Geo-Tracking Consumers and its Privacy Trade-offs

Unnati Narang and Fernando Luco

August 3, 2023

Abstract

In recent years, firms have become capable of constantly geo-tracking consumers' locations and movement patterns through mobile apps. Geo-tracking data can allow firms to better predict consumers' future actions. However, geo-tracking also raises privacy concerns among consumers and regulators. Using rich geo-tracking data with over 120 million driving instances for 38,980 app users, we quantify the extent to which these data allow restaurants to better predict their consumers' visits one week ahead relative to using consumer demographics and past behaviors. We also examine how restricting geo-tracking data under counterfactual privacy policies impacts the performance of prediction models that use these data. Using a machine learning framework, we show that geo-tracking data increase the prediction accuracy of our models by 3.1% relative to models that use detailed demographic and behavioral information on past visits and by 14.8% relative to models that use only baseline demographics that do not include any location information. The results from our counterfactual policy analyses further show that privacy restrictions that limit *what* geo-tracking data are tracked, *which* users are tracked, and *where* and *how frequently* they are tracked reduce the predictive performance of geo-tracking. However, the decrease in performance varies by the type of policy restriction; policies that restrict *what* data are geo-tracked (i.e., user-level summary of driving behaviors rather than geo-coordinates) and *where* users are geo-tracked (i.e., within a few miles of a business location), result in the largest decreases in predictive performance relative to complete geo-tracking. Overall, models with restricted geo-tracking still generally outperform models that do not use any geo-tracking information. Our research can assist managers and policymakers interested to assess the risks and benefits associated with the use of geo-tracking data for making predictions.

Keywords:
Mobile apps, geo-tracking, privacy, targeting, machine learning

*Unnati Narang (email: unnati@illinois.edu) is an Assistant Professor of Marketing at the University of Illinois at Urbana-Champaign. Fernando Luco (email: fluco@tamu.edu) is an Associate Professor of Economics at Texas A&M University.

INTRODUCTION

In recent years, most firms have become capable of tracking consumers' locations and movement patterns through their own or third-party mobile apps (Valentine-De et al. 2018). Firms use geo-tracking data to better predict consumers' future actions and make strategic decisions (Sun et al. 2022). Many restaurants, for example, use geo-tracking data to predict their customers' visits to improve their service experience and create local promotions (Dean 2023). Burger King implemented a well-known application of geo-tracking when it offered the Whopper burger for one cent to customers who ordered it through its app while located within 600 meters of a McDonald's (Clifford 2018).[1]

Despite the potential usefulness of geo-tracking data for firms, these data reveal sensitive personal information about consumers (Bleier, Goldfarb, and Tucker 2020; Choi, Jerath, and Sarvary 2023; Goldfarb and Tucker 2012). Canadian coffee chain Tim Hortons, for example, evoked a "mass invasion of privacy" by geo-tracking its millions of app users round-the-clock (Austen 2022). As a consequence, the use of geo-tracking data by firms has attracted regulatory action (Binns et al. 2018; Tau 2023). Recent privacy regulations, such as the California Privacy Rights Act (CPRA) explicitly recognize consumer location data as personal and sensitive information.

Though the use of geo-tracking data has attracted the attention of consumers, firms, and regulators, it is not clear to what extent geo-tracking data allow firms to better predict consumers' future actions, and how privacy policies impact the usefulness of these data for business applications. Our research has two main objectives. First, we examine the extent to which geo-tracking data are useful when predicting consumers' visits to restaurants one week ahead, relative to using only traditional metrics like demographics and past behaviors. Second, we examine how restricting geo-tracking data under potential privacy policies impacts the usefulness of these data in terms of their predictive performance.

[1] Many third-party firms, such as Radar and Bluedot enable businesses to build their geo-tracking capabilities and constitute the growing multi-billion dollar location data ecosystem (Macha et al. 2023)

While the potential usefulness of geo-tracking data for businesses depends on the application under study, in this paper, we identify one application in the restaurant industry that allows us to address our questions. Specifically, based on a set of structured interviews with restaurant managers, our application uses geo-tracking data from one week to predict customers' visits to restaurants in the subsequent week (see Web Appendix Table A1 for a summary of interview responses). Our empirical context is an ideal setting for this research because many restaurants have the ability to access geo-tracking data through their own apps or third-party delivery and deals apps. Furthermore, predicting consumers' visits can inform restaurant managers' decisions about their marketing and operations.

To address our research questions, we use proprietary data from an app that tracks individual-level driving. The primary purpose of the app is to encourage safe driving. To do this, the app partners with businesses and rewards safe drivers with points that they can redeem as discounts at participating locations. In total, our data set contains more than 120 million driving points tracked over a period of 60 weeks for 38,980 individuals in Texas. We extract useful features from consumers' driving trajectories, identify their visits to restaurants, train and evaluate machine learning (ML) models using separate users and time periods, and make predictions about whether a user will visit each of these restaurants in a given week using the previous week's geo-tracking data. Our findings show that geo-tracking data increase the prediction accuracy of our models by 3.1% relative to models that use detailed demographic and behavioral information on past visits and by 14.8% relative to models that use only baseline demographics without any location information.

After establishing that geo-tracking data have the potential to be useful to businesses for making predictions, we next assess how restricting geo-tracking data under counterfactual privacy policies impacts the predictions from our main model with complete geo-tracking. Motivated by the current policy landscape, industry practices, and the privacy literature, we identify and evaluate counterfactual policies of four types – policies that restrict *what* geo-tracking data are tracked (i.e., user-level summary of driving behaviors rather than geo-coordinates), *which* users are tracked (i.e., users above a certain age), *where* they are tracked (i.e., within a few miles of a restaurant), and *how frequently* they are tracked (i.e., with reduced frequency). We find that imposing privacy restrictions that limit geo-tracking reduces the usefulness of these data for predicting visits relative to models with complete geo-tracking by 2.35-14.15%. However, the extent of the decrease in predictive performance varies by the type of policy. Specifically, policies that restrict *what* data are geo-tracked and *where* users are geo-tracked result in the largest decreases in predictive performance while those that restrict *which* users are geo-tracked and *how frequently* result in smaller decreases. Importantly, models that use restricted geo-tracking data still generally outperform models that do not use any geo-tracking data.

Our paper contributes to the growing research on the value of consumer data for firms and the literature on privacy regulation and data governance . First, the existing work on the value of data primarily focuses on online platforms and understanding the value of online browsing data, search histories, and cross-platform data for personalization and targeting rather than offline geo-tracking data (e.g., Lei, Chen, and Sen 2023; Rafieian and Yoganarasimhan 2021; Wernerfelt et al. 2022; Yoganarasimhan 2020). In one exception that is closest to our research, Sun et al. (2022) quantify the value of omnichannel data, including data on consumers' offline visits for predicting *online* shopping two weeks ahead. In contrast, our application focuses on predicting *offline* visits using complete geo-tracking data (and informative features inferred from these data) and quantifying the implications of restricting geo-tracking under potential privacy policies. Second, the research on privacy in marketing has largely focused on either quantifying the impact of existing regulations for online firms (e.g., Goldberg, Johnson, and Shriver 2019; Johnson et al. 2023; Johnson, Shriver, and Goldberg 2022; Miller and Skiera 2017; Peukert et al. 2022, Zhao, Yildirim, and Chintagunta 2021) or proposing privacy-preservation frameworks (e.g., Li et al. 2022; Macha et al. 2023; Tian, Turjeman, and Levy 2023). We leverage this rich body of work as a starting point for identifying and evaluating potential privacy counterfactuals in the context of geo-tracking data, given the increasing importance of such data and a lack of understanding about their usefulness under privacy restrictions. Overall, our research is complementary to and extends these papers and the ongoing policy debate about the regulation of consumers' geo-tracking data (e.g., Austen 2022; Tau 2023).

Our research has several implications for firms and policymakers. First, our finding that geo-tracking data significantly improve the predictions of consumers' visits to a restaurant implies that managers may benefit from collecting and using these types of data. Second, one of our findings is that models that use restricted geo-tracking data generally outperform models that do not use geo-tracking at all. Therefore, if geo-tracking is restricted by policy or practical considerations, managers may still benefit from using some geo-tracking information over not using any. Finally, our finding that different types of restrictions on geo-tracking data impact their usefulness for prediction models differently is informative for managers and policymakers who may be interested in implementing such restrictions.

RELATED LITERATURE

Our research relates to the value of different types of data for firms and the privacy policies that govern data collection and use.

Value of Data for Firms

Researchers have long been interested in quantifying the relative value of different types of data for a variety of marketing applications. Consumers' purchase histories, for example, are more effective than demographic data for designing targeted pricing strategies (e.g., Acquisti and Varian 2005; Rossi, McCulloch, and Allenby 1996).

With the growth in digital and mobile technologies, the sources and types of consumer-level data available to firms have expanded rapidly (Lamberton and Stephen 2016). Recent research provides useful insights on how this expanded access to data helps digital platforms. Search platforms can improve the quality of their search results by adopting personalized rankings using data on individual consumer search histories (Yoganarasimhan 2020). Online retailers can increase their weekly revenues by 4-10% by adopting an analytics dashboard with descriptive data (Berman and Israeli 2022). Combined user data from online and offline channels is more effective for predicting users' next two weeks' online visits and actions on a website, outperforming single-channel predictions by 7.38% (Sun et al. 2022).

On the flip side, barriers to data access tend to hurt firms. Digital advertisers on Meta have to spend 37% more to acquire customers in the absence of cross-app data sharing (Wernerfelt et al. 2022). Similarly, the removal of a search engine's application programming interface (API) for accessing external data leads to a 4.6% decrease in the average click-through rates of its search suggestions (Lei, Chen, and Sen 2023).

