DE-TRANSFORMATION BIAS IN NON-LINEAR TRIP GENERATION MODELS

Liming Wang, Ph.D.[1]; and Kristina M. Currans, Ph.D., E.I.T.[2]

Abstract

In recent years, there have been substantial efforts from researchers and practitioners to improve the site-level trip generation estimation methods to address some of the pitfalls of conventional approaches for applications such as traffic impact analyses. These new trip generation models often adopt sophisticated non-linear model forms to utilize new information and incorporate new factors influencing trip generation. However, if sufficient caution is not taken in their application, these new predictive models may introduce severe bias. This manuscript focuses on a typical source of biases in the applications of such models arising from de-transformation of predictions from models with a non-linearly transformed dependent variables in the prediction process (for example, predicting from a semi-log model). While such biases are well-known and corrections have been proposed in other disciplines, they have not been adopted in the site-level trip generation models to our knowledge. The de-transformation bias is described and demonstrated—focusing on log-transformed models—with numeric simulations and empirical studies of trip generation models, before discuss their implications for trip generation applications and research.

Keywords: Bias, Predictive Model, Trip Generation, Land Use Development, Development-Level Estimation, Transportation Impact Analyses, Traffic Impact Analyses

---

[1] Toulan School of Urban Studies & Planning, Portland State University

[2] School of Landscape Architecture and Planning, University of Arizona

Introduction

Trip generation is the predominant metric used to assess the site-level traffic impacts of new development. In more recent years, as agencies have come to require tools that are more sensitive toward multimodal planning objectives at a site-level, the number and complexity of applied site-planning trip generation models has increased (Cervero and Arrington, 2008; Clifton et al., 2015; Ewing et al., 2011; Schneider et al., 2015). With the growing suite of prediction and estimation methods, it is worthwhile to examine potential statistical biases that have been identified in other travel demand and forecasting literatures that may occur when new approaches are developed and applied in practice. The authors examine one such bias in this manuscript: de-transformation bias, the bias arising when de-transforming predictions from models with a non-linearly transformed response variables. To the best of the authors knowledge, treatments for de-transformation bias has never been tested for trip generation models used in transportation impact analyses.

In the following sections, the authors first review the context of site-level trip generation estimations as well as the model forms of major trip generation models that are currently available. De-transformation biases are then discussed that may be introduced in applications of these models and explore their severity through simulation studies. Finally, the bias is demonstrated through analysis of two sets of empirical data before concluding with discussion and recommendations.

Model Form for Trip Generation Models

Trip generation can be modeled at various geographies and scales. Ortúzar and Willumsen (2011) provides a general review of various approaches to trip generation modeling for regional transportation demand models. But in this section, the authors focus on the type of models and data sources commonly used for development-level trip generation estimation. These data and methods are commonly available through the Institute of Transportation Engineers' (ITEs) *Trip Generation Handbook* (Institute of Transportation Engineers, 2014), but in recent years, a growing number of data and methods have been made available through studies in academia and practice, e.g., (Clifton et al., 2015; Ewing et al., 2011; Schneider et al., 2015). These data are used for a variety of purposes including, but not limited to: transportation impact analysis, transportation system development charges, impact or utility fees, re-zoning, scaling or scoping projects, and estimating greenhouse gas emission impacts of personal driving vehicles. Although this is not a comprehensive review, the purpose of this section is to orient the reader toward the wide range of approaches available. Whether the method is vulnerable to the de-transformation bias depends on the model form and is discussed in the next section.

Site-Based Direct Demand Models

Direct Demand Models (DDMs) utilizes site-based data as an alternative to full-fledged travel demand models. The main type of DDM in site-level traffic impact analyses is ITE's *Trip Generation Handbook*, which estimates vehicle trip rates for a variety of land uses and time periods. In a broad sense, a DDM is a model that collapses trip-generation and mode choice steps to directly predict vehicle trips, for example.

ITE Trip Generation Handbook

The *Handbook* (Institute of Transportation Engineers, 2014) has long been the authoritative source of determining site-specific trips generated—which include almost exclusively vehicle counts or count rates of all trips entering or existing the study establishment. For simplification, "trips" and "trip ends" are referred to interchangeably. For most land uses, the *Handbook* supplies either average trip rates or regressions that are used to predict trips as a function of the size of the development (e.g., dwelling units, square footage, and employees). While all data is presented in terms of vehicle trip rates (trips divided by the size of the establishment), if there exists more than four data points, and if regression results in a minimum R^2 of 0.5 (not adjusted for degrees of freedom), the coefficient and the univariate regression are provided. For some land uses, a log-log model form of the univariate trip/size regression is provided in lieu of the linear model, but only if it produces an improved R^2. It is worth emphasizing here that comparing R^2 for two models—one with a transformed dependent variable and one without—is the method used in ITE's approach. Generally, comparing the R^2 of two regressions requires the dependent variables of both models have the same variance; however, this issue is not the focus of this manuscript. The significance of the coefficient itself in any regression is not readily provided.

Although the simplicity of the ITE approach has an advantage as an off-the-shelf, nationally-available method for estimating vehicle demand, some of the major criticism of the ITE approach includes the its failure to consider factors other than size of the development (Clifton et al., 2013; Shoup, 2003). Alternative or supplementary approaches were developed as a response to accommodate these criticisms; they are discussed in the following subsections.

Alternative DDM Models

Currently, there are few alternative methods that directly estimate non-automobile trips (e.g., bike trips, or person trips). Washington, D.C. Department of Transportation (DDOT) is one of the few studies that had enough sample from one land use (residential and lodging) to estimate linear, multivariate multimodal DDMs (District Department of Transportation, 2015). Cervero and Arrington (2008) also presents an alternative DDM that utilize site-based linear models alternative to predict transit ridership. Although it depends upon the form of the model, DDMs require site-level data collected by intercept survey, which increases the cost and complexity of data collection, as well as the ability to control for a wide range of influences identified as influencing site-level transportation demand (Clifton et al., 2017, 2013). Studies that have collected multimodal person trip data have used that multi-land use, site-level data to create site-based adjustment (Clifton et al., 2015; Schneider et al., 2015; Bochner et al., 2011) to ITE's vehicle trip generation in the hopes that at some point there will exist enough multimodal data across a wide range of urban contexts and land uses to create one or many multimodal DDM (Clifton et al., 2013). Of these methods, Schneider et al. et al. (2015) uses a semi-log model to estimate vehicle trips in smart growth areas as an adjustment to ITE’s suburban rates; as such, it is vulnerable to a de-transformation bias when not corrected, and these data and regressions are revisited later in the Empirical Case Studies section.

Individual-Based Trip Rate Model

Cross classification analysis and regression models of trip making, especially trip productions of home-based trips, are usually developed using individual household or person as the unit of analysis, as the information of trip making for households and persons is easily available from common household travel surveys. Such models are routinely recommended for applications (Martin and McGuckin, 1998), although not without flaws (Guevara and Thomas, 2007). For example, a potential localized approach to estimate trips generated at residential locations (i.e., home-based trips) are commonly modeled at a household unit of analysis, as household-level demographic and travel behavior information are abundantly available from household travel surveys (Reid, 1982). The method regresses number of trips made by households upon household characteristics (e.g. household size, income, vehicle ownership) in a simple linear regression model. The main limitation in applying the individual-based trip rate models for development-level review is the limited residential-application, applied in the region in which it was developed (Planning Department: City and County of San Francisco, 2002).

