Spatial Transferability of Tour-Based Time-of-Day Choice Models: An Empirical Assessment

Sujan Sikder
Transportation Systems Modeler
Parsons Brinckerhoff
400 SW Sixth Avenue, Suite 802
Portland, OR 97204
Tel: (813) 507-3953; Fax: (813) 974-2957, Email: sikders@pbworld.com

Bertho Augustin
Graduate Research Assistant
Department of Civil & Environmental Engineering
University of South Florida
4202 E. Fowler Ave., Tampa, Fl 33620
Tel: (239) 285-3669; Fax: (813) 974-2957, Email: bertho@mail.usf.edu

Abdul Rawoof Pinjari*
Assistant Professor
Department of Civil & Environmental Engineering
University of South Florida, ENC 2503
4202 E. Fowler Ave., Tampa, Fl 33620
Tel: (813) 974- 9671; Fax: (813) 974-2957, Email: apinjari@usf.edu

Naveen Eluru
Assistant Professor
Department of Civil Engineering & Applied Mechanics
McGill University
Suite 483, 817 Sherbrooke St. W.
Montréal, Québec, H3A 2K6, Canada
Tel: (514) 398-6823, Fax: (514) 398-7361, Email: naveen.eluru@mcgill.ca

* Corresponding author
Submitted for Presentation and Publication
93rd Annual Meeting of the Transportation Research Board
Committee: ADB40 Travel Demand Forecasting

Word count: 6341 (text) + 3tables x 250 + 2figures x 250 = 7591 equivalent words

Revised: November 15, 2013

ABSTRACT

An empirical assessment is presented on the transferability of tour-based time-of-day choice models across different counties in the San Francisco Bay Area. Transferability is assessed using two different approaches: (1) application-based approach, and (2) estimation-based approach. The former approach tests the transferability of a model as a whole while the latter approach allows the analyst to test which specific parameters in the model are transferable. In addition, the hypothesis that pooling data from multiple geographical contexts helps in developing better transferable models than those estimated from a single context was tested. The estimation-based approach yields encouraging results in favor of time-of-day choice model transferability, with a majority of parameter estimates in a pooled model found to be transferrable. Pooling data from multiple geographical contexts appears to help in developing better transferable models. However, attention is needed in selecting the geographical contexts to pool data from. Specifically, the pooled data should exhibit same demographic characteristics and travel level-of-service conditions as in the application context.

1 INTRODUCTION

Spatial transferability of travel forecasting models can help in significant cost and time savings for transport planning agencies that cannot afford extensive data collection and model development. This issue is particularly relevant for tour-based/activity-based models (ABMs) whose development typically involves significant data inputs and long production times. However, only a handful of recent studies document transferability assessments of ABM model components [(*1-4*)]. Empirical assessments of tour-based time-of-day (TOD) choice models are even fewer [(*3-5*)].

TOD choice models are used to model individual-level travel timing choices for forecasting aggregate-level temporal variations in traffic volumes. Sound TOD choice models are paramount to travel forecasting models, because evaluations of several travel-demand management strategies rely on accurate predictions of the temporal variation in travel volumes. Therefore, several studies proposed appropriate methods to model individuals' travel timing within a tour-based approach [(*6-10*)]. Besides, recent studies [(*3,4*)] suggest that time-of-day choice models may be better transferable than other components of ABMs such as mode choice and location choice models. This is perhaps because individuals' time-of-day choices tend to be much less connected to land-use characteristics, spatial structure, and quality attributes of travel modes that tend to vary across geographical contexts but not easy to capture in empirical models. Therefore, transferability of time-of-day choice models is a fruitful avenue for research – for accumulating empirical evidence on what aspects of these models are transferable and for understanding how best to transfer such models.

1.1 Transferability Assessment Techniques

A variety of different approaches and metrics have been used in the literature to assess transferability of travel model parameters. Bowman et al. (*3*) classify the available transferability assessment approaches into two broad categories: (a) application-based approach, and (b) estimation-based approach. In the former approach, the model parameters are estimated using data from one region (the base context) and "applied" to data in other region (the application context) to assess how well the model predicts in the other region. This approach generally tests the transferability of models as a whole, without allowing an examination of which specific parameters are transferable. In the latter approach (also called joint-context estimation; see (*11, 12*)), data from both estimation and application contexts is combined to estimate a single model, while recognizing potential differences between the two contexts by estimating "difference" parameters. Simple t-tests on these difference parameters shed light on whether the parameter estimates are different between the two contexts. A particular advantage of this approach is that one can test if each (and every) parameter in a model is transferable (*3*).

The application-based approach has been the predominantly used approach, with most empirical evidence suggesting the difficulty of model transferability across geographical contexts. However, it is possible that the approaches used to assess model transferability may influence transferability results. For example, small sample sizes, if used in model estimation in the application-based approach, can easily confound the model transferability results toward less transferable. The estimation-based approach, on the other hand, helps alleviate such sample size issues (to an extent) by estimating common parameters where the data variability is small within a single context. Besides, many transferability assessment metrics are based on comparing the log-likelihood measures and predictions (in the application context) from the models estimated separately in the base and application contexts. Although many parameters may not be statistically different between the base- and application-context models, small numerical differences for a large number of parameters may lead to non-negligibly different log-likelihoods and predictions. In view of these issues, it would be useful to investigate the differences between the approaches on model transferability results.

1.2 Development of Better Transferable Models

As discussed earlier, joint-context estimation approach uses data from both the base and application contexts. An extension of the joint-context estimation approach is to pool data from multiple contexts, as opposed to only two contexts. In situations where the variation in important socio-demographic, land-use, or level-of-service variables is insufficient in either contexts, pooling data from multiple contexts can potentially help in achieving sufficient variation for better model specification and estimation. Several studies (*13-15*) allude to this strategy for developing better transferable models. Therefore, it would be useful to accumulate empirical evidence on the extent to which and the reasons for which pooling data helps in enhancing model transferability. It is also important to investigate if the improvement in transferability after pooling data is only due to sample size increase, or if the specific geographical contexts from which data is pooled have a bearing. If the geographical context has influence, it is useful to understand where to pool data from.

1.3 Current Research

The overarching aim of this paper is to investigate the spatial transferability of tour-based time-of-day choice models. To this end, an empirical assessment is conducted to assess the transferability of a commute tour start and end time choice model among different counties in the San Francisco Bay area. The transferability assessments aim to investigate what aspects of time-of-day choice models are transferable and what aspects are less transferable. In addition, the results from different transferability assessment approaches (i.e., application-based and estimation-based) are compared and contrasted. Furthermore, extensive explorations are performed to investigate the extent to which and the reasons why pooling data from multiple contexts helps in achieving better transferable models. These explorations resulted in useful guidelines on where to borrow data from, for developing transferable time-of-day choice models.