Relative to these papers, our research focuses on predicting offline store visits, quantifying the predictive performance of geo-tracking data relative to demographics and behavioral information, and understanding the implications of subjecting geo-tracking data to privacy restrictions. In addition to quantifying the incremental predictive performance of geo-tracking data for physical stores (i.e., restaurant locations), importantly, our work also relates to the current issues of data governance and privacy regulation in the context of protecting sensitive consumer location data.

Data Governance and Privacy Restrictions

In this section, we describe recent privacy measures that regulate firms' access and use of data, particularly geo-tracking data. We also derive potential privacy-safe ways to transform geo-tracking data based on these measures and the relevant literature on privacy.

Firms are often subject to regulations that impact how they collect and use their customers' data. Depending on their nature and scope, we broadly categorize these regulations into one or more of the following types: Restricting *what* data are tracked, restricting *which* users are tracked, restricting *where* or in what locations users are tracked, and restricting *how frequently* users are tracked. Table 1 summarizes these four types of restrictions with examples of regulations and/or industry practices and potential ways of applying these restrictions to geo-tracking data under counterfactual scenarios.

Table 1: Overview of Privacy Policy Restrictions and their Related Geo-Tracking Counterfactuals

Type of restriction	Example regulations/practices	Geo-tracking counterfactual
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 summary features of driving behaviors from geo-tracking data rather than geo-coordinates.
Which users are tracked	COPPA protects children under 13. AADC disallows targeting users under 18 with ads in EU.	Age restrictions: Geo-track only users above a certain age (e.g., 21).
Where users are tracked	Google does not allow ad targeting in many locations depending on the country's laws. Many data vendors only sell data to firms within close proximity of their stores using geofences.	Geographical restrictions: Geo-track users only within certain distances of a focal restaurant (e.g., 10 miles).
How frequently users are tracked	Users may disallow apps from constant geo-tracking and instead select into options, such as only track "while using the app" or temporarily when they want to use a location-based service at a certain place.	Frequency restrictions: Geo-track trips at lower frequency intervals (a) at random (i.e., allow tracking only at random points of a trip), (b) at start and end points of a trip (i.e., allow tracking when users are static at a place and not constantly), (c) random trip per week, and (d) at reduced frequencies.

Notes: GDPR = General Data Protection Regulation. COPPA = Children's Online Privacy Protection Act. AADC = Age Appropriate Design Code. EU = European Union.

What user data are tracked. Most privacy policies, such as the California Privacy Rights Act (CPRA) consider geo-tracking data about consumers' locations and movement to be sensitive personal data. One way in which privacy policies restrict geo-tracking is to completely ban any form of location tracking (Tau 2023). By comparing the predictive performance of models with geo-tracking data with those without any geo-tracking data (i.e., only demographics or past behavioral information), our research addresses this possibility. In practice, however, a complete ban is unlikely. Instead, regulations like the General Data Protection Regulation (GDPR) require anonymizing personally identifiable information (PII) about users (Wang, Jiang, and Yang 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.[2] Individual-level data generally perform better for targeting advertisements to consumers relative to aggregate data (Danaher 2023). As such, data brokers commonly use algorithms to create individual user profiles by combining data from multiple sources (Lin and Misra 2022; Neumann, Tucker, and Whitfield 2021; Yan, Miller, and Skiera 2022). However, individual-level geo-tracking data pose privacy risks by revealing exact home locations and trajectories (Macha et al. 2023) and when combined with other demographic information, pose additional reidentification risks (Li et al. 2022). Firms can use privacy-preservation frameworks to obfuscate individual identities (e.g., Tian, Turjeman, and Levy 2023). One way to make geo-tracking data privacy-safe is user-level summarization i.e., extract features from geo-tracking data that describe users' driving behaviors without recording latitude-longitude geo-coordinates or sensitive home location data. We address this possibility in our counterfactual on user-level summarization.

Which users are tracked. Another commonly employed privacy-protection strategy is to disallow tracking of vulnerable populations, such as users below a certain age (Singer 2023). Regulations like GDPR disallow ad targeting to users below the age of 18 in the EU, while COPPA requires parental consent for collecting data on children (Johnson et al. 2023). Many firms also self-regulate by employing age verifications on their websites (McCabe 2021). Similarly, the Information Commissioner's office in the UK bars geo-tracking of children and if location is tracked for a specific reason, requires turning off consent right after the app session that allowed tracking (ICO 2022, section 10). Many apps like Google services

[2] See, for example, Google's policy "Topics: Relevant ads without cookies". Accessed on July 21st, 2023.

for workplace and education require users to be 18 years of age and also follow the Age Appropriate Design Code (AADC) to be GDPR-compliant and not target ads to underage users in the EU.[3] One way to protect the privacy of younger audiences is to disallow the tracking of users below a certain age. We address this possibility in our counterfactual on age restrictions.

Where users are tracked. Most apps and firms rely on geo-tracking in confined geofenced areas for their marketing applications. The likelihood of shoppers to redeem 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 bar targeting in specific locations.[4] Geo-tracking in geofenced areas near the focal stores may be potentially more privacy-preserving for consumers relative to geo-tracking them anywhere in any location. One way to restrict geo-tracking data is to disallow the tracking of users beyond a certain distance of the store. We address this possibility in our counterfactual 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, Bradlow, and Iyengar 2022). In fact, just knowing two randomly sampled locations visited by a consumer poses the risk of identifying their entire trajectory with 49% success (Macha et al. 2023). Temporal limitations imposed on individual-level web tracking abilities impact online businesses (Trusov, Ma, and Jamal 2016). In the context of geo-tracking, temporal limitations would suggest tracking consumers' locations at larger time intervals. One way in which consumers themselves exercise this option in specific apps is to only allow the app to track their location "while using" the app – which limits the frequency

[3]See, for example, Google's policy "Control access to Google services by age." Accessed on July 14th, 2023.
[4]See, for example, Google's policy "Target ads to geographic locations." Accessed on July 14th, 2023.

of snapshots at which any one specific app can view consumer geo-locations and prevents complete geo-tracking. Data storage limitations may also prevent companies from saving very high frequency data, and they may themselves make the choice to store data at reduced frequencies. We address the possibility of reduced-frequency geo-tracking in our counterfactual on frequency restrictions.

Overall, recent privacy regulations and firm's own self-regulation practices 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 counterfactuals for geo-tracking data in our application.

DATA SOURCES, FEATURES, AND EMPIRICAL STRATEGY

Data Sources

Our primary source of data is a safe-driving app based in Texas, U.S. The app has over 200,000 users. For our research, we were able to access a random sample of about 20% of the app users (i.e., 38,980 users). 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 the user is 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 functions with other deal apps, such as Hooked by offering notifications 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, e.g., food delivery apps, navigation apps.

The app records the position (i.e., latitude and longitude) of a user at a fixed interval of 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. From this app, we have access to individual-level driving data for 60 weeks between September 2018 and October 2019, comprising over 120 million driving points. We use these data to identify driving trajectories as well as restaurant visits. We also have access to data on individual demographics (e.g., age, gender, zip code) for the app users.

The second source of our data is Safegraph.[5] This dataset consists of the polygons (i.e., geometries) that identify each restaurant in Texas. We use these locations together with individual-level driving data to determine if an individual visited a restaurant in a week.[6]

We supplement these data with additional data scraped from Yelp. The Yelp data help us identify the category of each restaurant using Yelp's category tags (Klopack 2022). We use these data when constructing our feature set relating to past visits by a user to a restaurant of the same category and brand as the focal restaurant for which we are interested to make 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 household income and education. These features are not directly recorded at the user level by the app but could contain important demographic information at CBG level.[7]

Features: Demographic, Behavioral, and Geo-Tracking Information

Recall that our goal is to examine how different sets of information, particularly geo-tracking data about each consumer, may allow firms to learn about consumers' likelihood of visiting a restaurant location. We use different sets of information as inputs in our prediction models. These information sets use demographic, behavioral, and geo-tracking information.

[5]See `https://www.safegraph.com/academics` "SafeGraph is 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 five devices visited an establishment in a month from a given census block group." In this paper, however, we only use Safegraph's geometry/polygon data to identify businesses' locations.

[6]In the Section Features: Demographic, Behavioral, and Geo-Tracking Information, we describe how visits are identified from the raw data of GPS points.

[7]The American Community Survey has been used extensively in academic research to extract demographic information at the Census Block Group and Census Tract levels. See, for example, Avenancio-León and Howard (2022); Bertrand, Kamenica, and Pan (2015); Chetty, Hendren, and Katz (2016); Landvoigt, Piazzesi, and Schneider (2015); Naik, Raskar, and Hidalgo (2016).

We report the information sets used in our prediction models in Table 2 and the summary statistics for each feature in Table 3.

Table 2: Information Sets for Predicting Restaurant Visits

Model	Feature set used as input
Model 1	Baseline demographics
Model 2	Baseline demographics + home-zip code distance
Model 3	Baseline demographics + home-zip code distance + behavioral information
Model 4	Baseline demographics + home-tracked distance + trip distance
Model 5	Baseline demographics + home-tracked distance + trip distance + behavioral information
Model 6	Baseline demographics + home-tracked distance + trip distance + geo-coordinates

Notes: Baseline demographics include age, gender, and census-block level data on education, income, etc. but no location information. *Home-zip code* distance refers to the distance between the centroid of the zip code (self-reported in the app) in which a customer resides and the focal restaurant in the prediction model. *Home-tracked* distance, *trip distance*, and *geo-coordinates* represent geo-tracking information; *Home-tracked* distance refers to the distance between the exact latitude-longitude of the customer's home (recovered from geo-tracking data) and the focal restaurant. *Trip distance* refers to the previous week's (i.e., week before the prediction week) minimum average distance between the user's trips and the focal restaurant. Behavioral information contains past number of visits to the focal restaurant, to the restaurants in the same category, and to the restaurants of the same brand as the focal restaurant.

