Privacy and Prediction: How Useful are Geo-Tracking Data for Predicting Consumer Visits?

Abstract

Accurately predicting consumer visits is critical for businesses to plan resources and serve their customers better. We examine the extent to which geo-tracking data about consumers' locations can improve the ability of businesses to predict total visits to their store. Given the sensitive nature of geo-tracking data, we also examine how potential privacy regulations that restrict such data may limit their usefulness for prediction. Using proprietary data from a safe-driving app with over 120 million driving instances across 38,980 app users and aggregate data on the total number of visits to over 400 restaurants in Texas, we quantify the value of geo-tracking data by training machine learning models to predict total visits to a restaurant one week ahead. Our results show that geo-tracking data improve the performance of prediction models by 10.76% relative to that of models that use demographic and behavioral data only. Simulation exercises that limit *what* data are tracked, *where* and *how frequently* these data are tracked show a decrease in the predictive performance of models that use geo-tracking data. However, the decrease varies by the type of restriction; regulations that restrict *what* data are geo-tracked (i.e., summaries of driving behaviors) result in the largest decreases in predictive performance (13.78%), while regulations that restrict *where* users are geo-tracked (i.e., within a few miles of a business location) and *how frequently* result in smaller decreases (5.52% and 3.15-3.41%, depending on the frequency). Importantly, models with geo-tracking generally outperform models that do not use any geo-tracking data by reducing the extent of overpredicting and underpredicting visits. These findings can assist managers and policymakers in assessing the risks and benefits associated with the use of geo-tracking data.

Keywords: Mobile apps, consumer location, geo-tracking, privacy, machine learning.

1 Introduction

Accurately predicting consumer visits allows businesses to better plan their resources and improve their customers' service experience. In recent years, many businesses have gained the ability to use their own or third-party apps to track their consumers' data (Bhargava et al. 2020; Hermalin and Katz 2006). Geo-tracking data about consumers' locations, in particular, may allow businesses to better predict their customers' future actions and plan for them (Dean 2023; Sun et al. 2022). However, despite the potential usefulness of geo-tracking data for businesses, these data reveal sensitive personal information about consumers (Bleier et al. 2020; Choi et al. 2023; Goldfarb and Tucker 2012). For example, Canadian coffee chain Tim Hortons evoked a "mass invasion of privacy" by geo-tracking its app users (Austen 2022). As a consequence, the use of geo-tracking data by firms has attracted legal and regulatory action (Tau 2023).[1]

Though geo-tracking has attracted the attention of consumers, firms, and regulators, it is not clear to what extent it allows firms to better predict future outcomes and how privacy regulations impact the usefulness of geo-tracking data. In this context, our research addresses two objectives. First, we quantify the value of geo-tracking data by examining the extent to which businesses may be able to better predict the total visits to their stores using geo-tracking data relative to using only consumer demographics and behavioral data on past visits to the store. Second, we examine how restricting geo-tracking data under potential privacy regulations impacts the usefulness of these data for prediction.

To address our research questions, we identify an application of geo-tracking data in the restaurant industry. Restaurants are interested in predicting the number of visits they expect each week to prevent under- or over-committing resources, such as staff and inventory (Oblander and McCarthy 2023), which are a major component of the operating costs for restaurants (Puranam et al. 2021). However, restaurants face significant uncertainty in demand (e.g., Fang 2022; Li and Srinivasan 2019). Therefore, we focus on quantifying the extent to which geo-tracking data can improve the prediction of customers' visits to a restaurant in a given week relative to the visits the restaurant would expect in a typical week.[2]

Our data come from a proprietary app in Texas that encourages safe driving. The app rewards points for driving safely without using one's phone. The points can be redeemed at partnering businesses. The app has over 200,000 users. For our research, we can access data from a random sample of 38,980 app users, including over 120 million driving instances for 60 weeks in 2018-2019. We combine these data with data from Safegraph on the total number of visits to 401 restaurants in 40 cities of Texas where the app is present, including data on chains like McDonald's and independent stores like Ozona Grill. We ensure that our set of restaurants is representative and, for generalizability, that it primarily includes businesses that are not listed on the app. We train Machine Learning (ML) models to predict the total visits one week ahead using consumers' demographic, behavioral, and geo-tracking information derived from the mobile app data. We compare the out-of-sample performance of the models using performance metrics, such as the mean absolute error of prediction.

[1] For example, data broker Kochava was sued by the Federal Trade Commission (FTC) for selling consumers' geolocation data that made it possible to identify their visits to sensitive locations (Federal Trade Commission 2022). The FTC also charged data vendors InMarket and X-Mode for selling consumers' raw location data (Federal Trade Commission 2024).

[2] We conducted a set of structured interviews with restaurant managers to identify this managerially-relevant application of predicting total visits one week ahead. Specifically, we predict deviations from baseline behavior rather than absolute levels. This adjustment ensures that predictions account for normal variation and focus on changes due to shocks or interventions. See Web Appendix Table A1 for a summary of the interviews.

To understand the privacy tradeoffs of geo-tracking data, we then examine how restricting geo-tracking data under potential privacy regulations impacts the predictive performance of models that use these data. Specifically, we simulate three types of regulations that restrict *what* geo-tracking data are tracked, and *where* and *how frequently* these data are tracked. We motivate and develop these simulations based on privacy regulations, industry practices, and recommendations from the privacy literature (e.g., Hermalin and Katz 2006; Neumann et al. 2023).[3] The types of restrictions we simulate reflect those that are either already implemented or have been formally proposed in regulatory settings. The proposed American Privacy Rights Act, for example, treats precise geolocation data as sensitive, aligning with our simulations that reduce precision by adding noise or aggregating location information (Congressional Research Service 2024).

We find that geo-tracking data improve the predictive performance of our models by 10.76% relative to that of models that use demographic and behavioral data only and by 31.66% relative to models that exclusively use demographic data. This improvement does not simply result from having "more data" in a mechanical way because our geo-tracking data come from drivers who use the app, while our prediction target of total visits to restaurants includes the broader universe of consumers, not just these app users. We also find that using geo-tracking data reduces the likelihood of both overpredicting and underpredicting future visits, which may allow restaurants to reduce the likelihood of both wasting supplies and staff hours (because of overprediction) and of impacting customers' experience (because of underprediction). We show that models that use geo-tracking data perform better than models that do not use these data, and that this difference in predictive performance increases with more extreme events (i.e., the prediction errors of the models that do not use geo-tracking data grow faster than those of the models that do use geo-tracking data). Our back-of-the-envelope calculations suggest that the range of savings from using geo-tracking data could be substantial in economic terms, representing about 5% of restaurant revenues.

Having established that geo-tracking data allow restaurants to better predict customers' total visits, we turn to quantify the impact of privacy regulations on predictive performance. Not surprisingly, we find that all forms of restrictions on geo-tracking data reduce their predictive value. Interestingly, the extent of the decrease varies significantly by the type of privacy restriction. Among our simulations, we find that restricting *what* data are geo-tracked (i.e., summaries of driving behaviors) results in the largest decreases in predictive performance (13.78%), while restricting *where* users are geo-tracked (i.e., within a few miles of a business location) and *how frequently* (i.e., at longer intervals) results in smaller decreases (5.52% and 3.15-3.41%, depending on the frequency) relative to models with complete geo-tracking data.

Our research contributes to two streams of literature on the value of data for firms and on privacy regulations. First, we contribute to the growing literature on the value of consumer data for firms. With the exception of Sun et al. (2022) that examine the value of omnichannel data for online behaviors, this body of research focuses on the

[3]We explore the face validity of our simulation exercises by conducting a survey about consumers' perceptions of these simulations (Jerath and Miller 2024; Lin and Strulov-Shlain 2024). See Web Appendix B for details.

value of online data only rather than offline geo-tracking data (e.g., Berman and Israeli 2022; Korganbekova and Zuber 2023; Lei et al. 2023; Rafieian and Yoganarasimhan 2021; Wernerfelt et al. 2024; Yoganarasimhan 2020). In one novel application, Katona et al. (2025) used satellite imagery data on retail stores and parking lots to show that access to such data benefits sophisticated investors. Second, we extend the literature on privacy and data obfuscation to geo-tracking data in our simulations (e.g., Aridor et al. 2025; Johnson et al. 2023a,b; Macha et al. 2024; Miller and Skiera 2023; Li et al. 2023).[4] Our research extends these papers and the ongoing policy debate about the governance of consumers' geo-tracking data (e.g., Federal Trade Commission 2024, Tau 2023).

Our research has several implications for firms and policymakers. First, we show that using geo-tracking data can improve a firm's ability to predict visits accurately by reducing both over- and under-prediction. Based on their context and the costs associated with over- and under-predicting, firms can evaluate their decision to collect and use geo-tracking data. Second, we propose practical ways in which firms can protect consumers' geo-tracking data. Third, we identify ways of restricting data that still allow firms to get predictive value from them. Specifically, we show that firms can restrict *how often* and *where* they geo-track consumers with relatively little loss in predictive value compared with restricting *what* data are tracked. Finally, our findings from the simulation exercises can allow policymakers to assess predictive losses from various types of regulatory simulations.

2 Data and Empirical Strategy

2.1 Data Sources and their Representativeness

Our primary data source is a safe-driving app based in Texas, U.S.[5] The app has over 200,000 users. For our research, we could access data for a random sample of about 20% of the app users (i.e., 38,980).

The purpose of the app is to encourage safe driving. The app can detect when a user is driving at a speed of over 10 miles per hour and if they are using their phone while driving. The app encourages safe driving by awarding users a fixed number of points for each mile driven when they do not use their phone while driving. The points can be redeemed at partnering firms, which include retailers, restaurants, and entertainment venues. Although the app is unique in its safe driving aspect, it shares some commonalities with deals and delivery apps by offering information about local businesses. The nature of data collected by the app is not unique, and most apps that access location tracking have similar data collection abilities.

The app records a user's position (i.e., latitude and longitude) every three minutes once it detects that a trip has begun. The app uses this information to record locations and driving speed with date and timestamps. We were able to access individual-level geo-tracking data for 60 weeks between September 2018 and October 2019, comprising over 120 million driving points. We use these data to identify driving trajectories and

[4] For details on the types of privacy restrictions simulated in our study, see Web Appendix Table A2

[5] The app is headquartered in College Station, Texas. During our sample period, the app had business partners and users in cities, such as Houston and San Antonio, among others.

restaurant visits. We also have access to data on app users' demographics (e.g., age, gender) self-reported by users at the time of signing up for the app. Other than age and gender, all other demographics are constructed from the American Community Survey (ACS) data we describe later in this section. The users of the safe-driving app seem to represent the general population in Texas. First, they are not patrons of a specific restaurant app. Second, we compare the age and the share of female population of the app users to that in the ACS data at the zipcode level. We find that the population in our sample is slightly younger than that in the ACS (33 years old on average, relative to 36 in the ACS data). The mean share of female population is 0.47 in the app data while it is 0.50 in the ACS data. Thus, though there are differences between the populations in the two samples, these do not appear to be dramatic, in particular in light of our data being provided by an app that targets general driving population. Unlike Texas, for places with majority non-driver population (e.g., New York), our approach may not be equally suitable.

The second source of our data is Safegraph.[6] This dataset consists of the number of visits to each restaurant location in Texas and allows us to define the outcome of prediction in terms of aggregate visits to a broad set of restaurants.

To assess the representativeness of Safegraph data, we examine three aspects of these data: overall coverage, demographic bias, and measurement error. First, we note that Safegraph accurately identifies about 95% of a brand's actual store locations when benchmarked against first-party and public sources and covers 15% of the U.S. population (Kendall 2021). Other studies also find Safegraph data to be evenly distributed across income deciles at the zip code level (Klopack and Luco 2025), suggesting strong national representativeness of the sample (Chen and Pope 2020). Second, there is work that points to potential underrepresentation of minority populations in Safegraph's panel (Coston et al. 2021; Li et al. 2024). We consider this possibility in our setting by examining whether model predictions systematically vary with local demographics, such as income and race. As shown in Web Appendix Figures D1 and D2, we find no such patterns. Finally, we also discuss the issue of measurement error in Safegraph data on our predictions in Appendix D.9. In addition, we also note that similar foot traffic data have been used in previous research (e.g., Allcott et al. 2020; Athey et al. 2018; Chen and Rohla 2018; Chiou and Tucker 2020; Goldfarb and Tucker 2020).

