TECHNICAL MEMORANDUM

Date: April 27, 2015

To: Rachel Schuett, San Francisco Planning Department, and Carli Paine, San Francisco Municipal Transportation Agency

From: Jennifer Ziebarth, Tien-Tien Chan, Chris Mitchell

Subject: Parking Analysis and Methodology Memo - Final

SF13-0666.02

OVERVIEW

This memo analyzes the relationship between auto mode share and parking availability for three different land use types in San Francisco: residential, retail, and office. Specifically, the models presented below describe how the observed auto mode share at specific sites in San Francisco is related to both the auto orientation of the site and to the availability of parking. Auto orientation of each site is represented by auto mode shares from the SF-CHAMP model (2012 Base year, model version 4.3); each of the three land use types uses a slightly different set of modeled trips (described with each model below) from which to calculate this "background" auto mode share.

The data used to develop the statistical models described in this memo was collected via surveys conducted at retail and residential sites, and, for the office model, by reviewing surveys conducted by the Transportation Management Association of San Francisco. Further details about the data used in this analysis, including the locations surveyed and the survey process, can be found in the Data Collection Memo and the Data Results memo.

All of the models which follow are calculated using a standard statistical technique called linear regression. The data collected includes a *dependent* variable: the observed driving behavior of each individual (eg yes they drove or no they did not), as well as *independent* variables: the auto orientation of the site (measured via 2012 base year auto mode share predictions from the SF-CHAMP travel demand model) and the parking status of the individual or site. The core idea of the analysis is to determine whether there is a relationship between the dependent and the independent variables, and if so to quantify it. In particular it is important to consider the possibility, called the *null hypothesis,* that there is no relationship between the dependent variable and one or more of the independent variables, but rather any apparent patterns connecting them could instead be explained as just "noise" in the variability of the data. The models below quantify the relationship between driving behavior and both parking and auto orientation of the site, and confirm that (with one exception explained below) both independent variables are in fact related to the dependent variable.

Key observations from these models are summarized below. Further details are given in the sections which follow.

- Residential: Both parking and auto orientation of the site are significant predictors of residential auto mode share, and residents with reserved parking are predicted to have substantially higher auto mode share than those without, particularly in the morning. By way of example, the AM residential model predicts that for a site with moderate auto orientation (eg at the median auto mode share), the absence of parking is associated with a 40% reduction in auto mode share. The PM residential model predicts that for a site with moderate auto orientation, the absence of parking is associated with a 35% reduction in auto mode share.
- Office: Both parking and auto orientation of the site are significant predictors of office auto mode share, and workers with free or subsidized parking have a substantially higher auto mode share than those without. The influence of the site's auto orientation is not as strong as in the residential case, while the influence of parking remains strong. As an example, the office model predicts that a for a site with moderate auto orientation, the absence of free or subsidized parking is associated with a 32% reduction in auto mode share.
- Retail: Auto orientation of the site is a significant predictor of retail auto mode share, while the relationship between auto mode share and parking is notably smaller than for the residential and office models, particularly in the morning. As an example, the AM retail model predicts that for a site with moderate auto orientation, the absence of parking is associated with a 20% reduction in auto mode share. The PM retail model predicts that for a site with moderate auto orientation, the absence of parking is associated with a 30% reduction in auto mode share.

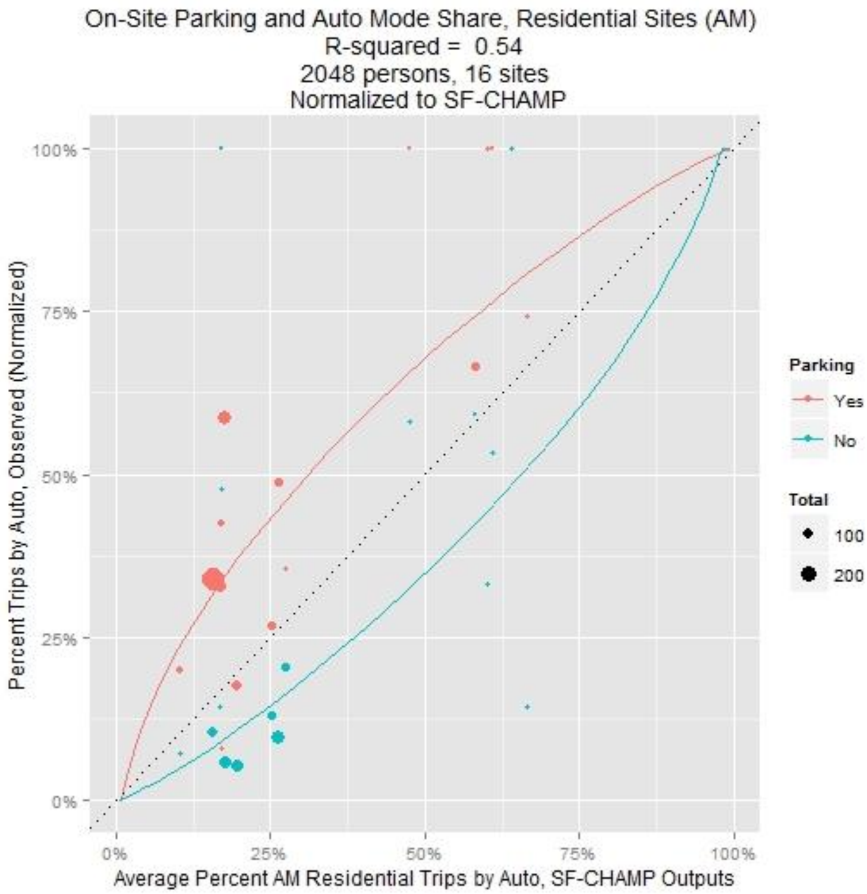
RESIDENTIAL MODELS

The residential models analyze the auto mode share of surveyed individuals during the AM and PM peak, as predicted by the auto orientation of the site and by the availability of parking. For the residential models, auto orientation of a site is represented by the peak period (AM or PM, respectively) auto mode share of home-originating trips (also referred to as "residential trips") from the site's TAZ, as predicted by SF-CHAMP for base year 2012. Availability of parking is measured by whether each resident surveyed has reserved parking available.