Hybrid Approach: Individual-Based Models for Site-Based Applications

To address the limitation of the ITE approach, individual-based models are adopted to estimate site-level trip generation. These methods are often developed as a stop-gap approach to fill the need for more robust methods of estimation that are sensitive to a wide range of policy objectives. These approaches use household travel surveys from one region (Daisa et al., 2013) or multiple regions (Currans and Clifton, 2015; Ewing et al., 2011), organizing the data into a trip-end data base where every trip is counted as both an origin and location. Contextual information about each trip-end environment is collected, such as the activity or land use in which the end is occurring, the built environment or some form of urban context area-type, or multimodal accessibility of the site. Travel outcomes can then be regressed upon the contextual variables (Currans and Clifton, 2015; Ewing et al., 2011) as well as characteristics of the trip-maker (Ewing et al., 2011).

The resulted models estimate mode shares, trip length, internal capture, and vehicle occupancy—using various forms of multivariate regression, including: hierarchical linear and nonlinear (Ewing et al., 2011), binary logistic (Currans and Clifton, 2015), and linear (Currans and Clifton, 2015) regression. Because household travel surveys capture household-level travel, they do not provide enough trip ends at any one non-residential land use or development to be able to estimate trip rates. Instead, these models are adjustment models to estimate relative differences in shares and distances. These approaches require a direct demand model, such as ITE's *Handbook*, to acquire some estimate of vehicle or person trip count for a single or multi-land use development. While the adjustment technique is documented (Institute of

Transportation Engineers, 2014), used (Bochner et al., 2011; Clifton et al., 2012; Currans and Clifton, 2015; Daisa et al., 2013; Ewing et al., 2011), and critiqued (Clifton et al., 2013; Currans and Clifton, 2015) in many methods, de-transformation bias, as it is described in the following section, may still be problematic for approaches that include a transformation of the error term—described in the following section—within the model form.

De-Transformation Bias and Correction

As new trip generation models more commonly take non-linear model form, applications of such models in prediction may suffer from a type of bias known as de-transformation bias. The de-transformation bias arises when predicting from models with non-linearly transformed dependent variables (e.g., a semi-log model)—for example, when the transformed responses ($\widehat{ln(Y)}$) are de-transformed ($\exp(\widehat{ln(Y)})$) to get the original response ($\hat{Y}$). The bias has long been discovered and corrections suggested in papers across a range of disciplines, such as, anthropology (Becker, 1965), economics (Wooldridge, 2012), ecology (Sprugel, 1983), forestry (Baskerville, 1972; Snowdon, 1991), and statistics (Finney, 1941; Miller, 1984). To the authors' knowledge, no prior research has investigated the de-transformation bias and its correction in the context of trip generation models used in development-level transportation impact analyses, even though non-linear models have been routinely applied in applications.

Complete derivation of the bias can be found in the literature (Finney, 1941; Miller, 1984). Here, the bias for log-transformed models is shown, briefly. In a linear model, the relationship between independent (X) and dependent (Y) variables can be expressed mathematically as:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where the error term $\varepsilon \sim N(0, \sigma^2)$, and

$$\boldsymbol{E(Y) = E(X\beta + \varepsilon) = E(X\beta) + E(\varepsilon) = E(X\beta).} \quad (2)$$

Thus, predicting Y from the linear model $\boldsymbol{X\beta}$ is not biased. However, in a semi-log model, the relationship is expressed as:

$$\mathbf{ln}(\boldsymbol{Y}) = \boldsymbol{X\beta + \varepsilon}, \quad (3)$$

where the error term $\varepsilon \sim N(0, \sigma^2)$,and

$$\boldsymbol{E(Y) = E(exp(X\beta + \varepsilon)) = E(exp(X\beta)exp(\varepsilon)) \neq E\big(exp(X\beta)\big)}, \quad (4)$$

as exp(ε) is log-normally distributed with mean = $\exp(\sigma^2/2)$ and thus $E(\exp(\varepsilon)) \neq 1$.

In other words, the results would be negatively biased (underestimated) if a semi-log model is estimated and then used to predict and de-transform the dependent variable without correcting for bias introduced in de-transformation.

Corrections

Three methods of bias correction are proposed in the literature. The first correction considered, which was proposed by Baskerville (1972), is $\exp(\hat{\sigma}^2/2)$—where $\hat{\sigma}$ is the estimator for σ in Equation (3), i.e., the standard deviation of the model residuals. While this correction term, here called the Baskerville correction, is consistent, it is itself biased (Miller, 1984; Snowdon, 1991). The second correction term considered is an unbiased correction term $\exp(g(\hat{\sigma}^2/2))$, where g is an infinite series approximation, originally proposed by Finney (1941). And third, a ratio correction term proposed by Snowdon (1991)—dividing the true values of the dependent variables by the estimated values—based on ratio-estimation techniques in sampling theory.

In the following sections of this paper, the severity of the de-transformation bias is assessed when not corrected compared with the performance of the three bias correction methods: Baskerville, Finney, and Snowdon. First, the magnitude of bias and performance of correction approaches is explored through Monte Carlo simulation, and then corrections with actual trip generation data in two case studies are examined.

Evaluation Criteria

Three criteria in evaluating the performance of the bias correction are used: bias, precision, and accuracy (Walther and Moore, 2005). *Bias* (or mean error) is the mean difference between the predicted values and the observed values:

$$\boldsymbol{Bias} = \frac{\sum_{i=1}^{n}(Y-\hat{Y})}{n}. \quad (5)$$

Mean error may be normalized by mean, standard deviation, or the range of observed values Y. Both mean error and percent mean error normalized by mean are used as measures of bias in this paper.

Precision is theoretically the deviations of predictions from their mean, estimated by the standard deviation of predictions:

$$\boldsymbol{PRECISION} = \boldsymbol{sd}(\hat{Y}). \quad (6)$$

Accuracy is a measure of discrepancy between the predicted values and the observed values, for which the commonly known root mean square error (RMSE) metric is used:

$$\boldsymbol{ACCURACY} = \sqrt{\frac{\sum_{i=1}^{n}(Y-\hat{Y})^2}{n}}. \quad (7)$$

Monte Carlo Simulation

Monte Carlo simulation provides comprehensive information of the magnitude of the bias and performance of the correction methods under ideal conditions. For the Monte Carlo simulation a simple semi-log model is used:

$$\boldsymbol{ln(Y) = b_0 + b_1 X + \varepsilon}, \tag{8}$$

where the coefficients have true values b_0=0.5 and b_1= 1.0, X is randomly drawn from a uniform distribution (0, 1), and the error term $\varepsilon \sim N(0, \sigma^2)$. Thus, the expectation $E\big(ln(Y)\big) = 1$.

In each iteration of the simulation, the following steps are taken:

1. Pick a σ from the range of [0.01, 2], drawing a sample of 1,000 observations of X and ε from their corresponding distributions;
2. Calculate $\ln(Y)$ with Equation (8) and the true Y by exponentiating $\ln(Y)$;
3. Combine X, Y, and ln(Y) to create a data sample with 1,000 observations;
4. For each data sample, repeat the following steps 1000 times:
   a. Randomly do a 50-50 split of the data sample into estimation set and validation set (i.e., 500 observations in each set);
   b. Regress $\ln(Y)$ with X to estimate coefficients b_0 and b_1 using the estimation set, and capture the standard deviation of the model residuals as an estimate of σ;

$$\boldsymbol{ln(Y) = \widehat{b_0} + \widehat{b_1} X} \tag{9}$$

   c. Using the validation set, Equation (9) is applied to predict $\widehat{ln(Y)}$ and transform it to get $\hat{Y}$ with and without correction for bias;
   d. Compare $\hat{Y}$with the true Y to assess the bias and the performance of bias correction methods.

