Predicting Consumer Visits using Geo-Tracking Data and its Privacy Trade-offs

Unnati Narang

University of Illinois at Urbana-Champaign, unnati@illinois.edu,

Fernando Luco

Texas A&M University, fluco@tamu.edu, https://flucoe.github.io

Accurately predicting consumer visits is critical for businesses to plan resources and serve their customers better. We investigate whether and to what extent geo-tracking data about consumers' locations can improve the ability of a business to predict consumer 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 the total number of visits to over 400 restaurants in Texas, we train machine learning models to predict consumer visits to a restaurant one week ahead. Our results show that geo-tracking data improve the performance of prediction models by 17.96% 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 (26.83%), 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 (2.12% and 4.74-5.21%, depending on the frequency). Importantly, models with restricted 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.

Key words: 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). Geo-tracking data about consumers' locations, in particular, may allow businesses to better predict their customers' future actions and plan for them (Sun et al. 2022, Dean 2023). However, despite the potential usefulness of geo-tracking data for businesses, these data reveal sensitive personal information about consumers (Goldfarb and Tucker

2012, Bleier et al. 2020, Choi et al. 2023). 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 (Binns et al. 2018, 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 examine the extent to which geo-tracking data are useful when predicting consumers' visits to a business location one week ahead relative to demographic and behavioral data. 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). However, they face significant uncertainty in demand (Li and Srinivasan 2019, Fang 2022). 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 as discounts at business locations, primarily restaurants with which the app has partnered. 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 train Machine Learning (ML) models to predict the total Safegraph visits one week ahead using different feature sets derived from the mobile app data (i.e., demographic, behavioral, geo-tracking data). We

---

[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 (FTC 2022). The FTC also charged data vendors InMarket and X-Mode for selling consumers' raw location data (FTC 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. See Web Appendix Section A for a summary of the interviews.

compare the out-of-sample performance of the models using performance metrics, such as the mean absolute error of prediction.

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.[3]

We find that geo-tracking data improve the predictive performance of our models by 17.96% relative to that of models that use demographic and behavioral data only and by 25.41% relative to models that exclusively use demographic data. We also find that using geo-tracking data reduces the likelihood of both overpredicting and under-predicting 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).

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 across the simulations. 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 (26.83%), 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 (2.12% and 4.74-5.21%, depending on the frequency) relative to models with complete geo-tracking data. Importantly, models with restricted geo-tracking generally outperform models that do not use any geo-tracking data.

---

[3]We verify that consumers perceive our simulation exercises to be privacy-preserving by conducting a survey about their perceptions (Lin and Strulov-Shlain 2023, Jerath and Miller 2024). See Web Appendix B for details.

Our research contributes to the growing literature on the value of consumer data for firms and to the literature on privacy regulations and data governance. First, the existing work on the value of data primarily focuses on the value of online data rather than offline geo-tracking data (e.g., Yoganarasimhan 2020, Rafieian and Yoganarasimhan 2021, Berman and Israeli 2022, Korganbekova and Zuber 2023, Lei et al. 2023, Wernerfelt et al. 2024). In one exception that is closest to our research, Sun et al. (2022) use omnichannel data, including data on consumers' offline visits for predicting each consumer's online activity. In contrast, we focus on a firm's ability to predict offline visits using complete geo-tracking data and on quantifying the implications of restricting geo-tracking under potential privacy regulations. Second, the research on privacy in marketing examines the impact of regulations for online firms (e.g., Zhao et al. 2021, Peukert et al. 2022, Miller and Skiera 2023, Johnson et al. 2023a, Johnson et al. 2023b, Goldberg et al. 2024), consumer perceptions of privacy (e.g., Lin and Strulov-Shlain 2023, Jerath and Miller 2024), and data obfuscation schemes (e.g., Li et al. 2023, Tian et al. 2023, Macha et al. 2024). We adapt recommendations from this literature for offline geo-tracking data in our simulations. In this way, our research extends these papers and the ongoing policy debate about the governance of consumers' geo-tracking data (e.g., Tau 2023, FTC 2024).

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. Related Literature

Our research relates to two streams of literature on the value of consumer data for firms and on privacy regulations and data governance.

2.1. Value of Consumer Data for Firms

Research on the value of data quantifies how different types of consumer data contribute to marketing outcomes of interest in various settings and applications. Historically, the marketing literature has documented the usefulness of consumers' purchase histories relative to their demographic data for designing targeted pricing strategies (e.g., Rossi et al. 1996, Acquisti and Varian 2005). More recently, Netzer et al. (2019) show significant additional value of individual-level textual data from loan applications for predicting loan default beyond just demographics and past financial data. In a similar predictive approach, Lee et al. (2024) show that retail transaction data about consumers can better predict their credit scores. With the growth in digital and mobile technologies, the sources and types of consumer data available to firms have expanded rapidly (Lamberton and Stephen 2016, Varga et al. 2024). However, most research in this domain focuses on consumers' online data and not their offline trajectories (e.g., Yoganarasimhan 2020, Berman and Israeli 2022, Lei et al. 2023, Wernerfelt et al. 2024). Research that does consider consumers' offline trajectories has focused on their walking path within a mall or their use of public transportation for one-time coupon targeting purposes (e.g., Ghose et al. 2019a,b).

The paper that comes closest to our research is Sun et al. (2022). Their research quantifies the usefulness of both online and offline data for predicting each consumer's online actions i.e., their likelihood of visiting, considering a purchase, and purchasing at multiple websites two weeks ahead. In their work, Sun et al. (2022) use data on past online and offline trajectories and show that omnichannel predictions outperform single-channel ones by 7.38%. In contrast, we focus on predicting the difference between the realized number of visits and the number of visits a restaurant expects in a typical week. We also examine how the accuracy of these predictions is impacted by privacy-preserving restrictions imposed on the data that firms can use. Such an inquiry is important for physical retail locations that are interested in predicting weekly visits but that may be subject to privacy regulations that restrict data tracking.

Overall, relative to the literature on the value of data, our research focuses on quantifying the predictive value of geo-tracking data beyond traditional metrics like demographics, and evaluating how this predictive value might change under potential privacy restrictions imposed on geo-tracking data.