The next section provides an overview of the data and geographical contexts considered in this study. Section 3 describes the model structure and the approach used to assess transferability. Section 4 presents empirical results. Section 5 concludes the paper.

2 DATA

The primary data source used for the analysis is the 2000 San Francisco Bay Area Travel Survey (BATS), an activity-based travel survey that collected information on all in-home and out-of-home activities over a two-day period from over 15,000 households in the San Francisco Bay Area. The geographical regions considered in this paper are different counties in the San Francisco Bay Area. For a majority of transferability assessments, we focus on the following six out of the nine counties in the Bay area: Alameda (AL), Contra Costa (CC), Santa Clara (SC), San Francisco (SF), San Mateo (SM), and Sonoma (SN). This helped in keeping the model estimation efforts manageable. Besides, the sample sizes from the other three counties were too small. Table 1 presents the descriptive information on the data in each of the six counties. It can be observed that the employed adults in Santa Clara are different from those in most other counties. For example, there appear to be greater proportions of full time workers, those with flexible work schedules, and with higher income levels in Santa Clara than in other counties. The closest County to Santa Clara, in terms of overall characteristics considered in this table, is San Mateo. Among the other counties, the employed adults from Contra Costa appear to be similar to those from Alameda County especially in their employment characteristics - employment type, status, and schedule flexibility. In the context of land use characteristics and household structures, San Francisco County appears to be different from the other counties in the Bay Area. Specifically, greater proportions of single person households and employed adults living in urban areas are observed in San Francisco County. Although descriptive statistics cannot shed full light on model transferability, the noted differences or similarities may have a bearing. Finally, note that the sample sizes for the San Francisco and Sonoma Counties are relatively small; caution is warranted in interpreting model transferability results for these Counties.

3 METHODOLOGY

The multinomial logit (MNL) structure was used for the tour-based time-of-day choice model. To define the choice alternatives, individuals' work tour start and end times were categorized into discrete, half-hour intervals in a day. Next, based on the observed tour start- and end-times in the data set, some of the consecutive half-hour intervals were aggregated into larger time intervals. As a result, a total of 25 different time-slots were used for tour start time choice and 21 time-slots were used for tour end time periods. Since the models are for predicting the joint choice of tour start-and end-times, the tour start time slots were combined with those of tour end time slots, resulting in a total 386 feasible alternatives, each representing a combination of tour start and tour end time intervals.

Following (*10*), the utility functions of the MNL models comprise three parts, one for start-time, one for end-time, and the other for duration:

$$U(s,e) = U^s(t_s) + U^e(t_e) + U^{dur}(t_e - t_s) \qquad (1)$$

In the above equation, $U(s,e)$ is the joint utility of starting the tour in time slot *s* and ending in time slot *e*, with t_s and t_e as the mid points of the time slots (measured from 3:00am as the beginning of the day), $U^s(t_s)$ is start-time function, $U^e(t_e)$ is end-time function, and $U^{dur}(t_e - t_s)$ is duration function.

Start-time function, $U^s(t_s) = \sum_r x_r f^s(t_s) + \beta_{thw}(\text{travel time})_{hw} + \beta_{chw}(\text{travel cost})_{hw} + \ln(\text{\# half-hour periods in slot s})$

End-time function, $U^e(t_e) = \sum_r x_r f^e(t_e) + \beta_{twh}(\text{travel time})_{wh} + \beta_{cwh}(\text{travel cost})_{wh} + \ln(\text{\# half-hour periods in slot e})$

Duration function, $U^{dur}(t_e - t_s) = \beta_1^{dur}(t_e - t_s) + \beta_2^{dur}(t_e - t_s)^2 + \beta_3^{dur}(t_e - t_s)^3 .... \beta_d^{dur}(t_e - t_s)^d \qquad (2)$

In the start- and end-time functions of the above equations, *r* is the number of demographic explanatory variables x_r used in the model, including constants; and $f^s(t_s)$ and $f^e(t_e)$ are specified as trigonometric cyclic functions of t_s and t_e, respectively, to ensure that the function value at a time period "*t*" is same as that at "*t+24*" (same time next day).

$$f^s(t_s) = \beta_1^s \sin\left(\frac{2\pi t_s}{24}\right) + \beta_2^s \sin\left(\frac{2.2\pi t_s}{24}\right) + \beta_3^s \sin\left(\frac{3.2\pi t_s}{24}\right) .... + \beta_n^s \sin\left(\frac{n.2\pi t_s}{24}\right)$$

$$f^{e}(t_e) = \beta_1^e \sin\left(\frac{2\pi t_e}{24}\right) + \beta_2^e \sin\left(\frac{2.2\pi t_e}{24}\right) + \beta_3^e \sin\left(\frac{3.2\pi t_e}{24}\right) .... + \beta_d^e \sin\left(\frac{d.2\pi t_e}{24}\right) \qquad (3)$$

In the above two equations, *n* and *d* are the number of terms in the start- and end-time cyclic functions, determined using statistical tests and the reasonableness of resulting temporal profiles of the functions. Note that the individual coefficients in these cyclic functions cannot be interpreted meaningfully. For interpreting the effect of a variable (say female with kids), all the corresponding coefficients in the cyclic function should be used to plot the utility profiles as a function of time-of-day.

The start- and end-time functions in the utility specification include time-varying travel conditions; i.e., travel times and travel costs for each of time-of-day alternative in the model. One potential source of this data is the travel time and cost skims from the regional travel model. However such skims were available only for very broad time periods in the day making it difficult to represent the variation in travel conditions within each broad time period. Therefore, following (*16*), an auxiliary travel duration regression model was developed to impute time-varying travel times using reported travel durations in the data. The regression model specifies the ratio of reported travel duration for a trip and the free-flow time (obtained from the travel time skims) between the origin-destination pair of the corresponding trip as the dependent variable, and zonal land-use characteristics, trip distance, and time-of-day as independent variables in the model. The model is formulated as below:

$$\frac{[\text{Travel Duration}]_{ijt}}{[\text{Free Flow Time}]_{ij}} = \text{intercept} + \sum_k \beta_k x_k + \alpha_n \exp[\sin^n(\frac{\pi t}{12})] + \gamma_n \exp[\cos^n(\frac{\pi t}{12})] \qquad (4)$$

In the above equation, $[\text{Travel Duration}]_{ijt}$ is the reported travel duration between zone *i* and zone *j* at time-of-day *t*, and $[\text{Free Flow Time}]_{ij}$ is the free flow travel time between zone *i* and zone *j*. The time-of-day "*t*" is measured in hours elapsed from an arbitrary time (3:00AM in this study). x_k includes the zonal land-use and OD distance variables, while the next two expressions with sine and cosine functions are cyclic functions of time-of-day *t*. The coefficients α_n and γ_n on the cyclic functions capture time-of-day effects on travel duration. The number of α_n and γ_n coefficients to be estimated is determined empirically based on statistical fit and intuitive considerations. In this paper, we used n = 3 since increasing n beyond 3 did not improve the statistical fit to data. Besides, the model with n = 3 resulted in the most reasonable travel time profiles reflecting the temporal variations observed in the data.