In our data, the number of visits to a restaurant is 334.10 on average each week across 401 restaurants. In addition to the visits data from Safegraph, we also collect the annual revenue data for each restaurant from Data Axle.

We supplement these data with additional data from Yelp to identify the category of each restaurant using Yelp's category tags (Klopack 2024).[7] We use these data when constructing our feature set relating to past user visits to a restaurant of the same category and brand as the target restaurant for which we are interested in making predictions.

[6] We received the total visits data as well as the data on polygons that identify a business location, from Safegraph, "a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the Safegraph Community. To enhance privacy, Safegraph excludes census block group information if fewer than two devices visited an establishment in a month from a given census block group."

[7] The Yelp category tags and the main category appear in Web Appendix C.

Finally, we also access the ACS 2016 5-year estimate for the Census Block Groups (CBG) in our sample to generate demographic features, such as household income and education. These features are not directly recorded at the user level by the app but could contain important demographic data at the CBG level.[8]

Our sample of restaurants for making predictions comprises 401 standalone restaurant locations. To arrive at this sample, we considered 10,582 standalone restaurant locations across the 40 cities in Texas in which the app was present during our data period. For our analysis, we use a sample of 401 restaurants (about 4% of the total restaurants) that at least 10 app users visited in our data period to avoid sparse outcomes. We use their corresponding aggregate number of visits in the Safegraph data for prediction. This sample includes a broad set of restaurants, including those that do not partner with the app. We verify that our sample of restaurants is representative of the larger set of 10,582 standalone restaurants. Specifically, we examined the type of restaurant (chain vs. non-chain) and the distribution of restaurant brands and categories overall and in our sample.

While we acknowledge the sparseness of the set of restaurants we are able to include in our main analysis, our sample of restaurants is representative and similar to the broader set of restaurants in terms of the proportion of chain stores (i.e., 75.05% overall vs. 73.69% in our sample), the brand of the restaurant (i.e, top chains are fast food brands like McDonald's, Starbucks, and Sonic overall and in our sample), and the food category of the restaurant (i.e., top categories are American, Burger, and Latin American overall and in our sample).[9]

2.2 Current Industry Practices and the Target of Prediction

In practice, managers typically use their own experience and basic sales data to gain a sense of future visits for planning staffing and inventory each week. For instance, in our interviews, managers at Culver's, Red Lobster, and a fine-dining restaurant in Atlanta noted that weekly projections help with planning staff schedules and ingredient purchases, but that they currently rely on basic counts and general intuition rather than sophisticated models. A few larger chains in our interviews (e.g., Baskin Robbins) reported integrating POS (point-of-sale) analytics and reservation systems for predictions. However, nearly everyone we interviewed said that predicting visits more accurately will be invaluable for their planning and decision-making.

From our interviews, it is clear that restaurants typically have some information to form an idea about future visits based on their category, location, seasonal trends, and other aggregate features (e.g., Varga et al. 2024). However, the true number of visits in a specific week could differ from this expectation. Depending on the magnitude of the difference between expected and realized visits, these differences may lead to inefficient stocking of inputs and poor staff scheduling, thus impacting customers' experience. Improving prediction, therefore, implies reducing the difference between expected and true visits, the ultimate object of interest for restaurants and retail managers.

[8]The American Community Survey has been used extensively in academic research to extract demographic data at the Census Block Group and Census Tract levels. See, for example, Bertrand et al. (2015), Chetty et al. (2016), Klopack et al. (2024), and Naik et al. (2016).

[9]Even though our 401 restaurants are a representative sample, in Web Appendix Table D1, we report the results for alternative samples of restaurants.

For these reasons, we focus on predicting the restaurant-week specific difference between the expected number of visits and the true number of visits as our target. We compute this target in two steps. In the first step, we estimate a linear regression that includes restaurant and month-of-year (i.e., June) fixed effects.[10] We use the estimates of this regression to compute restaurant-week deviations from expected visits (i.e., the "OLS residual"). This deviation is centered at zero over the entire estimation sample by construction. The summary statistics and distribution of the deviations from this step appear in Web Appendix Figure D3. In the second step, we use this measure of deviation between the realized and expected number of visits as our target variable of interest, which we train our models to predict using various information sets.

All our results report the out-of-sample performance of prediction models that use deviations as the target, with the exception of Web Appendix Table D2 that reports the results for total visits as an alternative target.

2.3 Relevant Information Sets and Features for Prediction

To construct predictors unique to a restaurant's consumer base, we need to aggregate our individual-level user data for the input features (i.e., predictors like demographics, behavioral, and geo-tracking data) to the restaurant-week level. To do this, we identify the relevant population for each restaurant as users who were within 30 miles of the restaurant in the previous week, the maximum distance users travel to visit restaurants in our data. We then aggregate the individual-level features of these users to the restaurant-week level. For example, if a Starbucks location has a subset of app users nearby in a given week, we average their user-week level data on each feature to create the data for that Starbucks location for that week. Note that in this approach, the relevant set of users for a restaurant varies by week depending on who was nearby and mimics the relevant target audience for a restaurant. Our models predict one week ahead to allow inter-temporal separation between the time period over which the feature set is constructed and the outcome (Lee et al. 2025).

Next, we describe three sets of information that we use as inputs in the models and that allow us to quantify the value of geo-tracking data beyond demographics and behavioral data. We report the information sets used in our prediction models in Table 1.

Model A: Demographic Data. In our first specification, we include information that is commonly available to restaurants about their customers' demographics. The demographic features include consumers' age, gender, and publicly-available ACS data that contain information about the population, race, employment, income, work commute, household size, and education at the census-block level.

Model B: Demographic and Behavioral Data. In our second specification, in addition to demographics, we also include behavioral information about the customers' number of past visits to the restaurant anytime in our data period prior to the prediction week.[11] Restaurant managers are typically able to access this type of historical

[10] Additionally, as a robustness check (reported in Table D3), we estimate a model pooling data across restaurants and include week fixed effects to capture week-specific shocks.

[11] Since behavioral data aims to capture consumers' past patronage of the restaurant, we reserve the first half of our data (i.e., weeks 1-30) to compute this information. We then make predictions for each week, starting week 31 till week 60. In this approach, past visits use the data from weeks 1-30 to predict for week

Table 1 Information Sets for Predicting Restaurant Visits

Model	Feature set used as input
Model A	Demographic data
Model B	Demographic + behavioral data
Model C	Demographic + behavioral + geo-tracking data

Notes: Demographic data include age, gender, and census-block level data on education, income, etc., but no location data. *Behavioral data* refer to consumers' number of past visits to the restaurant in the data period before the target week in the prediction model (e.g., 30 past weeks when predicting for week 31, 31 past weeks when predicting for week 32, and so on) and lagged total visits for the previous week. *Geo-tracking data* capture time-varying information recovered from geo-tracking data (e.g., proximate users, which is the number of users driving close to the restaurant, and proximity, which is the minimum average distance between the user's trips and the target restaurant overall and for different time-of-day windows). The geo-tracking data in Model C provide a broader view of consumer behaviors in terms of movement patterns beyond just visits to the restaurant captured in Model B.

information via point-of-sale (POS) systems, reservation data, or internal dashboards. For this reason, we treat overall past visits and lagged visits (i.e., visits in the prior week) to the focal restaurant as behavioral variables.

In contrast, variables based on visits to other restaurants of the same brand or category, as well as focal restaurant visits in just the previous week computed at the user level, rely on geo-tracking data and are included in the next specification (Model C). While it is possible that improved internal data systems could, in principle, capture some of this information, managers generally do not have access to user-level, cross-location, or competitor-level visit data without geo-tracking.

Model C: Demographic, Behavioral, and Geo-Tracking Data. In this specification, we build on Model B and add geo-tracking data to our prediction features. Our goal for extracting features from geo-tracking data is to create a set of features that managers and app platforms commonly use for predicting store visits as well as those that are conceptually relevant in the marketing literature.[12]

Next, we describe the geo-tracking features we use for prediction. First, geo-tracking data allow managers to observe visits not only to their own restaurant but also to competing restaurants, including those of the same brand or in the same category. Accordingly, we include metrics on both the number of visits and the number of unique users to the focal restaurant, same-brand restaurants, and same-category restaurants. These measures are constructed over two time windows: (i) all past weeks up to the prediction week (similar to the window used for behavioral data in Model B) and (ii) the most recent previous week only. A full list of these variables, along with their precise definitions and construction windows, is provided in the Web Appendix Table E1.

31, then weeks 1-31 to predict for week 32, and so on (consistent with Sun et al. 2022). Thus, behavioral data captures visits computed over an increasing number of weeks every week. This widening window reflects how managers in practice aim to use all available historical data to improve forecasting accuracy.

[12] The safe-driving app company we partnered with provides a dashboard to its partnering firms. The dashboard includes a subset of similar features, such as the number of visits to and the number of users driving within close proximity of a focal restaurant and other restaurants of the same brand and same category as the focal restaurant with day/timestamps. Managers we interviewed mentioned using data on similar select practical metrics, rather than a single metric or all geo-coordinates (see Web Appendix A).

Second, since geo-tracking data allow managers to identify timestamped data with date and time information, we also include the recency of visit to any restaurant of the same category (i.e., number and recency of category visits) and the same brand (i.e., number and recency of brand visits) as the target restaurant for prediction.

Third, geo-tracking data not only enable restaurants to track actual consumer visits but also track whether potential customers were driving nearby. This information on the number of *proximate users* could be useful for predicting how many consumers are likely to visit (Surange and Shankar 2024). Therefore, we include data on the number of consumers who drove within a certain radius of the restaurant (e.g., 17 miles, 30 miles) before the prediction week. Additionally, we include features about the demographics of these individuals, such as the proportion of males versus females, their age groups, and other relevant factors. These demographic and proximity metrics, unique to geo-tracking data, offer valuable insights for restaurants looking to predict visits from specific consumer segments.

Fourth, restaurant managers often rely on heuristics, such as how far their customers are relative to their store location and their co-locations (Shoshani et al. 2024; Zubcsek et al. 2017). Geo-tracking data allow restaurants to capture the information about *proximity* (i.e., the distance at which users are from their restaurant). We operationalize this metric as the mean of the closest distance of a user to the restaurant in the week before the prediction week. It provides time-varying information and considers a user's driving activity relative to the target restaurant.

Finally, since users' driving patterns vary depending on the time of day, we also compute the *proximity* by time-of-day to each restaurant for different time-of-day windows (e.g., 8 am to 12 pm, 12 pm to 4 pm) in the previous week. We report the summary statistics of the features at the restaurant-week level in Table 2.

To illustrate the patterns captured by the geo-tracking data, the panel at the center of Figure 1 presents a two-dimensional projection of all the driving features in a multi-dimensional space of our entire feature set (see Web Appendix F for details). Each point in this projection represents a user. The driving features determine the shape of this visual, but the points in the figure are color-coded by one feature, i.e., the distance of a user's trips for a specific restaurant-week. The users represented by purple dots travel shorter distances, while those represented by pink dots travel longer distances. To illustrate these patterns, we also show the heat maps of driving patterns for two users in different regions of the projections (Van der Maaten and Hinton 2008). The user from the purple region of the plot on the left, for example, has a lower driving distance and drove mostly within Katy. In contrast, the user from the pink region of the plot on the right has a higher trip distance spanning Killeen, Austin, and College Station in a week.