2.2.   Data Governance, Data Obfuscation, and Privacy

The research on data governance, data obfuscation, and privacy examines regulations and proposes restrictions that can impact how firms use customer data (e.g., Johnson et al. 2023a, Macha et al. 2024)

We broadly categorize these restrictions into the following types: Restricting *what* data are tracked, *where* and *how frequently* they are tracked. Table 1 summarizes these restrictions with motivating examples and their application to geo-tracking.

Table 1   Overview of Privacy Restrictions and their Related Geo-Tracking Simulations

Type of restriction	Motivation	Potential geo-tracking 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.	User-level summarization: Use summaries of driving behaviors extracted from geo-tracking data.	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. 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.	Trusov et al. (2016), Kim et al. (2022)

Notes: GDPR = General Data Protection Regulation. In Web Appendix B, we report consumer surveys to verify that our geo-tracking simulations are perceived as privacy-preserving.

What user data are tracked. Most privacy regulations, such as the California Privacy Rights Act (CPRA), consider geo-tracking data as sensitive personal data. One way in which privacy regulations restrict geo-tracking is to completely ban any form of location tracking (Tau 2023). In some cases, regulatory action also prohibits selling consumers' sensitive location data (FTC 2024). By comparing the predictive performance of models with geo-tracking data and those without any geo-tracking data (i.e., only customer demographics and behavioral data), our research addresses this possibility.

In practice, however, a complete ban is less likely. Instead, regulations typically restrict what user data can be tracked. For example, the General Data Protection Regulation

(GDPR) requires anonymizing personally identifiable information (PII) about users (Wang et al. 2024), which impacts digital firms, publishers, and web technology vendors (e.g., Johnson et al. 2023b, Miller and Skiera 2023). Similarly, Google's *Topics* API in its Privacy Sandbox hides the specific sites users visit and, instead, infers broad interest-based categories to serve relevant ads.[4]

Individual-level data generally perform better for targeting advertisements to consumers relative to aggregate data (Danaher 2023) but the costs vs. benefits of aggregation vary depending on the quality of training data used for ML applications (Simester et al. 2020). In practice, data brokers and third-party vendors commonly use algorithms to create individual user profiles by combining data from multiple sources (Neumann et al. 2021, Lin and Misra 2022, Yan et al. 2022).[5] Algorithms, such as k-anonymity can help transform consumer panel data to prevent their re-identification and mapping to other sensitive datasets (Li et al. 2023). However, geo-tracking data require a careful modification of the general privacy-preservation frameworks to make them privacy-safe in a way that is useful and can lend itself to managerial applications (see Macha et al. 2024).

One way to make geo-tracking data privacy-safe in the spirit of data obfuscation is summarization, i.e., to extract mobility features from geo-tracking data that describe users' driving behaviors without recording latitude-longitude coordinates or sensitive data about each location a consumer visited with timestamps (Pappalardo and Simini 2018). We address this possibility in our simulation on user-level summarization of geo-tracking data.

Where users are tracked. Most apps and firms rely on geo-tracking in confined geofenced areas for their marketing applications. The likelihood of shoppers redeeming mobile coupons for stores inside a shopping mall increases if they receive the coupons in close proximity to the focal stores (Danaher et al. 2015). Similarly, targeting based on real-time or historical consumer location within geofences is more effective for firms (Dubé et al. 2017). In practice, third-party data providers, such as Radar and BlueDot enable geofencing services for firms in a way that allows them to observe shoppers in their vicinity. Similarly, location-specific laws can allow firms to target users in select locations only.[6]

---

[4]See, for example, Google's policy https://privacysandbox.com/proposals/topics/ "Topics: Relevant ads without cookies". Accessed on October 8th, 2024.

[5]In practice, some firms (e.g., https://www.bridg.com/) create privacy-safe profiles for retail media networks to reach their audience.

[6]See, for example, Google's policy https://support.google.com/google-ads/answer/1722043 "Target ads to geographic locations." Accessed on October 8th, 2024.

One way to restrict geo-tracking data is to disallow the tracking of users beyond a certain distance of a store and only track users who enter the geofence. We address this possibility in our simulation exercise on geographical restrictions.

How frequently users are tracked. The frequency with which geo-tracking data can be collected may be consequential for consumer privacy as it is for firm decision-making (Kim et al. 2022). Temporal limitations imposed on individual-level web tracking abilities impact online businesses (Trusov et al. 2016). Regulatory actions that discourage or prohibit the sale of consumers' raw location data streams could deter firms from high-frequency tracking to avoid litigation and reputational damage, in addition to other deterrents like the cost of storing high-frequency data (FTC 2024). In the context of geo-tracking, temporal restrictions would suggest tracking consumers' locations at longer intervals or at static locations near points-of-interest (POIs) only.

One way to restrict the geo-tracking data is to reduce the frequency with which these data are tracked. We consider this possibility in our simulation exercise on frequency restrictions.

Overall, recent privacy regulations, firms' own self-regulation practices, and the data obfuscation literature seek to protect consumers' data and privacy in various ways. We leverage a few key privacy measures and the extant research on privacy to derive and evaluate privacy-safe simulation exercises for geo-tracking data in our application. Through a set of consumer surveys, we provide face validity that these simulation exercises are perceived as privacy-preserving by consumers (see Web Appendix B). By design, each of our proposed geo-tracking simulations restricts consumer geo-tracking data in a different way, either completely transforming the data (e.g., *what* data are tracked) or reducing the scope and extent of tracking (e.g., *where* and *how frequently* data are tracked). In this way, our proposed simulations vary in how they protect consumer geo-tracking data.

3. Data and Empirical Strategy

3.1. Data Sources

Our primary data source is a safe-driving app based in Texas, U.S.[7] The app has over 200,000 users. For our research, we accessed a random sample of about 20% of the app

---

[7]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.

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 incentivizes 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 are primarily restaurants. 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 also 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. The app company gave us access to 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 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.

The second source of our data is Safegraph.[8] This dataset consists of the number of visits to each restaurant location and allows us to define the outcome of prediction. The number of visits to a restaurant is 334.10 on average each week across 401 restaurants in our data. 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).[9] 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.

Finally, we also access the American Community Survey (ACS) 2016 5-year estimate for the Census Block Groups (CBG) in our sample to generate demographic features, such as

---

[8]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."