Unlike travel times, the survey did not collect information on travel costs for each trip. Therefore, as in previous studies (*6-10, 16*), travel cost information was extracted from the regional model's travel cost skims. Consequently, travel costs in the model vary across very broad time periods in the day.

3.1 Transferability Assessment Approaches

Transferability assessments were conducted using two approaches: (1) application-based approach and (2) estimation-based approach. For the application-based approach, separate models were estimated using data from each of the six counties and transferred to the other five counties. In addition, using the same approach, for each County, we assessed if a model built using data pooled from all eight Counties in the Bay area (other than the County to which the model is transferred) is better transferable than a model built with data from a single County. Further, joint-context models were estimated for each County to test if the behavior in each County was different from that in other eight counties.

From now on and throughout the paper, a model built using data from all nine counties will be denoted by the term "base" model, while the models using data from eight counties (i.e., the pooled-data model without a specific County "c" under consideration) will be indicated by the term "base-c" model, where "c" could be any of the six counties – AL, CC, SC, SF, SM, and SN. For example, "base-AL" indicates the model built using data from all eight counties in the Bay Area except Alameda.

To assess the prediction ability of the transferred models using the application-based approach, a transferability assessment metric called Transfer Index (TI) (*17*) was used.

$$TI_j(\beta_i) = \frac{L_j(\beta_i) - L_j(\beta_{reference,j})}{L_j(\beta_j) - L_j(\beta_{reference,j})} \qquad (5)$$

where, $L_j(\beta_i)$ is log-likelihood of the transferred model applied to the application context data, $L_j(\beta_j)$ is log-likelihood of the locally estimated model using data from the local/application context, and $L_j(\beta_{reference,j})$ is log-likelihood of a reference model (constants only model) in the local/application context. The closer the value of TI is to 1, the closer is the transferred models' performance to a locally estimated model (in terms of the information captured).

To assess transferability using the estimation-based approach, first a base model was estimated using data from all 9 counties in the San Francisco Bay area. Next, for each selected County, the dummy variable for that County was interacted with each of the variables in the base model to form the "difference" variables. The coefficients on these "difference" variables help in assessing if the corresponding variables are transferable between the specific County and the rest of the Bay area. Groups of such difference variables were included one by one in the model specification. For example, to test if females with kids in Santa Clara County had different time-of-day preferences from those in all other counties, the dummy variable for Santa Clara was interacted with all the variables in the cyclic functions for the female with kids demographic segment. All these interactions were introduced simultaneously over the base model. The resulting model would recognize any potential differences in the time of day preferences of females with kids between Santa Clara and other Counties in the Bay area. The decision of whether or not the preferences of females with kids were actually different between Santa Clara and other Counties was made based on a log-likelihood ratio (LLR) test for the entire set of Santa Clara specific variables just added. If the interaction variables (i.e., "difference" variables), as a set, are statistically different from zero (based on the LLR test), that indicates statistically significant differences in the time-of-day preferences between females with kids in Santa Clara and those in other Counties (hence the corresponding coefficients are NOT transferable). In addition to such statistical tests, the utility profiles were plotted as a function of time-of-day for females with kids in Santa Clara and for those in all other 8 counties to visually examine if the profiles appeared different. This approach was repeated for all demographic variables and level of service variables in the model specification until a final specification was arrived at. The final specification contains the specification for the base-SC model (that is the model for all 8 counties except Santa Clara) as well as the "difference" variables that were deemed to be statistically different from the base specification (the "difference" variables that were deemed insignificant were dropped from the model). Using the same approach, joint context specifications were developed for each of the six counties considered in this study to examine what types of coefficients were transferable and what were not. It is worth noting here that, in addition to allowing for County-specific differences in the deterministic utility functions, the differences in the unobserved factors were also allowed by estimating the ratio of the random error terms (i.e., scales) between the base and the application contexts.

Finally, the base-c specification and the County-specification (for all six counties) were extracted from the above-described joint-context models to compute a transfer index value using Equation (5). This TI value was compared with the TI value of transfers from a base-c model that was estimated by pooling data from all eight Counties except County “c”. This helps in comparing the extent to which the transferability results differ when models are estimated separately (in the estimation and application contexts) vis-à-vis when models are estimated using joint-context estimation. Similarly, the TI values from transferring the base-c models were compared with those from transferring the individual county models to assess the extent to which pooling data helps in estimating better transferable models.

4 EMPIRICAL RESULTS

The time-of-day choice model estimation results are not reported in the form of tables to conserve space, but the important variables in the model specifications are discussed here briefly. The statistically significant demographic variables in the models include age, gender, females with children, income levels, full-time/part-time work status, work schedule flexibility (flexible vs. inflexible), and employment type (government employees). The time-of-day preference profiles for some of these demographic variables are plotted in Figure 1 in the form time-varying utility profiles for each variable. As can be observed from the tour start time profiles in the top row of the figure, the tour start utility profile for full-time workers is toward the left compared to that for part-time workers. This suggests full-time workers are likely to start their work tours earlier than part-time workers. Further, in the bottom row of the figure, as expected, female workers with children show a higher propensity to start their work tours later in the day when compared to males or females without children. Similarly, the tour end-time profiles in Figure 1 show the influences of employment status (full time vs. part time), income levels, and presence of children at home on their work tour end-time choices. Full time workers and high income workers show a higher propensity to end their work tours after 5:00 PM when compared to their counterparts. Female workers with kids in households are found to end their work tours earlier than their counterparts. While not depicted in figures, all other demographic effects were reasonable and consistent with previous specifications in the literature.

The level-of-service variables that turned out statistically significant in the models include time-varying travel times (separately for home-work and work-home journeys) and travel costs, along with their interactions with work schedule flexibility (individuals with flexible work schedules were less sensitive to increases in travel times than those without flexible work schedules). The tour duration function also shows reasonable profiles with maximum utility for tour durations around 10 hours for full-time workers and smaller durations for part-time workers.

