a gain in efficiency in using data from both contexts.

  For more information on the joint context estimation, see Karasmaa (2007), pages 413-415).

Chapter 3 demonstrates the application of these transfer procedures through simple examples.

If no data at all are available for the application context, then only the simple transfer method would apply. As discussed later in this document, the important issues for the analyst to consider include:

- Determining the most suitable estimation context for model transfer; and
- Being aware of the limitations, associated assumptions, and risks of transferred parameters, and being prepared to consider other estimation contexts as necessary.

1.5 Research Summary

A literature review of research into spatial transferability of models was performed. This review built on a previous review documented in NCHRP Report 716. This synthesis of the spatial transferability literature has discussed the model transferability approach in the context of the trip-based modeling approach (Section A.2 and Appendix C) and the activity-based approach (Section A.3 and Appendix D), as well as the data transfer approach (see Sections A.4 and C.5).

There is also a clear indication that transferability improves with a better variable specification and with a disaggregate-level model in the estimation context. The results also emphasize that, whenever possible, some level of model updating should be undertaken using local data collected in the application context.

The synthesis of transferability studies in the context of trip-based models provides mixed results regarding the effectiveness and validity of transferability. However, there is also a clear indication that transferability improves with a better variable specification and with a disaggregate-level model (at the individual or household level) in the estimation context (thus capturing more behavioral determinants that effectively get controlled for in the application context). The results also emphasize that, whenever possible, some level of model updating should be undertaken using local data collected in the application context. While the collection of a small disaggregate-level data set in the application context would allow model updating using any of the methods described above (and the analyst can compare alternative updating methods), the synthesis suggests that even simple updating procedures such as a constants-only updating scheme using aggregate travel data in the application context typically provide superior transferability than the simple (no-update) transfer approach.

The synthesis of transferability studies in the context of activity-based models also provides mixed results regarding the effectiveness and validity of transferability. There is some evidence so far that the improved behavioral basis of ABMs does seem to manifest itself in the form of improved transferability potential, especially in those components that are not associated with travel mode and location choices. There is much more consistency in the transferability results of the very limited number of ABM transferability studies undertaken thus far than in the vast body of transferability literature on trip-based model components. Whether this is simply a chance occurrence in the limited ABM studies or a true improvement in transferability because of the improved behavioral basis of ABMs remains to be seen. In this regard, it is important to undertake more ABM transferability studies with a much more diverse set of regions than the set of regions used so far. Also, the updating methods used in current ABM transferability studies have been the simple transfer approach and the constants updating approach. There is

a need to extend transfer approaches to other updating methods and to consider the data transfer approach.

1.6 Glossary of Terms Used in This Report

The following terms used throughout the report are defined here as a reference:

- **Model transfer** – The process of using a model and its parameters from one context to develop a model for another context. This report focuses on "spatial transfer"—using a model from one geographic area to develop a model for another area.
- **Transferability** – How well the estimation context model is able to predict travel behavior in the application context.
- **Estimation context** and **application context** – If a model is being transferred from Area A to Area B, Area A is defined as the estimation context and Area B is the application context.
- **Updating** – The process of changing the estimation context model parameters using data collected in the application context, to improve the model for use in the application context.
- **Simple (naïve) transfer** – The use of model parameters from the estimation context without updating. The term "simple transfer" is used hereafter. This approach is also known as "borrowing" model parameters, but that terminology is not used in this report.

1.7 Report Organization

This chapter has presented some background on the rationale behind model transfer, issues with transferring models, and methods for model transfer. The remainder of this report is organized as follows. Chapter 2.0 presents recommended methods and guidance for model and data transfer. Chapter 3.0 provides examples of the application of these methods. The appendices present a summary of the research into model transferability. Appendix A summarizes the research while Appendix B presents a summary of model transferability measures that are appear throughout the literature. Appendix C provides more details about the transferability literature in the context of trip-based models, and Appendix D provides more details about the transferability literature in the context of activity-based models.

11

2.0 Guidance and Recommended Methods for Model Parameter Development and Transfer

This chapter presents guidance for analysts in the development of travel models in the methods for obtaining model parameters through estimation, model transfer, or some combination of these. First, evaluating the sufficiency of local data for model estimation is discussed. The remainder of the chapter focuses on optimizing the process of model transfer when sufficient local data for model estimation do not exist. This discussion includes the choice of the model estimation context, the effects of variable definition on model transfer, and selection of the model transfer procedure. The end of the chapter presents a discussion of the implications of model transfer for model validation and calibration.

The model transferability research summarized in the appendices is limited in the sense that the validity of spatial transferability of model parameters remains unknown in most situations. The literature provides little guidance in terms of the best options for transferring model parameters with an eye toward the use of existing information to optimize accuracy. Therefore, the research is incomplete, and it is impossible to recommend with certainty the optimal procedures for developing model parameters for all situations. The best that can be done here is to provide some basic concepts regarding model transferability based on the research that has been done. The following concepts are offered:

- **There is no basis in the research for defining situations in which model parameters are clearly and definitively spatially transferable.** While some studies show that model parameters seem to be transferable, others show the opposite result. Most of the studies cited in the appendices do not seem to be broadly applicable outside of the specific situations—the geographic areas studied, travel/activity purposes considered, and data sets used—in which the research was conducted.

- **Most of the research has concentrated on transferability of trip generation and mode choice, as well as a few recent studies on activity based modeling approaches** There has been almost no research on transferability of destination choice (trip distribution) and route choice (trip assignment) parameters. The lack of destination choice model spatial transferability studies may be an indication of a general recognition that such models may be inherently non-transferable due to their dependence on specific characteristics of urban areas. The sensitivity of destination choice to travel time must depend on the proximity of the universe of possible destinations, which in turn depends on such factors as the size of the region, the density of residential and commercial activity, the distances between activity centers, and levels of highway congestion. These factors can of course vary widely from one area to another.

- **While not a universal finding, a majority of the studies where updating procedures were considered seemed to show improvement in transferability when these procedures were used.** While such procedures require the availability of some data in the application context, which may not always be available, there seems to be value in having at least some application context data in terms of improving the validity of transferred parameters.

With these concepts in mind, this chapter provides guidance to modeling practitioners on the best ways to develop model parameters given their specific circumstances, including the amount of local data that is available or that could be collected.

Important issues for the analyst to consider include:

- Being aware of the limitations, associated assumptions, and risks of transferred parameters, especially during model validation and calibration, and being prepared to revise models or to consider other estimation contexts as necessary.
- Determining whether sufficient local data exists with which to reliably estimate or to update model parameters;
- Determining the most appropriate estimation context for model transfer;
- Considering the effects of variable specification on the parameter estimates;
- If it is decided to transfer model parameters, choosing the most appropriate transfer procedure for the situation; and
- Considering the effects of transferred parameters on model validation.

Generally speaking, given that the results from earlier research are at best mixed regarding the transferability of model parameters, it would seem that the most reliable model parameters would be obtained by estimating one's own parameters using local data if the necessary data and the resources for model estimation are available. However, this conclusion should not be construed to imply that estimating a model with inadequate local data is superior to transferring a model estimated using sufficient data from an estimation context, or that resources are better spent on model estimation than on other critical components of the model development process, such as data collection and model validation.

It should also be noted that performing model estimation does require specialized knowledge as well as substantial analyst time and is greatly enhanced by practical experience, and so it is recognized that many agencies may not have the resources to estimate their own models. This is especially true in smaller areas whose agency technical staff may not have adequate model estimation experience and who may not have the resources to hire consultants experienced in model estimation. (Smaller agencies, of course, are also less likely to have local household travel survey data sets of sufficient sample size for model estimation.)

If there is only a limited amount of survey data (or no data) in the application (local) context, then some type of model transfer will be necessary. Based on the research results described in the appendices, using updating procedures appears to be superior than a simple transfer approach though the simple transfer approach does not require any application context data (and therefore is the only possible approach when no application context data are available at all).

Depending on what is available in terms of data in the application context, different updating methods could be used; this choice of updating method is discussed in more detail in Section 2.4. In general, it makes sense to use the method that makes use of the most information possible from the estimation context. However, there are tradeoffs between the amount of information used and the level of effort and expertise required.

2.1   Determining the Sufficiency of Local Data for Model Estimation

It is impossible to develop a checklist for the sufficiency of model estimation data for various types of models. A number of variables enter into the process of determining sufficient sample sizes for model estimation, including the variability of observed behavior, the incidence of

relatively rare activity (such as, in many cases, use of transit or non-motorized modes or the decision not to own a vehicle), the desirability of segmentation of model parameters (for example, by trip or tour purpose, or by income level), and desired confidence intervals for parameter estimates. Stopher and Jones (2001) note that "sample size is probably the single most controversial item in household travel surveys and one on which there is virtually no agreement."

It would be desirable for model estimation, application, and validation considerations to dictate the sample sizes for household travel surveys, but this is very often not the case. Understandably, budget constraints often limit survey sample sizes, and political considerations can affect not only the sample sizes but also the segmentation of the sample (for example, ensuring that political jurisdictions have adequate sample sizes), especially when survey funding may be shared among different agencies.

Another issue for modelers is that survey data sets are sometimes "inherited," meaning that they may have been conducted years before the model development work begins, often even before the modeling objectives have been defined. This can limit the ability of the modeler to influence the sample size and result in sample sizes that are inadequate for certain modeling purposes.

If the modeler has the ability to influence the survey sampling plan, he or she can perform statistical analyses based on the means and variances of the variables of interest, the desired confidence levels and degrees of precision, and the size of the population (if small enough to affect the computations). A comprehensive description of this process can be found in *The On-Line Travel Survey Manual: A Dynamic Document for Transportation Professionals*, Chapter 5, which reproduces information first documented by Cambridge Systematics, Inc. (1996). This process can also be used to check whether the sample sizes of existing surveys are adequate.

2.2   Choosing the Estimation Context

If a model is to be transferred, it makes sense to choose the best available estimation context—that is, the context from which the model transfer is most likely to be valid. Unfortunately, the research summarized in the appendices provides insufficient information to determine the best estimation context for a particular case. Note that many of the studies cited in the appendices used data from only two contexts, and so conclusions about the quality of different types of estimation contexts were not possible in these cases. Some of the studies summarized in the appendices that provide some information about estimation context choice include Caldwell and Demetsky (1980), McComb (1986), Everett (2009), Nowrouzian and Srinivasan (2012), Sikder and Pinjari (2013), and Bowman et al. (2013).

The relatively small amount of information provided in the literature does not provide substantial guidance on the selection of the estimation context for transferring models. However, the following conclusions may be drawn from the literature review:

- There is some evidence that transferability is enhanced when the estimation context is more "similar" to the application context, in terms of area population/size and socioeconomic make-up. The problem in earlier studies is that "similarity" has not been rigorously defined, and objective measures of similarity have not been developed.
- Within the U.S., there is evidence that transferability is enhanced if the estimation context is in the same state as the application context, at least for relatively larger states.

2.2.1 Transferring Parameters from "Composite" Sources

It has been common practice since the late 1970s to use "transferable parameters" from national guidebooks. The latest in this series of guidebooks is NCHRP Report 716, Travel Demand Forecasting: Parameters and Techniques (Cambridge Systematics, Inc. et al., 2012).

NCHRP Report 716 provides a variety of transferable parameters for various components of the four-step modeling process, based on two main sources:

- The 2009 National Household Travel Survey (NHTS), administered by FHWA; and
- A database of information from model documentation from 69 metropolitan planning organizations (MPOs), referred to in the report as the "MPO Documentation Database."

In NCHRP Report 716, the weekday sample of the 2009 NHTS was used to obtain selected parameters including trip production rates, average trip lengths, vehicle occupancy rates, and time-of-day percentages. This information was estimated by urban area population range, using the urbanized area identifier in the data set. The population ranges available in the NHTS data set are as follows:

- Over 1 million population with subway/rail;
- Over 1 million population without subway/rail;
- 500,000 to 1 million population;
- 200,000 to 500,000 population;
- 50,000 to 200,000 population; and
- Not in an urban area.

It was found that many of the parameters estimated from NHTS data did not vary by population range, varied only between some ranges, or had only minor fluctuations that showed no trends and appeared to be related to survey sampling limitations. (This finding was consistent with the literature summary presented in the appendices, which showed mixed results for transferability of model parameters.) Therefore, some parameters were presented for aggregated population ranges or for all areas together.

The parameters reported in NCHRP Report 716 that were derived from the MPO Documentation Database included trip attraction rates, trip distribution gamma function parameters, mode choice utility function coefficients, and volume-delay function parameters for highway assignment. The report organized the metropolitan areas in the MPO Documentation Database by population range, as follows:

- Over 1 million population;
- 500,000 to 1 million population;
- 200,000 to 500,000 population; and
- 50,000 to 200,000 population.

While the list of MPOs included in the MPO Documentation Database is shown in NCHRP Report 716 (Table 4.1), when parameters from this database are provided from specific metropolitan areas, the areas are not identified; rather, specific characteristics of the areas are provided. For example, for vehicle availability models, model specifications from specific areas identified by descriptions such as "Western metro area, 1 to 2 million population range, about 1.9 vehicles per household" are provided.

The main advantage of using transferable parameters from NCHRP Report 716 is that they are easy to use and require very little time or resources from the modeler. Usually, these parameters have been used with a simple transfer approach (although transfer scaling might also be possible if a small sample of data for the application context is available). Other transfer approaches that require either the model estimation data set or the variances from model estimation, such as Bayesian updating, combined transfer approach, and joint context estimation, cannot be used with the NCHRP Report 716 transferable parameters. This means that the ability to take advantage of information from the application context is extremely limited.

As noted above, the transferable parameters are presented in NCHRP Report 716 by urban area size ranges, whether the parameters are derived from the NHTS or the MPO Documentation Database. This means that the user transfers the parameters for the urban area size range in which the region (the application context) falls. It should be noted that while a few of the research efforts cited in the appendices to this report indicate that choosing an estimation context of similar size to the application context improves transferability, there is no research indicating that all areas in a particular population range have parameters that are transferable among them. The research team for NCHRP Report 716 attempted to find other regional characteristics that enhanced transferability, but they were unsuccessful. The results were presented by population range because of the few studies that indicated some improvement in transferability, the ease for users in choosing transferable parameters, and the familiarity of readers with such segmentation based on earlier guidebooks.

Besides the limitations on the types of transfer approaches that can be used, the disadvantages of using the NCHRP Report 716 transferable parameters are that national averages rather than "similar" estimation contexts are used for parameters derived from the NHTS. For parameters obtained from the MPO Documentation Database, the complete documentation of the model specifications is not provided, and so the choice of estimation context is based on limited information, and other information about model estimation that might be useful in transferring parameters is unavailable. As an alternative, while it would require more effort and resources, an analyst could do his or her own research into potential estimation contexts, using a greater variety or more focused set of possibilities (including more recent models that were unavailable at the time NCHRP Report 716 was developed). For example, even if a modeler wanted to use a simple transfer approach, it is better to consider transferring parameters from a "good" model from a "similar" metropolitan area in the same state rather than using the parameters provided in NCHRP Report 716.

In summary, use of transferable parameters from sources such as NCHRP Report 716 can be quick and easy, but if time and resources are available, it may make sense for an analyst to consider using a transfer approach that uses updating and therefore uses more information from an estimation context than is provided in NCHRP Report 716, or finding a "better" estimation context that is more appropriate for the application context than national sources.

2.3   Variable Specification

There is, of course, no single set of variables that is universally used for a particular model component. In four-step models, cross-classification variables for trip production models vary from one region to another. In trip attraction models, employment categories vary among regions, and even the definitions of specific employment categories (for example, "service" employment) may not be the same everywhere. Various trip distribution models use different definitions for the impedance variables. Mode choice model variables, both those that represent transportation level of service and other variables, differ from one region to another. Activity-based models have many variables, and so far no two are alike (other than models that have been transferred using a simple approach without updating).

The wide variety of variables that can appear in models can make transfer difficult. Model parameters are generally associated with one or more specific variables; for example, trip attraction equations usually have coefficients associated with individual variables such as employment by type, and mode choice model parameters are coefficients associated with particular variables in utility functions. If a variable is defined differently between the estimation and application contexts, it is not correct to transfer the parameter(s) associated with the variable. Consider a trip production model that is a cross-classification of household size by income level. If the breakpoints that define income levels are not the same in the estimation and application contexts, the parameter estimates from the estimation context will not be valid for the application context model. Even if the breakpoints are the same in nominal dollars, the income levels may not be the same if the cost of living varies between the estimation and application contexts.

Even when a variable is defined identically in the application and estimation contexts, the associated parameters may not be transferable if the model specifications are not the same. For example, consider a linear regression model that has three variables in the estimation context, as shown in Equation 2-1.

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 \tag{2-1}$$

Consider the case where variable X_3 does not exist in the application context (perhaps because data are not available for model application). Since the presence of variable X_3 affects not only the estimate of B_3 but also the estimates of B_0, B_1, and B_2, these parameters are not transferable to the application context in the absence of variable X_3.

Taking into account the issues associated with variable definition, the following considerations in transferring parameters are offered:

1. When choosing the estimation context, consider whether the variables used in the estimation context models are available and are able to be defined similarly in the application context. If variable definitions cannot be reconciled between the application and estimation contexts, consider choosing a different estimation context.
2. Since variables may be able to be defined during the model development process, define variables in the application context to be the same or as similar as possible to the variable definitions in the estimation context.
3. Be prepared to adjust parameter estimates from the estimation context to account for differences in the application context. For example, transferred cost coefficients should consider the differences in the cost of living between the application and estimation contexts, as well as between the time of the data collection in the estimation context and the base year for the application context model.