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

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.[10]

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 chose the sample of 401 restaurants (about 4% of the total restaurants) that at least 10 app users visited in our data period and for which we have the aggregate number of visits in the Safegraph data. We verify that these locations are 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. Our sample is 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).

3.2. Defining the Target of Prediction

We now define the target of our prediction models. Restaurants typically have enough information to calculate the number of expected visits in a given period. For example, retailers commonly form an idea about the visits they expect based on their category, location, 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.

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 accounts for restaurant and seasonal effects. We use the estimates of this regression to compute

---

[10]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), Naik et al. (2016), and Klopack et al. (2024).

restaurant-week deviations from expected visits. By construction, this deviation will be centered at zero over the entire estimation sample. 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 the tables of results report the out-of-sample performance of prediction models that use deviations as the target, with the exception of Web Appendix Table D6 that reports the results for total visits as an alternative target.

3.3. 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.[11] 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. 2024).

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 2.

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.

---

[11]In Web Appendix Table D5, we report the results for an alternative threshold of 17 miles. We use 17 as our alternative threshold because it is the mean distance users travel to visit a restaurant, conditional on visiting.

Table 2   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). *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).

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.[12] Restaurant managers are likely able to access historical data on the total number of invoices, orders, or reservations for customers who use their app or reservation system. However, they cannot easily get this information for competing restaurants or for other locations of their own parent brand in the case of chains (Weis 2023). Therefore, we include past visits to the target restaurant as behavioral data when predicting visits to that restaurant for a given week. We treat visits to other restaurants of the same brand or category as geo-tracking data in the next set (i.e., Model C).

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 rely on for predicting store visits as well as those that are conceptually relevant in the marketing literature.[13]

---

[12]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 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.

[13]The app company we partnered with provides a dashboard of 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 Section A).

Next, we describe the geo-tracking features we use for prediction. First, since geo-tracking data allow managers to identify the visits to competing restaurants, including restaurants of the same category and brands as their restaurant, we include these metrics i.e., information about visits to the target, same-brand, and same-category restaurant. We include these data for all the past weeks (i.e., weeks 1-30) and we also include the visits from the most recent previous week before the prediction week. Second, since geo-tracking data allow managers to identify timestamped information with granular data and time stamps, 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 which we are predicting. 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 (Zubcsek et al. 2017, Shoshani et al. 2024). 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 3.

In our sample, the average age of app users is 32.7, 47% are female, and 72% are white. Additionally, 46% of users are in full-time employment, 44% are commuters, 24% live in

Table 3   Select Features for Prediction and Summary Statistics

| Feature	Description	Mean
Demographic/behavioral information		
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 ≥ 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 ≥ 25 year-old with highschool diploma	0.58
Income	Median household income ($) in the past 12 months	72,141
Past visits	No. of visits to the restaurant in the past	49.88
Geo-tracking information		
Proximity	Min. distance (in miles) between trip coordinates and restaurant	11.14
Proximate Users	No. of users within 30 miles of a restaurant	4,251
Past brand visits	No. of visits to same-brand restaurants in the past	445.3
Past category visits	No. of visits to same-category restaurants in the past	1,973
Recency of past visit	Days since last visit to the restaurant	123.50
Recency of past brand visit	Days since last visit to same-brand restaurants	118.20
Recency of past category visit	Days since last visit to same-category restaurants	115.60

Notes: N=12,030 restaurant-weeks. Geo-tracking features also 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) but these combinations of features are not reported to save space.

family households, and 58% have at least a high school diploma. The median household income is \$72,141. On average, a restaurant records 49.88 visits from an app user.

Among the geo-tracking features, the average driving proximity of consumers to a restaurant is 11.14 miles. The proximity varies based on the time of day, ranging from 10.50 miles during the afternoon (12 pm to 4 pm) to 11.02 miles in the late morning (8 am to 12 pm). There are on average 4,251 proximate users who were within 30 miles of the focal restaurant, of which 45.91% are female and 36.37% are younger than 25 years of age.

Restaurants of the same brand as the target restaurant recorded an average of 445.3 visits in the past weeks, while restaurants in the same category received 1,973 visits. The time since the last visit was 123 days for the target restaurant, 118 days for same-brand restaurants, and 116 days for same-category restaurants.

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 E 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.

Figure 1   Visualizing Geo-Tracking Data: Trip Distance
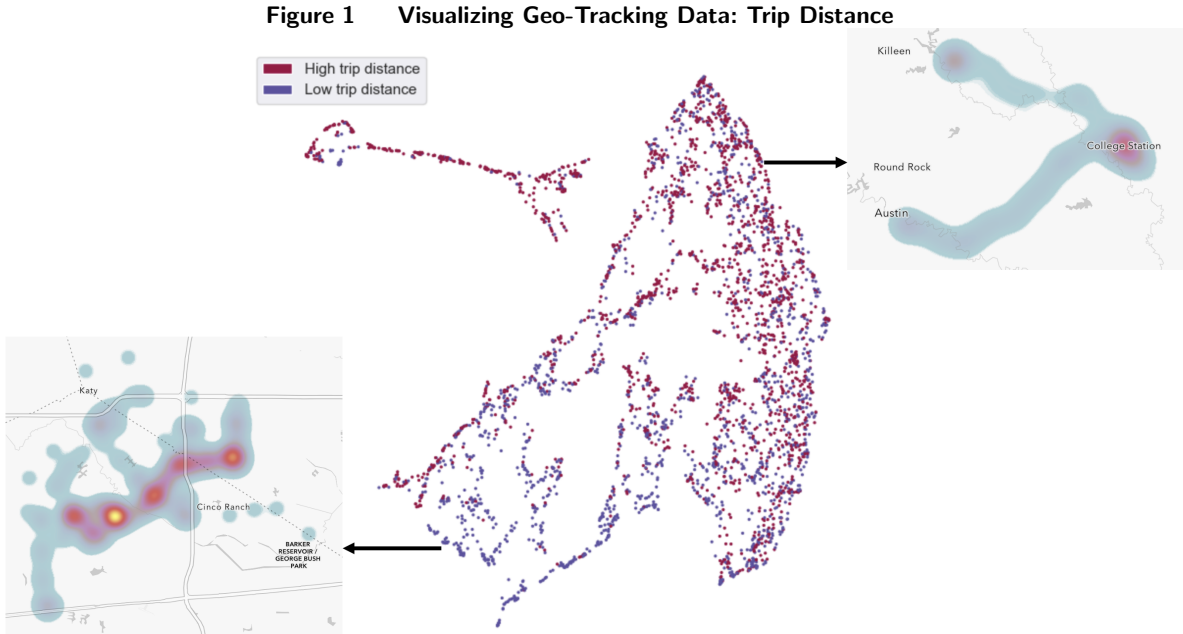
3.4.   Machine Learning Framework and Data Splits

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 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 D7). We find results that are consistent with those from our main model, although the magnitude varies.