Demographic Information. In our baseline specification (Model 1), we use the demographic features commonly available to managers: consumer's age, gender, and publicly-available ACS data that contain information about the population, race, employment, income, home-work commute, household size, and education at the census-block level.

Our feature set in the baseline specification does not contain any information about the consumer's location or how far they live relative to the restaurant. Many restaurants do not observe the home address of their customers. However, often restaurants can access consumers' home-zip code information either through loyalty programs or through their reservation systems. We next add this information to our set of demographic variables (Model 2). We operationalize home zip code as a distance metric i.e., the distance from the centroid of a zip code to the focal restaurant for which we are making the prediction.[8]

As shown in Table 3, the average age in our sample is 32.16 and 46% of the users are

[8] We also evaluated our models using an indicator variable for each zip code in our data, denoting whether or not the consumer resides in that zip code. However, the models with zip code dummies performed worse. For this reason, and because the distance metric came up in our interviews with restaurant managers, we use the distance metric.

Table 3: Information Sets for Prediction and Summary Statistics

Feature	Description	Mean
Demographic information		
Age	User's age in years	32.16
Gender (female)	User's gender =1 if female and =0 otherwise	0.46
Population	Total population in a census block	4,339
Race (white)	Total white population in a census block	2,941
Employment	No. of workers 16 years and over	1,998
Income	Median household income ($) in the past 12 months	74,674
Commuters	No. of workers 16 years and over who did not work at home	1,901
Family household	Total no. of households with > 1 person in a census block	1,064
Education	Population of age 25 or older with high school diploma	2,617
Home-zip code distance	Distance (in miles) between home zip code centroid and the focal restaurant across all restaurants	149.12
Behavioral information		
Past visits	No. of past visits to the focal restaurant (e.g., MOD Pizza at 6622 Fannin St, Houston)	0.004
Past category visits	No. of past visits to any restaurant of the focal restaurant's category (e.g., Pizza)	0.088
Past brand visits	No. of past visits to any restaurant of the same brand (e.g., any branch of MOD Pizza)	0.032
Geo-tracking information		
Home-tracked distance to the restaurant	Distance (in miles) between home coordinates inferred using geo-tracking trajectories and each focal restaurant across all restaurants	111.54
Trip distance to the restaurant	Minimum past distance (in miles) between trip coordinates and restaurant across all restaurant-weeks for the week preceding the target prediction week	136.20

Notes: Geo-tracking data are used to infer trip distances weekly for all users for the previous week's driving (i.e., the week prior to the target prediction week). Geo-tracking data also contains latitude and longitude trajectories for the previous week (Model 6). Web Appendix B reports the summary statistics for additional features computed from geo-tracking data about users' general driving behavior and mobility patterns for the counterfactual analyses.

female. The average population, white population, number of employed workers, number of commuters, number of family households, and number of people with high school diploma are 4,339, 2,941, 1,998, 1,901, and 2,617 in the consumers' census-block. The median household income is $74,674. Finally, based on their home zip code, users live at an average distance of 149 miles from a restaurant in our data. These distances are averaged across all restaurants in our data, which are located across Texas. For this reason, a user may live 2 miles from a Starbucks in their neighborhood but could be 200 miles away from another out-of-town Starbucks.

Behavioral Information. Behavioral information contains information about a consumer's past restaurant visits in relation to the focal restaurant for which we are making predictions. This set comprises of three features, i.e., each user's total number of past visits to the focal restaurant, to other restaurants in the same category as the focal restaurant, and to other restaurants of the same brand as the focal restaurant. We add this set of behavioral information features to our set of demographic variables in Model 3.

To select the time period for computing past visits, we reserved the first half of our data (i.e., weeks 1-30) in order to predict visits for each week starting week 31. 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. Thus, variables that constitute behavioral information include an increasing number of weeks every week. We do this because restaurant managers are likely to be able to access historical information on orders or reservations for repeated customers who use their app or reservation system. As shown in Table 3, a customer in our data has 0.004 visits to the focal restaurant, 0.088 visits to a restaurants in the same category, and 0.032 visits to a restaurant of the same brand.

In our application, we identify visits from the geo-tracking data by combining them with retail geometries. First, we define a stop as an instance when a driver is stationary. Based on the data, we operationalize stops as time gaps of at least 10 but at most 120 minutes when the driver is not on the move. Next, for each stop, we identify whether the location

Figure 1: Restaurant Polygons in Our Data

Notes: The polygons are geometric boundaries available for each store using satellite imagery.

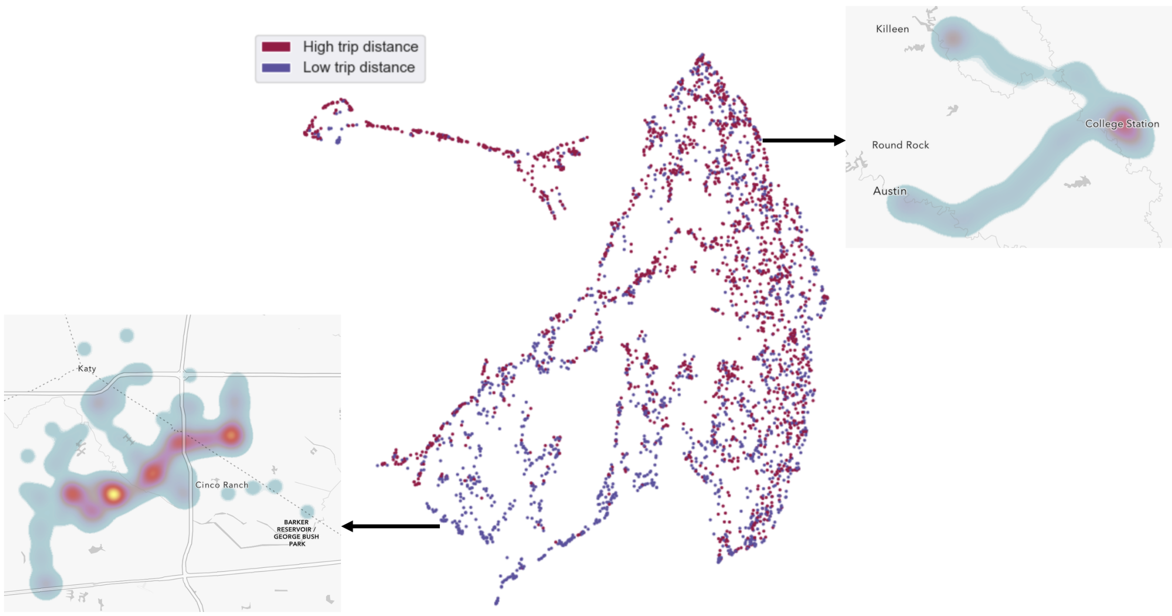
coordinates lie within a polygon defined by longitude-latitudes of each vertex of the physical location. Examples of polygons of restaurants in Bryan/College Station, Texas appear as shaded blue regions in the maps in Figure 1.

Our approach of identifying visits may not capture a fraction of the true visits to restaurants that are adjacent to one another or when a restaurant is within another store (e.g., Subway within Walmart). For this reason, we only use standalone restaurants in our data. Similarly, if the last point recorded for a specific trip lays outside the polygon of a restaurant, we would not identify this point as a visit if the user entered the restaurant just after the last GPS point was recorded. In these cases, training the model with the resulting data would lead to under-performance relative to a model trained with true visit data.

Geo-Tracking Information. Learning useful information from geo-tracking data to predict consumers' future restaurant visits is not a trivial exercise. First, geo-tracking data have spatio-temporal richness but may also be very noisy. Second, using the entire trajectory of consumer movement as inputs in our prediction models is computationally intensive and not readily scalable. We deal with these challenges in multiple ways. First, we follow Pappalardo et al. (2022) and use the latitude and longitude of geo-tracking data to identify the home location for each user. We use the home location to calculate a user-restaurant specific "distance from home" measure (i.e., *home-tracked distance* to restaurant). Second, for each user-week, we identify the minimum distance that the user was from each restaurant in our data in the previous week (i.e., *trip distance* to restaurant). This covariate provides additional information relative to home-tracked distance because it changes every week and considers a user's commuting during the training week relative to the focal restaurant's location. Restaurant managers often use these measures of how far consumers live and drive from their location as heuristics. In Model 4, we next add these home-tracked and trip-distances as geo-tracking features to our baseline demographics. Importantly, Model 4 uses home location coordinates inferred from geo-tracking data rather than home zip codes. Next, in Model 5, we add past behavioral information to Model 4. As shown in Table 3, users' home-tracked distance is 111.54 miles averaged across restaurants and their trip distance is 136.20 miles averaged across restaurant-weeks. Finally, in Model 6, we use latitude and longitude of geo-tracking data as direct inputs in addition to home- and trip- distance metrics.

To illustrate the patterns captured by the geo-tracking data, the panel at the center of Figure 2 presents a two-dimensional projection of our driving features in a multi-dimensional space of our entire feature set (i.e., features described in Web Appendix B). Each point in this projection represents a user. The shape of this visual is determined by the driving features, but the points in the figure are color coded by one feature, i.e., minimum distance between a user's trips to the focal restaurant for a specific restaurant-week. The users represented by purple dots have lower trip distances while those represented by pink dots have higher trip distances. To illustrate these two types of patterns, we also show the heat maps of actual driving for two users in different regions of the projections (Van der Maaten and Hinton 2008). Consider, for example, the user from the purple region of the plot on the left. This user's driving is concentrated in small geographical areas within Katy, Texas. In contrast, the user from the pink region of the plot on the right has a higher trip distance, driving in Killeen, Austin, and College Station within a week.