The table below shows some representative predictions for the AM residential model, applied to sites with low, medium, and high auto orientation. The graph below gives a visual representation of the AM peak residential model as well as the data used to develop the model. Each residential site is represented by two points: one for the individuals without reserved parking (red) and one for the individuals with reserved parking (blue). The size of each point reflects the number of individuals of each type surveyed at each site. The blue and red curves indicate the model's predictions for the relationship between auto mode share for residential trips and observed auto mode share for individuals without (red) or with (blue) reserved parking.

RESIDENTIAL AM MODEL PREDICTIONS

	Low auto orientation	**Medium auto orientation**	**High auto orientation**
Predicted AM residential auto mode share, with parking	15%	30%	70%
Predicted AM residential auto mode share, without parking	8%	18%	55%
Predicted AM reduction in auto mode share with no parking	48%	39%	22%

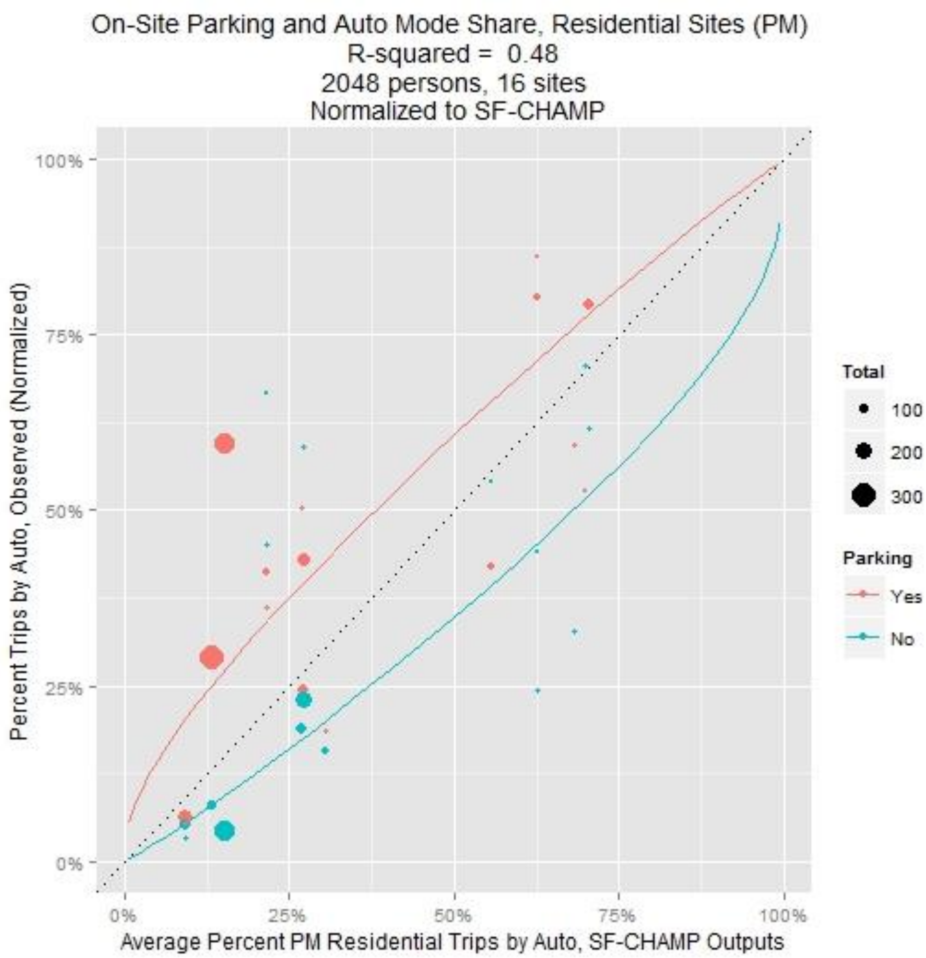


The model's fit can in part be evaluated by the R^2 value of 0.54, which indicates that the model explains 54% of the individual-level variability in choice whether or not to drive. This R^2 value should be understood from the perspective that individual choices of whether or not to drive are made based on a wide variety of factors. The fact that the two factors considered here (parking and site auto orientation) can explain over half of the variability in these individual choices indicates that these two factors together have a very strong relationship to residential AM auto mode share.

The PM peak residential model is not quite as strong as the AM peak model, but still explains close to half the variability in individual-level auto mode share.

RESIDENTIAL PM MODEL PREDICTIONS

	Low auto orientation	Medium auto orientation	High auto orientation
Predicted PM residential auto mode share, with parking	15%	30%	70%
Predicted PM residential auto mode share, without parking	9%	20%	52%
Predicted PM reduction in auto mode share with no parking	39%	35%	26%



The R^2 value indicates that the model explains just under 50% of the variability in individual-level choice of whether to drive. The relatively large R^2 values of both the AM and PM residential models, together with the substantial distance between the curves representing the presence or absence of reserved parking, provide evidence for a substantial reduction in auto mode share associated with the absence of residential reserved parking.