2.4   Selecting the Procedure for Obtaining Model Parameters

Table 2.1 summarizes the procedures that may be used for obtaining model parameters, with each column summarizing a particular procedure. All but the last column, estimation of parameters from local data, are model transfer procedures. Generally speaking, the procedures become "better" as one moves from left to right in the table, in terms of the validity of the transferability and the amount of information from the application context used. However, it is also notable that the levels of application context information, expertise, and resources required also increase moving from left to right, and so some agencies may be less able to choose "better" transfer procedures. Some procedures may be infeasible in some regions, if insufficient estimation or application context information, or resources, are available.

17

Table 2.1. Summary of Procedures to Obtain Model Parameters

	Simple Transfer	Transfer Scaling	Bayesian Updating	Combined Transfer	Joint Context Estimation	Estimation of Parameters from Local Data
Required Information from Estimation	Model structure, parameter estimates	Model structure, parameter estimates	Model structure, parameter estimates and variances	Model structure, parameter estimates and variances	Complete model estimation data set	(not applicable)
Application Context Data Required (Survey)	None	Small sample	Small sample	Small sample	Small sample	Complete survey data set
Level of Expertise Required	Low	Medium	Medium	Medium	High	High
Level of Effort	Low	Low/ Medium	Medium	Medium/ High	High	High
Expected Transferability	Low	Low/ Medium	Medium	Medium/ High	Medium/ High	High

In general, the guidance is to use the procedure for obtaining model parameters with the highest expected transferability that is feasible given the available resources, including:

- Estimation context survey data and models;
- Application context survey data (or resources to collect data);
- Level of expertise available for model parameter estimation/transfer; and
- Resources available to devote to development of model parameters.

An important point is that the procedure for obtaining model parameters can vary among the model components. For example, say there are sufficient local (application context) data with which to estimate trip production rates, but insufficient local data with which to estimate a mode choice model. In such a case, trip production model parameters might be estimated using the local data, but the mode choice model would be transferred (though the constant terms in the utility functions might be updated during model validation to approximate mode shares reported for work trips in ACS).

The remainder of this section describes the transfer procedures presented in Table 2.1 in more detail. Chapter 3.0 presents examples of the application of these transfer procedures.

2.4.1 Simple Transfer

The **simple transfer method** is *the only method available for obtaining parameters if no application context survey data are available at all*. Assuming the same model structure in the estimation and application contexts, only the parameter estimates from the estimation context are required; these parameters are simply asserted as the (pre-calibration) parameters for the application context model. The main attraction of the simple transfer method is the low level of effort and expertise required although, as noted in Section 2.3, spending time carefully considering the choice of estimation context can improve the model transfer.

It is common practice for transferred parameters, especially constant terms, to be adjusted during model calibration to provide a better fit between base year model results and observed data. While a good fit can often be achieved, this type of adjustment changes the sensitivity of the model to the input variables and may in fact simply compensate for model deficiencies, including parameters that are not truly transferable. To maximize the validity of the model transfer, it makes sense to concentrate on optimizing the choice of the estimation context.

The main issue with the simple transfer method is that the validity of the transferred parameters remains unknown, and, as the literature shows, in many cases model parameters may not be transferable by any reasonable standards. It is common practice for transferred parameters, especially constant terms, to be adjusted during model calibration to provide a better fit between base year model results and observed data (see Section 2.5). While a good fit can often be achieved, this type of adjustment changes the sensitivity of the model to the input variables and may in fact simply compensate for model deficiencies, including parameters that are (unknown to the analyst) not truly transferable.

It would be preferable if resources always allowed the collection of at least a small sample of data in the application context so that one of the other transfer methods could be considered. However, it is recognized that this is not always the case, and so some agencies will have to use the simple transfer method. In such cases, it does make sense to concentrate on optimizing the choice of the estimation context (see Section 2.3) to maximize the validity of the model transfer.

2.4.2 Use of Data from Other Sources and Combining with Update Procedures

As the literature review in the appendices shows, much of the research into model transferability indicates that the accuracy of transferability is enhanced when some data from the application context is used in the transfer process. When some data are available for the application context, but estimation of models directly is not done, either because of time and resource concerns or because the data are insufficient to estimate local models, the four methods described in this section may be considered. Note that all of these methods require some new model estimation although the transfer scaling approach requires the estimation of only alternative specific constants and scaling factors.

The general approach is that a model that has been estimated in another area (the estimation context) is used as the starting point, and model parameters are updated using data collected in the area for which the model is being developed (the application context). In most cases, the data source for the application context is a household travel/activity survey.

Section 1.4 introduced four types of parameter updating procedures which make use of application context data. These are:

- **Transfer Scaling**, which entails the estimation of the scale and constants in the application context and assumes that the rest of the model coefficients are transferable from the estimation context;
- **Bayesian Updating**, which involves the use of a classical Bayesian process to update the estimation-context parameters using the application-context data;

19

- **Combined Transfer Method**, which is a generalization of the Bayesian updating process that includes transfer scaling; and
- **Joint Context Estimation**, which involves the estimation of a joint estimation/application context model and therefore requires estimation data (e.g. the original survey data) from the estimation context.

Some characteristics of these different methods are described below (and are summarized in Table 2.1).

Data Required from the Estimation Context

In the joint context estimation method, the data from the estimation context (i.e., the household survey data set) is needed. While sharing of survey data sets is becoming more commonplace, the data for a particular estimation context might not be available.

For the Bayesian updating and combined transfer methods, the entire estimation context data set is not needed, but the coefficient estimates and their variances are required. This information may be available in documentation of the estimation context models.

For transfer scaling, only the coefficient estimates from the estimation context are required.

Amount of Information Used from the Estimation Context

The transfer method that makes use of the most data from the estimation context is the joint context estimation. This method uses the complete model estimation data set from the estimation context as well as the data from the application context to perform a joint model estimation.

The transfer methods that make the use of the next most data from the estimation context are the Bayesian updating and combined transfer methods. Both make use of both the coefficient estimates and their variances.

The transfer scaling method uses only the coefficient estimates from the estimation context, the same information used in simple transfer.

With these issues in mind, the four model transfer procedures are summarized below.

Transfer Scaling

If only the parameter estimates from the estimation context are available, then transfer scaling is the only method available to use with the data from the application context.

Because the relative values of parameters from the estimation context model remain unchanged, less information is used in this process than in the more labor intensive transfer processes described below, meaning that the validity of the model transfer is expected to be less as well.

Bayesian Updating

Bayesian updating can be performed when both parameter estimates and their variances are available for the estimation context. As demonstrated in the example shown in Section 3.3, the computations are fairly simple. As is the case with transfer scaling, the Bayesian updating process does involve estimation of a model using the data set for the application context, and then using the estimation results from both contexts to perform computations to produce the final estimates. Therefore, the level of expertise and level of effort required is only slightly higher than for transfer scaling.

Note that the variable definitions and model specifications must be identical for both the application and estimation contexts.

Combined Transfer Approach

The combined transfer approach uses both Bayesian updating and transfer scaling, and so the level of expertise required is about the same as for either Bayesian updating or transfer scaling. Because both of those processes are used, the level of effort is a bit higher than for either process by itself. But the level of information used from the estimation context is higher than in either Bayesian updating or transfer scaling.

Joint Context Estimation

A major difference between the joint context estimation transfer method and those described above is that the originally estimated model parameters from the estimation context are not used directly. Rather, the data set from which the estimation context model was estimated is used, along with the application context data set, to estimate a combined model using all of the data from both contexts.

Combining the data sets is not a trivial exercise, and so the level of effort and the level of expertise required is the highest among the transfer methods discussed here. Additionally, while the structure from the estimation context model could be used without change although that is not a requirement; additional variables, for example, that can be defined in the data sets could be introduced. Revising the model structure, however, would require more model estimation expertise.

2.5 Model Validation and Calibration

One of the reasons why model transfer has been a common practice is that a model with transferred parameters can be rapidly built and calibrated so that base year model results match observed data. For many years, this type of comparison was the main component of model "validation," and model parameters, including constants, were calibrated to obtain a better match between model results and observed data. This type of model calibration is easy to perform whether the original parameters are estimated or transferred; however, it is tempting to simply adjust parameters so that model results match observed data without truly trying to understand the reasons why the parameters needed to be adjusted.

Typically, calibrating model constants, whether alternative-specific constants or parameters of indicator variables representing specific segments of the population (for example, income level or area type indicators), has been done much more often than revising parameters that represent sensitivity to specific input variables (for example, time and cost coefficients). This makes sense since constants generally represent influences on travel behavior that are not captured by the model's input variables, and this is especially true in the case of transferred model parameters that may not reflect or reflect only in a limited way the observed behavior in the application context. It should be noted that while there are some published comparisons of model parameters showing some similarity among mode choice model parameter estimates in a variety of urban areas (see, for example, NCHRP Report 716, Tables 4.8, 4.11, and 4.14), there are some noticeable differences. So while an analyst may want to limit changes to these parameters, they are not "off limits" to calibration.

Model validation has always been considered a critical component of the model implementation process. But the concept of what constitutes a properly validated model has evolved, and simply matching observed base year data is no longer considered sufficient validation. As documented in the FHWA Travel Model Validation and Reasonableness Checking Manual, Second Edition (Cambridge Systematics, Inc., 2010), temporal validation, sensitivity testing, and

reasonableness and logic checks are now all considered important parts of the validation process. Indeed, the travel modeling profession continues to refine the definition of what is required in a complete and proper model validation effort.

The true values of model parameters are never definitively identified. They may be estimated based on data that constitute a sample of the population whose travel behavior is being modeled, but there is always error associated with the parameter estimates. While calibration of parameters can improve the model's ability to replicate certain observed results, it may reduce the errors in the parameter estimates. In fact, calibration can introduce predictive errors even as the match between modeled and observed results improves. In some cases, rather than addressing the errors in certain parameters, other "more correct" parameters may be adjusted instead (and made "less correct").

There is an additional source of error in transferred model parameters that is not present in locally estimated model parameters: Any differences in the behavioral responses or contextual factors between the application and estimation contexts may not be captured (although updating or joint context estimation procedures may mitigate this issue). It is important to recognize that model calibration may not reduce or eliminate these types of errors and can introduce further "compensating" errors.

While model validation is always essential, performing all components of validation, especially those beyond simply trying to match observed data, is especially important when dealing with transferred model parameters.

3.0 Examples of the Application of Model Transfer Methods

This chapter provides examples of the application of the five transfer methods introduced in this report. The examples presented assume that "Urban Area A" wants to develop a cross-classification trip production model for home based work (HBW) trips. It is assumed any household survey data for Urban Area A are inadequate for estimating this model using only this data set, nor is it possible to conduct a survey with sufficient data. So Urban Area A is considering transferring a model from elsewhere.

"Urban Area B," an area of similar size in the same state as Urban Area A, has conducted a household survey that was adequate for trip production model estimation and has estimated a cross-classification trip production model for HBW trips. *Table 3.1* shows this estimated model, which is segmented by number of workers and number of autos in the household. *Table 3.2* shows the sample size for the household survey for Urban Area B, and *Table 3.3* shows the variances for the estimated trip rates.

Table 3.1. HBW Trip Productions by Number of Workers by Number of Autos – Urban Area B

Autos	Workers			
	1	2	3+	Average
0	1.0	2.4	5.1	**0.7**
1	1.0	2.6	5.1	**0.9**
2	1.3	2.6	5.1	**2.0**
3+	1.3	2.6	5.1	**2.6**
Average	**1.1**	**2.6**	**5.1**	**1.5**

Table 3.2. Sample Sizes from Survey for Urban Area B

Autos	Workers				
	0	1	2	3+	Total
0	500	300	100	25	**925**
1	1,000	1,500	400	25	**2,925**
2	300	1,000	1,500	100	**2,900**
3+	50	200	350	150	**750**
Total	**1,850**	**3,000**	**2,350**	**300**	**7,500**

Table 3.3. Mean HBW Trip Rate Variances from Survey for Urban Area B

Autos	Workers		
	1	**2**	**3+**
0	2.00	4.00	5.00
1	0.10	0.05	0.20
2	0.20	0.01	0.05
3+	0.30	0.02	0.04

The following examples illustrate how the various transfer methods discussed in this report could be used to transfer (and update, as appropriate) the model from Urban Area B (the estimation context) to Urban Area A (the application context). (Note that, for simplicity, the HBW trip rate for households with zero workers is assumed to be zero in all cases.)

3.1 Simple Transfer

This method would be the only option if there were no possibility of obtaining any survey data for model estimation in Urban Area A. In this case, the parameters in Table 3.1 would be used.

3.2 Transfer Scaling

Use of transfer scaling assumes that Urban Area A is able to obtain a small household survey data set for its region. If this small sample shows an average of 1.6 home based work trips per household, the values in Table 3.1 would be multiplied by 1.07 (=1.6/1.5), as shown in *Table 3.4*.

Table 3.4. HBW Trip Productions by Number of Workers by Number of Autos for Urban Area A Using Transfer Scaling

Autos	Workers			
	1	**2**	**3+**	**Average**
0	1.1	2.6	5.6	**0.8**
1	1.1	2.8	5.6	**1.0**
2	1.4	2.8	5.6	**2.2**
3+	1.4	2.8	5.6	**2.8**
Average	**1.2**	**2.8**	**5.6**	**1.6**

As an additional example of transfer scaling, consider a simplified binary mode choice model estimated in Area A with utility functions:

$U_{Ea} =$ $\quad$ −0.02 (in-vehicle time) − 0.05(out-of-vehicle time) − 0.150(cost)

$U_{Et} =$ $\quad$ −1.0 − 0.02(in-vehicle time) − 0.05(out-of-vehicle time) − 0.150(cost)

Where:

$U_{Ea} =$ $\quad$ Utility of auto in the estimation context model

$U_{Et} =$ $\quad$ Utility of transit in the estimation context model

To develop the application context model with transfer scaling, a model is estimated, using data from the application context, with utility functions given by:

$U_{Aa} =$ B_{Aa} [−0.02 (in-vehicle time) −0.05(out-of-vehicle time) −0.150(cost)]

$U_{At} =$ $B_{A0t} + B_{At}$ [−0.02(in-vehicle time) −0.05(out-of-vehicle time) −0.150(cost)]

Where:

$U_{Aa} =$ Utility of auto in the application context model

$U_{At} =$ Utility of transit in the application context model

$B_{Aa} =$ Estimated scale for auto for the application context model

$B_{At} =$ Estimated scale for transit for the application context model

$B_{A0t} =$ Alternative-specific constant for transit in the application context model

Say that the estimated model parameters are $B_{Aa} = B_{At} = 1.5$ and $B_{A0t} = −0.5$. Then the utilities for the application context model are:

$U_{Aa} =$ −0.03 (in-vehicle time) −0.075(out-of-vehicle time) −0.225(cost)

$U_{At} =$ −0.5 −0.03(in-vehicle time) −0.075(out-of-vehicle time) −0.225(cost)

3.3 Bayesian Updating

To perform Bayesian updating, the HBW trip production model would be estimated using the small sample from Urban Area A. Assume that the sample sizes for this small survey are shown in *Table 3.5*, the estimated trip rates from this small sample are shown in *Table 3.6*, and the mean trip rate variances are shown in *Table 3.7*.

Table 3.5. Sample Sizes from Survey for Urban Area A

Autos	Workers				
	0	**1**	**2**	**3+**	**Total**
0	50	30	10	5	**95**
1	80	140	50	5	**275**
2	25	110	150	10	**295**
3+	10	25	30	20	**85**
Total	**165**	**305**	**240**	**40**	**750**

Table 3.6. HBW Trip Productions Estimated from the Small Survey Sample for Urban Area A

Autos	Workers			
	1	**2**	**3+**	**Average**
0	1.0	2.2	2.8	**0.7**
1	1.1	2.5	5.0	**1.1**
2	1.3	2.7	5.2	**2.0**
3+	1.4	3.0	5.2	**2.7**
Average	**1.2**	**2.7**	**4.9**	**1.6**

Table 3.7. Mean HBW Trip Rate Variances from Survey for Urban Area A

Autos	Workers		
	1	**2**	**3+**
0	5.00	25.0	50.0
1	2.00	1.00	1.20
2	5.00	0.50	2.00
3+	10.0	1.00	1.00

The updated trip rate parameter estimates can be obtained through Equation 3-1 (which was previously presented as Equation 1-3).

$$\hat{\theta}_{updated} = \frac{\left[\dfrac{\hat{\theta}_{prior}}{\left(\hat{\sigma}^2_{prior}\right)} + \dfrac{\hat{\theta}_{updating}}{\left(\hat{\sigma}^2_{updating}\right)} \right]}{\left[\dfrac{1}{\left(\hat{\sigma}^2_{prior}\right)} + \dfrac{1}{\left(\hat{\sigma}^2_{updating}\right)} \right]}$$

(3-1)

Where:

$\hat{\theta}_{updated}$ =   New estimated trip rate parameter for Urban Area A

$\hat{\theta}_{prior}$ =   Estimated trip rate parameter for Urban Area B (from household survey)

$\hat{\theta}_{updating}$ =   Estimated trip rate parameter for Urban Area A (from small survey sample)

$\hat{\sigma}^2_{prior}$ =   Estimated trip rate variance for Urban Area B (from household survey)

$\hat{\sigma}^2_{updating}$ =   Estimated trip rate variance for Urban Area A (from small survey sample)

For example, for the 0 auto-1 worker, combination, the trip rate computation is:

[(1.0/2.00) + (1.2/5.00)] / [1/2.00) + 1/5.00)] = 1.1.

Note that the if one were to compute all of the updated trip rates in the cross-classification, some of the updated estimates would remain the same to one decimal place because of the substantial precision of $\hat{\theta}_{prior}$ relative to $\hat{\theta}_{updating}$.