Steps 1-4 are repeated 2000 times for sufficient coverage of the range of [0.01, 2] for ε.

Figure 1 shows how the bias (left subfigure) and accuracy (right subfigure) vary with the standard deviation of the residuals. As expected, it indicates that there is severe bias when the predictions are not corrected (considering the largest possible value is -100% for negative bias since both and $\hat{Y}$and Y are positive numbers). Additionally, the bias increases very quickly with the residual standard deviation when not corrected. It also demonstrates that all three correction methods are successful in reducing the bias, keeps it small when the standard deviation is less than 1.5, and performs well across the whole range of 0.01 to 2. Among the three correction methods, the Snowdon method performs best and it also has better consistency than the Baskerville and Finney method when residual standard deviation is large. The Finney method performs slightly better in minimizing bias than the Baskerville method when the standard deviation of residuals is large ($\hat{\sigma}$ > 1.5), but the difference is small. All three methods consistently over-predict when $\hat{\sigma}$ > 1.0, although the magnitude of positive biases after correction are much smaller than the negative bias without correction. The bias correction also improves the accuracy of prediction by a small amount when the standard deviation of residuals is large. (Note that the difference in accuracy among these three methods is so small that the curves representing them overlap with each other).

The simulation study is informative, but it cannot tell us the severity of bias and how the correction methods work in real world. Two case studies are conducted with actual trip generation data and empirical models to demonstrate the severity of de-transformation bias in real trip generation applications and the performance of the three methods for bias correction.

Empirical Case Studies

Two trip generation methodologies are selected, both susceptible to de-transformation bias due to the non-linear model forms. The first case study uses data and models from the ITE's *Trip Generation Manual* (Institute of Transportation Engineers, 2012). Trip generation rates estimated using a log-log form are selected for three land use types based on data availability and sample size: High-Cube Warehouse/Distribution Center (ITE Land Use Code, LUC, 152), Low-Rise Apartment (LUC 221), and Mobile Home Park (LUC 240). The second case study uses data made available online from the California Smart Growth Trip Generation Rates Study (Schneider et al., 2015) to estimate a semi-log Post-Meridiem (PM) peak hour model to adjust ITE's estimates for a number of land uses.

In each of the case studies described in the following subsection, the data are randomly split into two parts: an estimation sample and a validation sample. The data in the estimation sample are used to estimate an appropriate model, obtain the model coefficients and calculate the standard deviation of the model residuals for use in prediction and bias correction later. For the Finney approximation, an approximation of $g(\hat{\sigma}^2/2)$ to order n^{-2} is applied (Finney, 1941; Snowdon, 1991):

$$\hat{\boldsymbol{\sigma}}^2/2\left[1-\frac{\hat{\boldsymbol{\sigma}}^2(\hat{\boldsymbol{\sigma}}^2+2)}{4n}+\frac{\hat{\boldsymbol{\sigma}}^4(3\hat{\boldsymbol{\sigma}}^4+44\hat{\boldsymbol{\sigma}}^2+84)}{96n^2}\right]. \tag{10}$$

The ratio between the observed values of the dependent variable is calculated, and its fitted values from the model are used in the ratio correction method.

The estimated model is applied to the validation sample: first by predicting $\widehat{ln(\boldsymbol{Y})}$ and then transforming it to get prediction for $\widehat{\boldsymbol{Y}}$ without correction. Each of the three methods for bias correction is applied to $\widehat{ln(\boldsymbol{Y})}$ to get the corresponding corrected predictions. The predicted values of the dependent variables, and the three corrected predictions, are paired with the observed values to compute bias, precision, and accuracy for each prediction.

ITE Trip Generation Log-Log Models

ITE *Manual* suggests log-log models for some land use codes (LUCs) and time periods with a general mathematical form expressed as:

$$\boldsymbol{ln(Y) = b_0 + b_1 ln(X) + \varepsilon}. \tag{11}$$

The summary statistics, estimated coefficients, residual standard deviation, and R^2 for each of the three land use types can be found in Table 1. Since half of the observations are reserved for validation, the statistics and estimation results differ from those reported in the ITE manual. Table 1 also includes bias, precision and accuracy for each of the three methods for bias correction along with those for predictions with no correction.

It can be seen from Table 1 that, with no bias correction, there is consistent negative bias across the three land use types and each of the three correction methods reduce the bias substantially. Except for LUC 152, the bias correction methods also improve prediction precision and accuracy. Among the three bias correction methods, the ratio correction approach has the lowest bias for LUC 152 and 221 but results in a higher bias for LUC 240. The performances of the Baskerville correction and the Finney approximation are almost identical, echoing the findings in this simulation study and those of Snowdon (1991).

For each of the three land uses, Figure 2 through Figure 4 show (left subfigure) a scatter plot of the data in both the estimation and validation sample with the regression curve with and without the Baskerville bias correction, and (right subfigure) as well as the predicted versus observed trips in the validation sample with and without bias correction.

Shapiro-Wilk normality tests are conducted for the residuals for each of the fitted models in Table 1. The p-value of the normality tests are 0.845, 0.695 and 0.033 for LUC 152, 221, and 240, respectively. The residuals for LUC 240 are not likely log-normally distributed, while those for the other two land use types are likely log-normally distributed. This case study demonstrates that even when the residuals deviate from the assumed log-normal distribution, the bias correction methods still performs well.

California Smart Growth Trip Generation (SGTG) Study Semi-Log Models

In the second case study, the semi-log PM peak model from the California SGTG study (Schneider et al., 2015) is evaluated using the data published online (Schneider et al., 2012). While this study collected and compiled a large number of data points, there were not enough of any one land use to estimate a DDM of actual trips during analysis (Schneider et al., 2012). However, this model reflects the overall trend in the state-of-the-art methods in moving toward more complicated multivariate regression—compared with the univariate case study evaluated previously—to control for a range of contextual characteristics. As a result, the predictive model was estimated using a semi-log with a natural log transformation of the dependent variable: the ratio of the observed "actual trips" at smart growth sites divided by the estimated trips predicted by ITE's suburban data and estimates.

The procedure followed for the first case study is extended for the second case study. The model structure and dependent and independent variables from the California Smart Growth Trip Generation Rates Study can be expressed as:

$$\boldsymbol{ln}\left(\frac{\text{Actual Trips}}{\text{ITE Trips}}\right) = \boldsymbol{b_0} + \boldsymbol{b_1}\mathrm{X} + \boldsymbol{\varepsilon}, \tag{12}$$

where X is a vector of variables including one continuous variable (Smart Growth Factor) and a series of dummy variables.

Table 2 shows the estimation results with the full sample and the estimation sample. Note since the estimation sample include only half of the observations (the other half reserved for validation), the estimation results differ. Since the purpose is to investigate the de-transformation bias, not fitting the best model, the results from the estimation sample serve this purpose. Also, note the residual standard deviations of the two estimations are close.