3.5. 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. 2024). Because Random Forest offers better performance, we report this model in the main results. We report the alternative models in Web Appendix Tables D8 and D9. The results from these models show the additional value of geo-tracking data similar to those from our main model.

3.6. 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:

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

Second, we also report the Root Mean Square Error (RMSE). The RMSE is the square root of the sum of squared differences between the predicted and observed number of visits for each restaurant and week combination rw in our data. It is computed as follows:

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{rw=1}^{N_{test}} (y_{rw} - \hat{y}_{rw})^2}$$

Finally, we report the Median Absolute Error (MedAE). The MedAE is the median of the absolute differences between the predicted and observed number of visits for each restaurant and week combination rw in our data. Unlike RMSE, which emphasizes larger errors due to squaring, MedAE also provides a more robust metric that is less sensitive to outliers. It is computed as follows:

$$MedAE = \text{median}(|y_{rw} - \hat{y}_{rw}|)$$

Because of the nature of our empirical exercise, in which we aim to predict the deviation in visits with a focus on overall performance across a diverse set of restaurants and periods, we consider MAE as our main performance metric (e.g., Varga et al. 2024, Willmott and Matsuura 2005).[14] 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.

3.7. 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. For each sample, we 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.

---

[14]In Web Appendix Tables D11 and D12, we also report RMSE and MedAE and show that Model C tends to perform better than Models A and B.

4.   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 4 reports the predictive performance of each of our models corresponding to the information sets in Table 3. 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 4 show that Model C, which includes the geo-tracking data, performs better than all other models. The MAE for Model C is 25.41% lower than that of Model A with only demographic information and 17.96% lower than that of Model B with demographic and behavioral information. The confidence intervals show that Model C outperforms models A and B, which do not use geo-tracking data as inputs.[15]

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 evaluate the performance of Model C separately for instances when the model overpredicts and for instances when it underpredicts and compare its performance to that of Models A and B, which do not use geo-tracking data. The idea behind this analysis is that over- and under-prediction may generate different costs to restaurants, and thus, it is relevant to determine whether Model C performs better across these instances or is better only in one of them.

We present our findings in Figure 2, which reports the deciles of the distribution of MAE, separately for cases of under- and over-prediction. The x-axis represents the deciles of the distribution of the error and the y-axis plots the mean error for those deciles. The figures show that while Model C underpredicts and overpredicts by less than the models that do not use geo-tracking data, this difference is larger in the extremes of the distribution of MAE. The difference in predictive performance increases with more extreme events (i.e.,

---

[15]Web Appendix Figures D1 and D2 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 trip distances and visits to other restaurants, replace many of the variables deemed important in Model B.

Table 4   Results: Predictive Performance by Information Set

| Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic information	54.516 [54.319, 54.713]		
Model B: Behavioral information	49.588 [49.408, 49.768]	4.928 [4.84, 5.016]	
Model C: Geo-tracking information	40.681 [40.512, 40.85]	13.835 [13.727, 13.943]	8.907 [8.810, 9.004]

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 text. N = 12,030 restaurant-weeks. Note that all feature sets include the features from the previous set. The outcome is deviation from expected visits. Results correspond to Random Forest models.

the prediction errors of the models that do not use geo-tracking data grow faster than the errors of the models that do use geo-tracking data).

These results show that geo-tracking data could be useful to managers. From our interviews, we know that managers would, for example, alter their staffing and hiring if the difference between expected and realized visits is large, i.e., 20 or more visits each week. Our findings in Figure 2 suggest that in such cases, the magnitude of differences is enough to justify hiring an additional employee. Figure 2 shows that this is the case in extreme cases (i.e., higher MAE regions).[16]

Together, our back-of-the-envelope calculations suggest that the range of savings across our restaurant sample could be substantial in economic terms, with about $646-$1,121 per week (or $33,592-$58,292 annually) potential savings for a restaurant. To put these numbers in context, we also obtained the annual revenue data for our sample of restaurants

---

[16]In extreme cases of overprediction, firms can save an additional employee's salary of $571 per week based on $29,730 annual salary (according to the Bureau of Labor Statistics) and in case of underprediction, firms can save potentially lost revenue from turning away customers (i.e., 3-22 more customers per week at an average check value of $25 as per our interviews would mean $75-$550 in a week).

Figure 2   Distribution of the Prediction Error With and Without Geo-Tracking Data
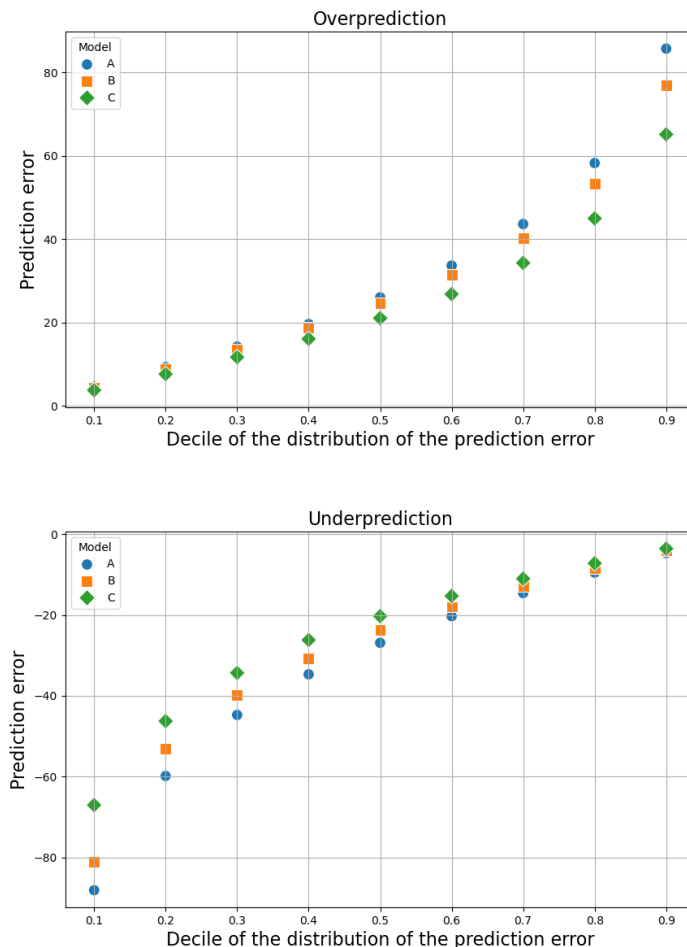
Notes: The plots report the prediction error for each model as a function of the deciles of the distributions of the prediction error for that model, separately for cases of over- and under- prediction. Model A uses only demographic data as input, Model B uses demographic and behavioral data, and Model 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.