3.4 Combined Transfer

The Bayesian updating procedure above assumes that the behavioral model parameters in the populations of urban areas A and B are identical. Thus, the procedure essentially obtains updated parameters for the urban area A (application area) as the weighted average of the estimates of θ_{prior} and $\theta_{updating}$, where the weights correspond to the inverse of the variances of the estimated parameters.

In most cases, though, the assumption of identical behavioral model parameters between urban areas may be too strong, and there may be a real difference in the model parameters in the populations (sometimes labeled as transfer bias). In this case, one may obtain updated parameters as the minimum mean squared error estimate of θ_{prior} and $\theta_{updating}$. In the one-dimensional case that we will present here, as Equation 3-2, the equivalent of Equation 3-1 in the combined transfer approach is as follows:

$$\hat{\theta}_{updated} = \frac{\left[\dfrac{\hat{\theta}_{prior}}{\left(\hat{\sigma}^2_{prior} + \hat{\Delta}^2\right)} + \dfrac{\hat{\theta}_{updating}}{\left(\hat{\sigma}^2_{updating}\right)}\right]}{\left[\dfrac{1}{\left(\hat{\sigma}^2_{prior} + \hat{\Delta}^2\right)} + \dfrac{1}{\left(\hat{\sigma}^2_{updating}\right)}\right]},$$

(3-2)

where $\hat{\Delta} = (\hat{\theta}_{updating} - \hat{\theta}_{prior})$

For example, for the 2 auto-3+ worker combination, the computation is:

$[(5.1/(0.05+0.1*0.1)) + (5.2/2.00)] / [(1/(0.05+0.1*0.1)) + (1/2.00)] = 5.1.$

In this specific computation, the Bayesian and combined transfer approaches provide the same result up to the first decimal point. This is once again because of the substantial precision of $\hat{\theta}_{prior}$, but also because of the closeness of the estimates of $\hat{\theta}_{updating}$ and $\hat{\theta}_{prior}$. In general, the combined transfer and Bayesian updating procedures will not provide the same estimates, except in the special case when $\hat{\theta}_{updating} = \hat{\theta}_{prior}$ (when $\hat{\Delta} = 0$).

3.5 Joint Context Estimation

Joint context estimation involves the estimation of a joint estimation/application context model, and generally requires individual-level disaggregate data from both the estimation and application contexts though the amount of data for the application context can be lesser. In such estimations, a variety of specifications can be used and tested. For example, in a linear regression context, if all variables are allowed to have different effects in the estimation and application contexts, and the error variances are also allowed to be different between the two contexts, joint context estimation is tantamount to using only the application area data set (because no value is gained from using the estimation area data set).

However, the basis of joint context estimation is that one or more variable coefficients may be the same across the estimation and application areas, and so there is a gain in efficiency in

using data from both contexts. In addition, if the analyst is willing to assume *a priori* that a specific variable has the same effect in both the estimation and application contexts, it is fine if no data on that specific variable is available in the application context (because the coefficient can be estimated from the estimation data portion of the joint estimation, and can be transferred directly as the prevailing parameter for the application context).

Further, when data on a suite of independent variables are available from both the estimation and application contexts, the analyst can statistically test for which coefficient effects may be constrained to be the same, and which to let free, so that the model parameters for the application area retain uniqueness when needed and combine with the estimation area parameters when appropriate (there are many nuances about joint context estimation in the context of specific model structures such as linear regression and discrete choice models).

In the context of the cross-classification example, the easiest way to undertake joint model estimation is to translate the cross-tabulation into an equivalent linear regression. For illustration and presentation ease, consider a simplified household-level trip production 2x2 cross-classification model with autos (say only taking the values 0 and 1+) and number of workers (say only taking the values 0 and 1+) as the cross-classification variables. (The concept can be readily extended to the case of the example used for the earlier methods, which has more segments for autos and workers, but the explanation gets cumbersome and hence the simplification). The use of joint context estimation can be demonstrated through the use of an equivalent zero-intercept linear regression model estimated from the estimation context data set, as shown in Equation 3-3.

$$Trips = B_1 (D_{00}) + B_2 (D_{10}) + B_3 (D_{01}) + B_4 (D_{11}) \tag{3-3}$$

Where:

v =	number of autos (0 or 1+)
w =	number of workers (0 or 1+)
$Trips$ =	Total trips
D_{vw} =	1 if there are v autos and w workers, 0 otherwise
B_k =	Estimated coefficients, i.e., the trip rates

The joint estimation would bring the household level observations from both the estimation and application contexts together, with an assumption that the overall magnitude of the effects of unobserved independent variables on trip production rates is the same in the estimation and application contexts (econometrically speaking, the assumption is that the regression error variances are the same across the estimation and application contexts; this assumption can be relaxed, but leads to additional nuances in discussion). Then, a linear regression model set-up is developed as shown in Equation 3-4, this time using the combined data set from *both* the estimation and application contexts.

$$Trips = B_1 (D_{00}) + B_2 (D_{10}) + B_3 (D_{01}) + B_4 (D_{11}) + B_5 (X_{00}) + B_6 (X_{10}) + B_7 (X_{01}) + B_8 (X_{11}) \tag{3-4}$$

Where:

v =	number of autos (0 or 1+)
w =	number of workers (0 or 1+)
$Trips$ =	Total trips
D_{vw} =	1 if there are v autos and w workers, 0 otherwise

$E =$		1 if the observation belongs to the estimation context, 0 otherwise
$X_{vw} =$		$D_{vw} * E$
$B_k =$		Estimated coefficients

So, for example, when $X_{10} = 1$ for an observation in the combined data set, this means that the household in the observation has 1+ autos and zero workers *and* that the observation belongs to the estimation context data set.

With this set-up, B_1 provides the expected trip production rate for a 0 auto-0 worker household in the application area, B_2 provides the expected trip production rate for a 0 auto-1+ worker household in the application area, B_3 provides the expected trip production rate for a 1+ auto-0 worker household in the application area, and B_4 provides the expected trip production rate for a 1+ auto-1+ worker household in the application area. The remaining four coefficients provide the differences in expected trip production rates for the auto-worker combination categories between the estimation and application areas (thus, B_5 provides the difference in expected trip production rates between a 0 auto-0 worker household in the application and estimation areas, and similarly for the other three coefficients). So the trip rates for the estimation area are given by:

- 0 autos, 0 workers: $B_5 + B_1$
- 0 autos, 1+ workers: $B_6 + B_2$
- 1+ autos, 0 workers: $B_7 + B_3$
- 1+ autos, 1+ workers: $B_8 + B_4$

Now suppose that B_5 is statistically insignificant. This implies that X_{00} can be dropped from the linear regression, and one cannot reject the hypothesis that the expected value of trip production rate for 0 auto-0 worker households is the same between the estimation and application contexts (i.e., $B_5=B_1$). The analyst can also test for the statistical significance of X_{10}, X_{01}, and X_{11} and develop an optimal specification. This demonstrates the gain in efficiency from joint context estimation, as only those application context parameters that are statistically different from the estimation context parameters need be estimated, and joint context estimation provides a test for statistical significance of these differences.

4.0 References

Abdelwahab, W.M. (1991). Transferability of intercity disaggregate mode choice models in Canada, *Canadian Journal of Civil Engineering*, 18(1), 20-26.

Agyemang-Duah, K and F.L. Hall (1997). Spatial transferability of an ordered response model of trip generation, *Transportation Research*, 31A, 389-402.

Arentze, T.A., and H.J.P Timmermans (2000). ALBATROSS: A Learning-based Transportation-oriented Simulation System, EIRASS, Eindhoven University of Technology, The Netherlands.

Arentze, T., F. Hofman, H. van Mourik, and H. Timmermans (2002). Spatial transferability of the Albatross model system: Empirical evidence from two case studies, *Transportation Research Record 1805*, 1-7.

Atherton, T.J., and M.E. Ben-Akiva (1976). Transferability and updating of disaggregate travel demand models, *Transportation Research Record 610*, 12-18.

Bekhor, S. and C.G. Prato (2009). Methodological transferability in route choice modeling, *Transportation Research* 43B, 422-437.

Bhat, C.R. (2005). A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions, *Transportation Research Part B* 39(8), 679-707.

Bhat, C.R. (2008). The multiple discrete–continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions, *Transportation Research Part B* 42(3), 274-303.

Bhat, C.R., K.G. Goulias, R.M. Pendyala, R. Paleti, R. Sidharthan, L. Schmitt, and H.H. Hu, (2013). A household-level activity pattern generation model with an application for Southern California, *Transportation*, 40(5), 1063-108.

Bowman, J.L., M. Bradley, and J. Castiglione (2013). *Making Advanced Travel Forecasting Models Affordable Through Model Transferability*, Final report submitted to the Federal Highway Administration under BAA DTFH61-10-R-00013.

Cambridge Systematics, Inc. (1996). *Travel Survey Manual*. Prepared for U.S. Department of Transportation and U.S. Environmental Protection Agency.

Cambridge Systematics, Inc. (2010). *Travel Model Validation and Reasonableness Checking Manual, Second Edition*. Prepared for Federal Highway Administration.

Cambridge Systematics, Inc., Vanasse Hangen Brustlin, Inc., Gallop Corporation, Chandra R. Bhat, Shapiro Transportation Consulting, LLC, and Martin/Alexiou/Bryson, PLLC. (2012). *NCHRP Report 716, Travel Demand Forecasting: Parameters and Techniques*, Transportation Research Board.

Caldwell, L.C. and M.J. Demetsky (1980). Transferability of trip generation models, *Transportation Research Record 751*, 56-62.

Cotrus, A.V., J.N. Prashker, and Y. Shiftan. (2005). Spatial and temporal transferability of trip generation demand models in Israel, *Journal of Transportation and Statistics*, 8, 1-25.

Everett, J.D. (2009). An investigation of the transferability of trip generation models and the utilization of a spatial context variable, PhD dissertation, University of Tennessee, 2009, http://trace.tennessee.edu/utk_graddiss/47.

Galbraith, R.A. and D.A. Hensher (1982). Intra-metropolitan transferability of mode choice models, *Journal of Transport Economics and Policy*, XVI, 7-29.

Greaves, S.P. and P.R. Stopher (2000). Creating a Synthetic Household Travel/Activity Survey - Rationale and Feasibility analysis, *Transportation Research Record 1706*, 82-91.

Gunn, R.F., M.E. Ben-Akiva, and M. Bradley (1985). Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Models, *Transportation Research Record 1037*, 21-30.

Guo, J. and Bhat, C.R. (2007). Population synthesis for microsimulating travel behavior, *Transportation Research Record 2014*, 92-101.

Hansen, S. (1981). In favor of cross-cultural transferability of travel demand modeling, in Stopher, P.R., Meyburg, A.H., and Brog, W. (eds.), *New Horizons in Travel Behavior Research*, D.C. Heath and Co., Lexington, Massachusetts, 637-651.

Karasmaa, N. (2001). The spatial transferability of the Helsinki Metropolitan Area mode choice models, presented at the 5th workshop of the Nordic research network on modeling transport, land-use, and the environment, 1-24.

Karasmaa, N. (2003). The transferability of travel demand models: An analysis of transfer methods, data quality and model estimation, PhD dissertation, Helsinki University of Technology Transportation Engineering Publication 106, http://lib.tkk.fi/Diss/2003/isbn9512268574/isbn9512268574.pdf.

Karasmaa, N. (2007). Evaluation of transfer methods for spatial travel demand models, *Transportation Research* 41A, 411-427.

Kawamoto, E. (2003). Transferability of standardized regression model applied to person-based trip generation, *Transportation Planning and Technology*, Vol. 26, No.4, 331-359.

Koppelman, F.S., G-K Kuah, and C.G. Wilmot (1985). Transfer model updating with disaggregate data, *Transportation Research Record 1037*, 102-107.

Koppelman, F. and E.I. Pas (1986). Multidimensional choice model transferability, *Transportation Research Part B*, Vol. 20B, No. 4, 321-330.

Koppelman, F.S. and J. Rose (1983). Geographic transfer of travel choice models: Evaluations and procedures, *Optimization and Discrete Choice in Urban Systems*, 272-309.

Koppelman, F.S. and C.G. Wilmot (1982). Transferability analysis of disaggregate choice models, *Transportation Research Record 895*, 18-24.

Le Vine, S., A. Sivakumar, and J. Polak (2010). Toward the implementation of an activity-based travel demand model for London: Development of the household activity scheduling module, presented at the 42nd conference of the Universities' transport study group, Plymouth, United Kingdom.

Lin, J. and L. Long (2007). Transferability of household travel data across neighbourhood types and geographic areas using NHTS, Technical Report, Civil and Materials Engineering Department, University of Illinois at Chicago.

Mamun M.S. and M.A. Sabbir (2012). Spatial transferability of mode choice model, *International Journal of Engineering Research and Development*, Vol. 2, No. 2, 65-71.

McComb, L.A. (1986). Analysis of the transferability of disaggregate demand models among ten Canadian cities, *Tribune Des Transports*, Vol. 3-1, 19-32.

Miller, E.J. and M.J. Roorda (2003). Prototype model of household activity-travel scheduling, *Transportation Research Record 1831*, 114-121.

Mohammadian, A. and Y. Zhang (2007). Investigating the Transferability of National Household Travel Survey Data, *Transportation Research Record 1993*, 67-79.

Nowrouzian, R. and S. Srinivasan (2012). Empirical analysis of spatial transferability of tour-generation models, *Transportation Research Record 2302*, 14-22.

Ortuzar, J. and L. Willumsen (2011). *Modelling Transport, 4th Edition*, John Wiley & Sons, Inc.

Pendyala, R.M., C.R. Bhat, K.G. Goulias, R. Paleti, K.C. Konduri, R. Sidharthan, H.H. Hu, G. Huang, and K.P. Christian (2012). Application of Socioeconomic Model System for Activity-Based Modeling, *Transportation Research Record, 2303*, 71-80

Pinjari, A.R., and C.R. Bhat (2011). Activity-based travel demand analysis in *A Handbook of Transport Economics*, Chapter 10, pp. 213-248, edited by A. de Palma, R. Lindsey, E. Quinet, and R. Vickerman, Edward Elgar Publishing Ltd.

Reuscher, T.R., R.L. Scmoyer, and P.S. Hu (2002). Transferability of Nationwide Personal Transportation Survey Data to Regional and Local Scales, *Transportation research Record 1817*, 25-32.

Sikder, S. and A.R. Pinjari (2013). Spatial transferability of person-level daily activity generation and time-use models: Empirical assessment, *Transportation Research Record 2343*, 95-104.

Stopher, P.R., P. Bullock, and S. Greaves (2003). Simulating household travel survey data: Application to two urban areas, Presented at the 82nd annual meeting of the Transportation Research Board, Washington, D.C..

Stopher, P.R. and P. Jones (2001). Developing standards of transport survey quality, Keynote Paper prepared for the International Conference on Transport Survey Quality and Innovation: How to Recognize It and How to Achieve It, Kruger Park, South Africa.

Talvitie, A. and D. Kirshner (1978). Specification, transferability and the effect of data outliers in modeling the choice of mode in urban travel, *Transportation 7*, 311-331.

Vovsha, P., S. Gupta, J. Freedman, W. Sun, and V. Livshits (2012). Workplace choice model: comparison of spatial patterns of commuting in four metropolitan regions, Presented at the 91st annual meeting of the Transportation Research Board, Washington, D.C.

Watson, P. L., and R.B. Westin (1975). Transferability of Disaggregate Mode Choice Models, *Regional Science and Urban Economics,* 5, 227-249.

Wilmot, C.G. (1995). Evidence of Transferability of Trip Generation Models. *Journal of Transportation Engineering 9*, 405-410.

Zhang, Y. and A. Mohammadian (2008a). Microsimulation of household travel survey data, presented at the 87th annual meeting of the Transportation Research Board, Washington, D.C.

Zhang, Y. and A. Mohammadian (2008b). Bayesian updating of transferred household travel survey data using MCMC simulation with Gibbs sampler, presented at the 87th annual meeting of the Transportation Research Board, Washington, D.C.

Appendix A Synthesis of Research on Transferability of Trip-Based Model Components

A.1 Introduction

This appendix provides a summary of a synthesis of research into the spatial transferability of travel model parameters. The complete synthesis is provided in Appendices C (for trip based models) and D (for activity based models).

The synthesis for trip-based models, presented in Section A.2, is categorized by the specific step within the four step modeling process. Summaries of research into the transferability of the parameters of trip generation, trip distribution/destination choice, mode choice, and trip assignment (route choice) are presented along with overall conclusions from the research on transferability for each component. Section A.3 presents a summary of a synthesis of transferability studies in the context of activity-based models. Four such studies were identified, one from the Netherlands from 2002 and three studies done in the U.S. over the past two years. Section A.4 presents a summary of data transferability studies. Conclusions from the literature review on model transferability are presented in Section A.5.

A.2 The Model Transferability Approach in the Context of the Four-Step Model

This section summarizes the research that has been performed regarding the model transferability of the parameters of four-step travel models. Specifically, there has been some research into the following types of model parameters:

- **Trip generation** – Eight studies in the literature were identified that have examined spatial transferability in the context of trip generation, published from 1980 to 2009. They all focused on trip productions; none examined trip attractions.
- **Trip distribution** - Only one study was identified that examined spatial transferability of trip distribution (destination choice), a nested logit joint mode-destination choice model from Finland in 2007.
- **Mode choice** - A total of 13 studies were identified in the literature since 1975, but only one study since 2001. All of the published studies focused on work trip mode choice, with no research found on non-work travel.
- **Route Choice** – One study was identified that examined spatial transferability in the context of route choice. This paper focused on testing the transferability of path generation techniques used within the traffic assignment step of the four step process. The study, however, did not examine transfer of volume-delay function parameters. Indeed, while it seems that transferring volume-delay function parameters is routine in the field, we could not identify any study assessing the validity of such a transfer.