2.4 Machine Learning Framework and Data Splits

Once we generate the total number of visits each week to our sample of restaurants and the input feature sets at the restaurant-week level, our final dataset for prediction has 12,030 restaurant-week level observations. This dataset represents a balanced panel of 401 restaurants observed over the 30 weeks for which we make predictions (i.e., the second half of the data). We split the data at random into 80% restaurant-weeks for

Table 2 Select Features for Prediction and Summary Statistics

Feature	Description	Mean
Demographic Data ***(N = 8 features)***		
Age	Average age of users	32.70
Gender (female)	Proportion of females	0.47
Race	Proportion of white population in a census block	0.72
Employment	Proportion of full-time employed people	0.46
Commuters	Proportion of $\geq$ 16 year-olds commuting to work	0.44
Family households	Proportion of households with >1 person in a census block	0.24
Education	Proportion of $\geq$ 25 year-old with high school diploma	0.58
Income	Median annual household income ($)	72,141
Behavioral Data ***(N = 3 features)***		
Past visits	No. of visits to the restaurant in the past weeks starting the first week in the data until the prediction week	49.88
Past visitors	No. of unique visitors to the restaurant in the past weeks starting the first week in the data until the prediction week	17.72
Lagged visits	No. of total visits (i.e., aggregate visits from Safegraph data) to the restaurant in the previous week	334
Geo-tracking Data ***(N = 87 features)***		
Proximity	Min. distance (in miles) between trip coordinates and restaurant in past weeks starting the first week in the data until the prediction week	11.14
Proximate users	No. of users within 30 miles of a restaurant in past weeks starting the first week in the data until the prediction week	4,251
Past brand visits	No. of visits to same-brand restaurants in the past weeks starting the first week in the data until the prediction week	1,232
Past category visits	No. of visits to same-category restaurants in the past weeks starting the first week in the data until the prediction week	4,815
Recency of past visit	Average number of days since last visit to the restaurant across all users	187.80
Recency of past brand visit	Average number of days since last visit to same-brand restaurants across all users	163.42
Recency of past category visit	Average number of days since last visit to same-category restaurants across all users	161.83

Notes: N=12,030 restaurant-weeks. We report a subset of geo-tracking features in this table. Other features include proximity and proximate users by demographics (i.e., these measures for female users only, for users between 25-35 years, and so on) and by time-of-day (i.e., these measures computed for morning, afternoon, and so on), and the complete list appears in Table E1. Average recency highlights visit frequency. Some places like Nitrogen Sub Zero see visits only every 30 weeks, while chains like Chick-fil-A are visited more than once per week. Lagged visits are computed using aggregate Safegraph data; all other visits and proximity metrics use the focal app's data.

Fig. 1 Visualizing Geo-Tracking Data: Trip Distance

training a model and use the remaining 20% restaurant-weeks as test data to evaluate the trained model. Thus, our main empirical strategy is to train one ML model at the restaurant-week level using the training data and evaluate and report the model's performance using the test data. The separation into training and test data ensures that the model is learning general patterns during training and is less likely to overfit to the same restaurant-weeks whose patterns it learns. We quantify the out-of-sample model performance for the test data using the Mean Absolute Error (MAE).

While our main empirical approach predicts deviations from expected visits at the restaurant-week level, we also report results from an alternative approach that splits the data by week within each restaurant to train and predict visits for that restaurant rather than at the restaurant-week level (see Web Appendix Table D3). We find results that are consistent with those from our main model, although the magnitude varies.

2.5 Model Training

To predict visits to each restaurant one week ahead, we train one model for all the restaurant-weeks in our training data. We then evaluate the performance of the model on the test data.

Because our goal is to quantify the predictive value of geo-tracking data over and above demographic and behavioral information, we use various ML algorithms, such as Random Forests, Boosted Regression Trees, and XGBoost that are able to handle non-linearities and interactions in features (Lee et al. 2025). Because Random Forest offers better performance, we report this model in the main results. We report the alternative models in Web Appendix Tables D4 and D5. The results from these models show the additional value of geo-tracking data similar to those from our main model.

2.6 Out-of-sample Evaluation

We evaluate the performance of each model using three performance metrics. First, we compute the mean absolute error (MAE), which is the average of the absolute differences between the predicted and observed values for each restaurant-week combination rw, computed as follows:

$$MAE = \frac{1}{N_{test}} \sum_{rw=1}^{N_{test}} |y_{rw} - \hat{y}_{rw}|$$

Our empirical application aims to predict the deviation in visits with a focus on overall performance across a diverse set of restaurants and periods. Given this objective, we use the MAE as our main performance metric (e.g., Varga et al. 2024; Willmott and Matsuura 2005).[13] Absolute errors provide a straightforward interpretation in the same units as the target variable and are less sensitive to outliers than RMSE. This is crucial in our case, where a few restaurant or weeks could have very high or very little deviation from the expected visits for a restaurant.

2.7 Bootstrapping Procedure

We use a bootstrap procedure to evaluate the differences in predictive performance across models with different information sets. We implement this procedure by generating 500 bootstrap samples from the training and test data with replacement for each restaurant-week combination in our data. In each bootstrap sample, we re-estimate the full model. This ensures that uncertainty from the entire estimation process is properly captured. We then re-compute the performance metrics for the models under consideration. By doing this over the bootstrap samples, we can construct a distribution for each measure of interest and model under consideration. We then use the percentile method to construct 95% bootstrapped confidence intervals of each evaluation metric. Our bootstrapping procedure allows us to quantify the uncertainty in our estimates.

3 Results: Predictive Performance by Information Set

We now present the findings related to our first research objective: examining the extent to which geo-tracking information improves the predictive performance of our models relative to consumers' demographics and behavioral information.

Table 3 reports the predictive performance of each of our models corresponding to the information sets in Table 2. We report the MAE for each model, the difference in MAE between models, and the bootstrapped confidence intervals computed using the test data at the restaurant-week level.

The results in Table 3 show that Model C, which includes the geo-tracking data, performs better than all other models. The MAE for Model C is 31.66% lower than that of Model A with only demographic information and 10.76% lower than that of Model B with demographic and behavioral information. The confidence intervals show

[13] In Web Appendix Tables D6 and D7, we also report the Root Mean Square Error (RMSE) and the Median Absolute Error (MedAE), and show that Model C tends to perform better than Models A and B.

Table 3 Results: Predictive Performance of Random Forest with Information Set

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	54.505 [54.298, 54.712]		
Model B: Behavioral data	41.739 [41.578, 41.900]	12.766 [12.635, 12.897]	
Model C: Geo-tracking data	37.249 [37.093, 37.405]	17.256 [17.125, 17.387]	4.490 [4.417, 4.563]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

that Model C outperforms models A and B, which do not use geo-tracking data as inputs.[14] It is important to note that the additional predictive value of geo-tracking data is not simply a function of adding more data in a mechanical way because the geo-tracking data come from drivers using the app, while our prediction target (i.e., total visits to restaurants) includes a broader universe of consumers.

While the performance metrics improve when we include geo-tracking data, the percentage improvement in the performance metric does not easily lend itself to business objectives. To address this, we use the MAE reported in Table 3 and combine it with auxiliary data to provide a back-of-the-envelope calculation of the managerial implications associated with these differences in predictive performance.[15] In Table 3, we show that Model C improves upon Model A by 17.26 visits (i.e., 54.51-37.25 visits) on average. When scaled to the full population of the United States (using SafeGraph's 15% coverage of the U.S. population), this difference implies an average prediction improvement of about 115 visits per week (17.26 / 0.15 $\approx$ 115). Lower over-prediction reduces unnecessary staffing by about 0.82 employees (equivalent to about \$469 per

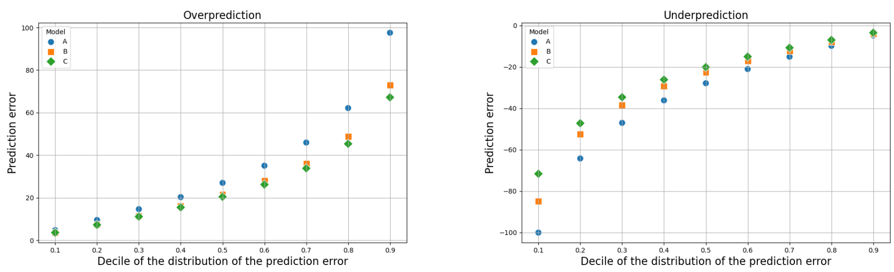
[14]Web Appendix Figures D4 and D5 present the variable importance plots for models B and C. While both figures indicate that the number of visitors and unique visitors from the previous week are the strongest predictors of week-specific deviations, incorporating geo-tracking data in Model C significantly alters the variable rankings. A large set of variables considered important predictors in Model B are replaced by features engineered from the geo-tracking data in Model C. For instance, in Model B, the third most important predictor is the mean age, but this variable does not rank among the top predictors in Model C. Instead, features derived from geo-tracking data, such as proximity and visits to other restaurants, replace many of the variables deemed important in Model B.

[15]We use information from two sources to make these calculations. First, from our interviews, we learned that restaurant managers would change their staffing scheduling and hiring if visits were to change by about 20 visits per shift or 140 per week. The Occupational Employment and Wage Statistics survey (U.S. Bureau of Labor Statistics 2024) for the restaurant industry in Texas suggest that the average annual salary of an employee is about \$29,730, or \$571.73 per week. We use this number as the cost of one additional employee. Second, we assume that in the case of underpredicting visits, a restaurant would need to respond by turning customers away, thus losing the associated revenues (i.e., we abstract from long-term impacts associated with turning customers away due to underpredicting visits). We combine data from Data Axle and SafeGraph to construct a mean ticket size of about \$25.00 per visit.

week), while lower under-prediction results in lower lost sales of about \$2,875, resulting in expected weekly gains of about \$1,672, which is approximately 5% of average annual sales for the median restaurant in our sample.[16] These findings show that incorporating geo-tracking data not only reduces average prediction error but also mitigates the financial consequences of both over-staffing and lost sales due to inaccurate demand forecasts. While these estimates are back-of-the-envelope and assume flexible staffing and uniform customer value, they suggest that improving predictive accuracy can translate into tangible financial benefits for firms.

Finally, in Figure 2, we report the deciles of the distribution of mean absolute error (MAE) separately for cases of under- and over-prediction. This figure shows that Model C consistently yields smaller prediction errors than Models A and B across the entire distribution, meaning that the model that incorporates geo-tracking data over- and under- predicts by less than the competing models, thus being more accurate when predicting the first stage deviation in both instances.

Fig. 2 Distribution of the Prediction Error With and Without Geo-Tracking Data



Notes: The plots report the prediction error for each model as a function of the deciles of the distributions of the prediction error for that model, separately for cases of over- and under- prediction. Model A uses only demographic data as input, Model B uses demographic and behavioral data, and Model C uses demographic, behavioral, and geo-tracking data. The sample consists of all observations in the test data across bootstrap replications. Deciles are computed for each model's prediction errors separately.

4 Simulation Exercises: Results Under Restricted Geo-Tracking

In this section, we report our findings related to our second research question: how does restricting geo-tracking data under different types of privacy regulations impact the

[16] Over- and under-predictions occur equally in our data, so the weekly expected cost can be computed as the average of both. The gains from reduced over-prediction are about \$469 per week (\$0.82 × \$29,730 ÷ 52 ≈ \$469). The gains from reduced under-prediction are given by the product of the number of visits and the ticket size (115 × \$25 per visit i.e., \$2,875). The total expected weekly gain is approximately (\$469 + \$2,875) ÷ 2 i.e., \$1,672.

predictive performance of models that use these data? This question is important from a policy perspective because of the recent emergence of regulations restricting data tracking (Klosowski 2021; Federal Trade Commission 2024). If policymakers regulate consumer geo-tracking by requiring, for example, data summarization so firms cannot access coordinate-level data, or by imposing geographical restrictions on where users are geo-tracked, what would the prediction performance be for models that use these restricted data relative to models that use unrestricted geo-tracking data?