Figure 2: Visualizing Geo-Tracking Data: Trip Distance



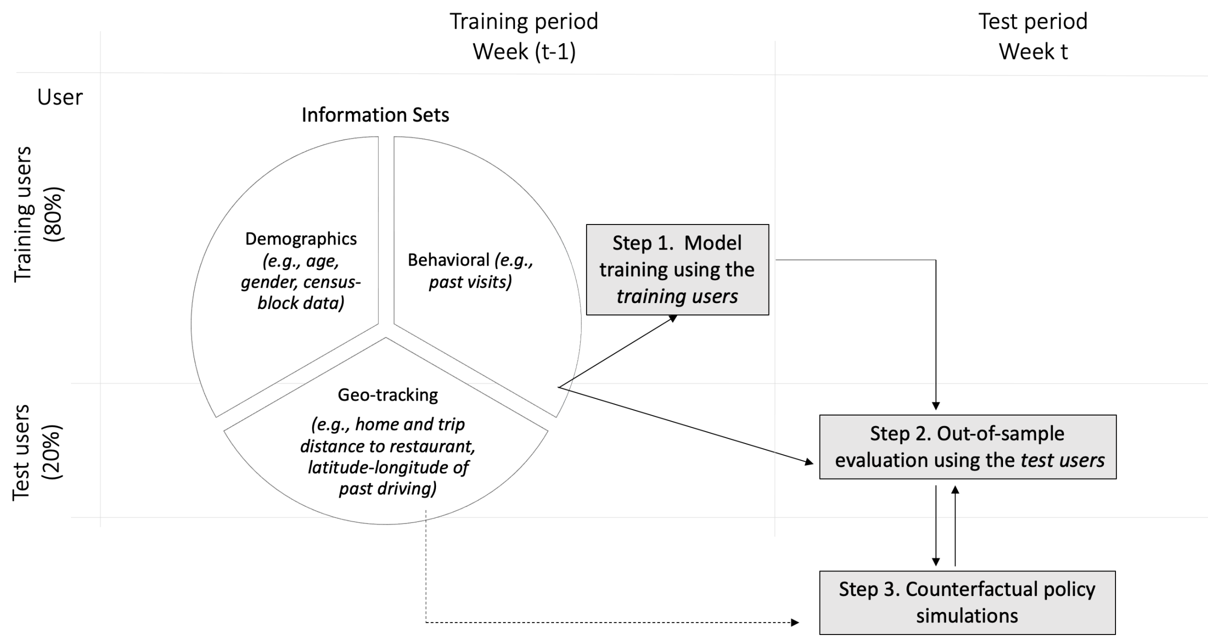
Empirical Strategy and ML Framework

In this section, we describe our ML framework to (a) train a model for predicting consumers' visits to a restaurant using our information sets, including demographic, behavioral, and geo-tracking information, (b) evaluate the out-of-sample predictive performance, and (c) undertake various counterfactual exercises that restrict geo-tracking and compare model performance. Our ML framework appears in Figure 3.

As shown in our ML framework, we split the data both by time periods and users. The time and user splits ensure that we train our model using a different time period for a different set of users compared to the time period and users for which we predict the visits (e.g., Lee, Yang, and Anderson 2021; Rafieian and Yoganarasimhan 2021).

The user separation in our ML framework ensures that the model is learning general patterns during training and is less likely to overfit to the same users whose patterns it learns. Specifically, we train our models using a random sample of 80% of our users in each training period. We evaluate and report the predictive performance of the models estimated for the 20% test users.

Figure 3: Machine Learning (ML) Framework: Predicting Consumer Visits using Demographic, Past Behavioral, and Geo-Tracking (Complete vs. Restricted) Information



Notes: In Step 1, we train an ML model using the information sets (i.e., combinations of demographic, behavioral, and geo-tracking data described in Models 1-6 in Table 2) for the users in the training sample for the training period. In Step 2, we evaluate the out-of-sample performance using the data for the users in the test sample for the test period for the trained model. In Step 3, we train counterfactual models under a hypothetical policy regime with restricted geo-tracking. The approach we follow is the same as for the main model (i.e., Step 1 and Step 2) but we limit the extent of geo-tracking according to the policy under study (see Section Counterfactual Policy Analysis). The training and test periods are defined using a "rolling window" i.e., we compute and use geo-tracking information based on the first week to predict the target visits in the second week, and then we use geo-tracking information based on the second week to predict the target visits in the third week, and so on (consistent with Sun et al. 2022's rolling windows of past five weeks to predict next two weeks' website visits).

The time separation in our ML framework ensures that we are not capturing contemporaneous correlations between geo-tracked driving and visit behavior. For example, shorter driving distances to a restaurant in a week likely entail more visits but may not be informative about consumers' preferences.

Overall, our data span 60 weeks. We reserve the first half of our data for computing behavioral information on past visits (i.e., weeks 1-30). We use a rolling window for the remaining periods (i.e., weeks 31-60) to make weekly predictions i.e., we use geo-tracking information $week_{(t-1)}$ to predict the visit for $week_{(t)}$. Thus, the time separation is between each block of $week_{(t-1)}$ and $week_{(t)}$ for weeks 31-60. Note that the past behavioral information uses the entire past period of weeks 1-30 as well as any weeks before the target week for prediction to capture most recent visit behaviors (see details in the Section Features: Demographic, Behavioral, and Geo-Tracking Information).

For our prediction models, we use standalone restaurants in Texas with non-zero visits during weeks 31-60 by the users in our sample. To train our models, we need to have sufficient visits in a given restaurant-week. For this reason, we select top restaurant-weeks that have at least four visits, resulting in a set of 31 unique restaurants over 107 restaurant-week combinations. The mean number of visits across the restaurant-weeks in our data is 9.74.

Model Training

Because we want to predict visits to a specific restaurant one week ahead, we train a separate model for each restaurant-week and we evaluate the performance of the model on the test users during the test week. To illustrate, the prediction from one specific model estimation will give us the likelihood or probability that a user visits a specific restaurant (e.g., MOD Pizza at 6622 Fannin St, Houston) in a given week. We discretize the predicted probability to a visit or non-visit using a probability threshold of 0.50 and compare it with our observed outcome to compute model performance metrics.[9]

One empirical challenge in our setting is the sparsity of the target outcome that our model is predicting (i.e., our dataset has more non-visitors than visitors for each restaurant-week). The challenge of imbalanced target outcome is common in many ML applications in marketing, e.g., detecting fake reviews or detecting missed credit card payments as in cases of He et al. (2022) and Lee, Yang, and Anderson (2021). To handle class imbalance, we follow the oversampling strategies used in prior literature in marketing and computer science for class-balancing before training our ML models (e.g., Lee, Yang, and Anderson 2021;

[9]We consider two alternative random thresholds, 0.68 and 0.39. Our results are robust to these alternative thresholds and are reported in the Web Appendix Section Alternative Probability Thresholds.

Mohammed, Rawashdeh, and Abdullah 2020).

Because a goal from our analysis is to quantify the predictive value of geo-tracking data over and above demographics and behavioral information on past visits, we use various ML algorithms, such as Extreme Gradient Boosting (XGBoost), Least Absolute Shrinkage and Selection Operator (Lasso), Logit, and Random Forest. Because XGBoost has the highest accuracy for our data, we report this model in the main results, and other models in the Web Appendix Section Alternative Models. Importantly, the results from the other models follow similar patterns as those from our main model XGBoost (Chen and Guestrin 2016).

Out-of-sample Evaluation

In the second step of our framework, we evaluate the performance of our trained ML model. We do this in two ways. First, we compute the model performance metrics for each model specification in Table 2 across restaurant-weeks. Second, we quantify the relative predictive performance of models that use geo-tracking information (Models 4-6) over and above using traditional metrics, such as demographic and behavioral information only (Models 1-3). To generalize our framework and comparisons of various information sets, we report the average performance metrics across the restaurant-weeks.

Performance Metrics

We evaluate the predictive performance of our models using four commonly used metrics:

- *Accuracy.* Accuracy is the proportion of visits and non-visits that are correctly identified as visits and non-visits by our model. Accuracy measures the total number of correctly classified observations divided by the total number of observations.

- *Recall.* Recall is the proportion of correctly classified visits of the total number of visits in the data. This metric is important in our setting because accuracy alone does not reflect a higher proportion of correctly identified visits. With recall, we can trace how many true visits were actually identified among all the observations in our data when there was a visit. It is the number of true positives divided by the sum of true positives and false negatives.

- *Precision.* Precision is the proportion of correctly classified visits of the total number of visits predicted by our model to be visits. With precision, we can trace how many true visits were actually identified among all the observations our model predicts as being positive. It is the number of true positives divided by the sum of true positives and false positives. Note that precision will not be defined if the model does not identify any instances as positives, since the sum of true positives and false positives in the denominator is zero for such cases. We report precision conditional on at least one positive prediction.

- *F1.* F1 is the harmonic mean of precision and recall. In practice, managers may care about balancing both precision and recall because they want to accurately identify the visits that they predict as being true as well as those that are actually true visits. Note that F1 will not be defined if the model does not identify any instances as positives so we report it conditional on at least one positive prediction.

$$\text{F1} = \frac{2 * Precision * Recall}{Precision + Recall}$$

Bootstrapping Procedure

We use a bootstrap procedure to evaluate the differences in predictive performance across models. We implement this procedure by generating 50 bootstrap samples from the training and test data with replacement for each restaurant-week combination in our data. For each sample, we compute the accuracy, recall, precision, and F1 score 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 these distributions to report the standard error for each metric.

RESULTS: PREDICTIVE PERFORMANCE BY INFORMATION SET

In this section, we 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 traditional metrics such as consumers' demographics and past behaviors. To do this, we compare the predictive performance across our four metrics (i.e., accuracy, precision, recall, and F1 score) for the six information sets described in Table 2.