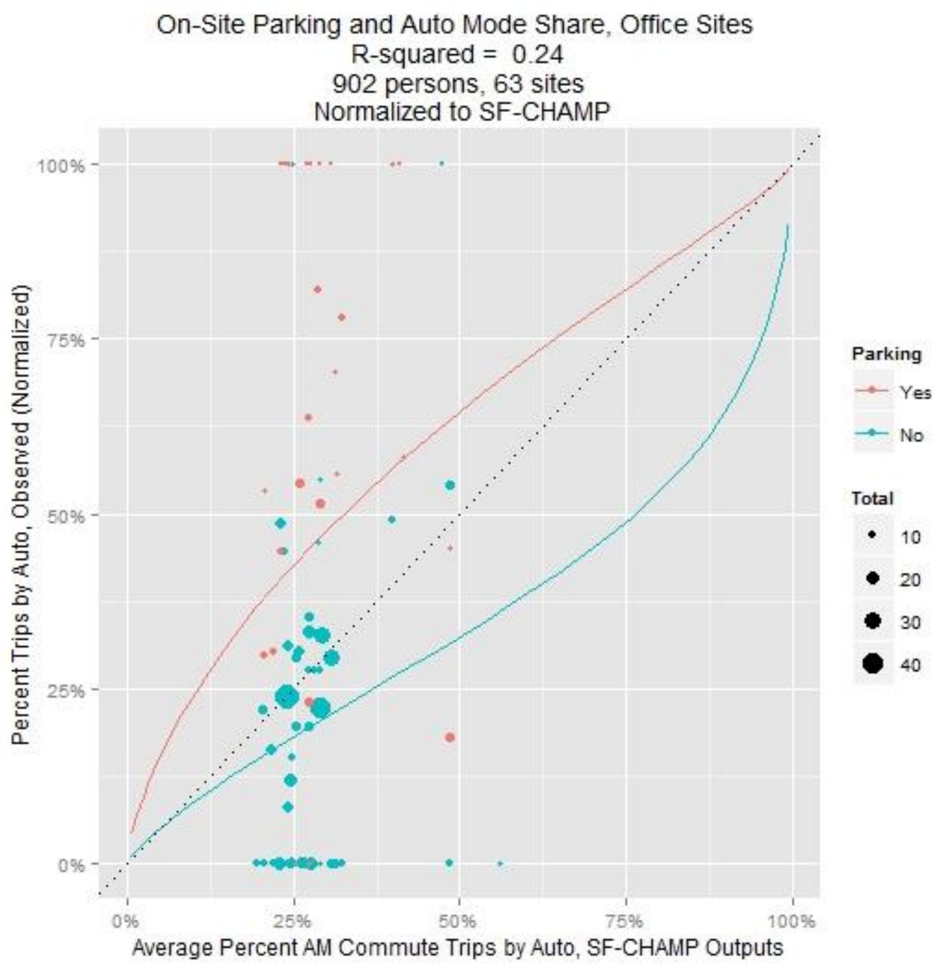
OFFICE MODEL

The office model analyzes the commuting auto mode share of individuals surveyed in the 2014 TMASF survey, as predicted by the auto orientation of the site and the availability of parking. Unlike the residential model, the office model is not separated into AM and PM peak versions. For the office model, auto orientation of a site is represented by the AM peak period auto mode share of commute trips with destination in the site's TAZ, as predicted by SF-CHAMP for base year 2012. Availability of parking is measured by whether each worker surveyed has free or subsidized parking provided by the employer. However, not all respondents to the TMASF poll answered this question; only workers who responded to the parking question are included in the model below.

In this model, the site's auto orientation is not as strong a predictor as it is in the residential models. The table below shows some representative predictions for the office model, applied to sites with low, medium, and high auto orientation. The graph gives a visual representation of the office model as well as the data used to develop the model. Each office site is represented by two points: one for the individuals without free/subsidized parking (red) and one for the individuals with free/subsidized parking (blue). The size of each point reflects the number of individuals of each type at each site who responded to the survey. The blue and red curves indicate the model's predictions for the relationship between auto mode share for commute trips and observed auto mode share for individuals without (red) or with (blue) free or subsidized parking.

OFFICE MODEL PREDICTIONS

	Low auto orientation	Medium auto orientation	High auto orientation
Predicted office auto mode share, with subsidized parking	20%	35%	70%
Predicted office auto mode share, without subsidized parking	15%	24%	45%
Predicted reduction in auto mode share without subsidized parking	24%	32%	36%



The office model's R^2 value indicates that the model explains 24% of the individual-level variability in choice whether or not to drive. This model does not explain as much of the individual-level variability in choice whether to drive as the residential model does; in part this is because the site's auto orientation does not add as much predictive power as it did in the residential case. One likely reason for this is that auto mode share for commute trips to San Francisco destinations does not vary as widely as residential auto mode share, with the middle half between 25% and 30% auto mode share . Nevertheless, the model's ability to explain one quarter of the variability based on only parking and the site's auto orientation indicates that these factors – primarily parking – have a moderately strong relationship to office auto mode share.