To answer this question, we outline an approach to quantify the extent to which restricting geo-tracking under alternative privacy regulations impacts predictive performance. In Table A2, we motivated three types of regulatory restrictions and how they can be applied to geo-tracking data. Next, we describe how we simulate each of these regulations in our setting and discuss how they impact predictions from our models relative to unrestricted geo-tracking.

User-level summarization. In the first category of simulations, we consider regulations requiring data to be anonymized with respect to the user that generates that data through summarization. We implement a version of this following Pappalardo and Simini (2018) and construct driving summary features for each user-week. These summary features contain aggregated information about a user's driving behavior in a week, such as the total distance traveled, entropy (i.e., variability of the locations visited relative to their past distribution of visited locations), time of driving, number of days driven, and number of trips each week. These features are independent of the target restaurant's location and capture general mobility patterns about a user in a privacy-preserving way (e.g., see consumer surveys in Web Appendix B). The technical details of computing these features and their summary statistics appear in Web Appendix F. Under this simulation, we train a model with the demographic and behavioral information, and these summary features. Unlike Model C that uses geo-tracking features, this model uses aggregated driving summaries.

Geographical restrictions. In the second category of simulations, we explore how geographical restrictions impact prediction outcomes. We implement this simulation using geofences that restrict firms to observe only users who entered a certain radius (e.g., one mile) of their location. Any users that are outside this distance are, therefore, not observable to the firms and are excluded from our simulated prediction model. Under this simulation, we train Model C using geo-tracking data from only those users who entered the geofence.

Frequency restrictions. In the third category of simulations, we consider regulations that restrict *how often* users may be tracked but still allow firms to use the data at the coordinate level. In practice, we consider regulations that record the data less often than what our focal app does. We implement two versions of this restriction.

In the first version, we keep the first point of each trip but systematically drop the data within a trip. Specifically, we re-construct the geo-tracking data at lower frequencies, assuming they are collected at one-half frequency of the original three-minute interval with which our data provider records the data. These exercises are meant to represent choices firms might make about temporal granularity when deciding how often to record data.

In the second version, we implement simulations that reduce geo-tracking frequency in ways that replicate static geo-tracking regulations. These exercises represent scenarios in which firms may track users at specific points of interest, e.g., at the start and end of their trips at places of business, recreation, and so on. In practice, we implement this simulation by keeping only the first and last points of a trip and dropping all other records.

Under each simulation exercise, we use the restricted geo-tracking data for both training and test purposes. Even though we implement three types of privacy restrictions based on Table A2, it is important to note that these restrictions are qualitatively different in how and how much they may protect consumer privacy. While some restrictions completely prevent user identification, others may simply hide some features of their data. Next, we discuss the results of our simulations.

4.1 Results: Predictive Performance under Varying Restrictions on Geo-Tracking

Table 4 presents our findings for the simulation analysis. In each simulation, we report the results for Model C after re-training and evaluating it using restricted geo-tracking data.

Table 4 Results: Simulation Exercises with Varying Restrictions on Geo-Tracking

Simulation	Mean MAE	Difference	Percentage
Complete geo-tracking (Model C in Table 3)	37.249 [37.093, 37.405]		
User-level summarization			
Summary features instead of geo-tracking	42.381 [42.226, 42.536]	5.132 [4.911, 5.353]	13.778%
Geographical restrictions			
Geofenced users within 1 mile of restaurant	39.304 [39.158, 39.304]	2.055 [1.972, 2.138]	5.517%
Frequency restrictions			
Reduced frequency of geo-tracking			
1/2 frequency	38.424 [38.259, 38.589]	1.174 [1.221, 1.290]	3.152%
Static geo-tracking			
First- and last- trip points only	38.519 [38.364, 38.674]	1.270 [1.216, 1.324]	3.410%

Notes: MAE = Mean absolute error. The complete geo-tracking model includes demographics, behavioral, and geo-tracking data (i.e., Model C). This table reports the mean and bootstrapped confidence intervals (in square brackets) of the MAE of each model and the difference in MAE between the complete geo-tracking model and each simulation using the test data. Results correspond to Random Forest models.

Our main finding in this section is that all the restrictions that we evaluate result in a lower predictive performance than that of Model C with complete geo-tracking. However, the decrease in performance varies by the type of restriction. Importantly, we find that even under various policy restrictions, models with geo-tracking information generally perform better than those that do not use geo-tracking information at all.

Next, we describe the results in detail. First, we find that regulations that use summaries of geo-tracking data at the user level result in the largest decrease in predictive performance relative to Model C. Specifically, the MAE is 13.78% larger than the one in Table 3.

Second, we find that regulations that restrict firms to observe data only for users who were within a mile of their locations result in a 5.52% decrease in predictive performance relative to Model C with complete geo-tracking data. This simulation shows that though geofences impose a significant restriction to firms that collect geo-tracking data, the restricted data are still useful when predicting customer visits relative to a context in which these data are not available at all or if they are restricted in other ways (e.g., limiting *what* data are tracked).

Third, we consider simulations that restrict the frequency with which geo-tracking data are recorded, including reduced frequency and static geo-tracking at specific times only. We find that using reduced-frequency data decreases the predictive performance by 3.15% under one-half tracking frequency, while using static geo-tracking decreases the predictive performance by around 3.41% relative to complete geo-tracking.

Overall, our findings show that regulations that restrict the data that can be used to predict total visits reduce the performance of the model relative to when unrestricted geo-tracking data are used. Though the direction of effects is somewhat expected, the relative magnitude of prediction losses provides us with rich insights and shows that not all regulatory restrictions are equal in their implications for prediction. Specifically, we find that the largest reductions in predictive performance are associated with restrictions that fully transform the data. For example, user-level summarization, which seeks to prevent user identification, results in the largest decreases in predictive performance. On the other extreme, frequency restrictions, which retain coordinates in their raw form, reduce predictive performance the least. Between these extremes, we consider a number of regulations that vary in how they restrict geo-tracking data.

5 Robustness Checks

We conduct a series of robustness checks to confirm the validity of our main results. First, we test the sensitivity of our findings to alternative restaurant samples. Second, we assess alternative model setups, including predicting total visits instead of deviations and estimating a pooled model across restaurant-weeks. Third, we validate that our results are not driven by the specific machine learning algorithm by replicating our analyses using Boosted Regression Trees and XGBoost. We also consider alternative accuracy metrics (RMSE, MedAE). Our main findings are robust to all these checks. Full details and tables for each of these analyses are reported in Web Appendix Tables D1-D7.

6 Conclusion

In recent years, many businesses have begun using geo-tracking data to inform marketing and operational decisions. However, these practices have raised privacy concerns among users and regulators. Companies like Tim Hortons have faced public scrutiny for their use of geo-tracking, and emerging privacy regulations, such as the California Privacy Rights Act, classify consumer location data as sensitive and restrict its use.

This research investigates the value of geo-tracking data in predicting visits to business locations and how privacy regulations impact the accuracy of these predictions. Using data from a Texas-based safe-driving app, we applied machine learning to predict restaurant visits one week ahead. Our results show that using geo-tracking data improves the out-of-sample predictive accuracy by 10.76% compared to using only demographic and behavioral data, and by 31.66% compared to using demographic data alone. However, privacy restrictions on geo-tracking reduce prediction performance by 3-14%, depending on the type of regulation. The largest decline in accuracy (13.78%) occurs when restrictions limit the type of data collected, such as summarizing driving behaviors. Despite these reductions, models with restricted geo-tracking generally outperform models that do not use geo-tracking data at all.

Our findings offer several practical implications for firms and policymakers. First, our results show that geo-tracking data substantially improve visit predictions, helping firms better plan staffing and inventory and potentially yielding savings of about 5% of median revenues. Second, the improved accuracy from these data reduces both over- and under-prediction, offering firms a clearer basis to weigh the costs and benefits of using geo-tracking. Third, our simulation exercises suggest that firms can adopt privacy-preserving practices, such as limiting where or how frequently data are tracked, without large losses in predictive value. This has practical implications for firms purchasing geo-tracking data or seeking to reduce data storage costs. Finally, our framework can also aid regulators by quantifying the predictive trade-offs of different privacy restrictions.

Declarations

- **Funding:** The authors have no external funding to report. The project was supported by an internal research grant from the institution of one of the authors.
- **Conflict of interest/Competing interests:** All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. No funding was received to assist with this manuscript.

- **Consent for publication:** Both authors provide their consent.
- **Data availability:** The geo-tracking data used in this research are from a proprietary source. The nature of the non-disclosure agreement with the app precludes us from sharing the app data.
- **Materials availability:** The authors will provide simulated data to replicate the results.
- **Code availability:** The authors will provide the code to replicate the results. Both authors contributed to the project.

References

Allcott H, Boxell L, Conway JC, et al (2020) What Explains Temporal and Geographic Variation in the Early US Coronavirus Pandemic? National Bureau of Economic Research Working Paper, accessed July 2025

Aridor G, Che YK, Hollenbeck B, et al (2025) Evaluating the impact of privacy regulation on e-commerce firms: Evidence from apple's app tracking transparency. Management Science Forthcoming

Athey S, Catalini C, Tucker C (2017) The digital privacy paradox: Small money, small costs, small talk. Tech. rep., National Bureau of Economic Research

Athey S, Blei D, Donnelly R, et al (2018) Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In: AEA Papers and Proceedings, pp 64–67

Austen I (2022) 'A Mass Invasion of Privacy' but No Penalties for Tim Hortons. http://tinyurl.com/33x6ujz5, new York Times, accessed June 11, 2022

Berman R, Israeli A (2022) The value of descriptive analytics: Evidence from online retailers. Marketing Science 41(6):1074–1096

Bertrand M, Kamenica E, Pan J (2015) Gender identity and relative income within households. The Quarterly Journal of Economics 130(2):571–614

Bhargava HK, Rubel O, Altman EJ, et al (2020) Platform data strategy. Marketing Letters 31:323–334

Bleier A, Goldfarb A, Tucker C (2020) Consumer privacy and the future of data-based innovation and marketing. International Journal of Research in Marketing 37(3):466–480

Chen MK, Pope DG (2020) Geographic Mobility in America: Evidence from Cell Phone Data. National Bureau of Economic Research Working Paper, accessed July 2025

Chen MK, Rohla R (2018) The effect of partisanship and political advertising on close family ties. Science 360(6392):1020–1024

Chetty R, Hendren N, Katz LF (2016) The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. American Economic Review 106(4):855–902

Chiou L, Tucker C (2020) Social Distancing, Internet Access and Inequality. National Bureau of Economic Research Working Paper, accessed July 2025

Choi WJ, Jerath K, Sarvary M (2023) Consumer privacy choices and (un) targeted advertising along the purchase journey. Journal of Marketing Research 60(5):889–907

Congressional Research Service (2024) The American Privacy Rights Act of 2024. Congressional Research Service Legal Sidebar, https://tinyurl.com/b4a93h3w

Coston A, Guha N, Ouyang D, et al (2021) Leveraging administrative data for bias audits: Assessing disparate coverage with mobility data for covid-19 policy. In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, pp 173–184

Danaher PJ, Smith MS, Ranasinghe K, et al (2015) Where, when, and how long: Factors that influence the redemption of mobile phone coupons. Journal of Marketing Research 52(5):710–725

Dean G (2023) Mcdonald's and Chick-fil-A are tracking your location to make sure your fries are perfectly crispy when you come collect your mobile order. Business Insider http://tinyurl.com/4zkr622h

Dubé JP, Fang Z, Fong N, et al (2017) Competitive price targeting with smartphone coupons. Marketing Science 36(6):944–975

Fang L (2022) The effects of online review platforms on restaurant revenue, consumer learning, and welfare. Management Science 68(11):8116–8143

Federal Trade Commission (2022) FTC Sues Kochava for Selling Data that Tracks People at Reproductive Health Clinics, Places of Worship, and Other Sensitive Locations. http://tinyurl.com/pwbedwke, accessed August 22, 2022

Federal Trade Commission (2024) FTC Order Will Ban InMarket from Selling Precise Consumer Location Data. https://tinyurl.com/3mh4nzw7, accessed January 18, 2024