Table 4 reports the predictive performance of these models for the test data averaged over the restaurant-weeks in our data. We first describe the results for the overall prediction accuracy of each model, which is the percentage of correctly classified visits and non-visits for each restaurant.

Table 4: Results: Predictive Performance of XGBoost Models for Various Information Sets

Model	Accuracy	Recall	Precision	F1
Model 1	68.50%	56.47%	75.17%	62.39%
	(1.12%)	(0.84%)	(2.61%)	(2.19%)
Model 2	76.13%	59.93%	88.05%	69.24%
	(1.32%)	(0.53%)	(2.77%)	(2.23%)
Model 3	76.70%	60.44%	89.14%	69.97%
	(1.30%)	(0.50%)	(2.70%)	(2.24%)
Model 4	78.61%	62.76%	91.79%	72.19%
	(1.34%)	(0.34%)	(2.78%)	(2.28%)
Model 5	79.07%	63.57%	92.05%	72.86%
	(1.34%)	(0.34%)	(2.78%)	(2.27%)
Model 6	77.87%	58.52%	91.88%	67.11%
	(6.03%)	(12.16%)	(12.68%)	(8.26%)

Notes: The table reports the mean and standard errors (in parentheses) produced from 50 bootstrap simulations for each of the four evaluation metrics using the test data.

As show in Table 4, Model 1 that uses only the baseline demographic information, such as age, gender, and census block- level features has a mean accuracy of 68.50%. Model 2, which adds home-zip code distances from each focal restaurant for which we are predicting, has a mean accuracy of 76.13% and performs better than Model 1. Thus, adding zip-code level information about where consumers live relative to the restaurant's location results in an increase in accuracy of 11.14%. Model 3, which adds behavioral information on past visits, performs similar to Model 2 with 76.70% accuracy.

Next, we discuss the model performance for Models 4 and 5 that include geo-tracking information. Model 4 introduces home-tracked locations inferred from the geo-tracking data to compute how far consumers live from the focal restaurant as well as trip distances of how close they drove to the focal restaurant in the previous week. With this information, Model 4 predicts visits with an average accuracy of 78.61%. Thus, using geo-tracking data improves the prediction accuracy by 14.76% over Model 1 with only baseline demographics, 3.26% over Model 2 with demographics and home-zip code distances, and 2.49% over Model 3 with baseline demographics, home-zip code distances, and past behavioral information.

Model 5 uses the same geo-tracking inputs as Model 4 but adds behavioral information to the model. Model 5 has an accuracy of 79.07%, which is 15.43% higher than Model 1 with only baseline demographics, 3.86% higher than Model 2 with demographics and home-zip code distances, and 3.10% higher than Model 3 with baseline demographics, home-zip code distances, and past behavioral information (but no geo-tracking information). The gains in prediction accuracy upon including geo-tracking information suggest additional value of these data for predicting consumer visits.

While home and trip distance metrics in Models 4 and 5 provide a useful way to capture geo-tracking features, they might miss additional information available in geo-tracked trajectories. In Model 6, we use latitude and longitude of the user's past week's trips as inputs in addition to the distance metrics. However, using geo-tracked coordinates in the analyses does not improve predictive performance but increases the computational cost significantly. For this reason, we use Model 5 as our preferred specification in the rest of the paper.[10]

[10] In addition, it is important to note that including the trajectory-level data not only results in worse and noisier outcomes for all our metrics, but also that using distances based on the trajectory data allows us, and firms, to compute these models significantly faster than what it takes to include the geo-coordinates. For example, while estimating Model 4 or Model 5 takes a few minutes for each restaurant-week using an iMac computer with 3.6 GHz 8-core i9 and 64 GB 2667 MHz RAM, estimating Model 6 with latitude-longitudes takes 2+ hours per restaurant-week in the same machine. Therefore, because using the trajectory-level data results in worse outcomes and require more

While the inclusion of geo-tracking data in Models 4 and 5 improves the overall accuracy of prediction, it is possible that the model with this information is simply predicting more non-visits as non-visits, but not predicting the actual visits as such. To assess how our models perform in identifying visits correctly as a proportion of the total observed visits, we next examine the recall metric. Our preferred Model 5 has a recall of 63.57%, which is 12.57% higher than Model 1, 6.07% higher than Model 2, and 5.18% higher than Model 3.

The results for precision, which identifies how many of our predicted visits were indeed visits, follows a similar pattern. Model 5 has a precision of 92.05%, which is 22.45% higher than Model 1, 4.54% higher than Model 2, and 3.26% higher than Model 3.

Firms may also care about the balance between precision and recall metrics because both precision and recall matter in practice. The F1 score takes the harmonic mean of precision and recall. The F1 score is 16.78% higher for Model 5 relative to the Model 1 with baseline demographics, 5.23% relative to the Model 2 with demographics and home-zip code information and 4.13% higher relative to Model 3 with both demographic and behavioral information.[11]

COUNTERFACTUAL POLICY ANALYSIS

The findings reported in the previous section showed that geo-tracking data are useful when predicting future restaurant visits. However, these data also raise privacy concerns because of the detailed knowledge that firms may gain from such data about their customers' movement and locations. In this context, we now address 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 policy perspective because of the recent emergence of regulations that restrict data tracking

computing resources, we prefer specifications that use distance measures derived from the geo-tracking data rather than using the geo-coordinates as inputs for our models.

[11]In Web Appendix Figure C1, we examine the variable importance scores of the different features and find that trip distance (i.e., the average minimum distance between trips in the previous week and the focal restaurant of interest), a geo-tracking feature, has the highest importance in prediction followed by home-tracked distance.

(Klosowski 2021). If policymakers regulate consuser geo-tracking, for example, by requiring data aggregation so firms cannot access coordinate-level data, imposing geographical restrictions on where users are geo-tracked, or imposing age or recording frequency restrictions, how would such restrictions impact the benefits of prediction from such tracking?

To answer this question, we outline an approach to quantify the extent to which alternative privacy policies impact prediction performance. Recall that in Table 1, we motivated four types of policy restrictions and ways in which they can be applied to geo-tracking data. Next, we describe how we implement each of these policies in our setting.

User-level summarization. In the first category of counterfactuals, we consider policies that require the data to be summarized at the user level rather than allowing firms to use latitude-longitude geo-coordinate information. We implement a version of this following Pappalardo and Simini (2018). We use the raw coordinates for each user to construct summary features for each user-week that we use in the prediction models instead of raw geo-coordinates, or exact home and trip information. 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), time of driving, number of days driven, and number of trips each week. Note that these features are independent of the focal restaurant's location and capture general mobility patterns about a user in a privacy-preserving way (e.g., Macha et al. 2023). The technical details and summary statistics of these features appear in Web Appendix B: User-level Summarization of Geo-Tracking Data: Technical Details. Under this counterfactual, we re-estimate Model 5 with these summary features instead of the home-tracked and trip distances from the focal restaurant.

Geographical restrictions. In the second category of counterfactuals, we explore how geographical restrictions impact prediction outcomes. We implement this counterfactual using geofences that restrict firms to observe only those trips that take place within 10 miles of their location. Any trip that takes place outside this distance is therefore not observable to the firms and is excluded from our counterfactual prediction model. Under this counterfactual, we re-estimate Model 5 with using these geographically restricted geo-tracking data.[12]

Age restrictions. In the third category of counterfactuals, we consider policies that restrict *who* may be tracked, i.e., prevent tracking of user below a certain age. Because our data comes from a safe-driving app, we implement a counterfactual in which firms are restricted to use data of users who are over 21 years old. Under this counterfactual, we re-estimate Model 5 only for those users who are 21 and above.

Frequency restrictions. In the fourth category of counterfactuals, we consider policies that restrict *how often* users may be tracked, but still allow firms to use the data at the coordinate level. In practice, we consider policies 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 1/2, 1/3, and 1/10 frequency of the original three-minute interval with which our data provider records the data. We also implement a variation that keeps the overall *amount* of data the same as the 1/2 geo-tracking frequency, but drops data at random rather than systematically. These exercises are meant to represent choices firms might make about temporal granularity when deciding how often to record and store data.

In the second version, we implement counterfactuals that reduce geo-tracking frequency in ways that replicate static geo-tracking policies. These exercises are meant to represent scenarios in which users may enable tracking at specific moments, for example when they want to use a location-based service at a point of interest, or at specific moments during a trip when they check a navigation app. In practice, we consider two policies that differ along this dimension. In the first one, we record the first and last points of a trip, and drop all other records. In the second one, we record data for one random trip per user-week, keep all

[12] In practice, businesses could use smaller geofences around their store locations. We also implemented another version of this counterfactual with a one-mile radius instead of 10 miles. The predictive performance of the models reduces to 63.44% accuracy, which is lower than our baseline Model 1 with only demographics. This is likely due to both home- and trip- distances, which play an important role in predictions, being inaccurate for very small geofences.

records of it, and drop any other trips that week.

Under each of the policy restrictions, we re-estimate our preferred model, Model 5, using restricted geo-tracking data for both the training and test users.

Results: Counterfactual Analyses under Varying Restrictions on Geo-Tracking

Table 5 presents our findings for the counterfactual analysis. In each counterfactual, we report the results for Model 5 after re-estimating it using restricted geo-tracking data.

Our findings show that all the restrictions that we impose result in lower prediction accuracy than the accuracy of our preferred Model 5 with complete geo-tracking. However, we also find that even under various policy restrictions, models with geo-tracking information generally perform better than or similar to the specifications that do not use geo-tracking information at all. Below we describe the results for each of our policy counterfactuals.