The literature on each of these four model components is described in detail in Appendix C.

A.2.1 Trip Generation

There have been several studies in the literature that have examined spatial transferability in the context of trip generation, as discussed in more detail in Section C.1. Table A.1 summarizes the results of the research on trip generation models. The results of the spatial transferability of trip generation models have been rather mixed. Two studies indicated that transferability was

valid under the conditions of their research while two others indicated valid transferability if scaling or updating was used. Two studies indicated valid transferability only under certain conditions while the remaining two studies rejected transferability while finding that the predictive ability of the transferred models was still reasonable. It is significant that all studies where updating or scaling of parameters was examined indicated that this updating improved the transferability of the models.

Unfortunately, it is difficult to synthesize these results to provide any conclusive guidelines for transferability because of the different variable specifications used, the different dependent variables adopted (some of which are at the person level and some at the household level), the different trip purposes considered, the different geographic and temporal periods of the studies, the different model forms employed, and the different independent variable specifications in the models. Besides, most of the trip generation studies have not controlled for land-use, accessibility, and transportation system characteristics when studying spatial transferability. A study by Lin and Long (2007) highlights this issue, and suggests that including these additional variables can enhance spatial transferability. The study by Everett (2009) reinforces the notion that even simple ways of characterizing the spatial context (such as using an area type measure) can improve transferability rather considerably.

Trip generation transferability will be improved with better variable specifications, a disaggregate-level analysis at the household or person level rather than at an aggregate zonal level, a model structure that reflects the ordinal and discrete nature of trips, and a transfer approach that involves transfer scaling of coefficients.

In general, however, it appears safe to say that trip generation transferability will be improved with better variable specifications, a disaggregate-level analysis at the household or person level rather than at an aggregate zonal level, a model structure that reflects the ordinal and discrete nature of trips, and a transfer approach that involves transfer scaling of coefficients. In the context of transfer scaling, it should be pointed out that most trip generation analyses of transferability have focused on a simple transfer approach, rather than on a transfer approach that combines some limited information from the application context to update the estimation context relationships for use in the application area.

Another important issue to note in the earlier trip generation studies is that they have all been trip-based, and do not consider trip chaining and the more general interdependence among trips of individuals. Thus, separate models for home based trips and non-home based trips are developed, without any consideration of the dependence between these categories of trips. Consequently, differences in trip chaining tendencies from one area to another, or from one time period to another, could immediately result in findings of poor trip generation transferability, even if models of the number of stops (out-of-home activity participations) have good transferability. This is an issue that we will return to in the context of activity-based frameworks for travel demand modeling (Section A.3).

Table A.1. Summary of Trip Generation Transferability Literature

Authors	Year	Model Type	Analysis Level	Trip Purposes*	Validity of Transfer?	Scaling/Updating
Caldwell and Demetsky	1980	Linear regression	Household, zone	HBW, HBNW, NHB	Yes, if similar cities	Not considered
Gunn et al	1985	Logit	Person	HBSh, HBPB	Yes, if scaled	simple uniform
Koppelman and Rose	1983	Linear regression	Household		Rejected (though predictive ability reasonable)	Not considered
Wilmot	1995	Linear regression	Household		Influenced by model specification, income	Improves transferability
Agyemang-Duah and Hall	1997	Ordered response	Household	HBSh	Simple transfer works OK	Improves transferability
Kawamoto	2003	Linear regression	Person	All HB	Yes (with updating)	Scaling variable-by-variable
Cotrus et al	2005	Linear regression, Tobit	Person		Rejected (though predictive ability reasonable)	Not considered
Everett	2009	Cross-classification	Household	HBW, HBNW, NHB	In some cases only	Not considered

* Trip purpose abbreviations:
    HBW – Home based work
    HBNW – Home based non-work
    HBSh – Home based shopping
    HBPB – Home based personal business
    NHB – Non-home based

A.2.2 Trip Distribution/Destination Choice

The literature on transferability of trip distribution/destination choice is relatively limited, and has been focused on temporal transferability, not spatial transferability that is the focus of this report. Only one study, by Karasmaa (2007) was identified, which is summarized in Section C.2.

The study by Karasmaa suggests that trip distribution/destination choice models transfer reasonably well over time in terms of predictive fit and forecast errors, though the behavioral parameters do show instability. However, there seems to be no clear indication of which type of updating method would be best suited for what type of transfer contexts. Of course, the trip-based nature of earlier studies ignores issues of destination linkages of stops, and identifies the

need for transferability analysis in the context of tour-based and activity-based frameworks for travel demand modeling.

A.2.3 Mode Choice

Mode choice has been perhaps one of the most studied trip dimensions in the context of spatial transferability. Several studies were identified and are discussed in Section C.3. There does not appear to be any published literature on transferability for non-work mode choice. Table A.2 summarizes the results of the research on mode choice models.

The literature on work mode choice transferability in space is mixed. However, some general conclusions are as follows:

- Coefficient equality between the estimation and application contexts should not be used as the sole yardstick for assessing transferability; rather disaggregate and aggregate prediction measures that provide an assessment of the amount of information provided by the transferred model should also be considered.
- Transferability improves with improved variable specification.
- Model updating leads to a substantial improvement in transferability relative to a simple model transfer, even if the updating is simply a constants-only updating to reflect the aggregate mode shares in the application context.
- There is no consensus regarding which update method is best, and it would behoove the analyst to consider all the updating procedures that are possible to assess which performs best in any given context.

It is interesting to note that most of the mode choice transferability studies have been undertaken in the 70s and 80s, with much fewer studies undertaken recently. Also, while there has been substantial focus on tour-based mode choice and activity-based modeling in general in the past two decades, there has been relatively limited analysis of transferability in the context of tour-based mode choice modeling.

Table A.2. Summary of Mode Choice Transferability Literature

Authors	Year	Model Type	Modes	Validity of transfer?	Scaling/Updating
Watson and Westin	1975	Binary logit	Auto, train	In some cases	Not considered
Atherton and Ben-Akiva	1978	Multinomial logit	Drive alone, shared ride, transit	Yes	Bayesian works best, but little improvement
Talvitie and Kirshner	1978	Multinomial logit	Drive alone, shared ride, transit	Rejected	Not considered
Galbraith and Hensher	1982	Binary logit	Auto, rail	Rejected, but predictive ability OK	Bayesian updating improves model
Koppelman and Wilmot	1982	Multinomial logit	Drive alone, shared ride, transit	Rejected, but predictive ability OK	Not considered
Koppelman and Rose	1983	Multinomial logit	Drive alone, shared ride, transit	Rejected	Updating improves transferability
Koppelman et al	1985	Multinomial logit	Drive alone, shared ride, transit	Rejected	Updating improves transferability
Gunn et al	1985	Nested logit	Drive alone, shared ride, transit	Rejected	Updating improves transferability
McComb	1986	Multinomial logit	Drive alone, shared ride driver, shared ride passenger, transit	Yes, for similar cities	Not considered
Koppelman and Wilmot	1986	Multinomial logit	Drive alone, shared ride, transit	Rejected, but better specifications improve transferability	Not considered
Koppelman and Pas	1986	Multinomial and nested logit		Yes	Not considered
Abdelwahab	1991			Poor	Transferability poor even with updating
Karasmaa	2001				Joint context generally best
Mamun and Sabbir	2012			Yes, but only aggregate shares compared	Not considered

A.2.4 Route Choice

Only one study of route choice was identified (see Section C.4). There is a clear need for further investigation of transferability in a route choice context to make any conclusions.

A.3 Spatial Transferability of Activity-Based Travel Models

This section summarizes a synthesis of research into the transferability of travel model parameters in the context of activity-based models. Over the past few decades, there has been an increasing interest in travel demand management strategies, such as mixed land-use development, parking pricing, and congestion pricing, all of which attempt to change land-use and transport service characteristics to influence individual travel behavior and control aggregate travel demand. The evaluation of such demand management strategies using travel demand models places more emphasis on the realistic representation of behavior to accurately reflect traveler responses to management policies. This realization has led to the consideration of a fundamental behavioral paradigm referred to as an activity-based approach to travel demand modeling (see Pinjari and Bhat (2011) for a recent review).

The fundamental difference between the trip-based and activity-based approaches is that the former approach directly focuses on "travel participation behavior" as the decision entity of interest, while the activity-based approach views travel as a demand derived from the need to pursue activities and focuses on "activity participation behavior." The underlying philosophy of the activity-based approach is to better understand the behavioral basis for individual decisions regarding participation in activities in certain places at given times, and hence the resulting travel needs. At a fundamental level, therefore, the activity-based approach emphasizes the point that the needs of the households are likely to be translated into a certain number of total activity stops by purpose followed by (or jointly with) decisions regarding how the stops are best organized.

In this regard, the activity-based approach uses "tours" as the basic element to represent and model travel patterns. Tours are chains of trips beginning and ending at a same location, say, home or work. The tour-based representation helps maintain consistency across, and capture the interdependency (and consistency) of the modeled choice attributes among, the activity episodes (and related travel characteristics) undertaken in the same tour. This is in contrast to the trip-based approach that considers travel as a collection of "trips", each trip being considered independent of other trips.

Further, unlike in the trip-based approach where time is included as a "cost" of making a trip and a day is viewed as a combination of broadly defined peak and off-peak time periods, the activity-based approach views individuals' activity-travel patterns as a result of their time-use decisions within a near-continuous time domain. At the same time, the activity-based approach also recognizes interactions among household members, which leads to the accommodation of linkages among trips of household members.

Overall, using microsimulation techniques, activity-based models predict entire activity-travel patterns at the level of individuals (while recognizing temporal/spatial constraints across individuals of a household due to joint activity participations activities to serve passengers). Such a methodology ensures a realistic and consistent prediction of activity-travel patterns, which should lead to the better aggregate prediction of travel flows on the network in response to demographic changes or policy scenarios. Thus the activity-based models (ABMs) are well equipped to forecast the longer-term changes in travel demand in response to the changes in the socio-demographic composition and the activity-travel environment of urban areas, as well as in response to land-use and transportation policies.

While, from a conceptual standpoint, the strengths of the activity-based modeling system are clear, the claim of improved behavioral basis should imply that ABMs should be more spatially transferable than trip-based models. To examine this issue, a few recent studies have begun to examine transferability in an ABM context. These studies are discussed in detail in Appendix C.

For ease in exposition, we discuss each ABM transferability study by positioning it along three dimensions:

1. The activity dimensions examined (including the overall model background and the specific modeling structures used for the ABM components);
2. The data context (including the estimation and application contexts, and the explanatory variables used); and
3. The transferability approach and results (including the metrics used and the conclusions).

In this section, we do not discuss efforts that simply transfer models from one location to another, without undertaking any quantitative evaluation (such as those by Vine et al., 2010 and Vovsha et al., 2012). But, very briefly, Vine et al. (2010) transfer model components of an ABM labeled as TASHA (Miller and Roorda, 2003) estimated in Toronto to London, assuming that the model components are transferable. Vovsha et al. (2012) informally compare estimation results from four locations for work location choice (including the binary choice of working from home versus working away from home, and a multinomial logit choice of the traffic analysis zone of work given it is away from home). The four locations are San Diego, Tucson, Phoenix, and Chicago. Vovsha et al. indicate that the work location coefficients appear to be qualitatively quite different across regions, and recommend local data estimation for the work location choice model.

For the synthesis, four studies were examined, three in the U.S. and one from Europe. The three American studies used different ABM structures, but all used data from the 2009 NHTS, and all were based in Florida and/or California. Table A.3 summarizes the results of the research on ABMs. None of the four studies is definitively conclusive about the transferability of ABM parameters; all can be classified as having mixed results.

The number of transferability studies in the context of activity-based models has been very limited, and it is difficult to make unequivocal conclusions at this time given that we have just about scratched the surface on this topic. There is clearly a need for many more transferability studies (and using a more diverse spatial setting than regions drawn only from California and/or Florida) before a clear picture of the effectiveness and validity of transferability can be obtained. With that said, it does seem that there are some consistent findings from the very limited number of ABM transferability studies undertaken so far.

- The updating constants method provides much improved transferability results when viewed from the perspective of data fit at both a disaggregate level (predicted individual-level choices) and an aggregate level (predicted choice shares). However, this improvement does not necessarily translate to an improvement in representing the responsiveness to exogenous variables in models of activity generation and sequencing (these are generally the higher-level models in activity-based platforms that tend to be less sensitive to the spatial and travel environment contexts compared to mode and location choice decisions). Given that transferred models are used primarily to examine policy scenarios and/or changes in demographics over time, there is evidence so far that constants updating may not provide as much benefit in the activity generation and sequencing components of the activity-based travel models as in other model components.
- Closely related to the first point is that models that are associated with social organization (activity generation and sequencing) appear to be more transferable in

general than those that relate to spatial organization (travel mode and location choice). Thus, models associated with work durations, timing of activity episodes and their durations, auto-ownership, and activity participations and sequencing generally are more transferable than mode and location choice decisions. A practical implication is that much more effort may need to be expended to update parameters (with local samples) when transferring mode and destination choice model components than when transferring other components of an ABM.

- There is consistent evidence thus far, even if in the very limited spatial context of the current ABM transferability studies, that transferability between regions within a state is much more appropriate than transferability across state boundaries. Within a state, urban-to-urban transfers appear much more effective than urban-to-non-urban transfers, which highlights, as in trip-based models, the importance of area type in making transfers.

- Further, even within a state, there is evidence of asymmetry in transferability, so that the typical practice of identifying two regions that appear to be "similar" based on, for example, population size, spatial context, or other demographics may not be sufficient to guarantee good transferability in both directions. Additional examination of the attributes that contribute to this asymmetry would be particularly helpful in selecting regions from which to transfer a model.

- Transferring ABMs from a comparable region with a large sample may be better than estimating ABMs from a much smaller local sample. This is particularly so because there is some evidence in the results that introducing a rich specification of individual-related variables that capture taste variations and preferences results in better transferability than models that ignore such effects. And estimating stable coefficients for specific segment groups (such as low income or high income or carless households) implies the need for a large overall sample size. But there is still value in having a small local sample or aggregate data in the application context, so that constant updating can be pursued as needed and especially in the ABM components associated with mode and location choice.

Table A.3. Summary of ABM Transferability Literature

Authors	Year	ABM Studied	Location	Dimensions	Validity of transfer?	Scaling/ Updating
Arentze et al	2002	ALBATROSS	Netherlands	Mode choice, activity location, accompaniment, sequencing of activities, associated travel linkages	"Encouraging," "substantial evidence"	Not considered
Nowrouzian and Srinivasan	2012	Original	Florida	Tour generation by purpose/ accompaniment/ flexibility	Mixed results, asymmetry noted	Not considered
Sikder and Pinjari	2013	Original (MDCEV)	California, Florida	Daily activity generation, time use	Mixed results, asymmetry noted	Constants updating improved results
Bowman et al	2013	DaySim	California, Florida	All ABM components	Mixed results, comparable regions better	Not considered

A.4   Synthesis of Data Transferability Studies

Data transferability studies can be used for any step of the trip-based system or for all steps at once. They are discussed separately here because, unlike the model transferability studies, they do not categorize well into each step of the four step process.

To provide context, in the typical model transferability approach, the emphasis is on transferring model coefficients from the estimation context to the application context. Then, predictions are undertaken in the application context based on the transferred model parameters. The data transferability approach is a variant of this model transferability approach, in which households in the estimation context are clustered into a limited number of classes based on sociodemographics and built environment variables of the household's residence, and the travel profiles of these households are transferred to the application context by synthesizing the population of households there, clustering the households based on the same demographic/built environment factors as in the estimation context, and assigning to each application area household a draw from the travel profiles of the estimation area households in the same cluster (see Mohammadian and Zhang, 2007 and Stopher et al., 2003).

The structure of the data transferability approach typically involves five steps:

1. In the first step, some form of clustering analysis is used to *group households into a limited number of classes* based on sociodemographics and built environment variables of the household's residence. The number and characterization of clusters either varies by the travel characteristic under consideration (for example, the trip production rate for work may involve a different clustering scheme than the trip production rate for shopping, such as in Greaves and Stopher, 2000 and Stopher et al., 2003) or may be based on a single overarching clustering scheme that is fixed regardless of the travel characteristic of interest (see, for example, Mohammadian

and Zhang, 2007). This clustering analysis is undertaken using a national or other large scale survey data set such as the National Household Travel Survey (NHTS).

2. The second step entails *developing distributions for the dependent variable of interest within each cluster* identified in the first step. The distributions can be in the form of an empirical frequency distribution based on the large survey data set (see Greaves and Stopher, 2000) or in the form of a fitted continuous distribution that tracks the empirical distribution from the large survey data set well (see Zhang and Mohammadian, 2008b).

3. The third step involves *generating a synthetic sample* of a predefined size or even the entire population for the application (or transfer) location. This can be based on the known aggregate demographic and built environment characteristics for the application area, which can be combined with small disaggregate-level samples collected in the application context (for example, a small sample survey conducted in the application context or the small application area sample available from the Public Use Micro-data Sample (PUMS) and other similar sources in other countries). This approach, traditionally referred to as synthetic population generation, is now quite well established (see Guo and Bhat, 2007, and Pendyala et al., 2012).

4. The fourth step involves *assigning households* synthesized in the application context to one of the clusters identified in the first step.