Figure 5 shows (left subfigure) a scatter plot of the data in the estimation and validation sample, the regression curve with and without bias correction (Baskerville method), as well as (right subfigure) the predicted versus observed ratio in the validation sample with and without bias correction.

Table 3 shows the bias, precision, and accuracy of predictions before and after correction for bias by the three methods. A similar pattern as the ITE trip generation case study is identified: with no bias correction, there is sizable negative bias in the predictions and each of the three bias correction methods reduces the bias substantially. All three bias correction methods also improve prediction precision and accuracy. Among the three methods, the Baskerville method performs as well as the Finney method, and slightly better than the Snowdon method. Shapiro-Wilk normality test of the model residuals has a p-value of 0.23, indicating the hypothesis that the residuals are log-normally distributed cannot be rejected.

Discussion

The simulation study demonstrates that the de-transformation bias is substantial and grows quickly as the residual standard deviation increases and the empirical case studies show that actual trip generation applications suffer from persistent negative bias when not corrected.

The simulations and case studies also provide evidence that the correction methods work well in reducing or eliminating the negative de-transformation bias. Among the three correction methods, the Baskerville's method and the Finney's approximation result in almost identical bias, precision and accuracy, corroborating earlier research (Snowdon, 1991). Considering this, there would be little reason of using the more complicated Finney approximation. The ratio correction method produces the least bias for some of the cases, but slightly higher one for some other. One advantage of the ratio correction method is that it can also correct other sources of bias than those from de-transformation. The case studies further demonstrate that the correction method can still work even when the assumption of log-normally distributed error term may be violated.

As a side effect of correcting for bias, the correction methods may help improve the accuracy of model predictions, but the improvement is usually small, as the bulk of the accuracy comes from precision, unless when the bias is substantial compared with precision.

A potential limitation of this study is that the empirical case studies rely on data with relatively small sample size, especially as the data are split into estimation and validation sample for rigorousness. Because of the small sample size, the results may vary with the composition of estimation and validation sample, even though based on our tests of different compositions of estimation and validation sample, the results hold except for rare cases. The correction methods also work well when using the whole sample for both estimation and validation. However, site-level trip generation data are typically made up of relatively small sample sizes; all it takes is four data points and an R^2 of 0.5 (not adjusted for degrees of freedom) to include a univariate regression within ITE's *Handbook* (Institute of Transportation Engineers, 2014). The magnitude of the bias from a log-transformed model is related to the standard deviation of the model residuals—which means that as the variation in the residuals increases, or the sample size decreases, this bias becomes larger.

Conclusion

The log-transformed model is one of the most commonly used model form in development-level trip generation modeling. Based on existing literature and through numeric simulation and empirical case studies, this research demonstrates that log-transformed models introduce negative de-transformation bias when not corrected. This analysis demonstrates that existing bias correction methods work well in simulation, as well as with real data.

Based on this simulation and empirical case studies, the Snowdon and Baskerville methods perform as well as the more complicated Finney method. Both methods are easy to apply if the necessary information for correction is available. The authors recommend these two methods for applications.

This research has important implications for practitioners and researchers of trip generation. To the authors' knowledge, these types of trip generation applications have ignored the de-transformation bias, and researchers have not provided enough guidance and sufficient information for correcting such bias. Research papers and reports using log-transformed models should include the bias correction procedure in their application recommendation or supplementary toolkits, and incorporate the information, including the residuals standard deviation and/or correction ratio, necessary for users to apply correction for de-transformation bias, neither of which is commonly included in modeling results. Even though an imperfect approximation exists if the original data or residual standard deviation is not available (Strimbu, 2012), it is at best an approximation with extra complexity.

Like the log-transformed models, other models with non-linearly transformed dependent variable also suffer from de-transformation bias. Miller (1984) derives correction terms for some of the common non-linear transformations, including square root, fractional powers, and inverse. These correction terms may become more useful as the additional methods and approaches are developed and tested to further improve site-level transportation impacts estimation for trip rate as well as alternative travel outcomes (e.g., multimodal travel, vehicle occupancy or ownership, trip length, or vehicle miles traveled). Even though the focus of this paper is trip generation models, these findings and suggestions are not limited to trip generation models alone, as any model with non-linearly transformed response used in predictions suffer from de-transformation bias.

While the bias discussed in this manuscript suggests semi-log or log-log models of vehicle trip counts are negatively biased—meaning they under-predict vehicle trips—this is merely the statistical relationship between the observed and predicted values. This does not take into account the suburban bias of the data, nor does it account for the overall bias towards over-predicting "new" trips (instead of pass-by or diverted traffic)—all of which have been discussed at length in the literature, e.g., (Bochner et al., 2011; Clifton et al., 2013; Ewing et al., 2011; Shoup, 2003). The negative bias in ITE's models, for many land uses, is likely to be masked behind over-sampling of development in single-use, vehicle-oriented, suburban locations. Instead, this paper hints at a much larger issue in the development of these approaches—the soundness of the statistical approach is often overlooked, instead relying on the "precedent" of methods developed more than forty years ago.

Performance aside, the practice for site-level transportation demand modeling is currently in the middle of a major evolution of both data and methods—becoming at once more multimodal and flexible, as well as more technologically complex. Few are focusing on the methods of estimation and prediction themselves. The *Handbook* will likely remain the predominate source of site-level predictions for much of the United States for some time—not including those few large metropolitan areas who have the resources to develop and refine more localized methods for evaluating new development. A more thorough review of the statistical techniques applied in this field is necessary to ensure the effects of such biases are known and understood.

Acknowledgements

Liming Wang would like to credit Michael Jensen at Puget Sound Region Council for first bringing the de-transformation bias into his attention. The research has been partially supported by National Institute of Transportation and Communities grants 881 and 1000 and the D. D. Eisenhower Graduate Transportation Fellowship Program grant DTFH6416G00057. The authors would also like to acknowledge ITE for making their data accessible to the researchers, as well as the Caltrans' Smart Growth Trip Generation study for publishing their data along with their method documentation.

References

Baskerville, G.L., 1972. Use of Logarithmic Regression in the Estimation of Plant Biomass. Can. J. For. Res. 2, 49–53. doi:10.1139/x72-009

Becker, G.S., 1965. A Theory of the Allocation of Time. Econ. J. 75, 493–517. doi:10.2307/2228949

Bochner, B.S., Hooper, K., Sperry, B., Dunphy, R., 2011. Enhancing Internal Capture Estimation for Mixed Use Developments (Research Report No. NCHRP Report 684). National Cooperative Highway Research Program: Transportation Research Board of the National Academies, Washington, D.C.

Cervero, R., Arrington, G.B., 2008. Effects of TOD on Housing, Parking, and Travel (Research Report No. TCRP Report 128). Transit Cooperative Research Program: Transportation Research Board of the National Academies, Washington, D.C.

Clifton, K.J., Currans, K.M., Cutter, A., Schneider, R., 2012. Household Travel Surveys in a Context-Based Approach for Adjusting ITE Trip Generation Rates in Urban Contexts. Transp. Res. Rec. J. Transp. Res. Board 2307, 108–119.

Clifton, K.J., Currans, K.M., Larco, N., Wettach-Glosser, J., 2017. Improving Trip Generation Methods for Livable Communities (Research Report). Transportation Research and Education Center, Portland, Oregon.

Clifton, K.J., Currans, K.M., Muhs, C.D., 2015. Adjusting ITE's Trip Generation Handbook for Urban Context. J. Transp. Land Use 8, 5–29.