Overall, the parameter estimates in most models have intuitive interpretations except in the County-specific models for San Francisco and Sonoma. In these two models, some of the coefficients and utility profiles appear unintuitive, perhaps due to small sample sizes, which will have a bearing on transferability assessments.

4.1 Transferability Assessment Results

4.1.1 Results from the Application-based Approach

The first part of Table 2 presents the TI values for inter-County transfers conducted using the application-based approach. As can be observed, the models transferred from and to Sonoma provide the lowest TI values, followed by those for the San Francisco. In most of the cases, higher TI values can be observed for models transferred from and to Santa Clara. Overall, Comparing the TI values to the corresponding sample sizes from each County suggest that the sample sizes have a strong influence on the TI values (e.g., San Francisco and Sonoma have the smallest samples, while Santa Clara has the largest sample). Therefore, it is difficult to draw inferences on what County characteristics help make the models more or less transferable.

4.1.2 Results from the Estimation-based Approach

As discussed earlier, in the estimation-based approach, County-specific "difference" variables were added to the base specification to explore any potential differences in the parameter estimates for each County under consideration and the rest of the Bay area. Important observations from these model results are discussed here without presenting the parameter estimates themselves (to conserve space). <u>First,</u> the "difference" variables for the time-of-day specific constants for most of the counties were not statistically significant; suggesting the potential transferability of model constants from a model built using data pooled from all other eight counties to a specific County. <u>Second,</u> among the level-of-service variables, while the travel time coefficient for the home to work journey was found to be statistically different (hence not transferable) between specific counties and the rest of the Bay area, the travel time coefficient for the work to home journey and the travel cost coefficients appear to be transferable. <u>Third,</u> in the context of other variables, e.g., socio-demographic variables, almost 95% of the coefficients (or more) in a County TOD model were not significantly different from the corresponding base-c model (when tested using "difference" variables). Overall, less than 5% of the coefficients were not found to be transferable. This provides encouraging empirical evidence of the transferability of time-of-day model coefficients from a pooled model. <u>Another finding</u> in favor of transferability is that no significant differences were found between the scales of the random error components for the base-c and county-specific utility functions (in all six models for six counties).

In addition to assessing transferability based on the statistical significance of "difference" variables in joint-context models (based on LLR tests), the time-varying utility profiles of the variables that were found to have statistically significant "difference" coefficients were compared. Some of them are presented in Figure 1 for illustration to the reader. One can observe from the profiles in the first row of the figure that the tour start- and end- time preference profiles of full-time employed adults in Santa Clara are visually different from those in other eight counties. Such differences in the time-of-day preference profiles are observed for the household income variable as well (profiles in the second row), however only for the tour end-time profiles. Similarly, the utility profiles in the third row show more discernible differences in the tour end-time profiles for females with kids between the San Mateo and other eight counties in the Bay area, than the differences in the tour start-time profiles. These and other similar results (not shown in figures) suggest that the demographic preference coefficients related to home-to-work journeys may be more transferable than those related to work-to-home journeys. The reason for this is not clear and needs further investigation.

4.1.3 Comparison of Results from Application-based and Estimation-based Approaches

To compare the performance of the application- and estimation-based approaches, the parameters estimated from base-c models (using both approaches discussed before) were transferred to each of the six counties considered in this study. As mentioned earlier, the base-c models are basically pooled models developed using data from eight counties (other than the County to which the model is transferred). For each County, two different "base-c" models are available – (1) the base-c model extracted from the joint-context estimation (where data from the specific County "c" was also included but "difference" variables were used to allow difference between the County "c" and other eight counties), and (2) the base-c model estimated separately using only the data from eight counties. The second part of Table 2 presents the TI values obtained from these transfers. As can be observed, the TI values of base-c models extracted from joint-context estimations are consistently higher than those from the application-based approach. This is because the estimation-based approach considers only the statistically different parameters that are different between the base and application contexts, while keeping all other parameters the same. The application-based approach, on the other hand, allows the values of all parameters to be different, even if they are not statistically different. All such numerical differences among the parameters add up to a substantial difference in terms of log-likelihood values (and even predictions and elasticity measures). Thus, whenever possible, it is useful to employ the estimation-based approach for transferability assessments.

4.1.4 Does Pooling data from Multiple Regions Result in Better Transferable Models?

Comparison of the TI values in the first part of Table 2 (i.e., for inter-county model transfers) to those in the second part (i.e., for transfers from pooled-data models, or base-c models, to specific counties) suggests that transferability improves significantly after pooling data from all other eight counties. These results point to the potential benefits of pooling data from multiple regions for developing better transferable models. However, it is not clear if the observed improvement in transferability is only due to increased sample sizes, or if the specific geographical contexts from which data comes also have an influence on this improvement. This section attempts to address this issue via carefully conducted experiments that control for sample size issues. Figure 2 presents the results of all the experiments conducted to understand the influence of, and the reasons why it helps to, pooling data from multiple regions on model transferability.

First, we investigated the influence of sample size (used to build the base context model) on model transferability, in the context of transfers from Santa Clara to Alameda. To do so, first 10 random samples, each of 500 individuals, were drawn from Santa Clara. Then time-of-day choice models were estimated with all these datasets and transferred to (i.e., applied to the data from) Alameda. Transfer Index (TI) values were computed for all these transfers and averaged over the 10 different random samples. This same procedure was repeated by gradually increasing the sample size of the transferred model, in increments of 500 individuals from the Santa Clara County. The TI values for each sample size – 500, 1000, 1500, 2000, 2500, 3000 (averaged over 10 different random samples) – are plotted in the first part of the graph on a solid curve. As can be observed, the TI values increase with sample size suggesting a strong influence of sample size on model transferability. Models estimated with small sample sizes have poor TI values (-0.59 at sample size = 500). The rate of increase of TI is steep in the beginning but slows down after a sample size of 1500.

To control for the effect of sample size, we conducted additional experiments by keeping the sample size constant. For example, at a sample size of 3000, we compared the TI values of a model built with all data from Santa Clara (SC) to those built using data for 1500 individuals from Santa Clara and the remaining 1500 individuals from other counties (i.e., pooling data from other counties). The TI values for all the combinations (averaged over 10 different random samples for each combination) for a total sample size of 3000 are plotted using diamond shaped dots in the figure. Similar exercise was undertaken for a total sample size of 2000 (circular dots in the figure) and 1000 (triangular dots in the figure). Several important observations can be made from all these TI values. First, for any sample size, the TI values of models built using data pooled from multiple counties are generally higher than those built using all data from a single County (Santa Clara). Second, the TI values strongly depend on the counties from which data is pooled. Specifically, at all sample sizes, models with data pooled (to Santa Clara data) from either Contra Costa or Sonoma or a combination of the two resulted in higher TI values than those with data pooled from San Mateo.