from Data Axle, a marketing data provider. For the median restaurant in our sample, these represent 1.9%-3.3% of the restaurant revenues.[17]

5.   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 predictive performance of models that use these data? This question is important from a

---

[17]For the mean restaurant in our sample, these represent 1.7%-2.9% of the restaurant revenues.

policy perspective because of the recent emergence of regulations restricting data tracking (Klosowski 2021, FTC 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 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 1, 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 E. 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.[18]

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.

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.[19]

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 1, 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.

5.1. Results: Predictive Performance under Varying Restrictions on Geo-Tracking

Table 5 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.

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

---

[18]We also consider the possibility that, in practice, businesses could use a different geofence distance around their store locations. We report alternative versions of this simulation with other radii in Web Appendix Table D10.

[19]In alternative implementations of the simulations, in Web Appendix Table D10, we also implement a variation that (a) collects data at one-third of the original frequency and (b) records data for one random trip per user-week, keep all records of it, and drop any other trips that week.

Table 5   Results: Simulation Exercises with Varying Restrictions on Geo-Tracking

Simulation	Mean MAE	Difference	Percentage
Complete geo-tracking (Model C in Table 4)	40.681		
	[40.512, 40.850]		
User-level summarization			
Summary features instead of geo-tracking	51.598	10.917	26.83%
	[51.401, 51.795]	[10.650, 11.183]	
Geographical restrictions			
Geofenced users within 1 mile of restaurant	41.542	0.861	2.12%
	[41.377, 41.707]	[0.640, 1.082]	
Frequency restrictions			
Reduced frequency of geo-tracking			
1/2 frequency	42.799	2.118	5.21%
	[42.614, 42.984]	[1.887, 2.348]	
Static geo-tracking			
First- and last- trip points only	42.609	1.927	4.74%
	[42.426, 42.792]	[1.695, 2.159]	

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.

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 26.83% larger than the one in Table 4.

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 2.12% 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 5.21% under one-half tracking frequency, while using static geo-tracking decreases the predictive performance by around 4.74% 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. Even in the most restrictive cases, the prediction model performs similarly to or better than when the models are restricted to not using geo-tracking data.

6.   Robustness Checks
6.1.   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 Tables D1 and D2, we report the results for alternative samples of restaurants based on more stringent sample selection criteria of at least 15 and 25 users visiting those restaurants anytime in our data period. It is also possible that our findings are generated by the sample of restaurants that have relatively few visits and that our model is unable to predict similarly for more popular restaurants. To examine this, we report the results for the top restaurants by number of visits during our prediction period in Web Appendix Tables D3 and D4. We find qualitatively similar results across all the samples, although the magnitude varies based on the number of visits in each sample.

6.2. Alternative Aggregation Thresholds

In our main analysis, we use a 30-mile distance threshold (i.e., the maximum distance that we observe users driving to visit a restaurant in our data) to identify a restaurant's potential customers and aggregate their demographic, behavioral, and geo-tracking features to predict the total visits to a restaurant each week. We also consider an alternative threshold of 17 miles, the mean distance consumers travel to visit a restaurant in our data. The results are reported in Web Appendix Table D5. We continue to find that models with geo-tracking data outperform models that do not use these data.

6.3. 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 D6. 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 D7. In both the setups, Model C with geo-tracking data outperforms the other models.

6.4. 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 4 and report their results for alternative models in Web Appendix Tables D8 and D9. The findings are consistent with those reported in the main analysis.

6.5. Alternative Implementation of Simulation Exercises

In addition to our main simulation exercises, we implemented alternative versions using different parameters for the simulations. Specifically, we trained the geofence models using alternative radii of two miles rather than the one-mile radius in our main analysis. Similarly, we extended the frequency simulation to one-third tracking and one random trip per week

tracking. We report the findings from these models in Web Appendix Table D10. We find support for the main result that the predictive performance of the models with these alternative thresholds is lower than that of models with complete geo-tracking.

6.6. 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 D11 and D12. Both sets of results show patterns that are similar to the MAE results in Table 4.

7. Conclusion, Implications, and Limitations

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 geo-tracking data improves the out-of-sample predictive accuracy by 17.96% compared to using only demographic and behavioral data, and by 25.41% compared to using demographic data alone.

However, privacy restrictions on geo-tracking reduce prediction performance by 2-27%, depending on the type of regulation. The largest decline in accuracy (26.83%) occurs when restrictions limit the type of data collected, such as summarizing driving behaviors. Despite these reductions, models with restricted geo-tracking still outperform models that use no geo-tracking data at all.

7.1. Managerial Implications

Our results have several implications for firms and policymakers. First, our finding that geo-tracking data allow firms to better predict the total visits from users can help managers plan their marketing and operations. Our interviews with restaurant managers revealed that they plan each week's resource commitments in advance (e.g., decisions to hire additional

staff for the upcoming week). For example, the manager of Culvers' and Red Lobster said that: "Knowing how many customers to expect the following week can reduce cost, increase profits, and help manage [staff and inventory] without waste." By improving the prediction of total visits, we show that geo-tracking data can help managers better plan ahead. Our back-of-the-envelope calculations in Section 4 suggest that using geo-tracking data for planning for consumer visits can result in potential annual savings of $33,000-$58,000 for a restaurant, representing 1.9%-3.3% of the median restaurant revenues.