Clifton, K.J., Currans, K.M., Muhs, C.D., 2013. Evolving the Institute of Transportation Engineers' Trip Generation Handbook: A Proposal for Collecting Multi-modal, Multicontext, Establishment-level Data. Transp. Res. Rec. J. Transp. Res. Board 2344, 107–117.

Currans, K.M., Clifton, K.J., 2015. Using Household Travel Surveys as a Method to Adjust ITE Trip Generation Rates. J. Transp. Land Use 8, 85–119.

Daisa, J.M., Schmitt, M., Reinhofer, P., Hooper, K., Bochner, B.S., Schwartz, L., 2013. Trip Generation Rates for Transportation Impact Analyses of Infill Developments (Research Report No. NCHRP Report 758). National Cooperative Highway Research Program: Transportation Research Board of the National Academies, Washington, D.C.

District Department of Transportation, 2015. Trip Generation and Data Analysis Study (Research Report). District Department of Transportation, Washington, D.C.

Ewing, R., Greenwald, M., Zhang, M., Walters, J., Feldman, M., Cervero, R., Frank, L., Thomas, J., 2011. Traffic Generated by Mixed-Use Developments—Six-Region Study Using Consistent Built Environmental Measures. J. Urban Plan. Dev. 137, 248–261.

Finney, D.J., 1941. On the Distribution of a Variate Whose Logarithm is Normally Distributed. Suppl. J. R. Stat. Soc. 7, 155–161. doi:10.2307/2983663

Guevara, C.A., Thomas, A., 2007. Multiple classification analysis in trip production models. Transp. Policy 14, 514–522. doi:10.1016/j.tranpol.2007.08.001

Institute of Transportation Engineers, 2014. Institute of Transportation Engineers' Trip Generation Handbook, 3rd ed. Institute of Transportation Engineers, Washington, D.C.

Institute of Transportation Engineers, 2012. Trip Generation Manual, 9th Edition. ed. Institute of Transportation Engineers, Washington, DC.

Martin, W.A., McGuckin, N.A., 1998. Travel Estimation Techniques for Urban Planning (Research Report No. NCHRP 365). National Cooperative Highway Research Program: Transportation Research Board of the National Academies, Washington, D.C.

Miller, D.M., 1984. Reducing Transformation Bias in Curve Fitting. Am. Stat. 38, 124–126.

Ortúzar, J. de D., Willumsen, L.G., 2011. Modelling Transport. John Wiley & Sons.

Planning Department: City and County of San Francisco, 2002. Transportation Impact Analysis Guidelines for Environmental Review (Guidelines). Planning Department: City and County of San Francisco, San Francisco, CA.

Reid, F.A., 1982. Critique of ITE Trip Generation Rates and an Alternative Basis for Estimating New Area Traffic. Transp. Res. Rec. J. Transp. Res. Board 874.

Schneider, R., Shafizadeh, K., Handy, S.L., 2012. California Smart-Growth Trip Generation Rates Study, Appendix F. Institute of Transportation Studies: University of California, Davis, Davis, CA.

Schneider, R.J., Shafizadeh, K., Handy, S.L., 2015. Method to Adjust Institute of Transportation Engineers Vehicle Trip-Generation Estimates in Smart-Growth Areas. J. Transp. Land Use 8, 69–83.

Shoup, D., 2003. Truth in Transportation Planning. J. Transp. Stat. 6, 107–123.

Snowdon, P., 1991. A ratio estimator for bias correction in logarithmic regressions. Can. J. For. Res. 21, 720–724. doi:10.1139/x91-101

Sprugel, D.G., 1983. Correcting for Bias in Log-Transformed Allometric Equations. Ecology 64, 209–210. doi:10.2307/1937343

Strimbu, B., 2012. Correction for bias of models with lognormal distributed variables in absence of original data. Ann. For. Res. 55, 265–279. doi:10.15287/afr.2012.66

Walther, B.A., Moore, J.L., 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography 28, 815–829. doi:10.1111/j.2005.0906-7590.04112.x

Wooldridge, J.M., 2012. Introductory Econometrics: A Modern Approach, 5th Edition. ed. South-Western Cengage Learning, Mason, Ohio.

Tables

Table 1. Bias, precision, and accuracy of predicted trips before and after correction for bias by the three methods, data source: (Institute of Transportation Engineers, 2012)

	No Correction	Baskerville	Finney	Snowdon
High-Cube Warehouse/Distribution Center[a]				
Bias	-4.83	0.44	0.43	***-0.26***
Precision	***67.40***	68.80	68.80	68.61
Accuracy	***67.57***	68.80	68.80	68.61
Low-Rise Apartment[b]				
Bias	-26.35	14.26	14.15	***5.98***
Precision	262.27	***256.30***	256.41	257.51
Accuracy	263.59	***256.79***	256.80	257.58
Mobile Home Park[c]				
Bias	-163.30	***-51.03***	-52.23	-78.35
Precision	268.89	268.00	267.89	***266.37***
Accuracy	314.59	***272.81***	272.94	277.65

Note*:*
For each land use type, the method produces the best results for each criterion (bias, precision, or accuracy) is ***emphasized***.
Y: Vehicle trip ends per peak hour
X: Size of development
Background information regarding ITE's data:
[a] ITE LUC 152; X = 1,000 square feet gross floor area; Equation: ln(Y) = -1.49 + 0.95 ln(X); Sample size = 19; $\bar{Y}$ (sd(Y)) = 163.39 (84.56); sd(ε) = 0.26; $R^2 = 0.82$.
[b] ITE LUC 221; X = Dwelling units; Equation: ln(Y) = 2.61 + 0.85 ln(X); Sample size = 11; $\bar{Y}$ (sd(Y)) = 1462.04 (781.69); sd(ε) = 0.24; $R^2 = 0.81$.
[c] ITE LUC 240; X = Acres; Equation: ln(Y) = 4.02+ 0.82 ln(X); Sample size = 14; $\bar{Y}$ (sd(Y)) = 980.46 (655.47); sd(ε) = 0.51; $R^2 = 0.54$.

Table 2. Estimation Results of the log-transformed trip ratio model, data source: (Schneider et al., 2015)

	Full Sample	Estimation Sample
Smart Growth Factor	-0.16 (0.10)	-0.13 (0.16)
Office[a]	-0.53 (0.21) **	-0.28 (0.32)
Coffee & Donut Shop[a]	-0.75 (0.32) **	-0.44 (0.45)
Mixed-Use[a]	-0.08 (0.21)	-0.05 (0.31)
University[a]	-0.31 (0.28)	-0.16 (0.44)
Constant	-0.49 (0.11) ***	-0.62 (0.19) ***
Observations	50	25
R^2	0.36	0.18
Adjusted R^2	0.29	-0.04
Residual Std. Error	0.51 (df = 44)	0.56 (df = 19)
F Statistic	4.98*** (df = 5; 44)	0.81 (df = 5; 19)

Note:
Standard error in parentheses.
*p-value < 0.1; **p-value < 0.05; ***p < 0.01
[a] Binary variable

Table 3. Bias, precision, and accuracy of predicted trip generation rates ratio before and after correction for bias by the three methods, data source: (Schneider et al., 2015)

	No Correction	Baskerville	Finney	Snowdon
Bias	-0.12	-0.04	-0.04	-0.06
Precision	0.35	0.35	0.35	0.35
Accuracy	0.38	0.35	0.35	0.35