The above observations are re-iterated in the second part of Figure 2 (dotted curve). This part of the graph is based on TI values of models built by simply adding, to the 3000 records from Santa Clara, more data from Contra Costa (1300 more records), and then from Sonoma (800 more records), and then from San Mateo (1200 more records). Specifically, adding data from Contra Costa and Sonoma resulted in an increased TI value, while doing so from San Mateo did not help increase the TI value. These results suggest that while pooling data from multiple regions can potentially help in building better transferable models, care must be taken in choosing the regions from which data is pooled. However, these results do not provide guidance on which regions to pool the data from (or which regions to transfer models from). To delve into these issues, we examined the descriptive statistics of the variables used in the model specification for the different counties, as described next.

First, from Table 1, comparing the demographic characteristics of different counties to those in Alameda (the region to which models were transferred in this exercise) suggests that Contra Costa comes closest to Alameda in its demographic makeup. While the land-use characteristics appear to be different between these two counties, the demographic characteristics relevant to the time-of-day choice model appear very similar while the commute travel conditions are not very different. Therefore pooling data from Contra Costa to data from Santa Clara resulted in a better TI value than using more data from Santa Clara itself. This may also be a reason why Table 2 shows a higher TI value (0.63) for a model transferred from Contra Costa to Alameda than most of the TI values observed in Figure 2. Second, again from Table 1, the demographic characteristics of Santa Clara are close to those in San Mateo. This is perhaps a reason why adding data from San Mateo to that from Santa Clara did not result in a discernible improvement in TI values (because it was not bringing any more variability to make it closer to the application context, Alameda). Third, Table 3 shows the descriptive statistics of the following pooled-data samples: (a) Santa Clara + San Mateo, (b) Santa Clara + Sonoma, (c) Santa Clara + Contra Costa (d) Santa Clara + Contra Costa + Sonoma, and (e) Santa Clara + Contra Costa + Sonoma + San Mateo. As can be observed from Table 3 and Figure 2, the combinations that are closest to Alameda in the composition of the explanatory variables used in the model are the combinations that exhibit higher TI values. These results suggest that for achieving better transferability of time-of-day choice models, one ought to build models using data that exhibits similar demographic characteristics (especially those relevant to time-of-day choice models) as in the application context.

5 CONCLUSIONS

This paper presents an empirical assessment of the spatial transferability of tour-based time-of-day choice models among different Counties in the San Francisco Bay area. Model transferability was assessed using two different approaches: (1) application-based approach and (2) estimation-based approach. The former approach tests the transferability of a model as a whole while the latter approach allows the analyst to test which specific coefficients in the model are transferable. Inter-County transferability assessments using the application-based approach did not reveal useful information because of confounding effects of sample sizes. Models built using small datasets resulted in poor transferability results, suggesting the importance of sufficient sample sizes in building models. Further, using the same (application) approach, models built using data pooled from multiple counties were found to be more transferable than models built using data from a single County. Such pooled-data models built using the estimation-based approach (or joint-context estimation) exhibited even better transferability indices than those built using the application-based approach. This is because the application-based approach allows all the parameters to be different between the model-lending context and the model-borrowing context, while the estimation-based approach allows only the statistically different parameters to be different while keeping all other parameters the same. Finally, the estimation-based approach provided encouraging results on the transferability of time-of-day choice models, with almost 95% of the model parameters found to be not statistically different (hence transferable) between any specific County and the rest of the Bay area. Among the non-transferable parameters include the home-work travel time coefficient and a few demographic preference coefficients. The transfer index values of such models ranged from 84% to 95% suggesting that such pooled models capture a large proportion of the behavior in the contexts to which they are transferred, when assessed using the joint context estimation approach. Therefore, whenever possible, joint-context estimation should be used to transfer models and to assess model transferability.

Additional transferability experiments were performed to understand the reasons why pooling data from multiple regions helps in achieving better transferability. Increased sample size was one reason why pooling data helped in achieving better transferability. Sample size increases resulted in steep increases in transferability at the beginning (i.e., at smaller sample sizes) but the rate of increase slows down after reaching a sample size of 1500. Additional transferability experiments controlling for sample size effects suggested that while pooling data from multiple regions can potentially help in building better transferable models, attention must be paid in choosing the regions from which data is pooled. Specifically, time-of-day choice models built using data that exhibits similar demographic characteristics as in the application context (especially the characteristics that are relevant to time-of-day choices) exhibit better transferability. Therefore, data should be borrowed in such a way that the resulting estimating data exhibits similar demographic characteristics to those in the application context. These findings provide empirical evidence in support of a widely held notion in the profession that models are likely to be better transferable between regions of similar characteristics.

The results in this study are based on transfers between counties within the San Francisco Bay area. It is not clear if the encouraging results on model transferability can be generalized to transfers between different metropolitan regions in the country. Additional empirical evidence is necessary to comment on transferability between more geographically dispersed contexts.

ACKNOWLEDGEMENTS
The corresponding author's efforts on this paper are dedicated to Siva Subramanyam Jonnavithula, a transportation engineer, who passed away in July 2013. Dr. Siva Srinivasan and Dr. Xuehao Chu encouraged us to undertake the data-pooling exercise. Peer-review panel discussions of the FHWA study conducted by Bowman et al. benefited this study. Anonymous reviewers provided helpful comments on an earlier manuscript. The second author's work on this paper is supported by the National Science Foundation under Grant No. 0965743.

REFERENCES

1. Sikder, S., A. R. Pinjari, S. Srinivasan, and R. Nowrouzian. Spatial Transferability of Travel Forecasting Models: A Review and Synthesis. *International Journal of Advances in Engineering and Applied Mathematics*, Vol. 5, No. 2-3, 2013, pp. 104-128. .

2. Sikder, S., and A.R. Pinjari. Spatial Transferability of Person-Level Daily Activity Generation and Time-Use Models: An Empirical Assessment. Forthcoming, *Transportation Research Record*, No. 2343, 2013, pp. 95-104.

3. Bowman, J.L., M. Bradley, and J. Castiglione. Making Advanced Travel Forecasting Models Affordable Through Model Transferability. Final Report, Prepared for Federal Highway Administration, US Department of Transportation, Washington, DC, 2013.

4. Picado, R. A Test of Transferability: The SE Florida Activity-based Travel Demand Model. Presented at the 14th National *Transportation Planning Applications Conference*, Houston, May 2013.