RETAIL MODELS

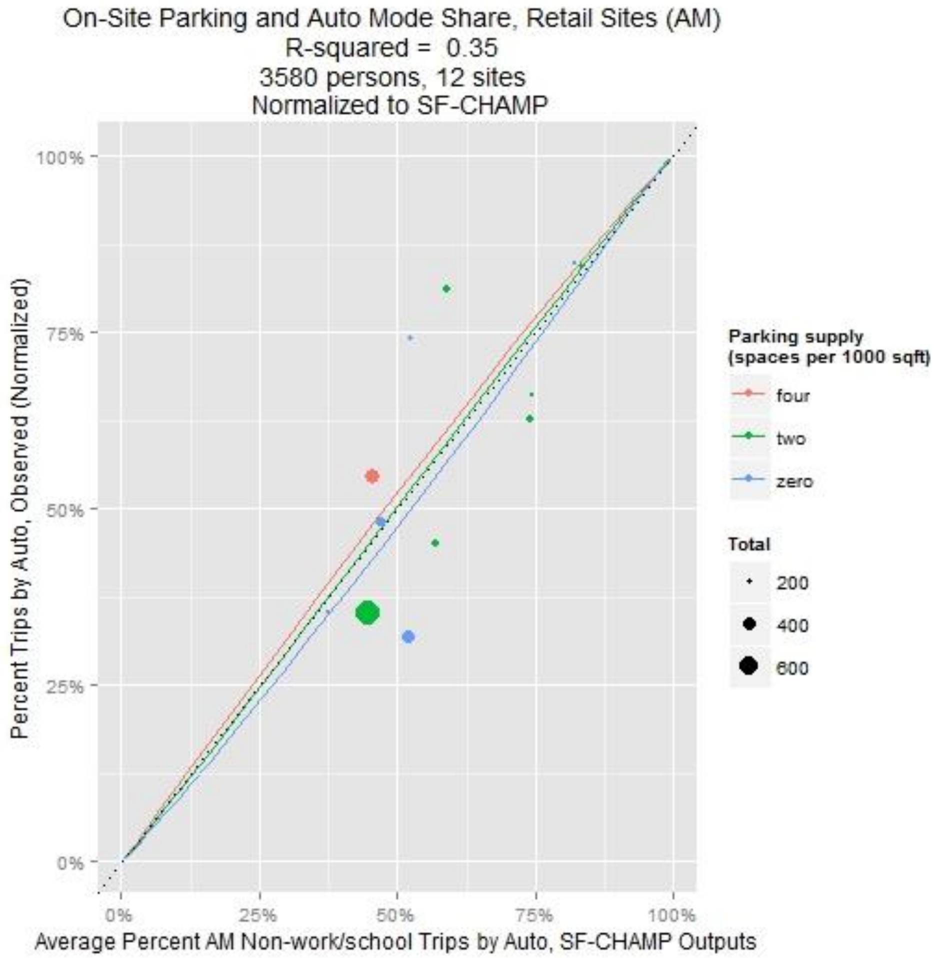
The retail models also seek to explain auto mode share behavior based on the auto orientation of each retail site and the on-site parking provided at each retail site. In this case, auto orientation of a site is represented by the peak period (AM or PM, respectively) auto mode share of all trips with destination in the TAZ whose purpose is neither commute nor school. Two related models for each time period were considered: one version considered parking as a numeric variable, measured as the number of parking stalls available on-site per 1,000 gross square feet of retail space, to test whether retailer sites with more parking have higher auto mode share than those who offer less parking. The other version considered parking as a yes/no variable to test whether the simple presence or absence of on-site parking would influence auto mode share. Note that unlike the residential and office models, this will not vary among different individuals at the same site.

PARKING AS NUMERIC

The table below shows some representative predictions for the AM retail models, applied to sites with low, medium, and high auto orientation. The graph gives a visual representation of the AM peak retail model as well as the data used to develop the model. Each retail site is represented by a single point. The size of each point reflects the number of individuals of each type surveyed at each site, and the color of each point reflects the (rounded) number of parking spaces per 1,000 gross square feet of retail space. The blue, green, and red curves indicate the model's predictions for the relationship between auto mode share for non-school, non-commute trips and observed auto mode share for individuals coming to sites with no parking (red), 2 parking spaces per 1,000 retail square feet (green), or 4 parking spaces per 1,000 retail square feet (blue).

RETAIL AM MODEL PREDICTIONS (PARKING AS NUMERIC)

	Low auto orientation	Medium auto orientation	High auto orientation
Predicted AM retail auto mode share, 4 parking spaces per 1000 sqft retail space	32%	52%	84%
Predicted AM retail auto mode share, 2 parking spaces per 1000 sqft retail space	30%	50%	85%
Predicted AM retail auto mode share, no parking	28%	48%	87%
Predicted AM reduction in auto mode share from 4 spaces per 1000 sqft to no parking	8%	5%	1%