First, we find that policies that force the summarization of geo-tracking data at the user-level result in the largest decrease in predictive performance relative to our preferred Model 5. While Model 5 has an accuracy of 79.07%, the model with user-level summarization has an accuracy of 67.88%, which compares to the performance of models that uses only baseline demographics (Model 1 in Table 4). The decreased accuracy under user-level summarization highlights that *what* geo-tracking data can be collected and used is an important consideration when predicting customers' visits one week ahead.

Second, policies that restrict firms to only observe data within 10 miles of their locations, result in the second largest decrease in prediction accuracy. While Model 5 has an accuracy of 79.07%, the model with geographical restrictions has an accuracy of 70.46%. However, even in this case, geo-tracking data are useful when predicting customer visits relative to a context in which these data are not available at all.

Third, age-based restrictions that prevent firms from geo-tracking younger users does not result in substantial changes in prediction accuracy. Note, however, that because of the nature of our data (i.e., driving), most app users are adults and by restricting to users who

Table 5: Results of the Counterfactual Analyses: Predictive Performance of Models under Restricted Geo-Tracking Data

Counterfactual	Accuracy	Recall	Precision	F1
Complete geo-tracking (Model 5 in Table 4)	79.07%	63.57%	92.05%	72.86%
	(1.34%)	(0.34%)	(2.78%)	(2.27%)
User-level summarization				
User-level summary features instead of geo-tracking	67.88%	51.75%	76.05%	59.73%
	(0.14%)	(0.24%)	(0.20%)	(0.26%)
Geographical restrictions				
Geofenced trips within 10 miles of the restaurant	70.46%	53.62%	80.86%	62.58%
	(0.61%)	(1.28%)	(1.02%)	(1.26%)
Age restrictions				
Users over 21-year old only	75.79%	55.59%	93.15%	67.75%
	(0.12%)	(0.24%)	(0.20%)	(0.26%)
Frequency restrictions				
Reduced frequency of geo-tracking				
1/2 frequency	75.53%	57.04%	90.66%	67.67%
	(0.16%)	(0.32%)	(0.29%)	(0.30%)
$1/3^{rd}$ frequency	77.21%	59.39%	92.53%	70.15%
	(0.25%)	(0.48%)	(0.33%)	(0.20%)
1/2 frequency at random	76.59%	57.98%	92.31%	69.07%
	(0.19%)	(0.37%)	(0.29%)	(0.34%)
$1/10^{th}$ frequency	75.43%	56.77%	90.66%	67.53%
	(0.19%)	(0.39%)	(0.27%)	(0.30%)
Static geo-tracking				
First- and last-trip points only	77.04%	59.48%	91.48%	69.48%
	(0.21%)	(0.42%)	(0.26%)	(0.21%)
One trip per week at random only	75.16%	51.21%	98.49%	67.36%
	(0.21%)	(0.41%)	(0.11%)	(0.28%)

Notes: The main model specification includes demographics, behavioral, and geo-tracking data (i.e., Model 5 of Table 4). The table reports the mean and standard errors (in parentheses) produced from bootstrap simulations for each of the four evaluation metrics using the test data under each counterfactual scenario.

are over 21 years old, our model still has access to the data for almost 70% of our users.

Fourth, we consider counterfactuals that restrict the frequency with which geo-tracking data are recorded including reduced frequency and static geo-tracking at specific times only. Across all these specifications, we find that the predictive performance of our model decreases relative to when all the data is available. Specifically, the prediction accuracy of models under frequency restrictions ranges from 75.16% to 77.21% in our data. However, the changes are relatively small and the models generally perform better than when geo-tracking data are not available at all.

Overall, our results suggest that ML models perform remarkably well even in settings where geo-tracking data are collected at longer intervals or restricted by users' age. In these cases, the loss in predictive accuracy ranges from 4.14% for age restrictions to 2.35-4.94% for frequency restrictions. We also find that, in our context, policies that restrict the geography within which businesses can collect data or that aggregate geo-tracking data to produce summaries of driving behaviors at the user-level, result in the largest decreases in predictive performance, which compares to that of models that do not user geo-tracking information at all. In these cases, the loss in predictive accuracy ranges from 10.89% under geographical restrictions to 14.15% under user-level summarization.

ROBUSTNESS CHECKS

Alternative Machine Learning Models

While our main models use an XGBoost machine learning algorithm, we also estimate alternative models, such as Logit, Random Forest, and Lasso to make sure the results are not specific to the type of 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 the Web Appendix Section Alternative Models. The findings are consistent with those reported in the previous section.

Alternative Time Periods

Our main model uses previous one week's geo-tracking information to predict one week ahead. However, restaurant managers may be interested in making farther out predictions for the next two or three weeks. In the Web Appendix Section Alternative Time Periods we report the outcome for predicting restaurant visits two- and three- weeks ahead. We also report the results of models that use a longer period of past 12 weeks instead of past one week of geo-tracking data. The results for these various temporal choices are consistent with the ones reported in the previous section.

Alternative Metrics

Our main results allow us to compare the accuracy and other performance metrics of various information sets (e.g., demographic, behavior, and geo-tracking information). However, they are less informative about which specific features within each set have higher predictive power. We compute the importance scores for the features in our preferred Model 5, and report them in Web Appendix Figure C1. The plot shows that the minimum weekly distance between the users' trips in the past week to the focal restaurant has the highest importance score. This is next followed by their home distance to the restaurant. Together, these two features have the highest predictive power followed by other behavioral and demographic features, including the ACS census block data. We also report the receiver operating characteristic (ROC) curve in Figure C2 to visually compare our models. The ROC curve plots true positive rate against the false positive rate at various thresholds. The greater the area under the curve, the higher is the overall predictive performance. This plots also shows that the models with geo-tracking information (purple and red dashed lines) outperform the models without it.

Outlier Drivers

It is possible that some users in our data may be restaurant delivery drivers or commercial taxi drivers. Such drivers are likely to have higher than typical levels of driving distances. To make sure our prediction results are not driven by learning these outlier drivers' patterns, we drop any user with driving distances of more than mean plus three times the standard deviation, and re-estimate our prediction model after excluding them. The results appear in Web Appendix Section Outlier Drivers and are consistent with those we reported earlier.

CONCLUSION

In recent years, many firms have started collecting geo-tracking data about consumers and using these data to inform their marketing and operational decisions (Clifford 2018). However, geo-tracking evokes privacy concerns among app users and regulators. For example, companies like Tim Hortons have attracted public scrutiny due to their geo-tracking practices (Austen 2022). Many emerging privacy regulations, such as the California Privacy Rights Act treat consumer location as sensitive data and restrict geo-tracking.

In this research, we examined two questions: First, to what extent are geo-tracking data useful relative to demographics and behavioral information for predicting consumers' future visits to a restaurant? Second, how does restricting geo-tracking data under various privacy policies impact the usefulness of these data in terms of their predictive performance?

We answered our research questions in the context of the restaurant industry using one application we identified through in-depth interviews with managers (Web Appendix Section Managerial Interviews), i.e., predicting consumers' visits to a restaurant one week ahead. Specifically, we used proprietary data from a safe-driving app in Texas with 120 million driving instances for 38,980 individual users to make predictions for the restaurant-weeks in our sample using a machine learning (ML) framework. From our interviews, we learned that predictions at a weekly frequency can allow managers to make better decisions about personnel and staffing (e.g., shifts), to plan weekly promotions, and to customize the restaurant experience. While the potential usefulness of geo-tracking data for businesses depends on the application under study, we focused on weekly prediction of visits, consistent with Sun et al. (2022) who predict consumers' online visits and purchases in future weeks as well.

Our research has several key results. First, we find that using geo-tracking data increases prediction accuracy by 3.1% when compared with using detailed demographic and behavioral information on past visits, and by 14.8% when considering only baseline demographics that do not include any location information. Second, imposing privacy restrictions that limit *what* geo-tracking data are tracked, *which* users are tracked, and *where* and *how frequently* users are tracked reduces the usefulness of geo-tracking for prediction by 2.35-14.15% relative to complete geo-tracking. Third, the extent of the decrease in predictive performance varies by the type of policy. Specifically, policies that restrict *what* data are geo-tracked (i.e., user-level summary of driving behaviors rather than geo-coordinates) and *where* where users are geo-tracked (i.e., within a few miles of a business location) are associated with the largest decreases in predictive performance and lead to a similar performance as that of models that do not use geo-tracking data. However, ML models perform remarkably well in our application even when geo-tracking data are collected with reduced tracking frequency or only for users above a certain age. Finally and importantly, models that use restricted geo-tracking still generally outperform models that do not use any geo-tracking information.

Managerial Implications

Our results have several implications for managers. First, our finding that geo-tracking data significantly improve the ability to predict consumers' visits each week implies that managers may be better off, when possible, by collecting and using these types of data to predict consumers' future actions.

Second, our finding that using geo-coordinates directly as input does not improve the model performance relative to models that use these data to infer distance measures implies that managers may not need to use detailed coordinate data and may focus on specific informative aspects of the data. Our finding is consistent with Sun et al. (2022), who find that using each location in a trajectory as a model input is less useful for predicting web visits than using less detailed information, such as categories of locations.

Third, our finding that using user-level summaries of driving behaviors results in lower prediction performance relative to using sensitive geo-coordinates, home- and trip- distance metrics, or even simply baseline demographic information implies that protecting consumers' privacy through these types of restrictions comes at a higher predictive cost for managers relative to other forms of restrictions.

Finally, our finding that models that use restricted geo-tracking data under various counterfactual policies still generally perform well suggests that collecting some geo-tracking information, even when restricted by policy or practical considerations, may be better for prediction purposes than not collecting these data at all. In fact, restricted geo-tracking could allow managers to manage these considerations while still getting the benefits of using such data for prediction. For example, managers could collect data lower frequency intervals using the first and last points of consumers' trip only, but still be better off than no geo-tracking at all while also saving in data storage and server costs.