Goldfarb A, Tucker C (2012) Shifts in privacy concerns. The American Economic Review 102(3):349–353

Goldfarb A, Tucker C (2020) Which retail outlets generate the most physical interactions? Working Paper

Gonzalez MC, Hidalgo CA, Barabasi AL (2008) Understanding individual human mobility patterns. Nature 453(7196):779–782

Hermalin BE, Katz ML (2006) Privacy, property rights and efficiency: The economics of privacy as secrecy. Quantitative Marketing and Economics 4:209–239

Hoteit S, Secci S, Sobolevsky S, et al (2014) Estimating human trajectories and hotspots through mobile phone data. Computer Networks 64:296–307

Jerath K, Miller KM (2024) Using the dual-privacy framework to understand consumers' perceived privacy violations under different firm practices in online advertising. Working Paper

Johnson G, Lin T, Cooper JC, et al (2023a) Coppacalypse? The Youtube settlement's impact on kids content. Working Paper

Johnson G, Shriver S, Goldberg S (2023b) Privacy & market concentration: Intended & unintended consequences of the GDPR. Management Science 69(10):5695–6415

Katona Z, Painter MO, Patatoukas PN, et al (2025) On the capital market consequences of big data: Evidence from outer space. Journal of Financial and Quantitative Analysis 60(2):551–579

Kendall S (2021) Reasons why data scientists choose safegraph data. https://tinyurl.com/tx62bdr5, accessed July 2025

Kim M, Bradlow ET, Iyengar R (2022) Selecting data granularity and model specification using the scaled power likelihood with multiple weights. Marketing Science 41(4):848–866

Klopack B (2024) One size fits all? The value of standardized retail chains. RAND Journal of Economics 55(1):55–86

Klopack B, Luco F (2025) Measuring Local Consumption with Payment Cards and Cell Phone Pings. https://flucoe.github.io/pdfs/Klopack_Luco_2025.pdf, working paper, accessed July 2025

Klopack B, Lewis E, Luco F (2024) Economic Consequences of Natural Disasters: The Impact of Hurricane Harvey on Local Retail and Consumer Welfare. Working Paper, accessed July 2025

Klosowski T (2021) The State of Consumer Data Privacy Laws in the US (And Why it Matters). http://tinyurl.com/ypjhssxj, new York Times, September 6, 2021

Korganbekova M, Zuber C (2023) Balancing user privacy and personalization. Working Paper

Lee JY, Yang J, Anderson ET (2025) Using Grocery Data for Credit Decisions. Management Science 71(4):2751–3636

Lei X, Chen Y, Sen A (2023) The Value of External Data for Digital Platforms: Evidence from a Field Experiment on Search Suggestions. Working Paper, accessed July 2025

Li H, Srinivasan K (2019) Competitive dynamics in the sharing economy: An analysis in the context of airbnb and hotels. Marketing Science 38(3):365–391

Li S, Schneider MJ, Yu Y, et al (2023) Reidentification risk in panel data: Protecting for k-anonymity. Information Systems Research 34(3):811–1319

Li Z, Ning H, Jing F, et al (2024) Understanding the bias of mobile location data across spatial scales and over time: a comprehensive analysis of safegraph data in the united states. Plos one 19(1):e0294430

Lin T, Strulov-Shlain A (2024) Choice architecture, privacy valuations, and selection bias in consumer data. Marketing Science Forthcoming

Van der Maaten L, Hinton G (2008) Visualizing data using t-sne. Journal of Machine Learning Research 9(11):2579–2605

Macha M, Foutz NZ, Li B, et al (2024) Personalized privacy preservation in consumer mobile trajectories. Information Systems Research 35(1):249–271

Miller KM, Skiera B (2023) Economic loss of cookie lifetime restrictions. International Journal of Research in Marketing 41(2):241–264

Naik N, Raskar R, Hidalgo CA (2016) Cities are physical too: Using computer vision to measure the quality and impact of urban appearance. American Economic Review 106(5):128–32

Neumann N, Tucker CE, Subramanyam K, et al (2023) Is first-or third-party audience data more effective for reaching the 'right'customers? the case of it decision-makers. Quantitative Marketing and Economics 21(4):519–571

Oblander S, McCarthy D (2023) Frontiers: Estimating the long-term impact of major events on consumption patterns: Evidence from covid-19. Marketing Science 42(5):839–1028

Pappalardo L, Simini F (2018) Data-driven generation of spatio-temporal routines in human mobility. Data Mining and Knowledge Discovery 32(3):787–829

Puranam D, Kadiyali V, Narayan V (2021) The impact of increase in minimum wages on consumer perceptions of service: A transformer model of online restaurant reviews. Marketing Science 40(5):985–1004

Rafieian O, Yoganarasimhan H (2021) Targeting and privacy in mobile advertising. Marketing Science 40(2):193–218

Shoshani T, Zubcsek PP, Reichman S (2024) Predicting Store Closures Using Urban Mobility Data and Network Analysis. Working Paper, accessed July 2025

Smith HJ, Milberg SJ, Burke SJ (1996) Information privacy: Measuring individuals' concerns about organizational practices. MIS Quarterly pp 167–196

Song C, Qu Z, Blumm N, et al (2010) Limits of predictability in human mobility. Science 327(5968):1018–1021

Sun C, Adamopoulos P, Ghose A, et al (2022) Predicting stages in omnichannel path to purchase: A deep learning model. Information Systems Research 33(2):429–445

Surange S, Shankar V (2024) Segmenting Consumers Based on their Mobility (Driving Patterns) and Predicting Store Vicinity Visit and Coupon Redemption. Working Paper, accessed July 2025

Tau B (2023) Selling Your Cellphone Location Data Might Soon Be Banned in U.S. for First Time. http://tinyurl.com/bdsa8khx, wall Street Journal, accessed July 10, 2023

Trusov M, Ma L, Jamal Z (2016) Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting. Marketing Science 35(3):405–426

U.S. Bureau of Labor Statistics (2024) Occupational employment and wages. URL https://www.bls.gov/oes/tables.htm, u.S. Department of Labor, Occupational Employment and Wage Statistics

Varga M, Tusche A, Albuquerque P, et al (2024) Does fMRI Data Improve Predictions of Product Adoption by Store Managers and Sales to Consumers for Consumer-Packaged Goods? Working Paper, accessed July 2025

Wang Y, Wu C, Zhu T (2019) Mobile hailing technology and taxi driving behaviors. Marketing Science 38(5):734–755

Wernerfelt N, Tuchman A, Shapiro B, et al (2024) Estimating the value of offsite data to advertisers on meta. Marketing Science Forthcoming

Willmott CJ, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research 30(1):79–82

Yoganarasimhan H (2020) Search personalization using machine learning. Management Science 66(3):1045–1070

Zubcsek PP, Katona Z, Sarvary M (2017) Predicting mobile advertising response using consumer colocation networks. Journal of Marketing 81(4):109–126

Appendix A Interview of Restaurant Managers

Table A1 Managerial Responses: How Predicting Consumer Visits May be Useful for Making Decisions

Manager Role	Experience	Gist of Interview Responses
Manager Jimmy John's (JJ)	5 years	Predicting consumer visits every week would help us with staffing. We can't fill staffing gaps immediately, so we receive applications weekly and look at their availability for the week. Typically, if we over or under predict visits by 20-30 customers, we need to hire or fire staff. Underpredicting costs us not only in lost potential revenue but also lost customers we can't serve. That's $25 per person check value gone. Customer demographics are important for us when planning. But I don't prefer to keep customer data for more than a few weeks or so at most.
Manager Asian restaurant in Chicago	2-3 years	We collect demand data and customer visits. A prediction algorithm that can tell us how likely a customer is to visit us each week would be helpful. This would help us to know who is coming and how often, which would allow us to plan and manage our resources effectively. Knowing what kind of food our customers like based on their past visits and/or to other restaurants also helps us make better decisions.
Manager and Chef Culvers' and Red Lobster	3+ years	It [predicting customer visits] would reduce costs greatly. If you can predict how many people are likely to visit and maybe even which meals they will need, you can plan which ingredients to order, or how much staff to hire. In my experience, you'd have half of the staff fulltime and half of the staff part-time. Schedules get set typically on Fri/Sat for the following week. Knowing how many customers to expect the following week can reduce cost, increase profits, and help manage all of those without waste. It would also be useful to see how many are local vs. out-of-town customers. Are they mostly going to be coming in for mornings or afternoons, then you could have specials around those meals and times instead of all day deals. Or even a special deal of the week.
Food Administrator University Dining	27+ years	If a prediction algorithm used data specific to my restaurant to make predictions about my consumers' visits, I would use it to make decisions. It would be important to ensure that the algorithm considers the demographics and needs of our customers. However, I would not rely solely on the algorithm. In terms of frequency, I would like to predict consumer visits, weekly. This would help tailor our offerings, staff, and marketing to meet customer needs. We'd hire more if we get 20 or so unplanned visits.

Manager Role	Experience	Gist of Interview Responses
Owner, Fine-dining restaurants in Atlanta	20+ years	I'd love to see the data on who comes back to our restaurant for repeat service. If I knew how many customers (and who) is coming, I can prepare my restaurant and work towards satisfying them, e.g., if I have mostly female or non-binary customers, or where they come from, maybe their relationship status. More detailed information about our customers is incredibly useful when it comes to satisfying them and growing our business. However, we also need to be careful not to overload ourselves by trying to predict customer visits every single day or by collecting too much data about them. It might be better to focus on weekly predictions, so that we can plan and make sure we are ready to provide the best possible experience for our customers. It can also help us plan ingredients, staffing, and store locations.
Manager Chilli's	18 months	The data I'd really be interested in is repeat customers. POS systems don't really track that and not everyone has rewards programs (or joins one). It would also be useful to know how many people are likely to come in and overlap that with community events they plan to ahead. Demographics about customers can also help look at trends about who they are, whether they are visiting if they are more proximate, how far they'd travel for a restaurant etc. Weekly or more regular data could also help especially for scheduling staff. If they can come and go quickly, it will improve the customer experience.
Hostess at a restaurant in Chicago	2 years	A prediction algorithm that can tell me how likely a customer is to visit my restaurant would be very helpful. I could use this information to adjust staffing schedules, move tables around, and plan weekly specials. I could also use it to determine when to promote happy hour specials and other promotions on social media. I would like to be able to predict consumer visits on a weekly basis. This is because most decisions, including staffing, scheduling, and social media marketing as well as supplies are planned weekly.
Manager, Crosby's Kitchen in Chicago	15+ years	The most important decisions I make are hiring and customer service. Predicting consumer visits can help because the problem is that sometimes it's slow and sometimes it's busy. I don't like to over- or under-staff. So I'd like to know the predictions a week in advance. Sometimes also a month in advance because people I work with have lives outside the restaurant, but usually a week.

Notes: The interviews were conducted after securing Institutional Review Boards (IRB) approval. Managers were recruited through a research database at a large public university. The only inclusion criterion was some experience in the restaurant industry in the U.S. Each interview lasted about 30 minutes via zoom.

Table A2 Overview of Privacy Restrictions and their Related Geo-Tracking Simulations

Type of restriction	**Motivation**	**Potential simulation**	**Example research**
What user data are tracked	GDPR protects user data that contain personally identifiable information. Google's Sandbox technology proposes anonymizing user browsing data within the Chrome browser instead of collecting raw browsing data using cookies. APRA classifies precise geolocation data as sensitive.	User-level summarization: Use summaries of driving behaviors extracted from geo-tracking data.	Aridor et al. (2025); Johnson et al. (2023b); Miller and Skiera (2023)
Where users are tracked	Google allows ad targeting depending on a country's laws. Many data vendors only sell consumer data tracked within specific geofences.	Geographical restrictions: Geo-track users only if they were within certain distances of a location (e.g., one mile).	Danaher et al. (2015); Dubé et al. (2017)
How frequently users are tracked	Regulatory actions discourage the sale of consumers' raw location data streams and could deter firms from high-frequency tracking. APRA and FTC actions motivate reduced frequency. Firms may elect to track lower-frequency data or data at only static locations to optimize storage costs.	Frequency restrictions: Geo-track trips at lower frequency intervals, at the start and end points of a trip, or for a randomly selected trip per week instead of constantly tracking all trips.	Kim et al. (2022); Trusov et al. (2016)

Notes: GDPR = General Data Protection Regulation. APRA = American Privacy Rights Act. GDPR is passed law in the EU; APRA is a proposed U.S. federal bill.