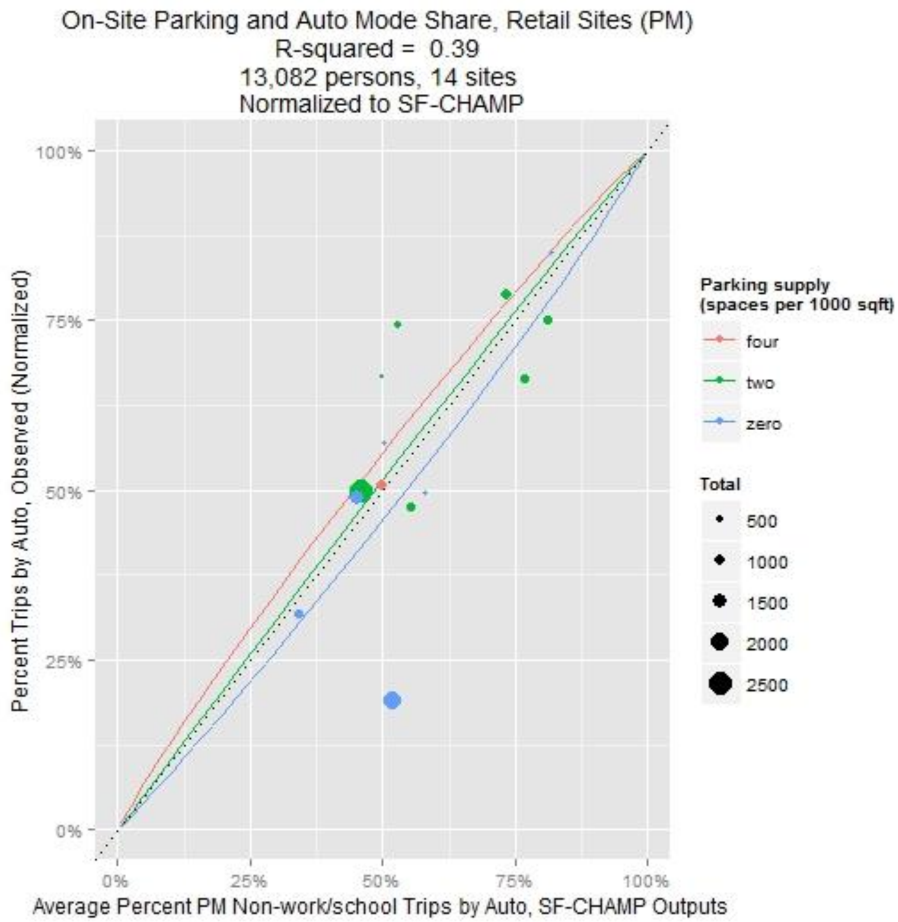


The data shows a very weak relationship between the provision of retail parking and auto mode share, as evidenced by the close proximity of the three curves in the graph. The model's fit, as evaluated by the R^2 value of 0.35, indicates that the model explains 35% of the individual-level variability in choice whether or not to drive.

The PM retail model shows a stronger effect on auto mode share from the amount of parking provided.

RETAIL PM MODEL PREDICTIONS (PARKING AS NUMERIC)

	Low auto orientation	Medium auto orientation	High auto orientation
Predicted PM retail auto mode share, 4 parking spaces per 1000 sqft retail space	35%	56%	88%
Predicted PM retail auto mode share, 2 parking spaces per 1000 sqft retail space	30%	50%	85%
Predicted PM retail auto mode share, no parking	28%	46%	82%
Predicted PM reduction in auto mode share from 4 spaces per 1000 sqft to no parking	12%	9%	4%



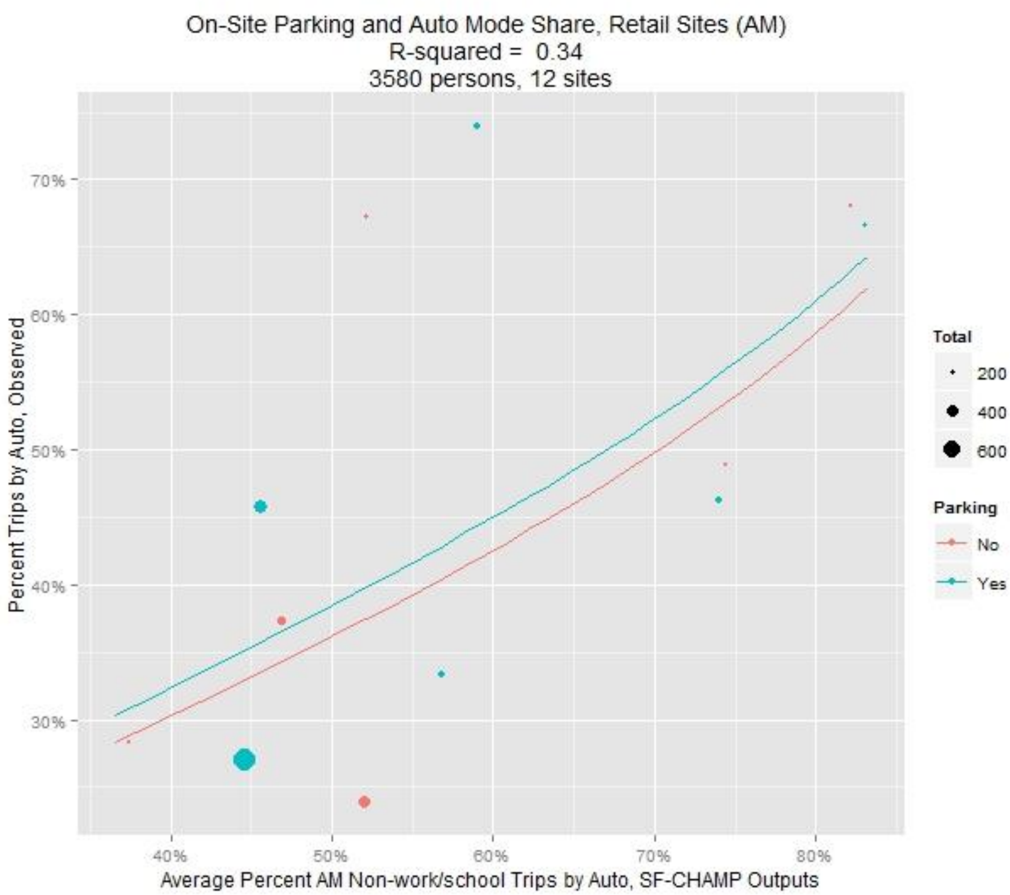
The model's R^2 values indicates that it explains almost 40% of the individual-level variability in choice whether or not to drive. As in the AM retail model, the influence of the site's auto orientation is more pronounced in the PM retail model than in the office or residential models.

PARKING AS BINARY

A variant of the retail model includes parking not as a numeric variable indicating the amount of parking, but rather as a binary variable indicating whether or not the retailer offers parking on site. These models provide a slightly simpler view of the effect of parking on retail auto mode share; however in the AM model the effect of parking is so small that it does not provide evidence against the null hypothesis (that there is no relationship between parking and observed driving behavior).