5. Jin, X., and A.J. Horowitz. Transferability of Time-of-Day Choice Modeling for Long-Distance Trips. In *Transportation Statistics*, Chapter 12, ed. Brian W. Sloboda J. Ross Publishing, Inc, 2009.

6. Vovsha, P., and M. Bradely. Hybrid Discrete Choice Departure-Time and Duration Model for Scheduling Travel Tours. *Transportation Research Record*, No. 1894, 2004, pp. 46-56.

7. Abou Zeid, M., T.F. Rossi, and B. Gardner. Modeling Time-of-Day Choice in Context of Tour-and Activity-Based Models. *Transportation Research Record*, No. 1981, 2006, pp. 42-49.

8. Popuri, Y., M. Ben-Akiva, and K. Proussaloglou. Time-of-Day Modeling in a Tour-Based Context: The Tel-Aviv Experience. *Transportation Research Record*, No. 2076, 2008, pp. 88-96.

9. Lemp, J.D., and K.M. Kockelman. Empirical Investigation of the Continuous Logit for Departure Time Choice using Bayesian Methods. *Transportation Research Record*, No. 2165, 2010, pp. 59-68.

10. Ben-Akiva, M., and M. Abou-Zeid. Methodological Issues in Modeling Time-of-Travel Preferences. *Transportmetrica,* 2012, pp.1-14.

11. Karasmaa, N. Evaluation of Transfer Methods for Spatial Travel Demand Models. *Transportation Research Part A,* Vol. 41, 2007, pp. 411-427.

12. Bradley, M.A. and A. J. Daly. Estimation of logit choice models using mixed stated preference and revealed preference information. In: P.R. Stopher and M. Lee-Gosselin,

Editors, Understanding Travel Behavior in an Era of Change, Pergamon, Oxford, 1997, pp. 209-232.

13. Ou, F-L, and J.C. Yu. Transferabilities of Urban Travel Demand Models. *Journal of Transportation Engineering.* Vol. 109, No 4, 1983, pp. 579-595.

14. Badoe, D.A. and E.J. Miller. Modeling Mode Choice with Data from Two Independent Cross-Sectional Surveys: An Investigation. *Transportation Planning and Technology*, Vol. 21, No. 4, 1998, pp. 235-261.

15. Hood, J. A Multilevel Analysis Framework for Developing Transferable Travel Demand Models. Presented at the *Innovations in Travel Modeling*, Tampa, Florida, May 2012.

16. Komma, A., and S. Sivaramakrishnan. Modeling Home-to-Work Commute-Timing Decisions of Workers with Flexible work Schedules. Proceedings of the *87th Annual Meeting of the Transportation Research Board*, 2008.

17. Koppelman, F.S., Kuah, G.K., and Wilmot, C.G. Transfer Model Updating with Disaggregate Data, *Transportation Research Record* No. 1037, 1985. pp. 102-107.

TABLE 1 Sample Characteristics

	Alameda	Contra Costa	Santa Clara	San Francisco	San Mateo	Sonoma
Sample Size (# employed adults)	1940	1348	3001	538	1209	881
Gender**						
Male	53.6(%)	52.4(%)	56.2(%)	56.9(%)	50.0(%)	48.1(%)
Female	46.4(%)	47.6(%)	43.8(%)	43.1(%)	50.0(%)	51.9(%)
Ethnicity						
Caucasian	73.4(%)	80.7(%)	72.8(%)	72.3(%)	77.9(%)	89.1(%)
African American	4.6(%)	3.3(%)	1.6(%)	2.8(%)	1.5(%)	0.7(%)
Asian/Pacific Islander	11.9(%)	5.9(%)	15.8(%)	13.8(%)	10.3(%)	2.0(%)
Other	10.1(%)	10.1(%)	9.8(%)	11.1(%)	10.3(%)	8.2(%)
Employment Type**						
Govt. Employees	18.2(%)	18.5(%)	10.1(%)	17.3(%)	14.0(%)	15.0(%)
Other	81.8(%)	81.5(%)	89.9(%)	82.7(%)	86.0(%)	85.0(%)
Employment Status**						
Full-Time	87.8(%)	88.1(%)	90.0(%)	93.7(%)	89.3(%)	85.9(%)
Part-Time	12.2(%)	11.9(%)	10.0(%)	6.3(%)	10.7(%)	14.1(%)
Employment Schedule**						
Flexible	63.8(%)	61.4(%)	74.4(%)	70.4(%)	68.7(%)	58.8(%)
Not Flexible	36.2(%)	38.6(%)	25.6(%)	29.6(%)	31.3(%)	41.2(%)
Household Size						
1	16.0(%)	13.6(%)	15.1(%)	28.6(%)	17.0(%)	13.5(%)
2	36.9(%)	39.9(%)	40.9(%)	43.7(%)	40.7(%)	41.1(%)
3+	47.1(%)	46.5(%)	44.0(%)	27.7(%)	42.3(%)	45.4(%)
Number of Children**						
0	63.4(%)	61.0(%)	64.3(%)	78.6(%)	66.8(%)	61.1(%)
1	16.4(%)	16.0(%)	15.7(%)	8.7(%)	14.1(%)	17.5(%)
2	15.5(%)	17.1(%)	15.1(%)	10.6(%)	15.1(%)	16.2(%)
3+	4.7(%)	5.9(%)	4.9(%)	2.1(%)	4.0(%)	5.2(%)
Household Income**						
Low(<=25K)	2.7(%)	2.7(%)	1.5(%)	2.8(%)	2.2(%)	4.8(%)
Medium(25K-75K)	40.1(%)	38.7(%)	28.1(%)	39.6(%)	32.8(%)	50.7(%)
High(>75K)	57.3(%)	58.6(%)	70.4(%)	57.6(%)	65.1(%)	44.4(%)
Number of Vehicles**						
1	81.1(%)	83.3(%)	83.1(%)	81.8(%)	84.2(%)	84.9(%)
2	13.9(%)	11.8(%)	12.5(%)	13.0(%)	12.3(%)	11.0(%)
3+	5.0(%)	4.9(%)	4.3(%)	5.2(%)	3.4(%)	4.1(%)
Area Type						
Home zone						
CBD	0.3(%)	0.0(%)	0.2(%)	14.7(%)	0.0(%)	0.0(%)
Urban	22.2(%)	5.6(%)	18.1(%)	85.3(%)	24.7(%)	4.9(%)
Suburban	74.4(%)	84.9(%)	79.7(%)	0.0(%)	71.2(%)	77.0(%)
Rural	3.1(%)	9.6(%)	2.0(%)	0.0(%)	4.1(%)	18.2(%)
Work zone						
CBD	10.6(%)	8.1(%)	5.5(%)	35.9(%)	9.4(%)	1.6(%)
Urban	39.8(%)	32.1(%)	54.4(%)	44.1(%)	52.4(%)	12.5(%)
Suburban	47.3(%)	55.3(%)	38.4(%)	18.2(%)	35.7(%)	76.5(%)
Rural	2.3(%)	4.5(%)	1.7(%)	1.9(%)	2.4(%)	9.4(%)
Commute Travel*						
Free Flow Time(minutes)**	18.0(10.5)	21.1(12.9)	16.1(9.0)	19.2(12.2)	19.4(11.3)	20.2(13.8)
Distance(miles)**	12.9(10.3)	15.4(12.4)	10.9(8.8)	11.9(11.9)	13.5(10.3)	12.5(12.8)
Tour Duration (hours)**	9.8(2.8)	9.7(2.9)	9.8(2.6)	9.8(2.9)	9.9(2.6)	9.2(2.9)