Note:
Sample size: 25
Mean (standard deviation) of observed actual trips/ITE trips ratio = 0.62 (0.41)

Figure 1

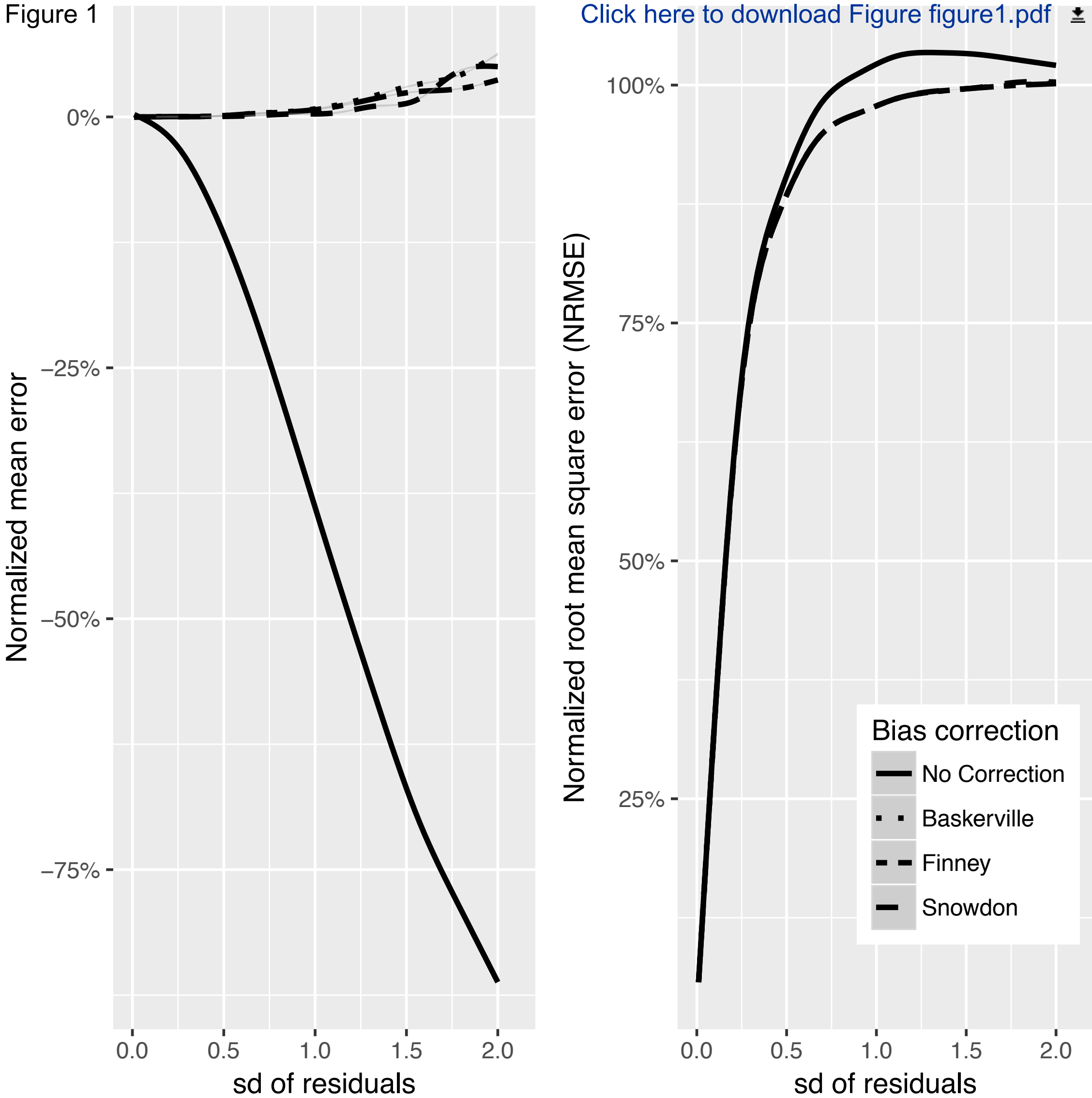

Normalized mean error
0%
−25%
−50%
−75%
0.0
0.5
1.0
1.5
2.0
sd of residuals
Normalized root mean square error (NRMSE)
100%
75%
50%
25%
0.0
0.5
1.0
1.5
2.0
sd of residuals
Bias correction
No Correction
Baskerville
Finney
Snowdon