RETAIL AM MODEL PREDICTIONS (PARKING AS BINARY) – NOT NORMALIZED

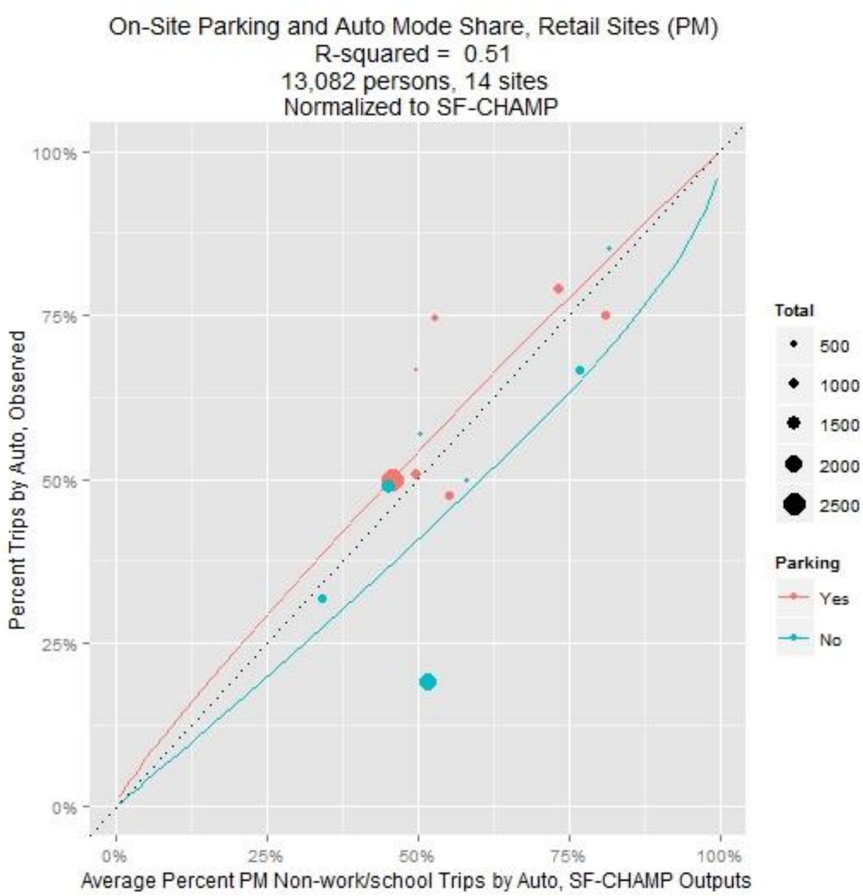
	Low auto orientation	Medium auto orientation	High auto orientation
Predicted AM retail auto mode share, with parking	30%	50%	85%
Predicted AM retail auto mode share, no parking	24%	41%	74%
Predicted AM reduction in auto mode share without parking	20%	19%	13%



The PM retail model which considers parking as a binary variable has a slightly better fit than its numeric counterpart, as measured by its R^2 value of 50%.

RETAIL PM MODEL PREDICTIONS (PARKING AS BINARY)

	Low auto orientation	Medium auto orientation	High auto orientation
Predicted PM retail auto mode share, with parking	30%	41%	62%
Predicted PM retail auto mode share, no parking	20%	29%	49%
Predicted PM reduction in auto mode share without parking	33%	30%	21%



CHOOSING BETWEEN THE TWO RETAIL VARIANTS

The choice of which variant of the retail model to use (parking as numeric or parking as binary) is somewhat muddied by the fact that among the AM models, the numeric model performs better, while among the PM models, the binary model performs better. Because the AM binary model does not provide evidence that parking availability predicts driving behavior, this model is the least usable of the four. If consistency between the AM and PM models is desired, it is therefore preferable to use the numeric models. However, if having a better-performing PM model is more important than keeping it consistent with the AM model, then the binary PM model would be a better choice. For purposes of the TDM tool, the binary model was selected because only the PM model was used.

STATISTICAL DETAILS

This section provides greater details about the statistical models behind the graphs shown in the previous sections. It is likely to be of interest primarily to the most statistically enthusiastic of readers.

BINOMIAL LOGIT MODELS

All of the models in this analysis are binomial logit models, with individual persons as the unit of analysis. Each model uses two independent variables: auto orientation of the site, as measured by the 2012 base year SF-CHAMP prediction of auto mode share for a particular set of trips defined above for each model, and parking, measured either as a binary variable (0=no, 1=yes) or as a numeric variable. In all of the models, the "background auto mode share" used to measure auto orientation is transformed via the logit transformation:

$$logitAMS = log\left(\frac{AMS}{1 - AMS}\right)$$

In a binomial logit model, the dependent variable is also transformed using a logit transformation, so that the model equation takes the form:

$$\log\left(\frac{\hat{y}}{1 - \hat{y}}\right) = b_{Intercept} + b_{logitAMS} * logitAMS + b_{Parking} * Parking$$

In each of the tables which follows, the following information is provided for each of the three coefficients b_i above:

- Estimate: The estimated value of the coefficient.
- Std. Error: A measure of the uncertainty in the value of estimate – essentially a margin of error for the estimate.
- z value: The ratio of the estimate to the standard error. The larger the z value, the further away from zero the estimate is relative to its margin of error, and therefore the more likely it is that the variable in question is adding genuine information to the model.
- p value: The likelihood of obtaining a z value this large or larger in absolute value, under the null hypothesis that the variable in question has no effect. A larger p value (above 0.01 for example) indicates that the null hypothesis is plausible, and the variability seen in the data is random rather than the systematic effect of the variable in question. Very small p values are reported using scientific notation, eg 2.71E-12 = 0.00000000000271.
- Significance: A quick visual representation of how small the p value is, and thus of how strong the evidence is against the null hypothesis.

  ***: Under 0.001 – very strong evidence against the null hypothesis

  **: Between 0.001 and 0.01 – strong evidence against the null hypothesis

  *: Between 0.01 and 0.05 – moderate evidence against the null hypothesis

  . : Between 0.05 and 0.1 –weak evidence against the null hypothesis

  : Above 0.1 – no evidence against the null hypothesis

RESIDENTIAL MODELS

Coefficients for the AM peak and PM peak residential models are shown in the tables below. For both models, all variables are highly statistically significant.