*Averages are reported, with standard deviations in the parentheses.

** All these variables were used as explanatory variables in the time-of-day choice models estimated in this study.

TABLE 2 Transfer Index

Application-based Approach (Inter-County Transfer)						
Transferred To / Transferred From	Alameda	Contra Costa	Santa Clara	San Francisco	San Mateo	Sonoma
Alameda	--	0.53	0.66	0.42	0.56	0.23
Contra Costa	0.63	--	0.58	0.28	0.58	0.24
Santa Clara	0.68	0.49	--	0.44	0.74	0.09
San Francisco	0.23	-0.13	0.42	--	0.40	-0.18
San Mateo	0.37	0.17	0.56	0.39	--	-0.03
Sonoma	-0.21	-0.26	-0.18	-0.23	-0.02	--
Application-based vs. Estimation-based Approach (Pooled Model Transfer)						
Base-c (Application-based) *	0.79	0.66	0.78	0.57	0.75	0.36
Base-c (Estimation-based) *	0.94	0.84	0.87	0.95	0.94	0.50

* Base-c models are basically pooled-data models for eight counties. In the term “Base-c”, c can be AL (Alameda), Contra Costa (CC), Santa Clara (SC), San Francisco (SF), San Mateo (SM), or Sonoma (SN). For instance, “Base-AL” indicates the model that includes all counties in the Bay Area except Alameda. Two types of “base-c” models were estimated: (1) base-c models estimated separately using only the data from eight counties, and (2) base-c model extracted from the joint context estimation (where data from the specific County “c” was also included) by setting the County-specific dummy variables to zero.

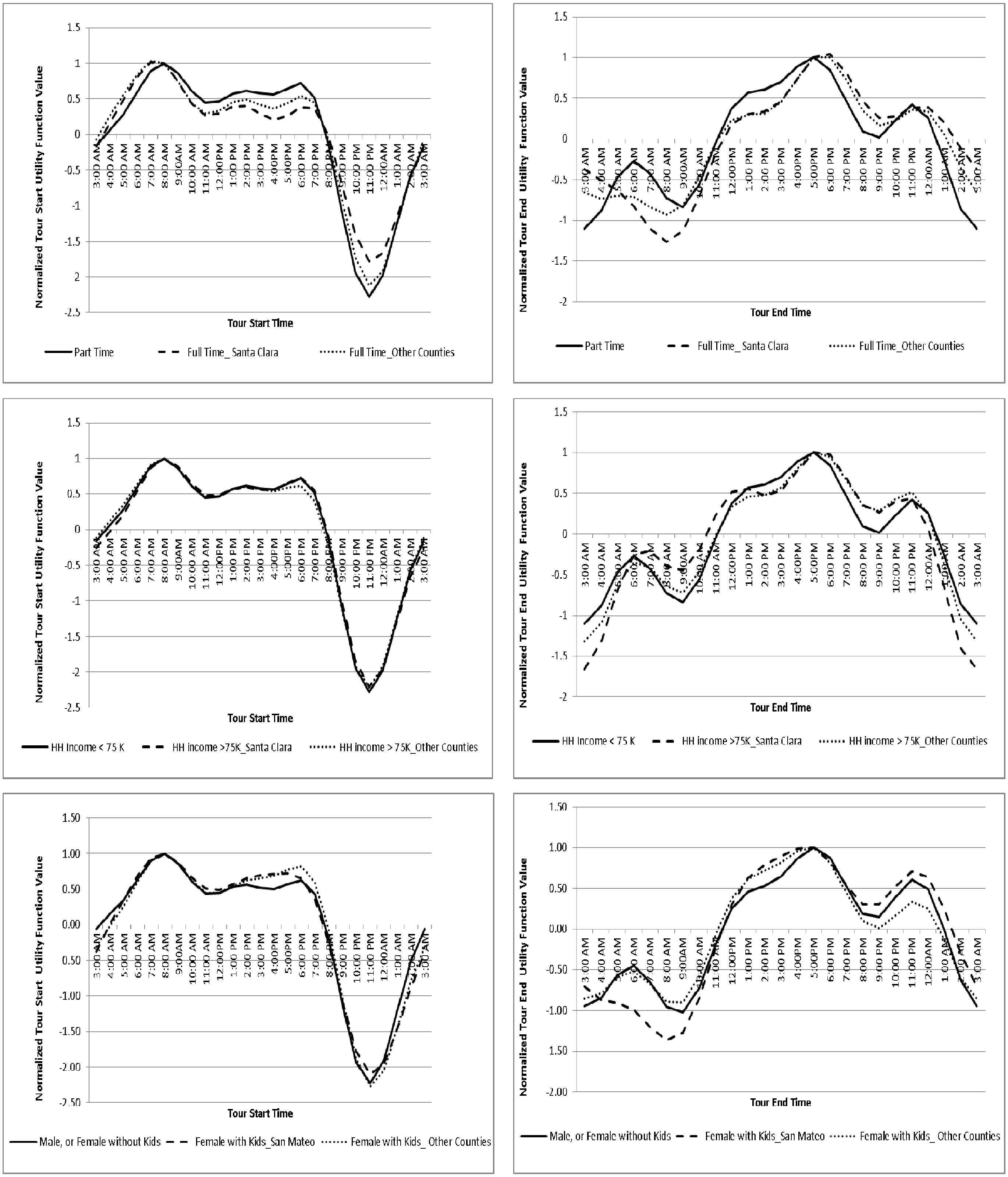

FIGURE 1 Utility profiles based on (a) Employment Status (Santa Clara and Other Counties), (b) Household Income (Santa Clara and Other Counties), and (c) Female with Kids (San Mateo and Other Counties)