Figure 2

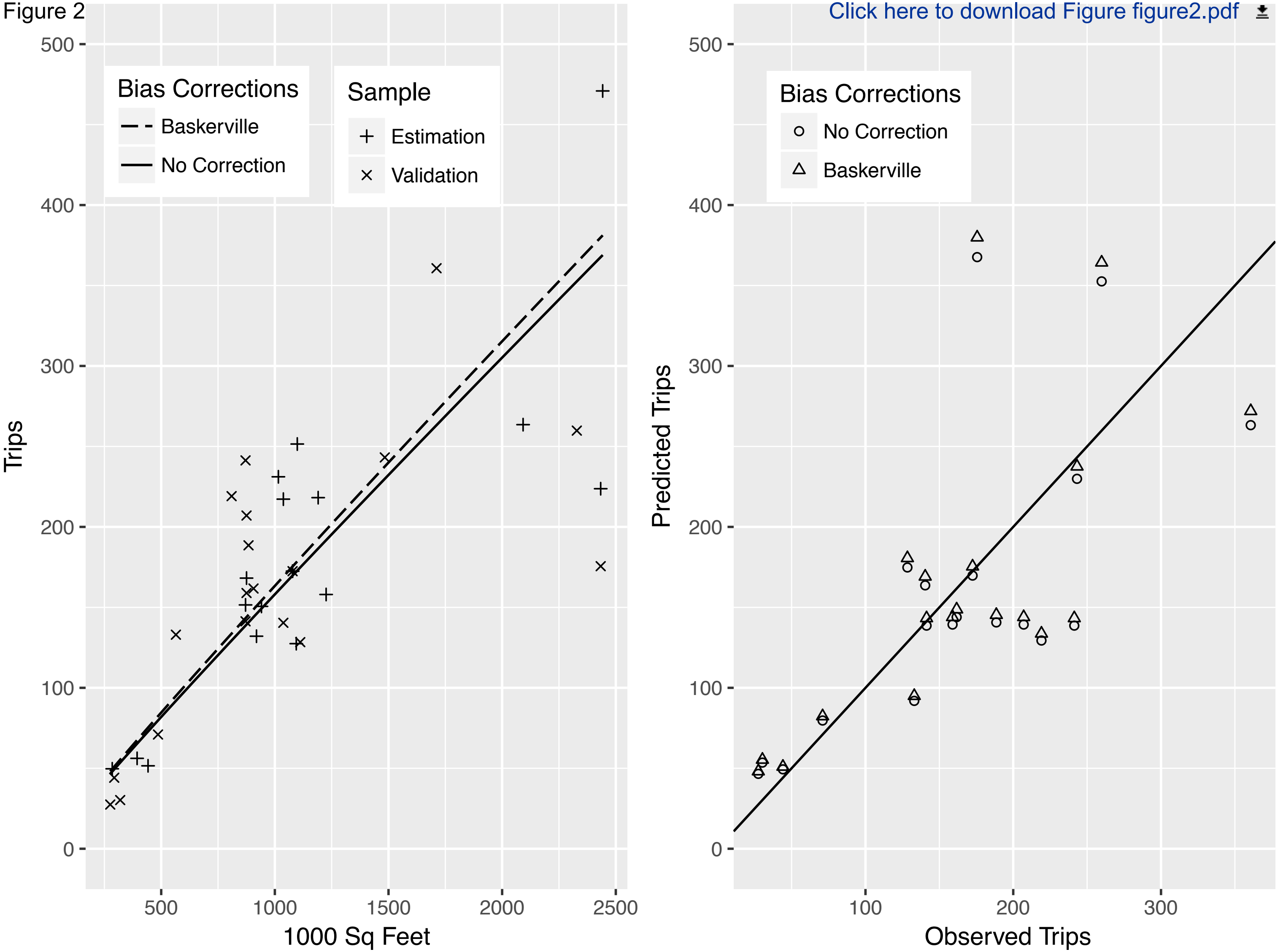
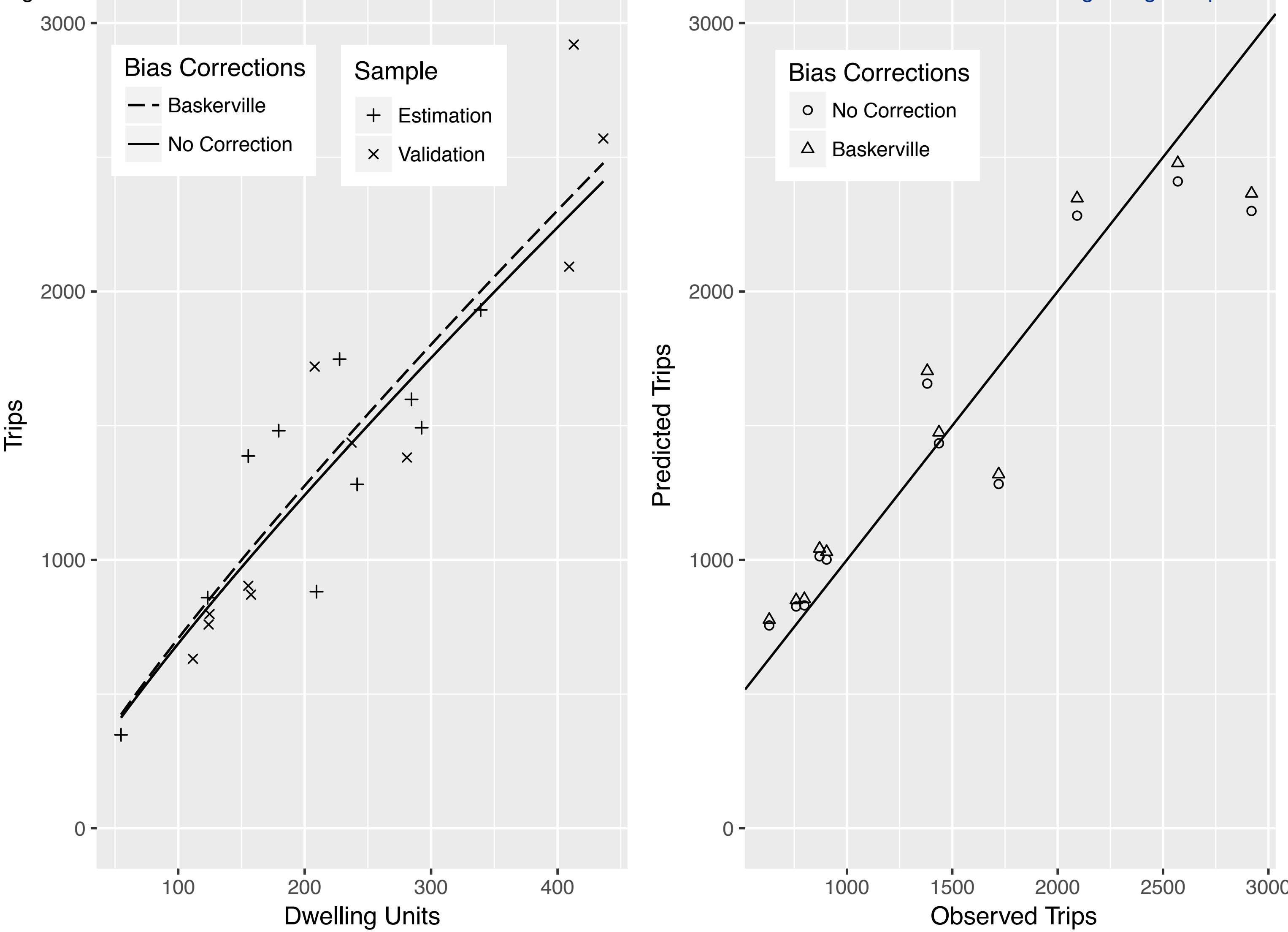

Bias Corrections
Baskerville
No Correction
Sample
Estimation
Validation
Trips
1000 Sq Feet
0
100
200
300
400
500
500
1000
1500
2000
2500
Bias Corrections
No Correction
Baskerville
Predicted Trips
Observed Trips
0
100
200
300
400
500
100
200
300

Figure 3



Bias Corrections
Baskerville
No Correction
Sample
Estimation
Validation
Trips
Dwelling Units
3000
2000
1000
0
100
200
300
400
Bias Corrections
No Correction
Baskerville
Predicted Trips
Observed Trips
3000
2000
1000
0
1000
1500
2000
2500
3000