AM RESIDENTIAL MODEL COEFFICIENTS

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.73256	0.10477	-6.992	2.71E-12	***
logitAMS	0.39824	0.07062	5.639	1.71E-08	***
Parking	1.38087	0.10061	13.725	< 2e-16	***

PM RESIDENTIAL MODEL COEFFICIENTS

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.64022	0.08652	-7.399	1.37E-13	***
logitAMS	0.37793	0.0454	8.324	< 2e-16	***
Parking	1.06604	0.08948	11.914	< 2e-16	***

OFFICE MODEL

Coefficients for the office model are shown in the table below. The parking variable is highly statistically significant, and the site auto orientation variable is moderately statistically significant.

OFFICE MODEL COEFFICIENTS

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-1.1459	0.2726	-4.203	2.63E-05	***
logitAMS	0.6705	0.2612	2.567	0.0103	*
Parking	1.3707	0.1809	7.576	3.55E-14	***

RETAIL MODELS

The retail models include both a variant in which parking is considered as a numeric variable (number of stalls per thousand square feet of retail space), and a variant in which parking is considered a binary variable. For both numeric models, site auto orientation is highly statistically significant. Parking is highly statistically significant in the PM model, and moderately statistically significant in the AM model.

AM RETAIL MODEL COEFFICIENTS (PARKING AS NUMERIC)

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.58512	0.05258	-11.128	<2e-16	***
logitAMS	0.69537	0.06141	11.324	<2e-16	***
Parking	0.05152	0.02396	2.151	0.03	*

PM RETAIL MODEL COEFFICIENTS (PARKING AS NUMERIC)

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.80339	0.02634	-30.498	<2e-16	***
logitAMS	0.64415	0.03075	20.95	<2e-16	***
Parking	0.11237	0.01377	8.159	3.37E-16	***

For the binary parking models, site auto orientation is highly statistically significant in both models, and parking is highly statistically significant in the PM model. In the AM model however, the effect of parking is not statistically significant, and there is no evidence in this model against the null hypothesis that the presence of on-site parking influences AM retail auto mode share.

AM RETAIL MODEL COEFFICIENTS (PARKING AS BINARY)

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.56753	0.058	-9.786	<2e-16	***
logitAMS	0.66124	0.05916	11.177	<2e-16	***
Parking	0.09887	0.07133	1.386	0.166	

PM RETAIL MODEL COEFFICIENTS (PARKING AS BINARY)

	Estimate	Std. Error	z value	p value	Significance
(Intercept)	-0.96627	0.02959	-32.66	<2e-16	***
logitAMS	0.60252	0.03086	19.52	<2e-16	***
Parking	0.53878	0.03745	14.39	<2e-16	***

CALCULATION OF R^2 VALUES

The R^2 value for each of the statistical models measures the amount of variability in the person-level decision whether or not to drive which can be explained by the model. For each model, R^2 is calculated using the following equation:

$$R^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2}$$

where:

- y is the observed driving behavior of each individual. Because each individual either drove or did not, it is always either 0 (for individuals who did not drive) or 1 (for individuals who did).
- $\hat{y}$ is the model prediction of the probability that the individual will drive, based on the individual or site parking status and the auto orientation of the site.

- $\bar{y}$ is the mean of all of the y values for all individuals in the model data set. Note that it is equal to the auto mode share for the model data set as a whole.

To understand this calculation, note that the numerator measures how far away the model's prediction for each individual is from the observed behavior of that individual, summed for every individual. As such, the numerator measures the amount of variability which is *not* explained by the model. The denominator measures how far away each individual is from the overall average of the entire group. As such, the denominator measures the total amount of variability present in the data. The fraction as a whole therefore represents the proportion of variability which is *not* explained by the model, and subtracting it from 1 results in the proportion of variability which *is* explained by the model.