5. The fifth step entails *drawing a travel value realization* for each travel characteristic of interest for each household in the application context from the associated travel characteristic distribution of the cluster to which the household is assigned. In this process of transferring travel characteristics, Zhang and Mohammadian (2008b) have an intermediate step between the fourth and fifth steps where they use a Bayesian procedure to update the parameters representing the cluster-wise travel characteristics distribution from the large scale survey using a smaller scale region-specific (i.e., local area) sample.

Overall, the "data transferability" approach discussed above implicitly develops a relationship between the travel characteristic of interest and independent variables using statistical factor/cluster analysis technique on a large scale survey, and then transfers this relationship to the application area to directly predict travel characteristics. Any updating is undertaken at the direct level of the travel characteristic in this approach.

A.5   Transferability Research Results

Unfortunately, the research summarized in this appendix provides insufficient information to determine the best estimation context for any particular case. The information that is provided by this research is summarized below. Note that many of the studies cited in Appendices C and D used data from only two contexts, and so conclusions about the quality of different types of estimation contexts were not possible in these cases.

- **Caldwell and Demetsky** (1980) concluded that trip generation models can be transferred between cities, at least as long as care is taken in selecting "similar" cities. "Similar" cities are implicitly defined in the study as those with similar household size, household auto ownership levels, and per capita income. However, data from only three relatively small cities in Virginia was used in the study.

- **McComb** (1986) concluded that that mode choice model parameter equality cannot be rejected between Canadian cities of similar socioeconomic make-up, size, and transportation system quality (such as Edmonton and Winnipeg, and Calgary and Winnipeg). However, coefficient equality was rejected for cities that are very different in character (such as Toronto and Winnipeg and Ottawa and Winnipeg). However, this study did not develop clear objective measures of what constitutes similarity in socioeconomic make up or transportation system quality.

- **Everett** (2009) examined the regional and spatial contexts influencing the spatial transferability of household-level trip production models using data drawn from 11 distinct metropolitan regions in two states, Ohio and Tennessee. The research results indicated that, within the same state, transferability of trip rates appeared to be valid for the small MPOs. Transferability results across states for all urban area sizes were mixed. Across metropolitan area sizes, the results indicated good transferability in Ohio, but not Tennessee.

- **Nowrouzian and Srinivasan** (2012) examined transferability assessments in the context of tour generation models within an activity-based approach. The study used household survey data from three regions in Florida: the Northeast (NE) region around Jacksonville, the Southeast (SE) region around Miami, and the Tampa Bay (TB) region. While the study generally found greater transferability between the NE and SE regions than between the TB region and either of the other two, it was not clear what characteristics of the regions led to these results—why, for example, the SE region provided a better estimation context than the TB region did for an application context of the NE region.

- **Sikder and Pinjari** (2013) examined the spatial transferability of person-level daily activity generation and time-use models within an activity-based approach. This study also used data from Florida, for five regions/area types: Southeast Florida, Central Florida, Tampa Bay, District 1 urban areas (D1U) comprising a group of less urbanized counties, and rural Florida. Additionally, inter-state transferability was studied by using the data from Florida for estimation and transfer to California, and using data from California for estimation and transfer to Florida. The results revealed better transferability between the urban areas than from the urban areas in Florida to the D1U area or to rural Florida. Transferability between states was found to be poor.

- **Bowman et al.** (2013) conducted a comprehensive analysis of transferability in the context of activity-based models. The study used data from the 2009 National Household Travel Survey (NHTS) "add-on" program drawn from four regions in California (Fresno, Northern San Joaquin Valley, Sacramento, and San Diego) and two regions in Florida (Jacksonville and Tampa). This study found that transferring ABMs from a "comparable" region with a large sample may be better than estimating ABMs from a much smaller local sample and that, in general, ABM components are more transferable between regions within a state than between regions in different states. However, Tampa Bay variable coefficients seemed to be rather different from those of the coefficients from other regions, including Jacksonville. For unknown reasons, it seems that there is something different about the Tampa Bay region compared to other regions even within Florida.

It is interesting that both Nowrouzian and Srinivasan and Bowman et al. found that, for unknown reasons, the Tampa Bay region's model parameters were not transferable to other regions in Florida. This finding points out the risk inherent in transferring parameters even when the estimation context appears to be similar to the application context.

Appendix B   Model Transferability Measures

Many measures of transferability have been suggested in the literature (as discussed in the appendices). Broadly speaking, transferability may be assessed by examining the performance of an estimated model in a specific region when applied to another region (sometimes referred to as application-based transferability analysis) or by estimating model coefficients based on observed choices and relevant exogenous variable data from two different regions and testing whether or not the coefficients on variables are statistically different between the two regions (estimation-based transferability analysis).

Almost all earlier studies have used an application based transferability analysis method, though Bowman et al (2013), for example, recently used the latter estimation-based method. In this latter method, the typical approach is to combine the data sets from the estimation and application contexts and estimate two separate sets of models: (1) A model that holds the coefficients on the variables to be the same across the two regions (estimation and application), and (2) Another model that keeps the variables from the first model, and adds another set of variables specific to the application region. That is, if there is a variable X in the first model, the second model has another variable (X*Application region dummy variable), where the "Application region dummy variable" takes a value '1' if the observation corresponds to the application region and '0' if it is from the estimation region. The coefficients on these additional variables in the second model provide the "difference" effect between the estimation and application contexts. The t-statistics of these variables immediately provide a sense of whether or not the coefficients are different, while a traditional F-test (for linear regression type models) or a likelihood ratio test (if the models are estimated using the maximum likelihood approach) provide an overall sense of whether or not there is a difference in the models between the two contexts.

This estimation-based transferability analysis method has the advantage (over the application-based transferability analysis method) of rigorously testing the equality of individual coefficients (and therefore ostensibly of transferability of responsiveness behavior to specific variables) rather than just an overall sense of model transferability. But it also constitutes a rather rigorous test and has an expectation of high data quality and consistency in how the variables were created/collected in the estimation and application contexts.

In this section, we discuss the many transferability measures that have been developed within the application-based method for transferability (because the estimation-based transferability analysis method entails only the usual likelihood ratio test). Two groups of measures are distinguished in the application-based method: **disaggregate** and **aggregate**. Further within each of these two groups, the measures can be based on **fit** or based on **policy response**.

The fit measures, broadly speaking, examine how well the model from the estimation context is able to predict the outcomes of interest in the application context (compared to the actual observed outcomes). The policy response measures examine, on the other hand, how well the model from the estimation context is able to predict the effect (on the dependent variable of interest) of a change in one or more explanatory variables (compared to the corresponding change that is predicted by a model estimated in the application context itself based on observations from the application context). In this latter "policy response" approach, if the ratio of the predicted changes (in the application context) from the model estimated from the estimation region and from the model estimated in the application region is close to 1.00 for many policy response scenarios, it is taken as a measure of good transferability. Of course, there is some subjectivity on how close the ratio should be to one to be considered as an indication of good transferability, as well as how many of the tested scenarios should meet the

ratio standard established for good transferability (see, for example, Sikder and Pinjari, 2013). The remainder of this section focuses on disaggregate and aggregate transferability measures within the class of fit measures.

B.1. Disaggregate Transferability Measures

Disaggregate transferability refers to the ability of the model from the estimation context (i.e., the transferred model) to predict the individual-level outcomes in the application context. Since most of the relevant models in the travel demand literature tend to be non-linear in nature and are estimated using maximum likelihood methods, the disaggregate transferability fit is based off the transfer log-likelihood. Accordingly, the common disaggregate fit measures of transferability include the transferability test statistic (*TTS*), the transfer index (*TI*), and the transfer rho-squared value.

The *TTS* tests the hypothesis that the transferred model is statistically equivalent to a model estimated in the application context. It takes the following form:

$$TTS = -2[L_i(\beta_j) - L_i(\beta_i)],$$

where $L_i(\beta_j)$ is the predictive log-likelihood of the transferred model from context *j* when applied to context *i*, and $L_i(\beta_i)$ is the estimated log-likelihood of a local model estimated from data in context *i*. Under the null hypothesis that the local and transferred model coefficients are essentially equivalent, the TTS statistic is chi-squared distributed with the degrees of freedom being equal to the number of coefficients in the transferred model (or equivalently in the locally estimated model). If this chi-squared value is higher than the table chi-squared value with the appropriate degrees of freedom at a certain level of significance, one can declare that the transferred model is not “good enough” at that level of significance. If this chi-squared value is lower than the table chi-squared value with the appropriate degrees of freedom at a certain level of significance, one can declare that the transferred model is “good enough” at that level of significance. Of course, there is subjectivity in what significance level to use to undertake this test. Given that the TTS statistic tends to be a rather stringent yardstick for declaring transferability success, it may not be inappropriate to use a low level of significance such as the 2% level (that is, a very high level of confidence such as the 98%) to reject equality of coefficients (that is, to reject transferability).

The *TI* statistic is a relative measure of strength of the transferred model, as opposed to an *TTS* statistic (which is an absolute measure of transferability). The *TI* statistic reflects the degree of fit improvement offered by the transferred model relative to a base model when compared to the degree of fit improvement offered by a locally estimated model relative to the same base model. Originally proposed by Koppelman and Wilmot (1982), it takes the following form:

$$TI = \frac{L_i(\beta_j) - L_i(C_i)}{L_i(\beta_i) - L_i(C_i)},$$

where $L_i(C_i)$ is the log-likelihood of a base model estimated in the application context and other notations are as defined earlier. In choice models, the base model is the model with only constants (that is, the log-likelihood evaluated at market shares in the application context). The *TI* statistic has an upper limit of 1.00, denoting that the transferred and locally estimated models are equally accurate in prediction in the application context. There is no lower bound for the *TI* statistic, with negative values denoting that the transferred model is worse than a naïve model in the application context. The analyst obtains a relative valuation of the performance of the

transferred model. Thus, a *TI* statistic of 0.8 may be interpreted as reflecting that the transferred model offers 80% of the accuracy of a locally estimated model.

The transfer rho-squared statistic takes a value between 0 and 1 (except if the transferred model log-likelihood value is worse than the log-likelihood of the base model in the application context, in which case the rho-squared statistic takes a negative value). Unlike the *TI* statistic, the transfer rho-squared statistic is an absolute index of how well the transferred model does relative to a base model in the application context. The transfer rho-squared statistic takes the following form:

$$\rho^2 = 1 - \frac{L_i(\beta_j)}{L_i(C_i)}.$$

The problem with the statistic above is that it does not provide an intuitive feel of fit. As a result, this statistic is less often reported and used in transferability studies.

B.2.   Aggregate Transferability Measures

The aggregate measures provide a measure of the ability of the transferred model to predict aggregate travel behaviors in the application context. The aggregate behavior may be the total number of trips or choice alternative share (such as modal shares or chaining shares or destination alternative shares, where the shares are dependent on how the alternatives are defined). The aggregate measure can be broken down by purpose or demographic group, but the distinguishing feature of these aggregate measures is that they are not developed at the individual decision-maker level. As a result, aggregate measures, by themselves, are less reliable for transferability analyses, though they are often used in combination with disaggregate measures.

The typical aggregate measures are the mean absolute relative error measure (MA-REM), the root mean square error (RMSE) measure, and the relative aggregate transfer error (RATE). In this section, we provide these measures in the context of a choice model, though application to other types of models is very straightforward.

To develop the measures above, a relative error measure is first defined for each alternative *m*.

$$REM_m = \frac{S_m - \hat{S}_m}{S_m},$$

where S_m is the actual share of alternative *m* in the application context and $\hat{S}_m$ is the predicted share of alternative m by the transferred model in the application context. Then, the various aggregate measures are computed as follows:

TMIP

$$MA - REM = \frac{1}{M}\sum_{m=1}^{M} REM_m$$

$$RMSE = \left(\frac{\sum_{m=1}^{M} \hat{S}_m \times REM_m^2}{\sum_{m=1}^{M} \hat{S}_m}\right)^{0.5} \text{, and}$$

$$RATE = \frac{RMSE_i(\beta_j)}{RMSE_i(\beta_i)}.$$

For the last measure (RATE), the numerator corresponds to the ratio of the RMSE of the transferred model to the RMSE of a locally estimated model in the application context. Unlike the RMSE measure, which is an absolute measure, RATE provides a relative measure.

There is subjectivity in what values of the aggregate statistics above constitute a reasonable range for transferability success, though, as emphasized earlier, the aggregate measures by themselves provide little information on transferability unless taken in combination with the disaggregate measures. While there are no clear norms of acceptable ranges for these measures, an RMSE value of 0.3 or less is generally considered acceptable.

Appendix C   Research on Transferability of Trip-Based Model Components

C.1   Trip Generation

There have been several studies in the literature that have examined spatial transferability in the context of trip generation, as discussed below. The results are summarized in Section A.2 (see Table A.1).

Caldwell and Demetsky (1980) evaluated spatial transferability of linear regression models of household-level trip generation and zonal-level trip generation, using data from three cities in Virginia: Roanoke, Harrisonburg, and Winchester. In the household-level model, they considered two explanatory variables (auto ownership and household size) and used total trip productions per household as the dependent variable. In the zonal-level model, they used a single explanatory variable (zonal level number of cars), with total zonal trip productions classified by home-based work, home-based non-work, and non-home based productions as the dependent variable.

Overall, the results of the study suggest that this type of trip generation models can be transferred between cities, at least as long as care is taken in selecting "similar" cities. "Similar" cities are implicitly defined in the study as those with similar household size, household auto ownership levels, and per capita income.

Gunn et al. (1985) examined the transfer scaling approach for spatial transferability using two adjacent urban regions of the Netherlands, one located around Rotterdam and The Hague and the other located around Utrecht. The transferability analysis was based on data collected at each of the two urban regions, though the data were collected at the two locations at different points in time, as well as at different times of year. To accommodate the intrinsic differences in background variables across the two spatial contexts due to different times of data collection and different periods within the year of data collection, the authors utilized a nationwide travel survey as a control data set and then examined the spatial transferability of a daily shopping trip generation model as well as a personal business trip generation model (which are parts of a linked disaggregate-level nested logit system of mode-destination and trip generation specific to each trip purpose).

The overall empirical results indicate that a simple uniform scaling of the coefficients between the joint model components of the base area and the transfer area is adequate relative to separate locally-estimated models for the two areas, both from a statistical log-likelihood ratio fit perspective and from a prediction perspective on a suite of predefined market segments. This is quite interesting, given that the specifications adopted in these joint models are not particularly comprehensive in trip determinant variables. Specifically, the independent variables included level of service variables, demographic variables (cars per licensed driver, gender, and a Central Business District destination dummy variable), and an intrazonal trip dummy variable.

Koppelman and Rose (1983) indicated that aggregate models are not likely to be spatially and temporally transferable, even in cases where the underlying disaggregate-level behavioral process is similar. This is because of differences in the distribution of variables within aggregate population groups in the estimation and application contexts.

In their empirical analysis, the authors, among other things, examined the intra-regional transferability of household-level linear regression trip generation models between two sectors of each of three urban areas: Baltimore, Minneapolis-St. Paul, and Washington, D.C. The dependent variables in the analysis included number of stops and number of tours. The results

indicate large differences in parameter estimates of the trip generation model between sectors in each urban region. However, the authors find reasonable predictive ability of the transferred models based on typical goodness-of-fit and prediction measure comparisons between the transferred models and locally estimated models. At the same time, their statistical tests reject transferability, despite the closeness of goodness-of-fit and prediction errors.

Wilmot (1995) also examined the transferability of household-based linear regression trip generation models. He used total trips per household as the dependent variable, and considered household size and number of workers as the independent variables. He examined transferability within cities, between areas in a city, and between several cities in South Africa. His results suggest that model specification does influence the level of transferability, as does the difference in average income between the estimation and application contexts. Wilmot also emphasizes the need to have quality data in the application context to evaluate transferability. In his study, he finds a substantial improvement in transferability when the constant in the linear regression model is updated based on application context data.

Agyemang-Duah and Hall (1997) build upon the earlier research in two ways. First, they used an ordered-response model that respects the discrete and ordinal nature of number of trips, and includes built-in upper limits for trip rates as the values of the explanatory variables increase. Second, they included variables related to cost of travel and accessibility in evaluating spatial transferability.

The research focused on weekday home-based shopping trips made by households with one or more vehicles in the metropolitan Toronto area, based on a 1986 travel survey. The independent variables included household size, number of children less than 16 years, number of vehicles in the household, number of full-time employed individuals working outside home, number of part-time employed individuals working outside home, number of individuals employed at home, number of unemployed individuals, and accessibility to shopping opportunities. Spatial transferability is examined by evaluating models estimated on a core area (estimation area) to predicting trip generation in a periphery area (application context). Similarly, spatial transferability was also examined between the east and west parts of the metropolitan area, and among three pairs of municipalities. The transferability was assessed for a simple transfer scheme as well as a transfer updating scheme where factors (or scales) are applied to the latent index contribution of socioeconomic variables and the accessibility variable (the model coefficients used here are as obtained in the estimation context).

Transferability was evaluated using a transferred pseudo R2 measure (or the fraction of the constants-only log-likelihood ratio value in the prediction context explained by the model coefficients obtained from the estimation context), comparison of predicted versus observed aggregate shares, weighted root mean square error (the average relative error in the aggregate predicted shares weighted by the predicted shares), and two other related measures. The results indicate that the simple transfer mechanism works quite well for model transfer, though the transfer updating procedure substantially improves the predictive ability of the transferred model.