Second, we find that using geo-tracking data for prediction reduces both over- and under-prediction of visits to a restaurant. The better performance of the model that uses geo-tracking data allows firms to evaluate, based on the cost of over- and under-prediction, their decision to collect and use geo-tracking data for prediction applications.

Third, in our simulation exercises, we propose practical ways in which firms can protect and restrict consumers' geo-tracking data, and show that limiting *what* data are tracked results in the highest predictive losses for firms in our application. However, other restrictions, such as *where* data are tracked and *how frequently*, still allow firms to derive relatively higher predictive value from geo-tracking, which implies that firms are better off restricting geo-tracking in these ways when possible. Our findings can also be useful when firms are purchasing geo-tracking data from third party vendors and need to make decisions about what data and how much data to access. Reduced frequency of tracking, for example, can allow firms to save on data storage and server costs relative to complete tracking. More importantly, our findings imply that there are ways in which firms can restrict geo-tracking data in privacy-safe ways (while still getting predictive value from them), which can allow them to potentially mitigate the risk of litigation and reputational damage, given recent regulatory action against the mishandling of consumer geolocation data (FTC 2024).

Finally, our simulation exercises are meaningful for regulators who often evaluate various types of restrictions to impose because they can better quantify the relative predictive losses of each in the context of our application.

7.2. Limitations and Future Research

Our research has limitations that future research can address. First, many restaurants use geo-tracking data to make real-time predictions. Our ML framework predicts visits one week ahead using past weeks' data (consistent with Sun et al. 2022). Future research may address real-time predictions under privacy restrictions for other relevant applications,

e.g., for targeting mobile coupons. Second, while our simulated regulations are generally perceived as privacy-preserving by consumers we surveyed, future research can better elicit consumers' willingness to share data under different geo-tracking data-sharing scenarios since consumers' stated and actual preferences for privacy may vary (Adjerid et al. 2018). Finally, while we examine simulated regulations drawn from the current policy landscape and privacy literature, there may be other regulations of interest we cannot study. For example, most app users in our setting are adults, so we cannot comment on privacy protection for children (e.g., Johnson et al. 2023a). Similarly, app users in our setting have opted-in to be tracked, so we cannot comment on consenting vs. non-consenting users or on the design of opt-in requests (e.g., Lin and Strulov-Shlain 2023, D'Assergio et al. 2024). Future research, if data are available, can specifically examine these additional important regulations.

References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Science* 24(3):367–381.

Adjerid I, Peer E, Acquisti A (2018) Beyond the privacy paradox. *MIS Quarterly* 42(2):465–488.

Austen I (2022) 'A mass invasion of privacy' but no penalties for Tim Hortons. *New York Times* http://tinyurl.com/33x6ujz5.

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, Arora R, Boehnke J, Daniels K, Derdenger T, Kirschner B, LaFramboise D, Loupos P, et al. (2020) Platform data strategy. *Marketing Letters* 31:323–334.

Binns R, Lyngs U, Van Kleek M, Zhao J, Libert T, Shadbolt N (2018) Third party tracking in the mobile ecosystem. *Proceedings of the 10th ACM Conference on Web Science* 23–31.

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.

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.

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.

Danaher PJ (2023) Optimal microtargeting of advertising. *Journal of Marketing Research* 60(3):564–584.

Danaher PJ, Smith MS, Ranasinghe K, Danaher TS (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, Luo X (2017) Competitive price targeting with smartphone coupons. *Marketing Science* 36(6):944–975.

D'Assergio C, Manchanda P, Montaguti E, Valentini S (2024) Express: The race for data: Utilizing informative or persuasive cues to gain opt-in? *Journal of Marketing* Forthcoming.

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

FTC (2024) FTC order will ban inmarket from selling precise consumer location data. *Federal Trade Commission* https://tinyurl.com/3mh4nzw7.

Ghose A, Kwon HE, Lee D, Oh W (2019a) Seizing the commuting moment: Contextual targeting based on mobile transportation apps. *Information Systems Research* 30(1):154–174.

Ghose A, Li B, Liu S (2019b) Mobile targeting using customer trajectory patterns. *Management Science* 65(11):5027–5049.

Goldberg SG, Johnson GA, Shriver SK (2024) Regulating privacy online: An economic evaluation of the GDPR. *American Economic Journal: Economic Policy* 16(1):325–358.

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

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

Hoteit S, Secci S, Sobolevsky S, Ratti C, Pujolle G (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, Zhong L (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.

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, Lewis E, Luco F (2024) Economic consequences of natural disasters: The impact of hurricane harvey on local retail and consumer welfare. *Working Paper* .

Klosowski T (2021) The state of consumer data privacy laws in the US (and why it matters) http://tinyurl.com/ypjhssxj.

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

Lamberton C, Stephen AT (2016) A thematic exploration of digital, social media, and mobile marketing: Research evolution from 2000 to 2015 and an agenda for future inquiry. *Journal of Marketing* 80(6):146–172.

Lee JY, Yang J, Anderson ET (2024) Using grocery data for credit decisions. *Management Science* .

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* .

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, Gupta S (2023) Reidentification risk in panel data: Protecting for k-anonymity. *Information Systems Research* 34(3):811–1319.

Lin T, Misra S (2022) Frontiers: The identity fragmentation bias. *Marketing Science* 41(3):433–440.

Lin T, Strulov-Shlain A (2023) Choice architecture, privacy valuations, and selection bias in consumer data. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2023-58).

Macha M, Foutz NZ, Li B, Ghose A (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.

Netzer O, Lemaire A, Herzenstein M (2019) When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research* 56(6):960–980.

Neumann N, Tucker CE, Whitfield T (2021) Frontiers: How effective is third-party consumer profiling? evidence from field studies. *Marketing Science* 38(6):918–926.

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.

Peukert C, Bechtold S, Batikas M, Kretschmer T (2022) Regulatory spillovers and data governance: Evidence from the GDPR. *Marketing Science* 41(4):746–768.

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

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Science* 15(4):321–340.

Shoshani T, Zubcsek PP, Reichman S (2024) Predicting store closures using urban mobility data and network analysis Working Paper.

Simester D, Timoshenko A, Zoumpoulis SI (2020) Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Science* 66(6):2495–2522.

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

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

Sun C, Adamopoulos P, Ghose A, Luo X (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.