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 the previous week's data for a different set of training users. Future research may address real-time predictions under privacy restrictions, e.g., for targeting mobile coupons.

Second, we acknowledge that the potential usefulness of geo-tracking data for businesses depends on the application under study. Our paper is able to speak to one application, which uses geo-tracking information from one week to predict customers' visits to a restaurant in the subsequent week. If data are available, future research can extend the usefulness of geo-tracking in other contexts and for other applications of interest. Given the heightened concerns with using personal consumer data, firms may also prefer to use such data in situations where it advances consumer interest, which is another potential avenue for future research (Soleymanian, Weinberg, and Zhu 2019).

Third, while we examine counterfactual policies drawn from the current policy landscape and privacy literature, there may be other regulations of interest we are not able to study. For example, most app users in our setting are above the age of 13, so we cannot comment on privacy protection for children. Future research, if data are available, can specifically examine these additional important policies.

REFERENCES

Acquisti, Alessandro and Hal R Varian (2005), "Conditioning Prices on Purchase History," *Marketing Science*, 24 (3), 367–381.

Austen, Ian "'A Mass Invasion of Privacy' but No Penalties for Tim Hortons," (2022) https://www.nytimes.com/2022/06/11/world/canada/tim-hortons-privacy-data.html.

Avenancio-León, Carlos F and Troup Howard (2022), "The Assessment Gap: Racial Inequalities in Property Taxation," *The Quarterly Journal of Economics*, 137 (3), 1383–1434.

Berman, Ron and Ayelet Israeli (2022), "The Value of Descriptive Analytics: Evidence from Online Retailers," *Marketing Science*, 41 (6), 1074–1096.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan (2015), "Gender Identity and Relative Income within Households," *The Quarterly Journal of Economics*, 130 (2), 571–614.

Binns, Reuben, Ulrik Lyngs, Max Van Kleek, Jun Zhao, Timothy Libert, and Nigel Shadbolt (2018), "Third Party Tracking in the Mobile Ecosystem," *Proceedings of the 10th ACM Conference on Web Science*, pages 23–31.

Bleier, Alexander, Avi Goldfarb, and Catherine Tucker (2020), "Consumer Privacy and the Future of Data-based Innovation and Marketing," *International Journal of Research in Marketing*, 37 (3), 466–480.

Chen, Tianqi and Carlos Guestrin "Xgboost: A Scalable Tree Boosting System," "Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining," pages 785–794 (2016).

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz (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, W Jason, Kinshuk Jerath, and Miklos Sarvary (2023), "Consumer Privacy Choices and (Un) Targeted Advertising along the Purchase Journey," *Journal of Marketing Research*, page 00222437221140052.

Clifford, Catherine "You can now head to McDonald's to get a Burger King Whopper for 1 cent — here's how," (2018) https://www.cnbc.com/2018/12/04/how-to-get-a-burger-king-whopper-for-1-cent-near-mcdonalds.html.

Danaher, Peter J (2023), "Optimal Microtargeting of Advertising," *Journal of Marketing Research*, 60 (3), 564–584.

Danaher, Peter J, Michael S Smith, Kulan Ranasinghe, and Tracey S Danaher (2015), "Where, When, and How Long: Factors that Influence the Redemption of Mobile Phone Coupons," *Journal of Marketing Research*, 52 (5), 710–725.

Dean, Grace "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," (2023) https://www.businessinsider.com/mcdonalds-chick-fil-a-apps-track-location-restaurant-mobile-order-2023-5.

Dubé, Jean-Pierre, Zheng Fang, Nathan Fong, and Xueming Luo (2017), "Competitive Price Targeting with Smartphone Coupons," *Marketing Science*, 36 (6), 944–975.

Goldberg, Samuel, Garrett Johnson, and Scott Shriver (2019), "Regulating Privacy Online: An Economic Evaluation of the GDPR," *Working paper.*

Goldfarb, Avi and Catherine Tucker (2012), "Shifts in Privacy Concerns," *The American Economic Review*, 102 (3), 349–353.

Gonzalez, Marta C, Cesar A Hidalgo, and Albert-Laszlo Barabasi (2008), "Understanding Individual Human Mobility Patterns," *Nature*, 453 (7196), 779–782.

He, Sherry, Brett Hollenbeck, Gijs Overgoor, Davide Proserpio, and Ali Tosyali (2022), "Detecting Fake-Review Buyers using Network structure: Direct Evidence from Amazon," *Proceedings of the National Academy of Sciences*, 119 (47), e2211932119.

Hoteit, Sahar, Stefano Secci, Stanislav Sobolevsky, Carlo Ratti, and Guy Pujolle (2014), "Estimating Human Trajectories and Hotspots through Mobile Phone Data," *Computer Networks*, 64, 296–307.

ICO, Information Commissioner's Office "Age Appropriate Design: a Code of Practice for Online Services," (2022).

Johnson, Garrett, Tesary Lin, James C Cooper, and Liang Zhong (2023), "COPPAcalypse? The Youtube Settlement's Impact on Kids Content," *Working paper.*

Johnson, Garrett, Scott Shriver, and Samuel Goldberg (2022), "Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR," *Management Science*, forthcoming.

Kim, Mingyung, Eric T Bradlow, and Raghuram Iyengar (2022), "Selecting Data Granularity and Model Specification using the Scaled Power Likelihood with Multiple Weights," *Marketing Science*, 41 (4), 848–866.

Klopack, Ben (2022), "One Size Fits All? The Value of Standardized Retail Chains," *RAND Journal of Economics*, forthcoming.

Klosowski, Thorin "The State of Consumer Data Privacy Laws in the US (And Why it Matters)," (2021) https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

Lamberton, Cait and Andrew T Stephen (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.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2015), "The Housing Market(s) of San Diego," *American Economic Review*, 105 (4), 1371–1407.

Lee, Jung Youn, Joonhyuk Yang, and Eric T. Anderson (2021), "Buying and Payment Habits: Using Grocery Data to Predict Credit Card Payments," *Working paper.*

Lei, Xiaoxia, Yixing Chen, and Ananya Sen (2023), "The Value of External Data for Digital Platforms: Evidence from a Field Experiment on Search Suggestions," *Available at SSRN.*

Li, Shaobo, Matthew J. Schneider, Yan Yu, and Sachin Gupta (2022), "Reidentification Risk in Panel Data: Protecting for K-Anonymity," *Information Systems Research*, forthcoming.

Lin, Tesary and Sanjog Misra (2022), "Frontiers: The Identity Fragmentation Bias," *Marketing Science*, 41 (3), 433–440.

Macha, Meghanath, Natasha Zhang Foutz, Beibei Li, and Anindya Ghose (2023), "Personalized Privacy Preservation in Consumer Mobile Trajectories," *Information Systems Research*, forthcoming.

McCabe, David "Anonymity No More? Age Checks Come to the Web.," (2021) https://www.nytimes.com/2021/10/27/technology/internet-age-check-proof.html.

Miller, Klaus M. and Bernd Skiera (2017), "Economic Loss of Cookie Lifetime Restrictions," *Working paper.*

Mohammed, Roweida, Jumanah Rawashdeh, and Malak Abdullah "Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results," "2020 11th international conference on information and communication systems (ICICS)," pages 243–248, IEEE (2020).

Naik, Nikhil, Ramesh Raskar, and César A. Hidalgo (2016), "Cities Are Physical Too: Using Computer Vision to Measure the Quality and Impact of Urban Appearance," *American Economic Review*, 106 (5), 128–32.

Neumann, Nico, Catherine E. Tucker, and Timothy Whitfield (2021), "Frontiers: How Effective is Third-party Consumer Profiling? Evidence from Field Studies," *Marketing Science*, 38 (6), 918–926.

Pappalardo, Luca and Filippo Simini (2018), "Data-driven Generation of Spatio-temporal Routines in Human Mobility," *Data Mining and Knowledge Discovery*, 32 (3), 787–829.

Pappalardo, Luca, Filippo Simini, Gianni Barlacchi, and Roberto Pellungrini (2022), "scikit-mobility: A Python Library for the Analysis, Generation, and Risk Assessment of Mobility Data," *Journal of Statistical Software*, 103 (1), 1–38.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer (2022), "Regulatory Spillovers and Data Governance: Evidence from the GDPR," *Marketing Science*, 41 (4), 746–768.

Rafieian, Omid and Hema Yoganarasimhan (2021), "Targeting and Privacy in Mobile Advertising," *Marketing Science*, 40 (2), 193–218.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15 (4), 321–340.

Singer, Natasha "F.T.C. Seeks 'Blanket' Ban on Meta's Use of Young Users' Data," (2023) https://www.nytimes.com/2023/05/03/technology/facebook-meta-ftc-data-ban-instagram.html.

Soleymanian, Miremad, Charles B Weinberg, and Ting Zhu (2019), "Sensor Data and Behavioral Tracking: Does Usage-based Auto Insurance Benefit Drivers?," *Marketing Science*, 38 (1), 21–43.

Song, Chaoming, Zehui Qu, Nicholas Blumm, and Albert-László Barabási (2010), "Limits of Predictability in Human Mobility," *Science*, 327 (5968), 1018–1021.

Sun, Chenshuo, Panagiotis Adamopoulos, Anindya Ghose, and Xueming Luo (2022), "Predicting Stages in Omnichannel Path to Purchase: A Deep Learning Model," *Information Systems Research*, 33 (2), 429–445.

Tau, Byron "Selling Your Cellphone Location Data Might Soon Be Banned in U.S. for First Time," (2023) https://www.wsj.com/articles/first-u-s-ban-on-sale-of-cellphone-location-data-might-be-coming-fbe47e53.

Tian, Longxiu, Dana Turjeman, and Samuel Levy (2023), "Privacy Preserving Data Fusion," *Working paper*.