Kawamoto (2003) examined the spatial transferability of a linear regression model of total home-based trip productions at the person-level between two urban areas is Brazil: Sao Paulo and Bauru. They used a standardized form of the regression model, where the dependent and independent variables are represented in standardized form and are unit free. This procedure requires the values of the mean and standard deviation of each model variable in the application area, and represents a transfer updating scheme where the scaling is done on a variable-by-variable basis. Transferability was evaluated based on a Wald test statistic of parameter equality in the regression models in the estimation and application contexts after accounting for

variance differences in the two contexts. The variables considered in the analysis included relationship with householder, educational attainment, number of cars in household, student status, employment status, and if the individual is a child younger than 11 years. The results indicate that the standardized regression models are transferable between the two cities though the unstandardized versions are not. This is interesting, especially given that the Sao Paulo data was collected in 1987, while the Bauru data was collected in 1998.

Cotrus et al. (2005) examined the spatial transferability of linear regression and Tobit models of person-level trip generation models, using data from Tel Aviv and Haifa in Israel. The data were drawn from the 1984 and 1996/97 Israeli National Travel Habits Survey. The models included age, car availability, possession of a driver's license, employment status, education level, and whether the individual defines herself/himself as the head of the household. The results indicate that the Tobit models fit the data better, but that equality of coefficients in the two areas is rejected for both the regression and Tobit models on the basis of statistical tests. In particular, the coefficients on the license holding and age variables are statistically different, while those of other coefficients are not. However, the transferred models appear to do well in terms of aggregate predictions.

Everett (2009) examined the regional and spatial contexts influencing the spatial transferability of household-level trip production models. The data used in this research was drawn from four different travel survey projects that collected data from 11 distinct metropolitan regions in two states (Ohio and Tennessee). Seven of the metropolitan regions were from Ohio, and the surveys for these regions were undertaken as part of an Ohio statewide data collection effort in 2003. Three separate survey efforts in Tennessee obtained data for the four metropolitan regions there. The 11 regions included five small metropolitan planning organizations (MPOs) (two in Tennessee and three in Ohio, with populations less than 150,000), three medium sized MPOs (one in Tennessee and two in Ohio, with populations in the range between 375,000 to 500,000), and three large MPOs (one in Tennessee and two in Ohio, with populations larger than 700,000).

The trip productions included all motorized vehicle person trips on a weekday by members of the household over 4 years of age for the Tennessee surveys and for members of the household of all ages for the Ohio survey. The trip production models were based on a cross-classification approach for four types of trips (total trips, home-based work (HBW) trips, home-based other (HBO) trips, and non-home-based (NHB) trips), and transferability was assessed by comparing the trip rate for each cell of the cross-tabulation across pairs of areas.

Except for HBW trips, the cross-classification variables included household size and the number of vehicles. For HBW trips, the control variables included the number of workers in the household and the number of vehicles. The results indicated that, within the same state, transferability of trip rates appeared to be valid for the small MPOs. Across states for the small MPOs, the results indicated that transferability was valid for HBW and HBO trips, but not for NHB and total trips. Across states for medium and large-sized MPOs, transferability appeared valid for total trips, but not for NHB trips (with mixed results for HBW and HBNW trips). Across metropolitan area sizes, the results indicated good transferability in Ohio, but not Tennessee.

Possible reasons suggested by the author for these results are that the Ohio areas were all surveyed as part of a single survey effort (unlike the Tennessee areas) and/or Ohio travel is more similar across MPOs in the state than across MPOs in Tennessee. The author then proceeded to examine transferability when controlling for an additional "area type" variable in the cross-classification to accommodate urban design features and development density differences between sub-regions within a region. The area type variable is based on partitioning each region into grid cells of about one square mile, and assigning a density index based on the

total population in each grid cell divided by the area of the grid cell. Based on this population density index, three area types (rural, suburban, and urban) were identified based on an "optimal thresholding" procedure. The net result was that transferability improved substantially from the case without including area type, and could be rejected mainly for the suburban area type and for NHB trips.

C.2   Trip Distribution/Destination Choice

The literature on transferability of trip distribution/destination choice is relatively limited, and has been focused on temporal transferability, not spatial transferability that is the focus of this report.

Karasmaa (2007) studied spatial transferability in the context of a disaggregate nested logit trip destination (upper level)-mode (lower level) choice model for home-based other (non-work and non-school) trips. The study specifically focused on a comparison of alternative methods of spatial transfer and assessed the effect of the size of the application context data on transferability effectiveness. Four transfer approaches were evaluated: transfer scaling, Bayesian updating, combined transfer, and joint context estimation. The estimation context in the paper is the Helsinki Metropolitan Area (HMA) and the application context is the Turku region. The study is based on a 1995 mobility survey undertaken in HMA and a 1997 mobility survey undertaken in Turku. For each transfer method and sample size, the transfer effectiveness is assessed at two levels: (1) the examination of coefficient ratios and elasticity effects of variables, and (2) the equality of coefficients using a transferability test statistic (TTS) that takes a likelihood ratio test form. In terms of elasticity effects, a cost increase of drive alone by 10 percent and a time increase by public transit by 30 percent are evaluated using the transferred model and the model estimated directly on the application context using the full sample collected there. The difference in predicted change computed by the transferred model relative to the estimated model in the application context (due to the change in cost/time) as a function of the predicted change by the estimated model is used as one measure of transfer effectiveness.

Overall, the results indicated that the joint context estimation provides the best transferability success on both of the assessment levels identified above, especially when the imprecisely estimated coefficients (from the estimation context) are estimated as being fixed across the estimation sample and the small application sample. The performance of the combined transfer method was highly sensitive to the sample size from the application context, and the performance of the transfer scaling approach was not good. The Bayesian updating procedure too did not perform well on the coefficient ratios and elasticity tests. Overall, the study also found that an assessment of transferability based solely on the TTS measure can be misleading, because two models that lead to very different coefficient ratios and elasticity effects can still have similar TTS values.

C.3   Mode Choice

Mode choice has been perhaps one of the most studied trip dimensions in the context of spatial transferability.

Watson and Westin (1975) studied the spatial transferability of binary logit intercity mode choice models among different subareas in the Edinburgh-Glasgow area of Scotland. Specifically, they identified six travel "corridors" in the Edinburg-Glasgow area based on whether the origin and destination ends were in the central city, the suburbs, or peripheral to the urban area. The modes considered were the automobile and train. They included level-of-service variables and a mode-specific constant, but no socioeconomic characteristics of the travelers. The models estimated in the six travel corridors were then compared for similarity in model

coefficients, and each model was also transferred to the other five corridors to evaluate modal split predictions.

Their findings indicate that there is a high level of model transferability between the three models estimated in the corridors with a trip-end in the central city. However, this is not the case for the models estimated in the remaining three corridors that did not have a trip-end in the central city.

Atherton and Ben-Akiva (1976) examined the spatial transferability of a home-to-work trip mode choice model estimated on data collected in Washington, D.C. in 1968 to New Bedford, Massachusetts and Los Angeles. Data from 1963 in New Bedford and 1967 in Los Angeles were available to test the extent of transferability of the multinomial logit model estimated from Washington. The alternatives considered in the mode choice model included driving alone, sharing a ride, and public transit.

The authors conclude, based on statistical tests of parameter equality and predictive ability in the transfer contexts, that the Washington model is transferable to the other two application areas. They further examined the benefit of updating approaches that (a) update the constants only based on aggregate shares of the alternative modes in the application area, (b) update the constants as well as estimate a single factor that scales the other coefficients, and (c) a Bayesian update method based on the inverse of the variance-covariance matrices of the coefficient estimates from the estimation context and the application context as weighting factors. The results indicate that the Bayesian update approach works best, especially when the disaggregate sample available from the application context is small in size and the original estimation context choice model is well specified. However, there is little difference in the extent of transferability between the model with no updating and that with even the Bayesian update.

Talvitie and Kirshner (1978), in their study of urban commute mode choice model transferability between Washington, Minneapolis-St. Paul, and San Francisco, used the same variable specification as that in Atherton and Ben-Akiva. The modal alternatives are drive alone, shared ride, and bus with walk access (the individuals choosing the Bay Area Rapid Transit System in the San Francisco Bay area were removed from the analysis).

The authors examined transferability both within each region and between regions. The within region transferability was examined by partitioning the sample from each region in three ways: (1) urban travel versus suburban travel (not done for the San Francisco sample), (2) Central Business District (CBD) travel versus Non-CBD travel, and (3) a random split of the sample into two sub-samples. Overall, the results of statistical tests of parameter equality between the samples within each region were mixed and inconclusive, though there was more evidence of non-equality of parameters than equality of parameters. The between-region transferability in terms of model parameter equality was also statistically rejected with a high level of confidence. These results are clearly different from the results of Atherton and Ben-Akiva. The authors suggest that several factors may have played a role in their findings, including variations in network coding routines and differential trimming of outlying data points across the data sets.

Galbraith and Hensher (1982) emphasized the need to consider both level-of-service variables as well as a reasonably extensive set of socioeconomic and contextual characteristics in mode choice models before evaluating transferability. They also identified the need to use consistent data (i.e., same measurement procedures, sampling procedures, variable definitions, questionnaire wording, etc.) in the estimation and application contexts to engage in any meaningful debates about the extent of model transferability.

Their empirical analysis of the spatial transferability of mode choice models involved examining the intra-urban transferability of commute binary mode choice coefficients from two suburban

areas in Sydney. The alternatives included car and rail. In addition to the usual level-of-service variables, the final specification used in the paper included variables representing gross annual individual income, number of licenses drivers in the household, and number of cars in the household. Their statistical tests reject parameter equality of the logit models in the two suburban regions, though they find that a specification that normalized travel cost by income transferred relatively better than a specification that used a non-normalized travel cost variable. However, in an evaluation of predictive ability at the mode share level, the simple transferred models without any updating performed quite adequately relative to the locally estimated model. They find a Bayesian transfer update approach to perform somewhat better than the approach without any updating and the approach that updates the constants/scale.

Koppelman and Wilmot (1982) focused on the intraregional transferability of a commute mode choice model for breadwinners who work in the Central Business District of Washington, D.C. They caution against the sole use of model parameter equality as an indicator of whether a model is transferable or not, indicating that model parameter equality is a symmetric property between two contexts, while transferability is a directional property.

In their empirical analysis, they used disaggregate measures of transferability (transfer log-likelihood ratio, transfer log-likelihood index, and the transfer rho-squared) as well as aggregate measures of transferability (root mean square error and relative root mean square error). The data sample was partitioned into three groups based on three predetermined geographic sectors in the Washington area, and model transferability is studied between the resulting three pairs of sectors. The alternatives included drive alone, shared-ride, and transit, and the variables included in the specification are level-of-service variables, income, vehicles per driver, a government worker dummy variable, and the number of workers in the household.

The results reject parameter equality across the models for the three pairs of sectors. Further, the disaggregate measures of transferability reject the hypothesis of intraurban transferability, even if the modal constants are adjusted to match the application area modal shares. However, the transferred models provide close to 80 percent of the information provided by local models, indicating that the extent of transferability is not bad from a non-statistical perspective. In addition, the transferred models perform quite well compared on the basis of aggregate modal share predictions. This seeming inconsistency between statistical tests and transfer errors is not uncommon, and the authors recommend that "although statistical tests can be used to alert the planner or analyst to differences between models, they must be considered with reference to the magnitude if errors that are acceptable in each application context."

Koppelman and Rose (1983) study the intra-regional transferability of a multinomial work mode choice model by partitioning the Baltimore region into a North sector and a South sector. The modal alternatives were drive alone, shared-ride, and transit, while the independent variables included level-of-service variables as well as socioeconomic variables such as income and cars per driver. The results reject transferability based on parameter equality, disaggregate measures of transferability, and aggregate measures of transferability, though there is substantial improvement in the aggregate measures of prediction when the estimated model constants are adjusted based on the aggregate modal shares in the applicant region.

The authors conduct a similar analysis of intra-regional transferability of mode choice models from the Washington, D.C. area and Minneapolis-St. Paul, and find that the transfer performance is much better in these other urban areas relative to Baltimore. However, even in these other areas, intra-regional transferability is rejected based on statistical tests.

Koppelman et al. (1985) examined the effectiveness of model updating using limited data from the application context on intraregional and interregional work travel mode choice transferability. Specifically, they studied the effect of updating alternative specific constants and the scale of

the model. The data used for the intraregional transferability analysis was from Washington, D.C., with the same use of three sectors as defined in Koppelman and Wilmot (1982). The data used for interregional transferability was from Washington, Minneapolis-St. Paul, and Baltimore. The independent variables used included three level-of-service variables, a car per driver variable specific to the drive-alone and shared-ride alternatives, and modal constants.

The same transferability measures as developed in Koppelman and Wilmot were used in evaluating transfer effectiveness. The results indicate that transferability is improved substantially when the constants are updated, and even more so when the constant and scale are updated. However, the returns from updating the constant and scale is not as high as with updating the constant only. This holds for both interregional and intraregional transferability.

Gunn et al. (1985) conducted a similar evaluation of the effect of model updating as Koppelman et al. (1985), using a joint system of mode, destination, and trip generation system (see discussion of this paper under Section C.1). Their results corroborate the findings of Koppelman et al. (1985) that updating constants and the scale leads to improved model transferability.

McComb (1986) assessed spatial transferability using data from a single "high quality" data source (the transportation supplement of the Canadian Labor Force Data) for ten cities in Canada. He used the same uniform model specification and consistent data collection and preparation across the cities, and examined socioeconomic moderating effects of sensitivities to level of service variables. The work trip mode choice model developed for the City of Winnipeg was used as the estimation context, while the other cities were considered as the application contexts. Four modal alternatives were considered: drive alone, driver in a shared ride arrangement, passenger in a shared ride arrangement, transit, and walk/other. The independent variables included level of service variables, sex of individual, family income, age, work trip distance, and peak versus off-peak work start time.

The author found that coefficient equality cannot be rejected between cities of similar socioeconomic make-up, size, and transportation system quality (such as Edmonton and Winnipeg, and Calgary and Winnipeg). However, coefficient equality is rejected for cities that are very different in character (such as Toronto and Winnipeg and Ottawa and Winnipeg).

Koppelman and Wilmot (1986) reported an analytic and empirical investigation of omission of variables on the spatial transferability of mode choice models using the same data set and procedures in Koppelman and Wilmot (1982). Three different specifications were considered to evaluate omitted variable effects on transferability, with each subsequent specification including the variables in the earlier specification and new variables as follows: (1) Three level of service variables and modal constants, (2) Addition of cars per driver variables specific to drive alone and shared ride, and (3) Addition of a government worker dummy variable and a number of workers in the household variable, both specific to the shared ride mode.

The results indicate substantial improvement in transferability with improved specifications, and with modal constant updating based on the aggregate share in the application context. The authors also indicate that models with only level-of-service variables and constants are unlikely to achieve adequate levels of transferability for practical use.

Koppelman and Pas (1986) also examined spatial transferability of a mode choice model using the Washington, D.C. data, but added a multidimensional element to the analysis. The main focus was on whether a nested logit model of auto ownership and mode choice is more or less transferable than a simpler joint multinomial logit model of auto ownership and mode choice. The nested logit model was estimated using a two-step sequential estimation approach, which can lead to a loss of efficiency.

In the empirical analysis, the nested logit model's logsum parameter is not statistically significantly different from 1 at the 0.05 level of significance. The results show that the transferred models without updating are able to capture more than 85% of the information obtained from locally estimated models for both the multinomial and nested logit models, indicating that both these models are transferable across three sectors in the Washington area. The multinomial logit model has a small advantage in the extent of transferability though this improvement over the nested logit model is marginal. However, this result is likely to be specific to the empirical context in the study because the nested logit specification essentially collapsed to the multinomial logit specification for all the three sectors. Further analysis is needed to examine the effect of model structure on transferability.

Abdelwahab (1991) examined spatial transferability of intercity mode choice models between two regions in Canada encompassing travel between 23 major metropolitan areas. He used the 1984 Canadian Travel Survey (CTS) in the analysis, and geographically divided the 23 metropolitan areas into two regions: an eastern region including Thunder Bay and cities east of Thunder Bay, and a western region including Winnipeg and cities west of Winnipeg.

The intercity travel in each of these regions was categorized based on trip length (short trips lesser than 600 miles and long trips) and purpose (recreational and business). The author used two transfer updating methods, one being the constant only update scheme and the second being the Bayesian update method that updates all model coefficients. The independent variables used in the analysis are not provided in the paper.

The results indicate that the transferred models explain about 50% to 93% of the information (i.e., the difference between the log-likelihood value at convergence and the log-likelihood value at market shares) provided by the locally estimated models. Overall, the findings indicate poor transferability, as measured by disaggregate predictive fit and aggregate error, for both updating methods considered.

Karasmaa (2001) explored the spatial transferability of work trip mode choice models in the Helsinki and Turku regions of Finland. The Helsinki region was used as the estimation context, and the Turku as the transfer context. Four transfer approaches were evaluated: transfer scaling with re-estimation of alternative specific constants and the scale, Bayesian updating, combined transfer, and joint context estimation. The influence of the size of the estimation context data on transferability was also examined by using four different sample sizes for estimation of the Helsinki mode choice model using a 1995 mobility survey.

The results show that the joint context estimation is generally the best method of transfer, especially when the estimated coefficients of the locally estimated models are quite different between the estimation and application contexts. The combined transfer estimation approach is best when there is a large estimation sample and the transfer bias is small between the estimation and application contexts.

It is interesting that there has been a lull in mode choice transferability studies since the early 2000s. In a recent paper, **Mamun and Sabbir** (2012) examined the impact of omitting travel cost during the transferring process. They used two datasets for their analysis, one originating from the Dallas-Fort Worth (DFW) area (based on a 1996 survey) and another from the San Francisco Bay area (based on the 2000 BATS survey). Each of the datasets was split into a 90% estimation sample and a 10% application sample. The alternatives in both the spatial contexts were consolidated into car, transit, and other.