Appendix B Consumer Surveys

In this section, we report the results of a survey we conducted to ask consumers about their privacy perceptions towards geo-tracking. While our privacy restrictions under the simulation exercises are primarily motivated by the current regulatory environment, industry practices, and recommendations from the data obfuscation literature, we further wanted to validate that the simulation exercises are privacy preserving from the perspective of end consumers who are subject to such tracking.

We recruited survey participants through a research database at a large public university in the U.S. after securing IRB approvals. The survey was administered using a Qualtrics link. Participation in the survey was voluntary. Participants were offered a Target gift card worth five dollars for completing the survey. Upon accessing Qualtrics, the participants were presented the following information:

"Imagine you are using an app on your phone that gives you points that you can redeem at local restaurants and businesses. Read the scenarios below about the kind of data the app tracks from your usage, then indicate the extent to which you agree with each statement that follows: <***insert scenario***>

We presented the scenarios listed in Table B1 in a random order. After each scenario was presented, we asked survey participants to indicate their privacy perceptions for that scenario across five dimensions adapted from the privacy literature (e.g., Smith et al. 1996).

Table B1 Scenarios

Text of Scenarios in the Survey
Complete tracking: The app collects data about your geo-location constantly after it has detected that driving has started.
Summary features: The app collects data about your driving summary (e.g., total distance you drive, time of day when you drive) rather than geo-location coordinates.
Geofence: The app collects data if you were near a local restaurant.
Frequency: The app collects data about your geo-location coordinates with a frequency of every half hour (i.e., 30 minutes) after it has detected that driving has started.
Static first- and last- points: The app collects data about your geo-location coordinates at the start and end of your trip.
Static random trip per week: The app collects data about your geo-location coordinates for one trip at random each week.

The privacy dimensions were:

1. It bothers me that the app collects these data.
2. I'm concerned that the app is collecting too much information about me.
3. I am concerned about my privacy and how the app might use my data.
4. It bothers me to give my information to this app.
5. I will stop using this app in the future.

Participants were asked to rank each of the five statements for each scenario they were presented on a scale of 1 (strongly disagree) to 5 (strongly agree).

B.1 Results of the Survey

Overall, 191 participants completed the survey. Of the participants, 60% were female. The average age was 44 years. Across the five privacy measures, participants' privacy concerns under complete geo-tracking had an average score of 4.18 out of 5. Relative to complete geo-tracking, user-level summarization had an average score of 3.17 and geographic restrictions had an average score of 3.22. Among frequency restrictions, the half-hour interval of tracking had an average score of 3.73, the static tracking of first- and last-trip points only had an average score of 3.77, and one trip per week at random had an average score of 3.46. We report the mean rating for each question and averages across the five questions in the survey for each scenario in Web Appendix Table B2.[17]

Table B2 Results: Mean Rating for Each Survey Question and Overall

Simulation	Q1	Q2	Q3	Q4	Q5	Mean
Complete geo-tracking	4.20	4.27	4.32	4.24	3.88	4.18
User-level summarization						
Summary features instead of geo-tracking	3.11	3.18	3.43	3.19	2.92	3.17
Geographical restrictions						
Geofenced users when near restaurant	3.21	3.20	3.45	3.25	2.99	3.22
Frequency restrictions						
Reduced frequency of geo-tracking						
Reduced frequency	3.71	3.78	3.93	3.76	3.49	3.73
Static geo-tracking						
First- and last- trip points only	3.78	3.87	4.04	3.81	3.34	3.77
One trip per week at random only	3.46	3.43	3.74	3.49	3.19	3.46

Notes: N = 191. Survey statements Q1-Q5 were as follows: Q1. It bothers me that the app collects these data. Q2. I'm concerned that the app is collecting too much information about me. Q3. I am concerned about my privacy and how the app might use my data. Q4. It bothers me to give my information to this app. Q5. I will stop using this app in the future. Participants were asked to rank each statement from 1 (strongly disagree) to 5 (strongly agree). Reduced frequency refers to half hour.

[17] The results presented here should be interpreted with caution. As highlighted by Athey et al. 2017, survey-based measures of privacy attitudes are often highly sensitive to question framing and context. Our intention is not to draw strong empirical conclusions from these responses, but rather to provide a qualitative sense of how individuals might perceive geo-tracking data in a commercial context.

Appendix C Categorization of Restaurants

Table C1 Categories of Restaurants using Yelp API

	Categories	Description	Popular tags
1	American	Restaurants serving American cuisine, but excluding restaurants specializing in burgers and sandwiches, and excluding restaurants that were also tagged as another type.	American (Traditional), American (new), breakfast, brunch, chicken wings, diners
2	Asian	Restaurants specializing in cuisines from South Asian, East Asian, and Southeast Asian countries, as well as Pacific islands.	Chinese, Japanese, sushi bars, Asian fusion, Thai, Indian, Hawaiian
3	Burgers	Restaurants with tag “Burgers”.	Burgers, hot dogs, sports bars, steakhouses
4	Coffee	Restaurants with tag “Coffee”.	Tea, coffee, café
5	Dessert	Restaurants with tag “Dessert”.	Dessert, frozen yogurt
6	European	Restaurants specializing in Italian, French, or other European cuisines, except for restaurants also tagged “Pizza”.	Italian, French, Irish, Wine Bars, Noodles, Mediterranean
7	Latin American	Restaurants specializing in cuisines from South and Central America and the Caribbean.	Mexicana, Tex-Mex, Latin American, seafood, Caribbean, Cuban
8	Pizza	Restaurants with tag “Pizza”.	Pizza, Italian, salad
9	Sandwiches	Restaurants with tag “Sandwiches”, “Deli”, or “Cheesesteaks”.	Sandwiches, deli, cheesesteaks
10	Other	Restaurants not tagged as any of the above categories.	

Appendix D Robustness Checks and Additional Analyses

In this appendix, we describe our robustness checks and present the results for an alternative sample of restaurants (Table D1), alternative outcome in the form of total visits, rather than deviation from visits (Table D2), alternative model setups (Table D3), alternative ML models (Tables D4-D5), and alternative metrics (Tables D6-D7).

D.1 Alternative Restaurants

In our main analysis, we report the results for a sample of 401 restaurants that at least 10 app users visited. Even though we verified that these 401 restaurants are representative of the broader set of restaurants, it is possible that the results are sensitive to the specific restaurant sample. Thus, in Web Appendix Table D1, we report the results for an alternative sample of restaurants based on more stringent sample selection criteria of at least 25 users visiting those restaurants anytime in our data period. We find qualitatively similar results across all the samples, although the magnitude varies based on the number of visits in each sample.

D.2 Alternative Model Setups

Our main model uses data at the restaurant-week level and makes predictions for the deviations in visits. Alternatively, we test the robustness of our results to different model setups based on alternative target and data splits. First, instead of predicting the difference between realized and expected visits, we predict the total realized number of visits. We report the results in Web Appendix Table D2. Second, we train and report a "pooled" model. Instead of splitting the restaurant-week level data into training and test sets, in this approach, we take each restaurant and split the weeks into training and test, and stack them. To train our prediction model, we also include week fixed effects as predictors. The results appear in Web Appendix Table D3. In both the setups, Model C with geo-tracking data outperforms the other models.

D.3 Alternative Machine Learning Models

While our main models use a Random Forest Regression, we also train alternative ML models, such Boosted Regression Trees and XGBoost, to make sure the results are not unique to the model we use. Since our main interest is in comparing various information sets, we repeat the specifications reported in the main results in Table 3 and report their results for alternative models in Web Appendix Tables D4 and D5. The findings are consistent with those reported in the main analysis.

D.4 Alternative Metrics

Our main results allow us to compare the performance of various information sets using MAE as the performance metric. In addition, we also report the results in our main analysis for alternative performance metrics, i.e., the Root Mean Square Error (RMSE) and the Median Absolute Error (MedAE). The results appear in Web

Appendix Tables D6 and D7. Both sets of results show patterns that are similar to the MAE results in Table 3.

Table D1 Results: Predictive Performance of Random Forest by Information Set for Restaurants with At Least 25 App Users Visiting

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	83.858 [83.296, 84.420]		
Model B: Behavioral data	61.866 [61.447, 62.285]	21.992 [21.683, 22.301]	
Model C: Geo-tracking data	53.111 [52.694, 53.528]	30.747 [30.395, 31.099]	8.755 [8.529, 8.981]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 3,990 for 133 restaurants across 30 prediction weeks. The restaurants are selected based on at least 25 app users visiting in the data period. All models include season fixed effects.

Table D2 Results: Predictive Performance of Random Forest by Information Set for Total Visits

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	215.102 [214.696, 215.508]		
Model B: Behavioral data	39.512 [39.376, 39.648]	175.590 [175.235, 175.945]	
Model C: Geo-tracking data	36.466 [36.328, 36.604]	178.636 [178.286, 178.986]	3.046 [2.978, 3.114]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

Table D3 Results: Predictive Performance of Random Forest by Information Set for a "Pooled" Model within Restaurants

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	53.771 [53.695, 53.847]		
Model B: Behavioral data	39.824 [39.755, 39.893]	13.947 [13.865, 14.029]	
Model C: Geo-tracking data	37.359 [37.287, 37.431]	16.412 [16.327, 16.497]	2.465 [2.408, 2.522]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. Note that all feature sets include controls for seasons as well as a week identifier. Compared to the main analysis with restaurant-week level splits, the data in this analysis is split within restaurant by week i.e., training data contain subset of weeks for a restaurant and the test data contain the remaining weeks for the same restaurant. However, we use the week prior to the prediction week to construct the input features, consistent with the main model.

Table D4 Results: Predictive Performance of Boosted Regression Trees by Information Set

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	52.255 [52.036, 52.474]		
Model B: Behavioral data	48.472 [48.284, 48.660]	3.783 [3.692, 3.874]	
Model C: Geo-tracking data	47.716 [47.532, 47.900]	4.539 [4.446, 4.632]	0.756 [0.712, 0.800]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

Table D5 Results: Predictive Performance of XGBoost Model by Information Set

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic	51.305 [51.080, 51.530]		
Model B: Behavioral data	41.739 [41.578, 41.900]	9.566 [9.436, 9.696]	
Model C: Geo-tracking data	37.250 [37.094, 37.406]	14.055 [13.924, 14.186]	4.489 [4.417, 4.561]

Notes: MAE = Mean absolute error. In column "Mean MAE," the confidence interval corresponds to that of the mean MAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

Table D6 Results: Predictive Performance of Random Forest by Information Set using Root Mean Squared Error (RMSE) Metric

Model	Mean RMSE	Difference from Model A	Difference from Model B
Model A: Demographic data	153.928 [152.527, 155.329]		
Model B: Behavioral data	109.931 [108.743, 111.119]	43.997 [42.885, 45.109]	
Model C: Geo-tracking data	99.408 [98.108, 100.708]	54.520 [53.274, 55.766]	10.523 [9.894, 11.152]

Notes: RMSE = Root mean squared error. In column "Mean RMSE," the confidence interval corresponds to that of the mean RMSE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

Table D7 Results: Predictive Performance of Random Forest by Information Set using Median Absolute Error (MedAE) Metric

Model	MedAE	Difference from Model A	Difference from Model B
Model A: Demographic data	27.464 [27.401, 27.527]		
Model B: Behavioral data	22.093 [22.046, 22.140]	5.371 [5.313, 5.429]	
Model C: Geo-tracking data	20.279 [20.237, 20.321]	7.185 [7.122, 7.248]	1.814 [1.769, 1.859]

Notes: MedAE = Median absolute error. In column "MedAE," the confidence interval corresponds to that of the MedAE for that model. In all other columns, the confidence interval corresponds to that of the mean difference between two models. We implement this using 500 bootstrap replications for each model, as described in the main text. N = 12,030 restaurant-weeks. All models include season fixed effects.