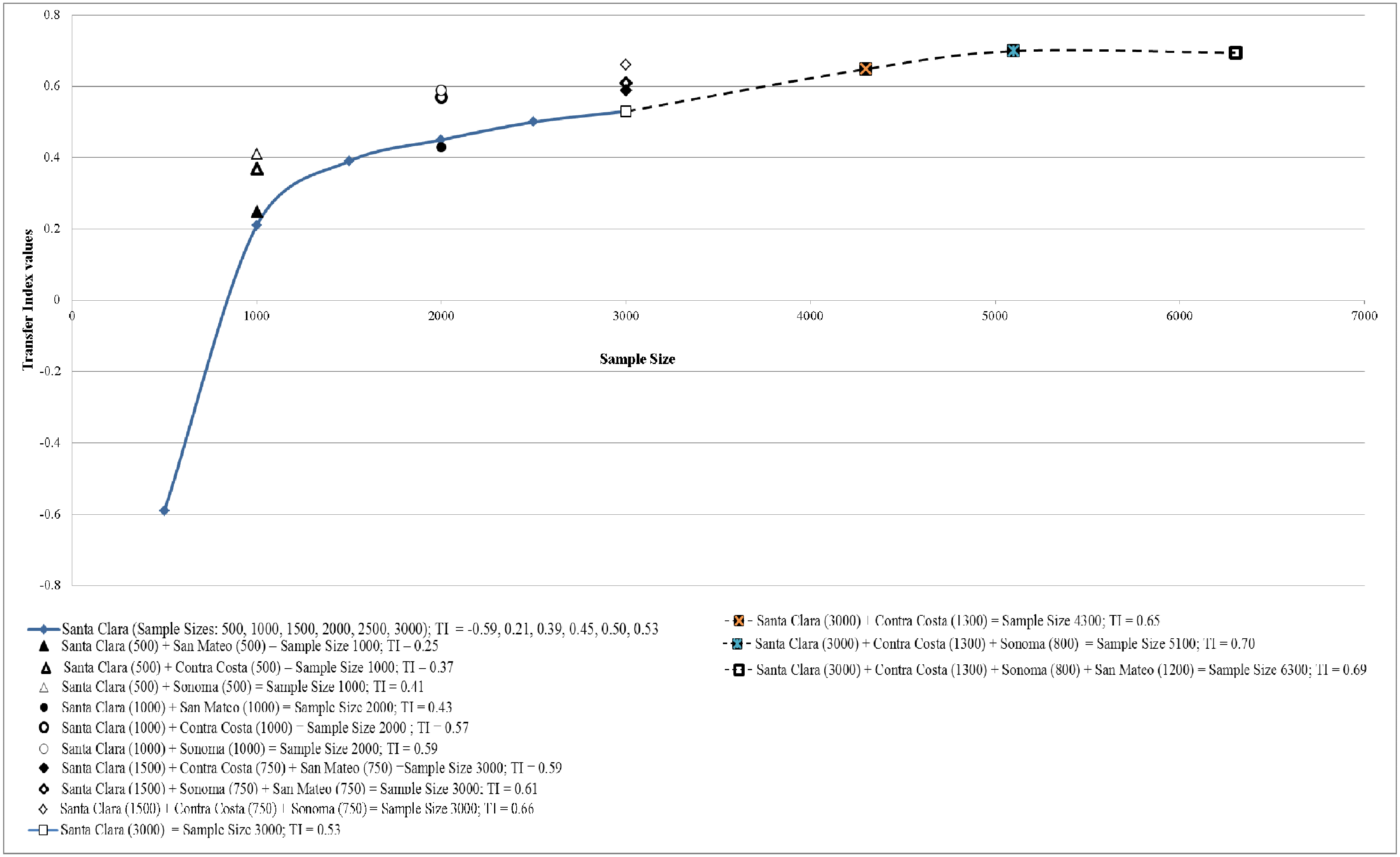

FIGURE 2 Results from Data Pooling Experiments

TABLE 3 Descriptive Statistics of the Pooled Data Samples

	Application Context	Initial Base Context	Pooled Data Samples in the First Part of Figure 2		Pooled Data Samples in the Second Part of Figure 2		
	Alameda	**Santa Clara**	**Santa Clara (1000) + San Mateo (1000)**	**Santa Clara (1000) + Sonoma (1000)**	**Santa Clara + Contra Costa**	**Santa Clara + Contra Costa + Sonoma**	**Santa Clara + Contra Costa+ Sonoma + San Mateo**
Total Sample Size	1940	3000	2000	2000	4300	5100	6300
Transfer Index (TI)	-	-	0.43	0.59	0.65	0.70	0.69
Household Income							
High income (>75K)	57.3%	70.4%	*67.6%*	**58.2%**	64.3%	**63.0%**	*63.4%*
(Difference)			+10.3% (67.6-57.3)	+ 0.9% (58.2-57.3)	+7.0% (64.3-57.3)	+5.7% (63.0-57.3)	+6.1% (63.4-57.3)
Household Structure							
Female with kids	15.8%	14.5%	*14.5%*	**16.7%**	15.3%	**15.6**	*15.4%*
(Difference)			-1.3% (14.5-15.8)	+0.9% (16.7-15.8)	-0.5% (15.3-15.8)	-0.2% (15.6-15.8)	-0.4% (15.4-15.8)
Employment Type							
Govt. Employees	18.2%	10.1%	*11.6%*	**12.3%**	**14.3%**	13.1	*13.3%*
(Difference)			-6.6% (11.6-18.2)	-5.9% (12.3-18.2)	-3.9% (14.3-18.2)	-5.1% (13.1-18.2)	-4.9% (13.3-18.2)
Employment Status							
Full-Time	87.8%	90.0%	*89.5%*	**87.9%**	89.1%	**88.9**	**88.9%**
(Difference)			+1.7% (89.5-87.8)	+0.1% (87.9 - 87.8)	+1.3% (89.1-87.8)	+1.1% (88.9-87.8)	+1.1% (88.9-87.8)
Employment Schedule							
Flexible	63.8%	74.4%	*71.8%*	**67.3%**	**68.3%**	**68.3**	68.5%
(Difference)			+8.0% (71.8-63.8)	***+3.5%*** (67.3- 63.8)	+4.5% (68.3-63.8)	+4.5% (68.3-63.8)	+4.7% (68.5-63.8)

*The proportions of the demographics in bold and italic respectively indicate the 1st and 2nd closest to those in the application context