The model estimated on the 90% BATS estimation sample was applied to predict the modal shares in the 10% BATS application sample as well as in the 90% DFW estimation sample. Similarly, the model estimated on the 90% DFW estimation sample was applied to predict the

modal shares in the 10% DFW application sample as well as in the 90% BATS estimation sample. The study noted that the modal shares are well predicted by the transferred model even in the absence of a critical variable such as travel cost. However, the study appears to miss the point that variables such as cost are the key to policy analyses. It also completely ignores trip purpose, as well as infers transferability success solely based on predicted aggregate sample shares of the modes (which is very heavily biased toward the car mode to begin with).

C.4 Route Choice

Bekhor and Prato (2009) is the only study that we are aware of that examined transferability in the context of route choice. The authors correctly identify the generation of alternatives as one of the most difficult tasks in route choice modeling and test the transferability of path generation techniques based on both "cost-effectiveness" (in the context of generating realistic routes in a route choice context) as well as transferability performance (in the context of replicating the actual routes chosen by individuals).

The authors used two measures of transferability performance. The first is based on estimating route choice models separately in the estimation and application contexts, then estimating another model by combining the two data sources, and finally examining coefficient differences from the estimations. The second is based on the familiar transferability test statistic (TTS) to examine predictive fit at the disaggregate level.

The data for the analysis is obtained from Boston and Turin for the route choice from home to work. The Boston survey was conducted in 1997 and collected information from MIT faculty and staff members, while the Turin survey collected route choice information from faculty and staff members of Turin Polytechnic. The selection of relevant routes for the choice situation was done in two ways: (1) Identify an alternative path that provides enough variability (in addition to the chosen path that was automatically included) and (2) Identify multiple paths based on major arterials connecting the origin and destination.

The model forms used for route choice included a simple multinomial logit (MNL) model, a variant of the MNL labeled as the path size logit (PSL) model that includes a correction factor in the deterministic part of the utility function that is equal to the path length, and another model form labeled as the logit kernel (LK) model that captures the correlation in paths based on the extent of overlapping lengths. The best path generation technique across both Boston and Turin was the branch and bound technique. In terms of transferability, the LK model form that accounts for similarity in the stochastic part of utility performed better than the PSL model form that adds correction terms in the deterministic utility part. But the transferability of model parameters in the route choice model was not verified for any model form.

C.5 Data Transferability Studies

Greaves and Stopher (2000) employed the data transferability approach to transfer trip production models. Specifically, they used the 1995 Nationwide Personal Transportation Survey (NPTS) data and clustered households into relatively homogenous groups for each of six trip purposes: home-work, home-school, home-shop, home-other, other-work, and other-other. A classification and regression tree method, combined with the standard analysis of variance procedure, was adopted to determine the clusters. The number of clusters varied from six groups for the home-work, home-school, and work-other purposes to 16 groups for the remaining purposes. The clustering variables included household size, number of workers, number of vehicles, and number of children and adults by age group. Within each cluster for each trip purpose, a cumulative frequency distribution was developed for number of trips

produced. They then applied the cluster scheme to predict the trip productions for a survey sample of households in the Baton Rouge MPO region.

For this process, they applied the clustering scheme to the add-on sample as developed earlier from the main NPTS sample and then drew a random realization from the cumulative trip production frequency distribution for each purpose and each Baton Rouge region sample household, based on the cluster to which the sample household is assigned. Next, they compared the trip production predictions from their method and from a transferred model that is based only on household size as the independent variable, using the survey-collected trip productions as "ground reality." They found that their approach does better than the transferred model, a result that is not surprising given that the transferred model is based only on a single household size variable, while the authors' approach effectively uses several independent variables. They also compared the model estimates obtained from estimating trip production models using their synthesized trip production data and the actual survey trip production data, and concluded that the trip production models for "home-work and home-school are well estimated, home-shop and work-other are acceptably estimated, and home-other and other-other are marginally well estimated."

Stopher et al. (2003) undertook a similar analysis as Greaves and Stopher, except that they examined the effectiveness of their approach in application areas (Dallas and Salt Lake City) where household travel surveys may not be based on the same survey collection methodology as NPTS (the Baton Rouge household travel survey used earlier was patterned after the 1995 NPTS). Specifically, household travel surveys are usually collected over the fall or spring of a year rather than the year-round data collection of NPTS and are based on an activity survey rather than the trip-based survey of NPTS. Stopher et al. (2003) also examined if the travel characteristics are a function of city characteristics in addition to demographic attributes that formed the clustering basis in the earlier work.

Their results showed that the simulation does not work well for the Dallas and Salt Lake City areas though this result may simply be an artifact of the way the survey questions were worded and interpreted by respondents. They also concluded that city characteristics do matter in trip production estimates, and they recommend using contextual variables such as city population size and transit service quality. In addition, they suggest the use of a Bayesian updating of the travel characteristics for the clusters using small samples from the application context.

Reuscher et al. (2002) also pursued a data transferability analysis of vehicle trips per household, vehicle miles of travel (VMT), person trips per household, and person miles of travel (PMT) rates. They used a combination of cluster/regression analysis, judgment, and well established relationships between VMT and area type and demographics. In particular, they first classified the census tracts in the U.S. into nine groups defined by area type (urban, suburban, and rural) and income (very low, very high, and other). Next, they developed household size-specific, number of vehicles-specific, and census tract cluster-specific vehicle trip, VMT, person trips, and PMT rate estimates (and standard error of estimates) using the 1995 NPTS data. Based on this initial classification, they subsequently undertook a clustering analysis procedure to determine the final clusters based on a combination of household size, number of vehicles, and the initial census tract clusters. Once this clustering was established, the travel characteristics for any census tract in the U.S. could be determined based on the cluster to which it belongs. The authors assessed their approach using data from Baton Rouge and three NPTS add-on samples from New York, Massachusetts, and Oklahoma, and found their approach to be better than other approaches that cluster census tracts based on Metropolitan Statistical Area size, census region, and census division.

Mohammadian and Zhang (2007) used methods similar to the earlier data transferability studies, but considered a more comprehensive set of variables to cluster households on, including demographics, pedestrian friendly environment characteristics (such as intersection density, road density, and block size), transit usage, and congestion factors (the Urban Mobility Index measure, total number of road users divided by road density, and the percentage of workers driving to work divided by road density). A combination of principal component analysis and cluster analysis was undertaken to define a total of 11 relatively homogenous groups of household types using the 2001 NHTS. This clustering scheme was then transferred to the NHTS add-on samples from New York, Wisconsin, Texas, Kentucky, and Hawaii. The transferred travel characteristics from the original NHTS survey were then compared to the actual travel characteristics directly collected in add-on samples, as a way of assessing the performance of transferability. They found reasonable transferability on such travel characteristics as person/vehicle trips and tours by purpose.

Zhang and Mohammadian (2008a) applied the data transferability approach by generating a synthetic population for the application context using well-established population generation methods. Their application context corresponds to the New York region. They classified the generated population using the approach in Mohammadian and Zhang (2007), and compared the mean values of trips per person and trip distance per person from the simulated data with the mean values from corresponding clusters from the actual observed survey data (from the New York NHTS add-on sample). The results showed good fit of the simulated and observed travel characteristics.

Zhang and Mohammadian (2008b) further improved upon Zhang and Mohammadian (2008a) by fitting a gamma distribution for the trip rate per person and trip distance per person for each cluster using the main NHTS survey, and next updated the parameters of this distribution using a small sample randomly selected from the NHTS add on for New York (as suggested by Stopher et al., 2003). The authors used a Bayesian approach to updating, and compared the parameters of the updated gamma distribution within each cluster with the equivalent best fit gamma distribution parameters from the corresponding cluster of households from the entire New York add-on sample. The authors note that the parameters of the updated gamma distribution are closer to those from the New York add-on sample compared to the un-updated gamma distribution parameters.

Overall, the data transferability approach uses statistical clustering mechanisms as the basis to predict travel characteristics in the application area. The relationship between independent variables and the dependent variable is implicitly captured in the clustering mechanism. The data transferability approach (once the clusters are determined and a synthetic sample/population is generated for the application area) has the appeal of being very easy to implement (it becomes almost like a look-up table). If a small sample is available in the application area, updating mechanisms are also possible as in Zhang and Mohammadian (2008b). However, the updating mechanisms do not allow as clear an analysis of the specific response behaviors that vary across the estimation and application areas as do the model transferability approaches. For instance, consider a situation where the pattern of cost sensitivity variations across demographic (cluster) groups is very different between the estimation area and the application area, but the patterns of time sensitivity variations across demographic groups are very stable between the two areas.

The data transferability approach can be more inaccurate (than the model transferability approach) in this case in terms of policy analysis of congestion pricing strategies. On the other hand, in the model transferability approach, there is the possibility of updating the cost coefficient for the application area while holding the time coefficient fixed, which may lead to improved policy analysis of congestion pricing strategies. In any event, there needs to be more

extensive research on comparing the data transferability and model transferability approaches, especially in the case of activity-based models where synthetic population generation is a part of the modeling process. Perhaps combinations of the two approaches can also be considered and can be effective in specific transfer contexts.

Appendix D   Research on Transferability of Activity-Based Models

A few recent studies have begun to examine transferability in an ABM context. These studies are discussed below. For ease in exposition, we discuss each ABM transferability study by positioning it along three dimensions:

1. The activity dimensions examined (including the overall model background and the specific modeling structures used for the ABM components);
2. The data context (including the estimation and application contexts, and the explanatory variables used); and
3. The transferability approach and results (including the metrics used and the conclusions).

In this appendix, we do not discuss efforts that simply transfer models from one location to another, without undertaking any quantitative evaluation (such as those by Vine et al., 2010 and Vovsha et al., 2012). But, very briefly, Vine et al. (2010) transfer model components of an ABM labeled as TASHA (Miller and Roorda, 2003) estimated in Toronto to London, assuming that the model components are transferable. Vovsha et al. (2012) informally compare estimation results from four locations for work location choice (including the binary choice of working from home versus working away from home, and a multinomial logit choice of the traffic analysis zone of work given it is away from home). The four locations are San Diego, Tucson, Phoenix, and Chicago. Vovsha et al. indicate that the work location coefficients appear to be qualitatively quite different across regions, and recommend local data estimation for the work location choice model.

We have identified four studies of the transferability of activity-based models. One of the studies is from the Netherlands and was published in 2002. The other three studies were done in the U.S. over the past two years.

D.1   Arentze et al., 2002

D.1.1 Activity-Travel Dimensions Examined

The first ABM transferability study appears to be the one by Arentze et al. (2002). In their study, they assumed that the activity agenda (or skeleton) was given, and the focus was on the scheduling of activity episodes, including mode choice decisions, activity episode location choices, accompaniment (the "with whom" dimension of activity participation), the chronological sequencing of activities, and associated travel linkages in time and space, using a household-day as the unit of analysis. In the scheduling, "fixed" activities (such as work, school, and medical purposes) are first placed in the sequence based on their given locations, durations, and time of day in the activity agenda, and then the modeling system adds flexible activities based on feasibility and designated priority in the activity agenda.

In terms of the actual mechanics of the modeling system, the first decision modeled is the travel mode to work, followed by the number of flexible activity episodes to add to the fixed activity schedule. The subsequent decisions include the time-of-day of activity episode participation, the sequence of episodes (and, thereby, implicitly the formation of tours), the mode for each tour (assuming no changes in mode within a tour), and finally the location of each flexible activity. All activity-based systems use some kind of such a sequence of building the overall pattern, and the advantages and limitations of each building sequence will not be discussed

61

here. In this study, the authors use an ABM system labeled as "Albatross" (see Arentze and Timmermans, 2000), which involves a series of decision rules representing choice heuristics at each of a series of decision steps, combined with models of dynamic constraints on choice options at each step.

D.1.2 The Data Context

To test the transferability of the specific ABM used, the authors employed three activity surveys (all of which collected activity-travel information over a 48-hour period over the week): one from the municipality of Voorhout, the other collected in the City of Apeldoorn, and a third collected in the South Rotterdam region. Models estimated using 75 percent of the sample from the South Rotterdam region were transferred (without any kind of updating) to the municipality of Voorhout, the City of Apeldoorn, and to the 25 percent of the sample from the South Rotterdam region itself. Overall, there were two spatial transferability instances in the study (South Rotterdam to Voorhout, and South Rotterdam to Apeldoorn). The review of the study did not reveal the explanatory variables used.

D.1.3 The Transferability Approach and Results

The transferability of each of the model components (travel mode to work, number of flexible episodes added to the fixed schedule, time-of-day of activity episode participation, the sequence of episodes, the mode for each tour, and the location of each flexible activity) at the level of individual episodes as well at the level of the entire activity pattern was examined at both a person level as well as at an aggregate level. The simple transfer approach was used, with no updating of any kind using small data samples from the application context.

At the person level and for individual episodes, the ratio of correct predictions along each of the many dimensions was used as a measure of performance (this measure first assigns a value of '1' for each individual if the chosen alternative for a specific dimension is the same as the predicted alternative in that dimension, and then taking the number of observations with a '1' assignment as a ratio of total individuals). The results indicated that the ratio of correct predictions did not reduce substantially (for each and every dimension) between using the locally estimated model to the 25 percent validation sample in the South Rotterdam region and transferring the estimated model to the other two spatial regions.

Of all the dimensions, location choice seemed to be the least transferable in space (location is represented in terms of distance bands from the home location of the individual). At the person level and for the entire activity pattern, transferability was assessed using sequence alignment techniques that measure the degree of similarity in terms of the number of episode deletions, additions, and substitutions in the comparison of the predicted overall sequence with the actual observed overall sequence. In this overall pattern-level assessment, the mode choice component appeared to be the most problematic. But the authors noted that the results were very encouraging at this person level.

At the aggregate level, the assessment of different dimensions was made by comparing the predicted (transferred) and observed frequency distributions. The results showed remarkable aggregate fit between the transferred and observed distributions for the number of activity episodes by purpose, the accompaniment for each episode, time-of-day distributions, chaining tendencies, and even location. Once again, the transport mode aggregate predictions were less than desirable, especially for the car mode and other non-public transport modes. At the overall pattern level for the aggregate assessment, the authors compared the origin-destination matrices as predicted for Voorhout and Apeldoorn based on the estimated model (i.e., the model estimated in the southern Rotterdam context) with the actual origin-destination matrices

for Voorhout and Apeldoorn, respectively. This was achieved through the development of correlation coefficients. The results reinforce earlier results of good transferability.

In all, the authors conclude that there is "substantial evidence for the spatial transferability of Albatross, except for transport mode."

D.2   Nowrouzian and Srinivasan (2012)

D.2.1 Activity-Travel Dimensions Examined

This paper examines transferability assessments in the context of tour generation models. The authors use a simple transfer approach with no updating of any kind. In the paper, tours are classified based on four considerations:

1. Purpose (defined in relation to the primary activity undertaken in the tour, with the primary activity being work for workers and escort (chauffeuring) for non-workers);
2. Complexity (defined as a simple tour with only one stop or a complex tour with multiple stops);
3. Accompaniment (designated as solo/alone or joint with other household members); and
4. Flexibility (no constrained activities such as work/appointments/drop-off and undertaken independently, or with constrained activities and/or undertaken jointly).

Based on these four dimensions, the tour generation framework was made operational as follows. A worker decides on a mandatory tour first. The alternatives for this mandatory tour are: (1) no work on that day, (2) simple work tour, (3) a complex work tour with an escort-related stop, and (4) a complex tour with no escort-related stops. A non-worker decides on escort tours first, with the alternatives being (1) no escort tour in the day, (2) simple tour (escort is the only activity episode), and (3) complex tour (escort is combined with one or more non-work activity episodes).

Next, joint tours are determined for households with two or more individuals. The alternatives here are (1) a single simple (one-stop) joint tour during the day, (2) a single complex (multi-stop) joint tour, (3) or multiple joint stop tours (can be simple or complex tours). Finally, decisions about independent non-mandatory tours are made, and the same set of alternatives as for joint tours are used. Overall, then, tour generation includes the modeling of four types of tours: mandatory tours (for workers), escort tours (for non-workers), joint tours, and independent non-mandatory tours. A multinomial logit (MNL) model structure is used in the study for each tour type.

D.2.2 The Data Context

The study uses household survey data from three regions in Florida: the Northeast (NE) region, the Southeast (SE) region, and the Tampa Bay (TB) region to the West. Models were estimated for each of the four tour types identified earlier, and for each of the three regions (leading to a total of 12 models). The models for each region were estimated using 85 percent of the data from the survey in the region, with 15 percent set aside for transferability assessments.

For each tour type, models estimated in each of the three regions are transferred to the other two regions (thus, across the four tour types, there are 24 (4 tour types x 6 spatial pairs) transferability instances). The explanatory variables in the model included variables associated with age, employment status, number of children in the household, number of workers in the household, number of vehicles relative to number of adults, and whether the individual lived in a single family home or other housing arrangements.

D.2.3 The Transferability Approach and Results

The transferability of each of the tour type models is assessed in three different ways: (1) examining disaggregate predictions, (2) examining aggregate predictions, and (3) studying elasticity effects of variables.

At the disaggregate level, for each of the four tour types and for each of the three regions, the log-likelihood on the validation sample (of 15 percent) is computed based on the model estimated in that region as well as the model estimated in the other two regions. This procedure assesses the ability of each of the estimated and transferred models to assign a high probability to the actual observed outcome in the validation sample. If a transferred model has a log-likelihood that is within 10 percent of the model estimated from data in the region, the model is deemed transferable.