Tau B (2023) Selling your cellphone location data might soon be banned in U.S. for first time. *Wall Street Journal* http://tinyurl.com/bdsa8khx.

Tian L, Turjeman D, Levy S (2023) Privacy preserving data fusion. *Working Paper* .

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.

FTC (2022) FTC sues kochava for selling data that tracks people at reproductive health clinics, places of worship, and other sensitive locations. *Federal Trade Commission* http://tinyurl.com/pwbedwke.

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

Varga M, Tusche A, Albuquerque P, Gier N, Weber B, Plassmann H (2024) Does FMRI data improve predictions of product adoption by store managers and sales to consumers for consumer-packaged goods? *Working Paper* .

Wang P, Jiang L, Yang J (2024) Express: The early impact of GDPR compliance on display advertising: The case of an ad publisher. *Journal of Marketing Research* 61(1):70–91.

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

Weis A (2023) What franchisors should know about data privacy compliance in 2023. *Franchise Wire* https://tinyurl.com/mwrvd5c8.

Wernerfelt N, Tuchman A, Shapiro B, Moakler R (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.

Yan S, Miller KM, Skiera B (2022) How does the adoption of ad blockers affect news consumption? *Journal of Marketing Research* 59(5):1002–1018.

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

Zhao Y, Yildirim P, Chintagunta PK (2021) Privacy regulations and online search friction: Evidence from GDPR. *Working Paper* .

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

Acknowledgments

Appendix A: Interview of Restaurant Managers and Owners

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 Services	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, Three 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 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 understaff. 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.

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 (p < .001) and geographic restrictions had an average score of 3.22 (p < .001). Among frequency restrictions, the half-hour interval of tracking had an average score of 3.73 (p < .001), the static tracking of first- and last-trip points only had an average score of 3.77 (p < .001), and one trip per week at random had an average score of 3.46 (p < .001). 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.

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.

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, Latín 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 section, we present the robustness checks for alternative sample of restaurants (Tables D1-D4), alternative aggregation approaches (Table D5), alternative outcome in the form of total visits, rather than deviation from visits (Table D6), alternative model setups (Table D7), alternative ML models (Tables D8-D9), alternative ways of implementing the simulation exercises (Table D10), and alternative metrics (Tables D11-D12).

Table D1   Results: Predictive Performance of Random Forest by Information Set for Restaurants with At Least 15 App Users Visiting

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	61.236		
	[60.958, 61.514]		
Model B: Behavioral data	55.741	5.495	
	[55.472, 56.01]	[5.395, 5.595]	
Model C: Geo-tracking data	45.911	15.325	9.83
	[45.656, 46.166]	[15.184, 15.466]	[9.702, 9.958]

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 = 4,756 for 299 restaurants across 30 prediction weeks. The restaurants are selected based on at least 15 app users visiting anytime in the data period. All models include season fixed effects.

Table D2   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	228.841 [226.36, 231.322]		
Model B: Behavioral data	220.948 [218.446, 223.45]	7.893 [7.455, 8.331]	
Model C: Geo-tracking data	179.142 [176.329, 181.955]	49.699 [47.758, 51.64]	41.806 [39.956, 43.656]

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 D3   Results: Predictive Performance of Random Forest by Information Set for Top 25 Restaurants by App Users' Visits

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	232.646 [230.352, 234.94]		
Model B: Behavioral data	164.642 [162.78, 166.504]	68.004 [66.732, 69.276]	
Model C: Geo-tracking data	162.187 [160.122, 164.252]	70.459 [68.905, 72.013]	2.455 [1.587, 3.323]

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 = 567 restaurant-weeks for a sample of 25 restaurants with the highest number visits in the prediction period. All models include season fixed effects.

Table D4   Results: Predictive Performance of Random Forest by Information Set
for Top 50 Restaurants by App Users' Visits

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	175.913		
	[174.512, 177.314]		
Model B: Behavioral data	126.652	49.261	
	[125.591, 127.713]	[48.433, 50.089]	
Model C: Geo-tracking data	115.293	60.62	11.359
	[114.226, 116.36]	[59.642, 61.598]	[10.659, 12.059]

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 = 1,152 restaurant-weeks for a sample of 50 restaurants with the highest number visits in the prediction period. All models include season fixed effects.

**Table D5   Results: Predictive Performance of Random Forest by Information Set
for Alternative Aggregation Threshold of 17 Miles**

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	54.154 [53.952, 54.356]		
Model B: Behavioral data	49.353 [49.160, 49.546]	4.801 [4.716, 4.884]	
Model C: Geo-tracking data	41.676 [41.495, 41.857]	12.478 [12.380, 12.574]	7.677 [7.588, 7.766]

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. Compared to the main model that aggregates over users within 30 miles (i.e., the maximum distance consumers travel to visit a restaurant in our data) of a restaurant, we use the alternative threshold of 17 miles (i.e., the mean distance consumers travel to visit a restaurant in our data) for this analysis.

Table D6   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	203.767 [203.328, 204.206]		
Model B: Behavioral data	144.108 [143.806, 144.41]	59.659 [59.344, 59.974]	
Model C: Geo-tracking data	70.045 [69.79, 70.3]	133.722 [133.348, 134.096]	74.063 [73.831, 74.295]

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 D7   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	53.800		
	[53.720, 53.880]		
Model B: Behavioral data	47.489	6.311	
	[47.414, 47.564]	[6.258, 6.364]	
Model C: Geo-tracking data	41.233	12.567	6.256
	[41.152, 41.314]	[12.486, 12.648]	[6.183, 6.329]

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 D8   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	51.671		
	[51.461, 51.881]		
Model B: Behavioral data	51.566	0.105	
	[51.356, 51.776]	[0.088, 0.122]	
Model C: Geo-tracking data	50.805	0.866	0.761
	[50.594, 51.016]	[0.819, 0.913]	[0.714, 0.699]

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 D9   Results: Predictive Performance of XGBoost Model
by Information Set**

Model	Mean MAE	Difference from Model A	Difference from Model B
Model A: Demographic data	51.122		
	[50.904, 51.34]		
Model B: Behavioral data	54.005	-2.883	
	[53.748, 54.262]	[-3.118, -2.648]	
Model C: Geo-tracking data	43.602	7.52	10.403
	[43.417, 43.787]	[7.342, 7.698]	[10.149, 10.657]