D.5 Representativeness of Safegraph Data

To address the possibility that Safegraph panel may underrepresent minority populations, we examined whether the predictions of our models follow a systematic pattern as a function of local demographics. For example, we examine whether the models' performance increases in income, as it would if low-income communities are underrepresented in the panel. We report these in Figures D1 and D2. We do not find patterns to this effect.

Fig. D1 Model C prediction error by income decile

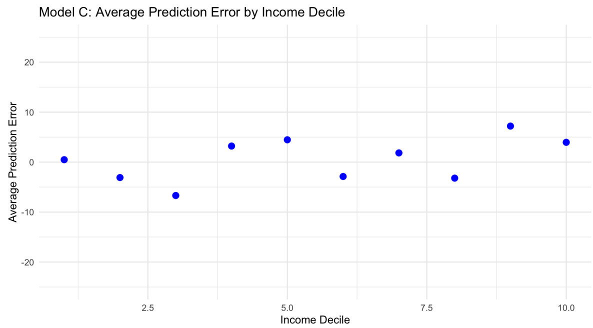
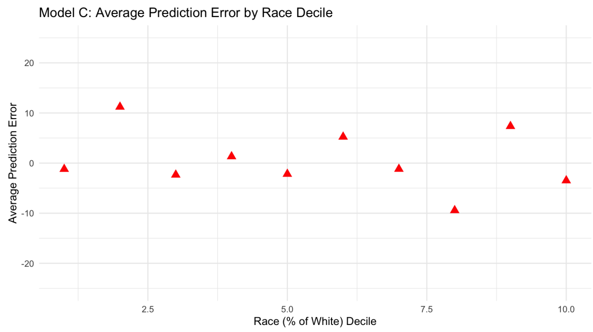


Fig. D2 Model C prediction error by race decile



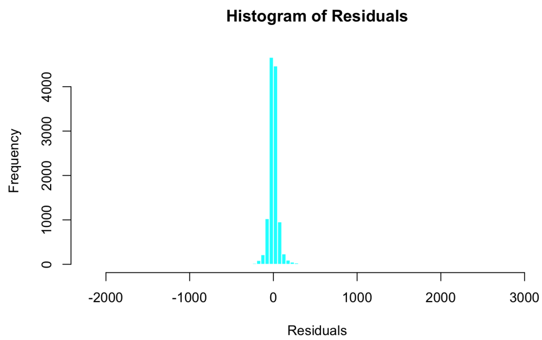
Another potential concern with Safegraph visits data could be measurement error, which is common in location and satellite imagery data of this type (e.g., Katona et al. 2025). Specifically, Safegraph visits may be measured with noise, as these data come from noisy cellphone Global Positioning System (GPS) pings that are overlapped with geometries of store locations to identify visits. When businesses are located close to each other, this may lead to incorrectly assigning visits due to the noisy nature of GPS points, and to missing true visits because the GPS point may be recorded as being outside a business. However, as long as measurement error is independent and identically distributed (iid) and has mean zero, which is a reasonable assumption in the context of GPS data, this leads to overstating the true MAE/RMSE. Informally, measurement issues lead to training the models with data that may misclassify points as visits that did not happen, and may miss visits that did happen. Because of this, both the training and testing stages of our empirical approach should perform worse than if we were to use the true visits data.

D.6 Distribution of Residual Deviations

To compute the deviations, we regress the total number of visits for a specific restaurant in a given week on restaurant and month-of-year fixed effects (i.e., June). The residuals from this regression are then used as the prediction target in the prediction models. This approach helps control for systematic variation across restaurants and seasonal effects before applying the models.

Since the residuals come from a linear model, they are centered at zero. The mean, median, min, max, and standard deviation are respectively 0, -0.82, -2178.35, 3235.29, and 160.33 for the residuals. We also show the histograms of the residuals in Figure D3.

Fig. D3 Distribution of Residuals



D.7 Alternative Metrics: Variable Importance

Fig. D4 Variable Importance Plot for Model C Across Restaurant-Weeks

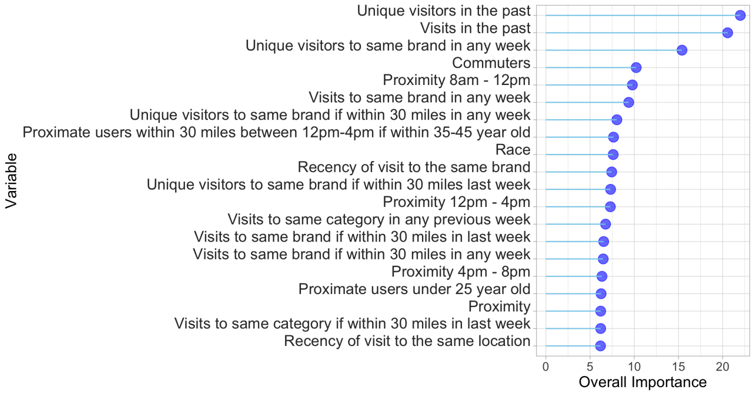
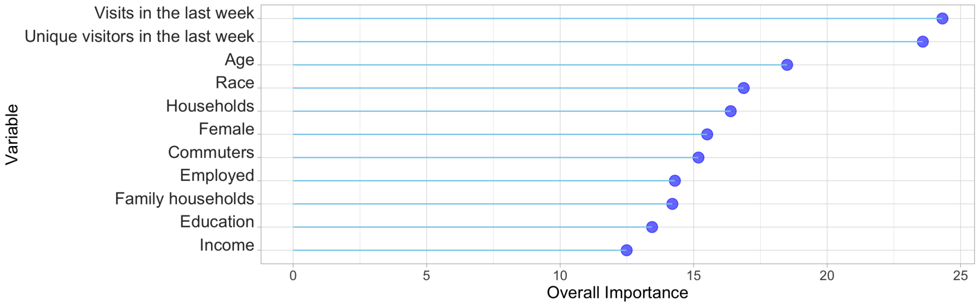


Fig. D5 Variable Importance Plot for Model B Across Restaurant-Weeks



Appendix E Complete Set of Geo-tracking Features

Table E1: Complete Set of Features included in Geo-tracking Data and their Summary Statistics

Feature	**Description**	**Mean**
Proximity	Min. distance (in miles) between trip coordinates and restaurant in past weeks starting the first week in the data until the prediction week	11.14
Proximate users	No. of users within 30 miles of a restaurant starting the first week in the data until the prediction week	4251.00
Past brand visits	No. of visits to same-brand restaurants in the past weeks starting the first week in the data until the prediction week	1231.98
Past category visits	No. of visits to same-category restaurants in the past weeks starting the first week in the data until the prediction week	4815.00
Recency of past visit	Average number of days since last visit to the restaurant across all users	187.80
Recency of past brand visit	Average number of days since last visit to same-brand restaurants across all users	163.42
Recency of past category visit	Average number of days since last visit to same-category restaurants across all users	161.83
Past brand visitors	No. of users visiting same-brand restaurants in the past weeks starting the first week in the data until the prediction week	445.30
Past category visitors	No. of users visiting same-category restaurants in the past weeks starting the first week in the data until the prediction week	1972.71
Previous week visitors	No. of users visiting the restaurants in the previous week	1.06
Previous week visits	No. of visits to the restaurants in the previous week	0.798
Previous week brand visitors	No. of users visiting same-brand restaurants in the previous week	4.68
Previous week brand visits	No. of visits to same-brand restaurants in the previous week	6.86
Previous week category visitors	No. of users visiting same-category restaurants in the previous week	19.25
Previous week category visits	No. of visits to same-category restaurants in the previous week	27.41
Proximity 12am - 4am	Mean proximity for trips between 12 am and 4 am	11.01

Continued on next page

Feature	Description	Mean
Proximity (Q1) 12am - 4am	First quartile of proximity for trips between 12 am and 4 am	5.72
Proximity (Q2) 12am - 4am	Second quartile of proximity for trips between 12 am and 4 am	11.28
Proximity (Q3) 12am - 4am	Third quartile of proximity for trips between 12 am and 4 am	18.82
Proximity 4am - 8am	Mean proximity for trips between 4 am and 8 am	10.67
Proximity (Q1) 4am - 8am	First quartile of proximity for trips between 4 am and 8 am	5.73
Proximity (Q2) 4am - 8am	Second quartile of proximity for trips between 4 am and 8 am	11.31
Proximity (Q3) 4am - 8am	Third quartile of proximity for trips between 4 am and 8 am	18.57
Proximity 8am - 12pm	Mean proximity for trips between 8 am and 12 pm	10.86
Proximity (Q1) 8am - 12pm	First quartile of proximity for trips between 8 am and 12 pm	5.56
Proximity (Q2) 8am - 12pm	Second quartile of proximity for trips between 8 am and 12 pm	11.24
Proximity (Q3) 8am - 12pm	Third quartile of proximity for trips between 8 am and 12 pm	18.26
Proximity 12pm - 4pm	Mean proximity for trips between 12 pm and 4 pm	10.78
Proximity (Q1) 12pm - 4pm	First quartile of proximity for trips between 12 pm and 4 pm	5.57
Proximity (Q2) 12pm - 4pm	Second quartile of proximity for trips between 12 pm and 4 pm	11.01
Proximity (Q3) 12pm - 4pm	Third quartile of proximity for trips between 12 pm and 4 pm	17.82
Proximity 4pm - 8pm	Mean proximity for trips between 4 pm and 8 pm	11.03
Proximity (Q1) 4pm - 8pm	First quartile of proximity for trips between 4 pm and 8 pm	5.74
Proximity (Q2) 4pm - 8pm	Second quartile of proximity for trips between 4 pm and 8 pm	11.23
Proximity (Q3) 4pm - 8pm	Third quartile of proximity for trips between 4 pm and 8 pm	19.41
Proximity 8pm - 12am	Mean proximity for trips between 8 pm and 12 am	10.98
Proximity (Q1) 8pm - 12am	First quartile of proximity for trips between 8 pm and 12 am	5.50
Proximity (Q2) 8pm - 12am	Second quartile of proximity for trips between 8 pm and 12 am	10.90
Proximity (Q3) 8pm - 12am	Third quartile of proximity for trips between 8 pm and 12 am	18.21

Continued on next page

Feature	Description	Mean
Proximate users female	Total proximate female users across all time periods	1953.51
Proximate users age <25	No. of unique users driving within 30 miles of the restaurant in the previous week of under age 25	1541.68
Proximate users age 25–35	No. of unique users driving within 30 miles of the restaurant in the previous week of ages 25-35	1374.83
Proximate users age 35–45	No. of unique users driving within 30 miles of the restaurant in the previous week of ages 35-34	577.38
Proximate users age 45–55	No. of unique users driving within 30 miles of the restaurant in the previous week of ages 45-55	464.80
Proximate users age >55	No. of unique users driving within 30 miles of the restaurant in the previous week over age 55	292.02
Proximate users 4am - 8am	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am	1446.34
Proximate users 4am - 8am female	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am for female users	632.44
Proximate users 4am - 8am age 25–35	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am of ages 25-35	471.48
Proximate users 4am - 8am age 35–45	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am of ages 35-45	139.49
Proximate users 4am - 8am age 45–55	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am of ages 45-55	102.95
Proximate users 4am - 8am age <25	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am under age 25	692.45
Proximate users 4am - 8am age >55	No. of unique users driving within 30 miles of the restaurant in the previous week between 4am and 8am over age 55	39.87
Proximate users 12am - 4am	Same as proximate users 4am - 8am but for the time 12am - 4am	3323.95
Proximate users 12am - 4am female	Same as proximate users 4am - 8am female but for the time 12am - 4am	1539.47
Proximate users 12am - 4am age <25	Same as proximate users 4am - 8am age <25 but for the time 12am - 4am	1275.25
Proximate users 12am - 4am age 25–35	Same as proximate users 4am - 8am age 25–35 but for the time 12am - 4am	1068.33