Figure 4

Click here to download Figure figure4.pdf

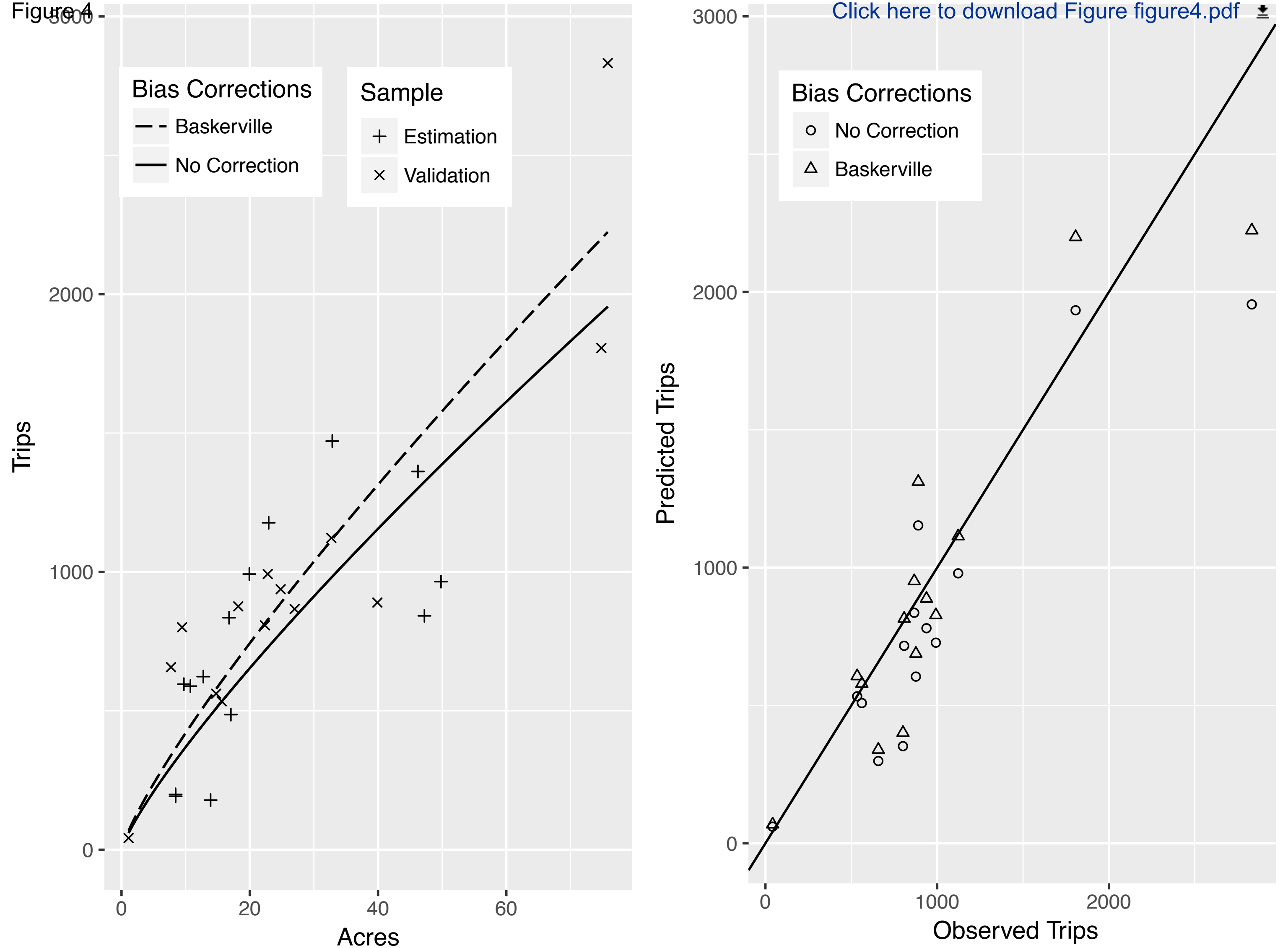


Figure 5

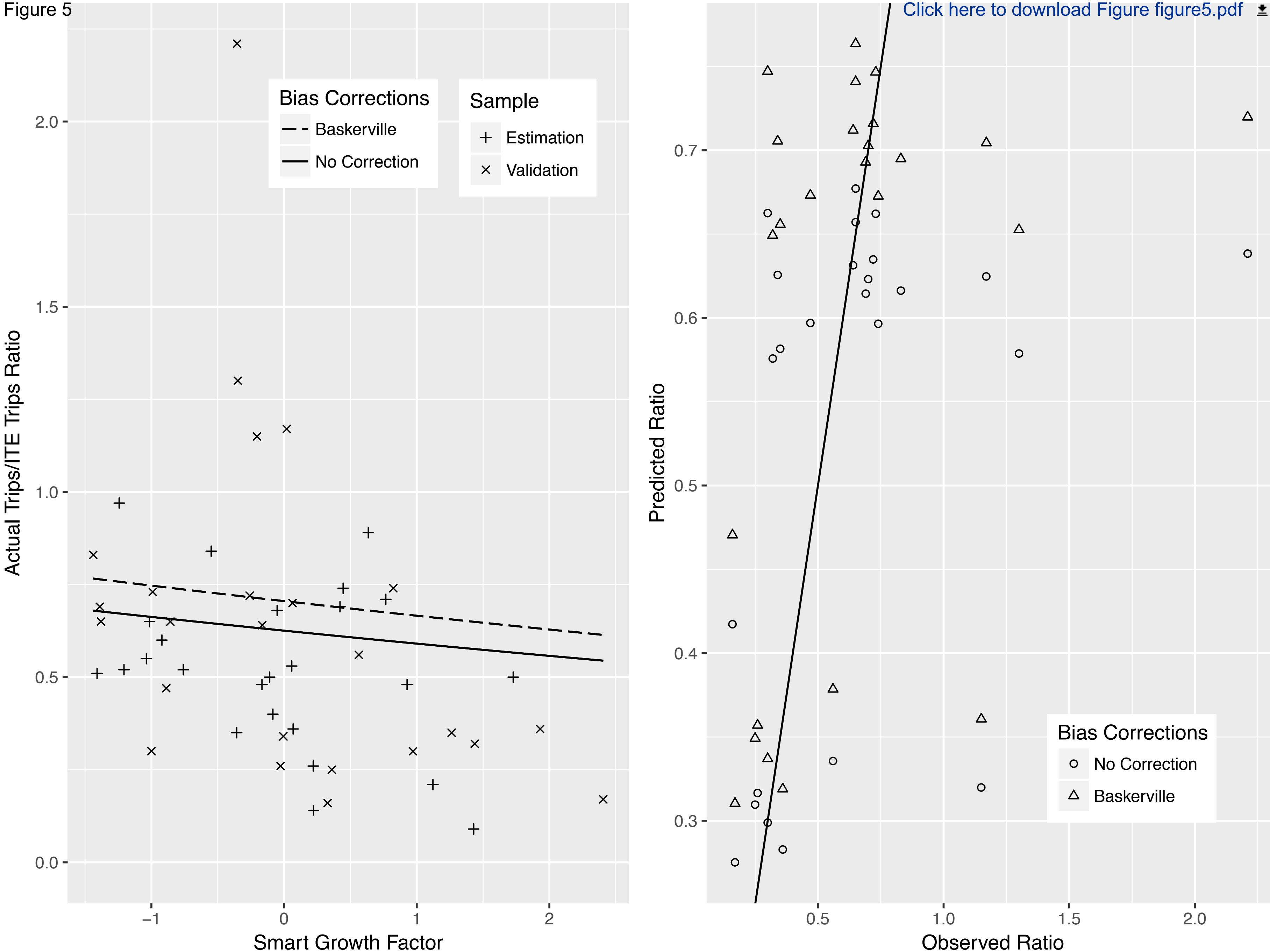

Bias Corrections
Baskerville
No Correction
Sample
Estimation
Validation
Actual Trips/ITE Trips Ratio
2.0
1.5
1.0
0.5
0.0
−1
0
1
2
Smart Growth Factor
Predicted Ratio
0.7
0.6
0.5
0.4
0.3
0.5
1.0
1.5
2.0
Observed Ratio
Bias Corrections
No Correction
Baskerville

Figure 1. (left) Normalized mean error (bias) and (right) normalized root mean square error (accuracy) in simulations with and without bias correction

Figure 2. Weekday P.M. Peak Hour Trips for High-Cube Warehouse/Distribution Center (LUC 152) with and without bias correction (Baskerville) for (left) observed trips versus establishments size and (right) predicted versus observed trips, data source: (Institute of Transportation Engineers, 2012)

Figure 3. Sunday trips for Low-Rise Apartment (LUC 221) with and without bias correction (Baskerville) for (left) observed trips versus dwelling units and (right) predicted versus observed trips, data source: (Institute of Transportation Engineers, 2012)

Figure 4. Sunday trips for Mobile Home Park (LUC 240) with and without bias correction (Baskerville) for (left) observed trips versus acreage and (right) predicted versus observed trips, data source: (Institute of Transportation Engineers, 2012)

Figure 5. California smart growth trip generation rates with and without bias correction (Baskerville) for (left) dependent variable versus smart growth factor and (right) predicted versus observed dependent variable, data source: (Schneider et al., 2015)