The authors observed that 20 of the 24 transferred models are successful based on this criterion. The models for mandatory tours and independent non-mandatory tours were more transferable than those for escort and joint tours. Models estimated from the NE region were very transferable to the SE region, and models estimated from the SE region were very transferable to the NE region. However, the transferability of the NE escort model to the TB area and the transferability of the SE joint tour model to the TB area were not very good. The worst performances were for the escort models estimated in the TB area when transferred to the NE and SE areas. This is because of the far fewer escort trips identified in the TB area relative to the other two areas, which may be a reflection of true differences or data consistency issues. There is also a suggestion that a simple constant updating procedure (rather than a simple transfer method) would have resolved this issue.

At the aggregate share level, for each of the four tour types and model application region using the 15 percent validation sample, each of the three estimated models were applied to obtain the average predicted probability (i.e., share) of each alternative. The predicted shares are compared with the actual observed shares to compute a mean absolute ratio error (MARE) and the root mean square error (RMSE). As would be expected, the application of the locally estimated model is generally superior to the transferred models. The best success in general was for the independent non-mandatory tour category. There was also good success for the SE escort tour model transferred to the NE region, and the NE joint tour model transferred to the SE region.

However, an interesting finding is that there can be asymmetry in transferability. That is, while the SE escort tour model transferred very well to the NE region, the NE escort tour model did not transfer very well to the SE region. Similarly, while the NE joint tour model transferred very well to the SE region, the SE joint tour model did not transfer very well to the NE region. Important to note also is that the disaggregate level analysis did not reveal this asymmetry in transferability potential, an issue that the authors did not point out in their analysis.

Further analysis of this issue is important, and again this may be related to sample share differences that can be addressed if a simple constant updating procedure were adopted. But, consistent with the disaggregate level analysis, transferability was the worst when the TB-estimated escort models were transferred to the other two regions. Using the criterion that a MARE value of 0.3 (no more than 30% error on average) or less represents a successful transfer, the authors noted that 15 of the 24 transferred models were successful. As in the case of the disaggregate level analysis, the models for independent non-mandatory tours were very transferable, and the worst transferability was for escort tours.

The authors also examined transferability on the basis of elasticity effects of specific variables. Two variable change scenarios were examined. The first scenario converted all car-sharing

households (a household is designated as a car-sharing household if the number of adults exceeds the number of motorized vehicles) in the validation sample to non-car sharing households (representing an increase in motorized vehicle ownership) while the second involved reclassifying single-child households to multiple child households (reflecting an increase in household size and number of children). The impacts of each of these changes on the predicted aggregate shifts of alternatives were obtained, and these were translated to a mean absolute change in aggregate shares across all alternatives for each tour type, each application region, and each estimated model. The results of this exercise did not reveal any clear "winners" and "losers" in terms of transferability potential. Indeed, in some cases, the TB models that performed quite poorly when transferred to the other regions in the disaggregate and aggregate analysis performed very well in the elasticity analysis.

Overall, the results indicate that models that may appear to be very transferable based on the traditional disaggregate and aggregate data fit analyses do not necessarily imply good transferability on the basis of elasticity effects. Also, there is evidence of asymmetry in transferability, so that the typical practice of identifying two regions that appear to be "similar" based on, for example, population size, spatial context, or other demographics may not be sufficient to guarantee transferability in both directions.

D.3   Sikder and Pinjari (2013)

D.3.1 Activity-Travel Dimensions Examined

Sikder and Pinjari focus on the spatial transferability of person-level daily activity generation and time-use models. They indicate that activity generation and time-use models are likely to be more transferable than those for travel choices (such as mode choice, location choice, and time-of-day choice) because of a comparatively lower influence of the spatial and transport context of a region on activity generation/time use than on travel scheduling choices.

Accordingly, they study individual-level activity generation/time-use using the multiple discrete-continuous extreme value (MDCEV) model developed by Bhat (2005, 2008) that is increasingly being incorporated today in several activity-based travel model systems (see, for example, Bhat et al., 2013). Specifically, weekday activity generation (whether or not an individual participates in one of several out-of-home activity purposes) and time-use (in-home time and out-of-home time investment in which the individual participates) is modeled for unemployed adults (age>18 years of age) using the MDCEV model. In addition to in-home activity, eight different types of out-of-home (OH) activities are identified (thus, the generation and time-use is by nine total activity purposes):

5. Shopping;
6. Other maintenance (buying gas, services, clothing, etc.);
7. Social and recreational;
8. Physically active recreation (gym activities, exercising, playing sports, etc.);
9. Medical;
10. Eating out;
11. Pick up/drop off; and
12. Other activities.

The daily time-use in each of these OH activity types is computed by aggregating the activity durations of all episodes (during the survey day) of each activity type. The time spent in in-home activity is computed as the total daily time minus the time allocated to OH activities, sleep time (assumed to be 8.7 hours), and travel time.

D.3.2 The Data Context

The study uses data from the 2009 National Household Travel Survey (NHTS) drawn from California and Florida. Within Florida, five specific regions/area types are identified to assess transferability within the state (no such intra-state transferability analysis is undertaken for California). The five regions/area types are: Southeast (SE) Florida, Central Florida, Tampa Bay (TB), District 1 urban areas (D1U) comprising a group of less urbanized counties, and rural Florida.

For the intra-state transferability exercise, models were estimated only from the SE Florida, Central Florida and TB (because of the small sample sizes from D1U and rural Florida), but the estimated models were transferred to all the four other regions/area types in Florida (for a total of 12 transferability instances). For the inter-state transferability, the entire data from Florida was used for estimation and transferred to California, and similarly the entire data from California was used for estimation and transferred (for a total of two transferability instances). Explanatory variables included gender, age, race, driver license holding status, education level, income, household size, and number of drivers in the household.

D.3.3 The Transferability Approach and Results

The transferability of models is undertaken using two different methods: the simple transfer approach (used in the two ABM transferability studies discussed earlier) and the constants updating approach. This leads to a total of 28 transferability instance-approach (14 transferability instances times 2 approaches per instance) combinations. The transferability of a model is assessed, similar to Nowrouzian and Srinivasan, in three different ways: (1) examining disaggregate predictions, (2) examining aggregate predictions, and (3) assessing policy prediction performance.

At the disaggregate level, two metrics are used to assess transferability. The first is the transferability test statistic (TTS), which tests the hypothesis that the predictive log-likelihood of the transferred model to the application context is the same as the log-likelihood of the estimated model, and takes the usual likelihood ratio test statistic form. The second is the transfer index (TI) measure that is the ratio of the difference in the log-likelihoods of the transferred model and the constants only model estimated in the application region to the difference in the log-likelihoods of the locally estimated model and the constants only model estimated in the application region.

The authors indicated that all 28 transferability instance-approach combinations were rejected based on the TTS test. On the TI measure, the intra-state transferability for the 12 instances were quite poor when using the simple transfer approach, with transferability being better between the major urban regions (SE Florida, Central Florida, and TB) than from these three urban regions to the other two non-urban area types (D1U and rural Florida). But there was asymmetry in transferability success even between the three major urban regions (as opposed to the study of Nowrouzian and Srinivasan, who only found asymmetry in the aggregate-level predictions, but not at the disaggregate-level predictions).

However, there was substantial improvement in the TI measure for intra-state transferability for all 12 instances when using the constants updating approach, with the improvement ranging from a 16% improvement in the index to as much as a 400% improvement in the index. For inter-state transferability, transferability was very poor for the simple transfer approach, with the TI index taking negative values (that is, the transferred model does worse than a constants only model estimated in the region). However, when using the constants updating procedure, the transferability improved very dramatically, with the TI index taking the value of close to 0.8 for both the directions of inter-state transfer instances.

At the aggregate share level, the metrics used included the root mean square error (RMSE) and relative aggregate transfer error (RATE). The RMSE measure used by the authors is the root of the mean square of the differences between the aggregate predicted and observed shares, weighted by the predicted share of each alternative. The RATE measure is a derivative of the RMSE measure that takes the ratio of the RMSE measure for the transferred model relative to that of a locally estimated model. The results, consistent with the disaggregate model, revealed:

- Better transferability between the urban areas than from the urban areas to the D1U area or to rural Florida;
- Relatively poor transferability across states; and
- Substantial improvements across the board due to constants updating.

The authors also examined transferability on the basis of the changes in the discrete choice probabilities as well as time investment due to an increase in age (for every individual by 10 years) in the application region . The change was measured based on the predicted shifts in shares and overall time use invested in each alternative between the base case and the policy case. These predicted shifts from the transferred models were compared with the predicted shifts from a locally estimated model. Very interestingly, in this case, updating constants did not improve transferability, and, in fact, even degraded transferability in some cases.

Overall, the results emphasize the point that, while updating constants can improve data fit in terms of shares and predicted time-use, this is not guaranteed to improve the forecasting performance of the model in response to changes in specific variables. The authors also suggest that there is better correlation between the transferability performance as captured by the TI index and the policy response of variables, rather than as captured by the aggregate prediction indices and the policy response of variables. That is, the disaggregate measures and the policy prediction measures may be more appropriate metrics to assess transferability than the aggregate prediction metrics. From a more practical standpoint, the study concludes that transferability between regions within a state appears to be more preferable than transferability across state boundaries.

D.4   Bowman et al. (2013)

D.4.1 Activity-Travel Dimensions Examined

The latest, and the most comprehensive analysis of transferability in the context of ABMs, is the report by Bowman et al., for an FHWA study. They used an activity-based model structure labeled DaySim (for Daily Simulator) that has been employed in Sacramento and Seattle among other locations.

A detailed description of the DaySim framework is beyond the scope of this report. Very broadly, DaySim consists of an econometric micro-simulation system with a three-tier hierarchy of: (1) Day-level activity pattern choice models (or, simply, pattern-level choice models), (2) Tour-level choice models, and (3) Trip/Stop-level choice models.

The pattern-level models consist of the daily activity pattern model and the number of tours model. These models predict: (a) the occurrence (and the number) of home-based tours (i.e., tours that originate and end at home) specifically for each of the following seven activity purposes during a day: work, school, escort, personal business, shopping, meal, and social/recreational, and (b) the occurrence of additional stops/trips that may occur (in other tours) for these seven purposes.

The tour-level models predict the primary destination (i.e., the destination of the primary stop for which the tour is made), travel mode, time-of-day of travel (i.e., time of arrival at, and time of departure from, the primary destination), and the number of additional stops by purpose (other than the primary stop) for all tours. Tour-level models also include a work-based tour (i.e., a tour that originates and ends at work) generation model that predicts the number (and purpose) of work-based tours for each home-based work tour predicted by models in the pattern level.

The stop-level models predict the stop location (or destination), mode choice, and time-of-day of travel for each of the stops (other than the primary stops) generated in the previous steps. The entire hierarchy of models just discussed is preceded by models for long term choice, including work location choice and auto ownership choice.

In the study, the model components examined in the context of transferability included 14 of those in the overall DaySim framework:

13. Regular work location choice
14. Auto ownership
15. Daily activity pattern (that jointly predicts whether or not a person participates in tours and extra stops for 7 activity purposes in a day)
16. Number of tours (that predicts the number of tours for each of the 7 activity purposes for which tour making is predicted in the daily activity pattern model)
17. Non-work tour primary destination choice for each non-work tour predicted in #4
18. Work-based sub-tour generation (that predicts the number and purpose of work-based sub tours that originate for each home-based work tour predicted by earlier models)
19. Work tour main mode choice
20. School tour main mode choice
21. Other tour main mode choice
22. Work tour time period choice (predicts half-hour time periods of arrival at and departure from the regular work location)
23. Other tour time period choice
24. Intermediate stop generation (predicts the exact number and purpose of stops for the half-tours leading to and from the primary destination of the tour)
25. Intermediate stop location (predicts the destination zone and parcel of each intermediate stop, conditional on tour origin and primary destination, and location of previous stops)
26. Trip departure time (predicts arrival time (departure time) choice for stops in first (second) half tour, conditional on the time windows remaining from previous choices).

All models were estimated as MNL models. In addition to these 14 models, an additional 15th model was considered that is a variant of the work tour mode choice model in DaySim. In particular, the DaySim work tour mode choice model included a generalized cost coefficient with pre-specified coefficients on the many time and cost coefficients. The 15th model included a work tour mode choice model with freely estimated coefficients on the many time and cost variables.

D.4.2 The Data Context

The study used data from the 2009 National Household Travel Survey (NHTS) "add-on" program drawn from four regions in California (Fresno, Northern San Joaquin Valley or NSJV, Sacramento, and San Diego) and two regions in Florida (Jacksonville and Tampa). The authors employed a variable specification derived directly from the Sacramento model specification,

though some simplifications were made to make the models easier to estimate. The transferability instances in this study are closely related to the transferability approach used, and so are discussed in the next section.

D.4.3 The Transferability Approach and Results

The transferability of models was assessed by estimating a base model using the Sacramento specification for the estimation context, and then estimating another "difference" model after combining data from both the estimation and application contexts. The "difference" model variable specification included the same set of variables as in the base model plus a set of additional variables that are interaction variables. These interaction variables are constructed by multiplying each of the base model variables with an application context dummy variable that takes a value of '1' for the application context observations and zero for the estimation context observations. Then, the coefficients on these interaction variables represent differences in the coefficient estimates between the estimation context and the application context.

If a "difference" coefficient corresponding to a variable is statistically significant, then it implies that that variable's impact is not transferable between the estimation and application contexts. This is similar to the "elasticity effect" measure of transferability used in Nowrouzian and Srinivasan (2012) and the "policy response" measure used in Sikder and Pinjari (2013), except the sensitivity differences in the variable effects are captured directly through estimation. In this set-up, one only needs to define the "estimation context" and the "application" context.

In Bowman et al.'s study, several base models are estimated, each corresponding to a specific "estimation context". There are twelve base models, six of which are estimated models based on data solely from each region (the estimation context corresponding to each of these base models is one of the six regions considered in the study; these are used to examine the estimability and significance of variables when interpreting the results of the transferability analysis though they are not directly used in the transferability analysis). The remaining six base models are as follows:

- A single "2-state" base model that includes data from all the six regions in the study (the estimation context is the entire set of six regions).
- Two "1-state" base models, one that is the California state base model and includes the four regions of California (the estimation context is the State of California, as represented in the four regions in California), and a second that is the Florida state base model and includes the two regions from Florida (the estimation context is the State of Florida, as represented by the two regions in Florida).
- A single "2-state+ASC" base model that has the same specification as the "2-state" base model, but includes an additional set of alternative specific constants for one of the regions (the estimation context is the entire set of six regions, with constant updating to allow the constants to vary across states).
- Two "1-state+ASC" base models, one that is the California state base model and includes data from the four regions of California with a set of alternative specific constants for each of the four regions in the state (the estimation context is the State of California with constant updating for each region), and a second that is the Florida state base model and includes data from the two regions from Florida with a set of alternative specific constants for each of the two regions in the state (the estimation context is the State of Florida with constant updating for each region).

Using the above six base models, 24 difference models were estimated as follows:

- Six difference models, one for each region, with the "2-state" model as the base;
- Four difference models, one for each California region, with the "1-state" California model as the base;
- Two difference models, one for each Florida region, with the "1-state" Florida model as the base;
- Six difference models, one for each region, with the "2-state+ASC" model as the base;
- Four difference models, one for each California region, with the "1-state+ASC" California model as the base; and
- Two difference models, one for each Florida region, with the "1-state+ASC" Florida model as the base.

The "difference" variable coefficients provide a sense of transferability of each variable effect. Another way to examine transferability is to undertake a standard likelihood ratio test between the difference model and the base model though this test tends to be very stringent and generally rejects transferability. The analysis in the study considered several explanatory variables, including person and household-level characteristics, land-use and accessibility variables, zone-to-zone transportation system variables.

The analysis provided several results, with the obvious caveat that these results are but general guidelines that have to be more rigorously examined through additional analyses on transferability from other regions, and using other model structures than the multinomial logit structures used in the DaySim structure components.

- A sample size of 6000 households for a region appears to be a good first target for survey data collection to support the development of an activity-based travel model, with potential benefits accruing from larger sample sizes. However, caution should be exercised in this result, simply because the study did not have a region with a sample size between 2500 households and 6000 households that may have provided additional information on the relative benefit as the sample size is increased from 2500 to 6000 households.
- Transferring ABMs from a "comparable" region with a large sample may be better than estimating ABMs from a much smaller local sample. This is particularly so because there is some evidence in the results that introducing a rich specification of individual-related variables that accommodates preference heterogeneity and response heterogeneity to spatial/transport context variables results in better transferability than models that ignore such effects. And estimating stable coefficients for specific segment groups (such as low income or high income or carless households) implies the need for a large overall sample size in the region from which to transfer. But there is still value in having a small local sample or aggregate data in the application context, so that constant updating can be pursued to improve transferability.
- Models that are associated with social organization (activity generation and sequencing) appear to be more transferable in general than those that deal with spatial organization (travel mode and location choice). Thus, models associated with work durations, timing and durations of activity participation, auto-ownership, and activity participations and chronological sequencing generally are more transferable than mode and location

choice decisions. A practical implication is that much more effort may need to be expended to update parameters (with local samples) when transferring mode and destination choice model components than when transferring other components of an ABM.

The Tampa Bay variable coefficients, in general, seemed to be rather different from those of the coefficients from other regions, including Jacksonville. When combined with the results of the earlier studies, it does seem that there is something very different about the Tampa Bay region compared to other regions even within Florida. This once again raises the problem that it is not always clear what may be a "comparable" region for model transferability. However, in general, ABM components are more transferable between regions within a state than between regions in different states.

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United State Government assumes no liability for its contents or use thereof.

The United States Government does not endorse manufacturers or products. Trade names appear in the document only because they are essential to the content of the report.

The opinions expressed in this report belong to the authors and do not constitute an endorsement or recommendation by FHWA.

This report is being distributed through the Travel Model Improvement Program (TMIP).

Federal Highway Administration