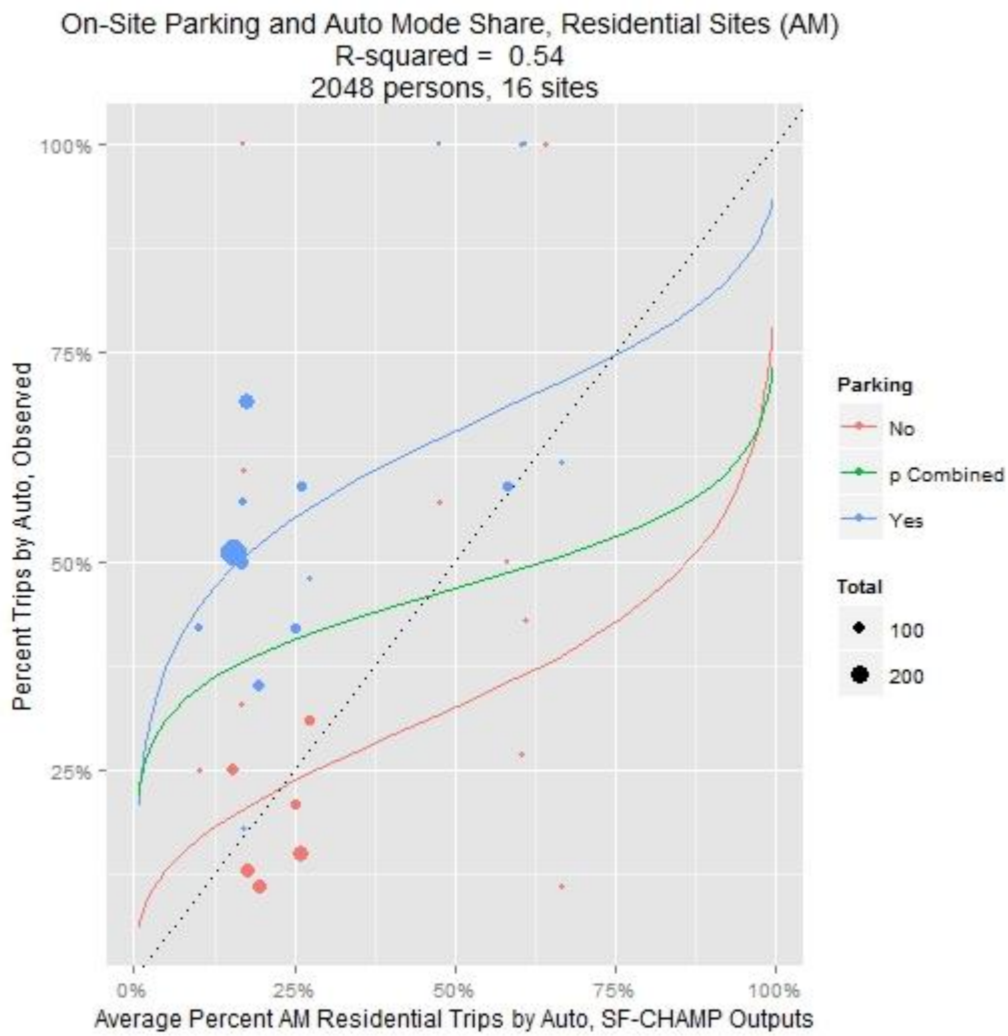
NORMALIZATION TO SF-CHAMP

Once each statistical model is estimated, the resulting predictions are normalized so that the resulting auto mode shares match predictions from SF-CHAMP. The first step in performing this normalization is to estimate a statistical model for the relationship between auto mode share as reported by SF-CHAMP and auto mode share as observed in the survey. This "combined" model incorporates both the observations with parking and those without. The graph below illustrates the pre-normalized AM residential models, including the with and without parking models as well as the combined model (in green).

The intent of the normalization process is to adjust the combined model (green) so that it matches the auto mode shares from SF-CHAMP (dotted black). To accomplish the normalization, points on or above the green line are scaled vertically to be on or above the dotted line, while points at or below the green line are scaled vertically to be on or below the dotted line. In equation form, this looks like:

$$y_{cal} = x + \frac{y_{raw} - y_{combined}}{1 - y_{combined}} * (1 - x) \quad \text{(Above)}$$

$$y_{cal} = x - \frac{y_{combined} - y_{raw}}{y_{combined}} * (x) \quad \text{(Below)}$$


On-Site Parking and Auto Mode Share, Residential Sites (AM)
R-squared = 0.54
2048 persons, 16 sites
100%
75%
50%
25%
Percent Trips by Auto, Observed
0%
25%
50%
75%
100%
Average Percent AM Residential Trips by Auto, SF-CHAMP Outputs
Parking
No
p Combined
Yes
Total
100
200

APPENDIX: SAMPLE PARKING APPLICATION SPREADSHEET

The *sampleParkingApplication.xlsx* spreadsheet provides a simple tool for applying the parking models described in this memo. The spreadsheet is divided into five tabs, corresponding to the residential AM and PM, retail AM and PM, and office models. For the retail models, the binary version of the model is provided in the sample spreadsheet.

AM residential models	Parking model			Combined model	
	Coeff	NO parking	Parking	Coeff	Value
(Intercept)	-0.733	1	1	-0.131	1
Background AMS	0.398	0.847	0.847	0.227	0.847
Parking	1.381	0	1	0	0
Prediction (log odds)		-0.40	0.99		0.06

Existing parking ratio (per DU)	1.00
Planned parking ratio (per DU)	0.00
TAZ residential auto mode share	70%
Minimum pred AMS raw	*40%*
Combined pred AMS raw	*52%*
Maximum pred AMS raw	*73%*
Minimum pred AMS normalized	55%
Combined pred AMS normalized	70%
Maximum pred AMS normalized	83%
Expected mode share	55%
VMT reduction	22%

Input: change me!
Calculation (don't change)
Final VMT reduction

Snapshot of AM residential tab in *sampleParkingApplication.xlsx*

The orange cells indicate input which the user can change. These inputs are:

1. The **existing parking ratio**. For the residential model, this is measured in parking spaces per **dwelling unit**; for office it is parking spaces per **employee**, and for retail it is parking spaces per **1,000 retail square feet**. This existing ratio is the baseline against which the VMT reduction is calculated. (Note that in the TDM tool itself, the square footage

assumed per employee can be changed; here the input to the office model is solely spaces per employee.)

2. The **planned parking ratio,** measured in the same units as the maximum parking ratio. This is the parking ratio whose VMT we wish to compare to the maximum.
3. The **background auto mode share** for the appropriate type of trip (residential in the example above)

The grey cells represent intermediate steps in the calculation; the upper box contains the coefficients of the statistical models relating background auto mode share and parking to observed auto mode share. The final line of the grey table, labeled **Prediction (log odds)**, represents the models' predictions in logit-transformed form (see the statistical details section for details). These predictions are transformed to mode shares in the first three grey cells labeled **Maximum pred AMS raw**, **Minimum pred AMS raw,** and **Combined pred AMS raw**. The normalization to SF-CHAMP takes place in the next three grey cells, labeled **Minimum pred AMS normalized, Combined pred AMS normalized,** and **Maximum pred AMS normalized.**

The **Expected mode share** is calculated by comparing the planned parking ratio to the existing parking ratio, then scaling this difference to fit between the minimum and combined normalized AMS predictions. For example, if the existing parking ratio is 1 and the planned parking ratio is 0.25, then the expected mode share should be one quarter of the way between the minimum and combined AMS predictions.

The final pink cell shows **the expected reduction in VMT** resulting from the planned parking ratio. It shows the percentage difference between the existing predicted mode share and the expected mode share resulting from the planned parking ratio. In the event that the planned ratio is actually larger than the entered maximum, as above, then the VMT reduction will show as a negative number.