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 D10   Results: Predictive Performance of Random Forest under
Alternative Implementation of Simulation Exercises**

Simulation	Mean MAE	Difference
Complete geo-tracking (Model C in Table 4)	40.681	
	[40.512, 40.85]	
Alternative geographical restrictions		
Geofenced users within 2 miles of restaurant	41.906	1.225
	[41.746, 42.066]	[1.008, 1.443]
Alternative frequency restrictions		
$1/3^{rd}$ frequency	42.799	2.117
	[42.614, 42.984]	[2.886, 2.349]
One trip per week at random only	42.194	1.512
	[42.019, 42.369]	[1.288, 1.737]

Notes: MAE = Mean absolute error. The complete geo-tracking model includes demographics, behavioral, and geo-tracking data (i.e., Model C of Table 4). The table reports the mean and bootstrapped confidence intervals (in square brackets) of the MAE of each model and the difference in the MAE between the complete geo-tracking model and each simulation using the test data, similar to the results reported in Table 5.

**Table D11   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	153.666		
	[152.319, 155.013]		
Model B: Behavioral data	140.911	12.755	
	[139.585, 142.237]	[12.478, 13.032]	
Model C: Geo-tracking data	116.739	36.927	24.172
	[115.379, 118.099]	[35.974, 37.880]	[23.294, 25.050]

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 D12   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	27.340		
	[27.280, 27.400]		
Model B: Behavioral data	24.941	2.399	
	[24.887, 24.995]	[2.345, 2.453]	
Model C: Geo-tracking data	20.912	6.428	4.029
	[20.866, 20.958]	[6.371, 6.485]	[3.979, 4.079]

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.1. Alternative Metrics: Variable Importance

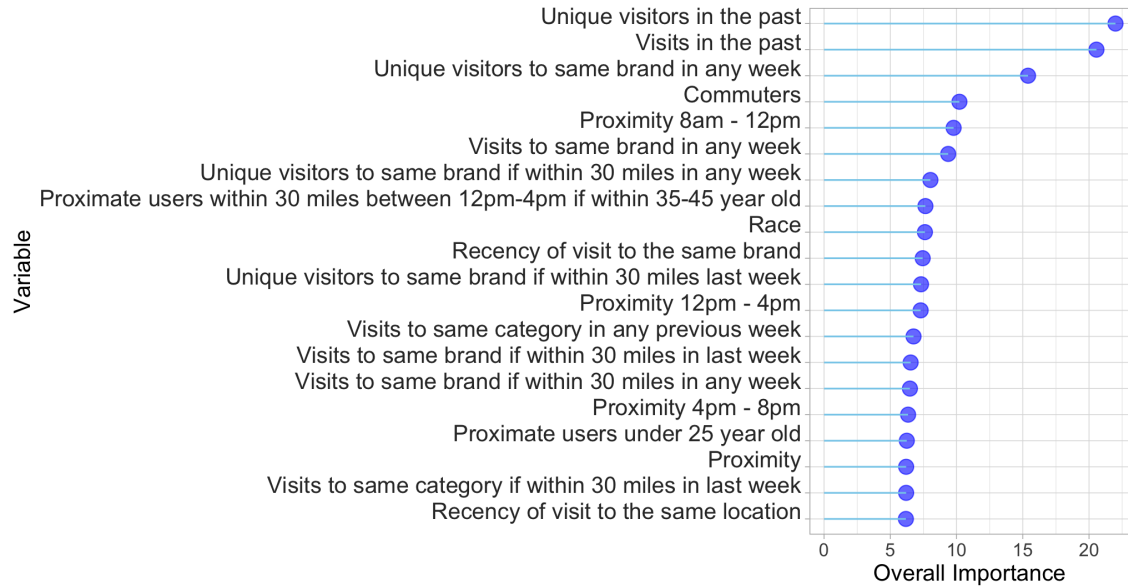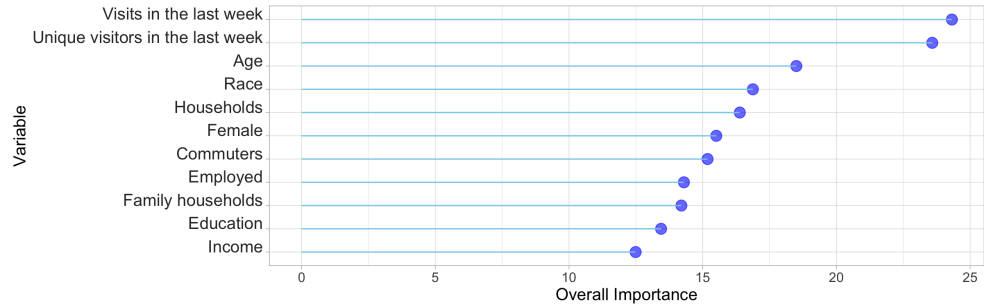**Figure D1   Variable Importance Plot for Model C Across Restaurant-Weeks**

Figure D2   Variable Importance Plot for Model B Across Restaurant-Weeks

Appendix E: 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 tab:summary$_d$$itras$.

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 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).

$$\Pr(\Delta_r) = (\Delta_r + \Delta_{r_0})^{-\beta} \exp(\frac{-\Delta_r}{k}) \tag{1}$$

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:

Table E1   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 ≤ 10 minutes	15.37
Short stops at restaurants	No. of stops at restaurants for ≤ 10 minutes	2.26
Short stops at unique restaurants	No. of stops at unique restaurants for ≤ 10 minutes	1.77
Long stops	No. of stops for ≥ 60 minutes	8.32
Long stops at restaurants	No. of stops at restaurants for ≥ 60 minutes	.94
Long stops at unique restaurants	No. of stops at a unique restaurants for ≥ 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.

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

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.

$$S_i^{rand} = \log_2 N_i \tag{3}$$

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.

$$S_i^{unc} = -\sum_{j=1}^{N_i} p(j) \log_2 p_i(j) \tag{4}$$

where $p(j)$ 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.

$$S_i^{real} = -\sum_{T_i' \subset T_i} P\left(T_i'\right) \log_2 \left[P\left(T_i'\right)\right] \tag{5}$$

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.

$$T_i = \{X_1, X_2, X_3, .., X_L\} \tag{6}$$

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