Continued on next page

Feature	Description	Mean
Proximate users 12am - 4am age 35–45	Same as proximate users 4am - 8am age 35–45 but for the time 12am - 4am	436.56
Proximate users 12am - 4am age 45–55	Same as proximate users 4am - 8am age 45–55 but for the time 12am - 4am	346.29
Proximate users 12am - 4am age >55	Same as proximate users 4am - 8am age >55 but for the time 12am - 4am	197.27
Proximate users 8am - 12pm	Same as proximate users 4am - 8am but for the time 8am - 12pm	1460.11
Proximate users 8am - 12pm female	Same as proximate users 4am - 8am female but for the time 8am - 12pm	612.85
Proximate users 8am - 12pm age <25	Same as proximate users 4am - 8am age <25 but for the time 8am - 12pm	488.32
Proximate users 8am - 12pm age 25–35	Same as proximate users 4am - 8am age 25–35 but for the time 8am - 12pm	469.77
Proximate users 8am - 12pm age 35–45	Same as proximate users 4am - 8am age 35–45 but for the time 8am - 12pm	215.57
Proximate users 8am - 12pm age 45–55	Same as proximate users 4am - 8am age 45–55 but for the time 8am - 12pm	189.52
Proximate users 8am - 12pm age >55	Same as proximate users 4am - 8am age >55 but for the time 8am - 12pm	96.84
Proximate users 12pm - 4pm	Same as proximate users 4am - 8am but for the time 12pm - 4pm	3290.62
Proximate users 12pm - 4pm female	Same as proximate users 4am - 8am female but for the time 12pm - 4pm	1524.43
Proximate users 12pm - 4pm age <25	Same as proximate users 4am - 8am age <25 but for the time 12pm - 4pm	1151.59
Proximate users 12pm - 4pm age 25–35	Same as proximate users 4am - 8am age 25–35 but for the time 12pm - 4pm	1071.70
Proximate users 12pm - 4pm age 35–45	Same as proximate users 4am - 8am age 35–45 but for the time 12pm - 4pm	464.77
Proximate users 12pm - 4pm age 45–55	Same as proximate users 4am - 8am age 45–55 but for the time 12pm - 4pm	373.22
Proximate users 12pm - 4pm age >55	Same as proximate users 4am - 8am age >55 but for the time 12pm - 4pm	229.09
Proximate users 4pm - 8pm	Same as proximate users 4am - 8am but for the time 4pm - 8pm	3443.55
Proximate users 4pm - 8pm female	Same as proximate users 4am - 8am female but for the time 4pm - 8pm	1596.49
Proximate users 4pm - 8pm age <25	Same as proximate users 4am - 8am age <25 but for the time 4pm - 8pm	1258.86
Proximate users 4pm - 8pm age 25–35	Same as proximate users 4am - 8am age 25–35 but for the time 4pm - 8pm	1106.66

Continued on next page

Feature	Description	Mean
Proximate users 4pm - 8pm age 35–45	Same as proximate users 4am - 8am age 35–45 but for the time 4pm - 8pm	467.98
Proximate users 4pm - 8pm age 45–55	Same as proximate users 4am - 8am age 45–55 but for the time 4pm - 8pm	373.79
Proximate users 4pm - 8pm age >55	Same as proximate users 4am - 8am age >55 but for the time 4pm - 8pm	236.09
Proximate users 8pm - 12am	Same as proximate users 4am - 8am but for the time 8pm - 12am	3704.23
Proximate users 8pm - 12am female	Same as proximate users 4am - 8am female but for the time 8pm - 12am	1711.34
Proximate users 8pm - 12am age <25	Same as proximate users 4am - 8am age <25 but for the time 8pm - 12am	1350.55
Proximate users 8pm - 12am age 25–35	Same as proximate users 4am - 8am age 25–35 but for the time 8pm - 12am	1192.33
Proximate users 8pm - 12am age 35–45	Same as proximate users 4am - 8am age 35–45 but for the time 8pm - 12am	505.46
Proximate users 8pm - 12am age 45–55	Same as proximate users 4am - 8am age 45–55 but for the time 8pm - 12am	405.60
Proximate users 8pm - 12am age >55	Same as proximate users 4am - 8am age >55 but for the time 8pm - 12am	249.95

Appendix F User-level Summarization: Technical Details

We follow Pappalardo and Simini (2018) to extract mobility patterns for the users in our data from their geo-coordinates. Because coordinate-level data may be considered sensitive, and policymakers may restrict its use, in our simulations, we explore how summarizing these raw data at the user level without revealing exact locations might perform when used as inputs in our analyses. Because human mobility follows remarkably consistent patterns (Gonzalez et al. 2008), using aggregated data rather than the geo-tracking data does not mean necessarily mean that our predictive performance will be hurt. To implement this approach, we rely on the DIary-based TRAjectory Simulator (DITRAS) framework (Pappalardo and Simini 2018). This framework separates the temporal characteristics of human mobility from its spatial characteristics. It turns mobility data into a diary generator represented as a Markov model, which we can use to generate features of interest.

Next, we describe the features that we compute at the user-week level to summarize geo-tracking data, following Pappalardo and Simini (2018)'s feature set. The first set of features relates to the randomness of consumers' driving trajectories. This is important because as the randomness of driving behavior increases, the likelihood of an algorithm being able to learn from past information to predict future visits decreases. Consumers whose driving patterns have a lower degree of randomness may be driving similar routes, e.g., they may have the same commute from home to work. As a result, they may be exposed to the same set of restaurants along their route. However, those with a higher degree of randomness, e.g., who may be driving out of town more frequently, may be exposed to a different set of routes and restaurants more. Thus, capturing the randomness in driving patterns can be informative of visitation decisions.

We use three measures of randomness: random, uncorrelated, and real entropy based on the mobility literature. Entropy is the informational value of past driving behavior when trying to predict future behavior (Pappalardo and Simini 2018). Random entropy measures the uncertainty of an individual's next location, assuming that this individual's movement is completely random among N possible locations (Wang et al. 2019). Uncorrelated entropy captures the heterogeneity of locations visited by the user. Real entropy additionally accounts for the order in which different locations are visited by users and their time spent at each location, thus capturing the user's full spatiotemporal mobility (Song et al. 2010).

The second set of features computed from geo-tracking data relates to how much the app users drive. We take this into account by computing the radius of gyration, which is the characteristic distance traveled by the driver (Gonzalez et al. 2008). The radius of gyration allows us to identify how far consumers typically drive, thus providing useful information related to each consumer. In addition to trajectory-based characteristics, we also capture the overall number of days on which a consumer drives, how many coordinates lie along their trip routes, and the maximum distance they traveled away from the focal location where they spend the most time during the training period.

Finally, we focus on the specific characteristics of each trip, for example, the number of quick stops (of $\leq$ 10 minutes) and long stops (of $\geq$ 60 minutes, Hoteit et al. 2014). We also include a measure of the proportion of trips completed by a user in various windows of time to account for specific time-of-day effects, and the average number of trips per week (Pappalardo and Simini 2018). The summary statistics of these features at the user-week level appear in Table F1.

Next, we describe the technical details of computing these features. Human mobility tends to display a great degree of spatial and temporal regularity. Driving points

Table F1 Summary Statistics of DITRAS Features of Geo-Tracking Data

Feature	Description	Mean
Random entropy	Variability of a user's visited locations if each location is visited with equal probability	5.916
Uncorrelated entropy	Variability of a user's visited locations based on probabilities of past visits	.99
Real entropy	Variability of a user's visited locations based on probabilities and order of past visits	5.802
Radius of gyration (miles)	Characteristic distance traveled by a user	17.69
Unique days	Average no. of unique days of driving	3.87
Locations	Average no. of unique points in a user's trip trajectory	84.21
Max distance (miles)	Maximum distance traveled by users from their home	51.44
Short stops	No. of stops of $\leq$10 minutes	15.37
Short stops at restaurants	No. of stops at restaurants for $\leq$10 minutes	2.26
Short stops at unique restaurants	No. of stops at unique restaurants for $\leq$10 minutes	1.77
Long stops	No. of stops for $\geq$60 minutes	8.32
Long stops at restaurants	No. of stops at restaurants for $\geq$60 minutes	.94
Long stops at unique restaurants	No. of stops at a unique restaurants for $\geq$60 minutes	.72
Morning driving	Proportion of trips in the morning (before 11am)	.33
Afternoon driving	Proportion of trips in the afternoon (11am to 5pm)	.33
Evening driving	Proportion of trips in the evening (after 5pm)	.34
Trip frequency	Average number of trips by a user	12.57

Notes: The summary geo-tracking features are computed for all users using their raw geo-coordinates each week. The reported numbers are aggregated over 12,030 restaurant-weeks for our sample of 401 restaurants. The number of stops at unique restaurants are computed as the count of unique restaurants a consumer stops at, e.g., if a consumer stops at twice at the same Pizza Hut location and once at a Starbucks location, the number of stops at unique restaurants will be two. DITRAS = DIary-based TRAjectory Simulator.

that follow a spatial distribution of displacements over all the users can be well approximated by a truncated power-law with random walk pattern of step size Δ_r (Gonzalez et al. 2008).

$$\Pr(\Delta_r) = (\Delta_r + \Delta_{r_0})^{-\beta} \exp(\frac{-\Delta_r}{k}) \tag{F1}$$

where $\beta = 1.75 \pm .15$, $\Delta_{r_0} = 1.5$km, and cutoff distance of $k|_{D1} \approx 400$km and $k|_{D2} \approx 80$km.

1. Radius of gyration:

   By this formulation, human motion follows a truncated Levy flight random walk with a probability distribution that is heavy-tailed. We can recover the radius of gyration, the characteristic distance travelled by user a when observed up to time t, as follows:

$$r_g^a(t) = \sqrt{\frac{1}{n_c^a}\sum_{i=1}^{n_c^a}\left(\overrightarrow{r_i^a} - \overrightarrow{r_{cm}^a}\right)^2} \tag{F2}$$

   where $\overrightarrow{r_i^a}$ represents the $i = 1, 2, .., n_c^a(t)$ positions recorded by user a and $\overrightarrow{r_{cm}^a}$ is the center of mass of the trajectory.

2. Entropy:

   Entropy is a measure of variability in a users' mobility. We compute three types of entropy: random, uncorreleted and real entropy (Song et al. 2010).

   (a) Random entropy captures the degree of predictability of the user's whereabouts if each location is visited with equal probability.

$$S_i^{rand} = \log_2 N_i \tag{F3}$$

   where N_i is the number of distinct locations visited by user i,

   (b) Uncorrelated entropy captures the degree of predictability of the user's whereabouts taking into account past visitation patterns.

$$S_i^{unc} = -\sum_{j=1}^{N_i} p\left(j\right) \log_2 p_i\left(j\right) \tag{F4}$$

   where $p\left(j\right)$ is the historical probability that location j was visited by user i characterizing the heterogeneity of visit patterns.

   (c) Real entropy captures the degree of predictability of the user's whereabouts taking into account past visitation patterns as well as the order in which a user visits a location. It captures the full spatiotemporal order in a user's mobility pattern.

$$S_i^{real} = -\sum_{T_i' \subset T_i} P\left(T_i'\right) \log_2\left[P\left(T_i'\right)\right] \tag{F5}$$

where $P\left(T_i^{'}\right)$ is the probability of finding a particular time-ordered sub sequence and $T_i^{'}$in the trajectory T_i.

$$T_i = \{X_1, X_2, X_3, .., X_L\} \tag{F6}$$

which denotes the sequence of locations at which user i was observed at each time interval.