Trusov, Michael, Liye Ma, and Zainab Jamal (2016), "Crumbs of the Cookie: User Profiling in Customer-base Analysis and Behavioral Targeting," *Marketing Science*, 35 (3), 405–426.

Valentine-De, Jennifer, Natasha Singer, Michael H Keller, and Aaron Krolik "Your Apps Know Where You Were Last Night, and They're Not Keeping it Secret," (2018) https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html.

Van der Maaten, Laurens and Geoffrey Hinton (2008), "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, 9 (11), 2579–2605.

Wang, Pengyuan, Li Jiang, and Jian Yang (2023), "EXPRESS: The Early Impact of GDPR Compliance on Display Advertising: The Case of an Ad Publisher," *Journal of Marketing Research*, forthcoming.

Wang, Yanwen, Chunhua Wu, and Ting Zhu (2019), "Mobile Hailing Technology and Taxi Driving Behaviors," *Marketing Science*, 38 (5), 734–755.

Wernerfelt, Nils, Anna Tuchman, Bradley Shapiro, and Robert Moakler "Estimating the Value of Offsite Data to Advertisers on Meta," (2022).

Yan, Shunyao, Klaus M Miller, and Bernd Skiera (2022), "How Does the Adoption of Ad Blockers Affect News Consumption?," *Journal of Marketing Research*, 59 (5), 1002–1018.

Yoganarasimhan, Hema (2020), "Search Personalization using Machine Learning," *Management Science*, 66 (3), 1045–1070.

Zhao, Yu, Pinar Yildirim, and Pradeep K. Chintagunta (2021), "Privacy Regulations and Online Search Friction: Evidence from GDPR," *Working paper*.

WEB APPENDIX A
MANAGERIAL INTERVIEWS

Table A1: Managerial Responses on How Predictions of Consumer Visits May be Useful, the Consumer Information they Collect/Wish to Collect, and at what Time Intervals

Manager Role	**Experience**	**Gist of Interview Responses**
Manager Jimmy John's (JJ)	5 years	Predicting consumer visits every week would help us with staffing and figuring out if we're overstaffed or understaffed. We can't fill staffing gaps immediately, so we receive staff applications on a weekly basis and look at their availability for the week. Customer demographics are important and will also tell us where our customers reside. Our app asks for an address for delivery, so we know how close customers live to JJ's store. Knowing consumers' location and visit patterns to other restaurants could help us target ads. I wouldn't prefer to keep data for more than a few weeks or six months at most to not alarm customers.
Manager Asian restaurant in Chicago	2-3 years	At our small Chinese restaurant in Chicago, 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 factors, such as the demographics and needs of our customers. However, I would not rely solely on the algorithm as it is important to stay in touch with what customers want. In terms of the frequency at which I would like to predict consumer visits, I would say weekly would be the most useful. This would help me to tailor our offerings and marketing strategies to better meet the needs of our customers.

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 i.e., where to locate based on what kind of demographics and competitors are there and what information we have on them.
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. Monthly predictions are fine, but for some decisions, weekly or more regular data could also help especially for scheduling. For example, if you know Thursday is going to be busy for family customers, hopefully you can increase your customer accounts too, you can move people in and out faster. 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 ordering supplies are planned on a weekly basis.
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 other people I work with have lives outside the restaurant, but usually a week is fine.

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. We include restaurant name and location only if the managers agreed to share it.

WEB APPENDIX B
USER-LEVEL SUMMARIZATION OF GEO-TRACKING DATA: 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 policy makers may restrict its use, in our counterfactuals 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, Hidalgo, and Barabasi 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 relate to the randomness of consumers' driving trajectories. This is important because as the randomness of driving behavior increases, the likelihood of an algorithm to be able to learn from past information to predict future visits decreases. Consumers whose driving patterns have 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 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, Wu, and Zhu 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 spatio-temporal mobility (Song et al. 2010).

The second set of features computed from geo-tracking data relate 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, Hidalgo, and Barabasi 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 most time during the training period.

Finally, we focus on the specific characteristics of each trip, for example, the number of quick stops (of $\leq$ 10 minutes) and long stops (of $\geq$ 60 minutes, Hoteit et al. 2014). We also include a measure of the proportion of trips completed by a user in various windows of time to account for specific time-of-day effects, and the average number of trips per week (Pappalardo and Simini 2018). The summary statistics of these features at the user-week level appear in Table B1.

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, Hidalgo, and Barabasi 2008).

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

where $\beta = 1.75 \pm 0.15$, $\Delta_{r_0} = 1.5$km, and cutoff distance of $k|_{D1} \approx 400$km and $k|_{D2} \approx 80$km.

Table B1: 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	5.814
Uncorrelated entropy	Variability of a user's visited locations based on probabilities of past visits	0.99
Real entropy	Variability of a user's visited locations based on probabilities and order of past visits	5.692
Radius of gyration (miles)	Characteristic distance traveled by a user	40.484
Unique days	Average no. of unique days of driving	3.77
Locations	Average no. of unique points in a user's trip trajectory	80.11
Max distance (miles)	Maximum distance traveled by users from their home	109.11
Short stops	No. of stops of ≤10 minutes	14.30
Short stops at restaurants	No. of stops at restaurants for ≤10 minutes	1.776
Short stops at unique restaurants	No. of stops at unique restaurants for ≤10 minutes	1.395
Long stops	No. of stops for ≥60 minutes	7.851
Long stops at restaurants	No. of stops at restaurants for ≥60 minutes	0.760
Long stops at unique restaurants	No. of stops at a unique restaurants for ≥60 minutes	0.585
Morning driving	Proportion of trips in the morning (before 11am)	0.33
Afternoon driving	Proportion of trips in the afternoon (11am to 5pm)	0.33
Evening driving	Proportion of trips in the evening (after 5pm)	0.34
Trip frequency	Average number of trips by a user	12.04

Notes: The summary geo-tracking features are computed for all users using their raw geo-coordinates each week. The reported numbers are averaged over all the user-weeks. 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.

1. Radius of gyration:

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

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

   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.

WEB APPENDIX C

ROBUSTNESS CHECKS AND ADDITIONAL ANALYSES

In this section, we present the robustness checks for alternative models, alternative time periods, outlier drivers, alternative probability thresholds, and alternative metrics. The tables report the Models 1-5 described in Table 2. Model 6 using geo-coordinates has worse performance than Model 5 and is computationally expensive, so we do not include it in these exercises.

Alternative Models

In this section, we present the main results using alternative models instead of XGBoost.

Table C1: Predictive Performance of Logit Models by Information Set

Model	Accuracy	Recall	Precision	F1 Score
Model 1	66.97%	56.70%	71.59%	61.77%
Model 2	74.09%	62.36%	82.04%	68.86%
Model 3	76.45%	62.69%	86.66%	70.33%
Model 4	78.16%	64.07%	89.33%	72.24%
Model 5	77.75%	62.70%	90.57%	71.28%

Table C2: Predictive Performance of Random Forest Models by Information Set

Model	Accuracy	Recall	Precision	F1 Score
Model 1	69.01%	53.71%	76.55%	61.26%
Model 2	74.50%	56.63%	87.79%	66.47%
Model 3	76.84%	61.70%	87.79%	70.18%
Model 4	78.72%	63.33%	91.07%	72.37%
Model 5	79.15%	64.02%	91.38%	72.99%

Table C3: Predictive Performance of Lasso Models by Information Set

Model	Accuracy	Recall	Precision	F1 Score
Model 1	67.35%	57.03%	72.06%	62.08%
Model 2	74.82%	64.77%	82.10%	70.06%
Model 3	78.16%	65.60%	87.49%	72.68%
Model 4	81.40%	71.18%	89.85%	76.85%
Model 5	79.82%	66.73%	90.56%	74.17%

Alternative Time Periods

In this section, we present the main results for predicting visits two- and three- weeks ahead instead of one, and using past 12 week's geo-tracking instead of previous one week.

Table C4: Predictive Performance of XGBoost for Predicting Visits in the Next Two Weeks using the Previous Week's Geo-Tracking

Model	Accuracy	Recall	Precision	F1 Score
Model 1	66.28%	53.82%	71.91%	59.45%
Model 2	77.64%	64.96%	86.50%	72.17%
Model 3	78.24%	65.54%	87.49%	72.69%
Model 4	79.38%	65.23%	90.91%	73.45%
Model 5	79.16%	64.81%	90.81%	73.05%

Table C5: Predictive Performance of XGBoost for Predicting Visits in the Next Three Weeks using the Previous Week's Geo-Tracking

Model	Accuracy	Recall	Precision	F1 Score
Model 1	65.98%	55.88%	69.95%	59.81%
Model 2	76.30%	63.36%	85.08%	70.39%
Model 3	76.73%	63.31%	86.03%	70.70%
Model 4	79.18%	66.01%	89.67%	73.27%
Model 5	79.28%	66.26%	89.67%	73.46%

Table C6: Predictive Performance of XGBoost for Predicting Visits in the Next One Week using Previous 12 Weeks' Geo-Tracking

Model	Accuracy	Recall	Precision	F1 Score
Model 1	68.50%	56.47%	75.17%	62.39%
Model 2	76.13%	59.93%	88.05%	69.24%
Model 3	76.70%	60.44%	89.14%	69.97%
Model 4	77.87%	61.70%	91.03%	71.12%
Model 5	78.01%	61.79%	91.24%	71.24%

Outlier Drivers

Table C7: Predictive Performance of XGBoost for Predicting Visits after Excluding Outlier Drivers

Model	Accuracy	Recall	Precision	F1 Score
Model 1	66.95%	51.58%	75.23%	59.44%
Model 2	75.10%	57.16%	88.88%	67.68%
Model 3	76.41%	59.15%	90.18%	69.53%
Model 4	77.76%	60.50%	92.52%	70.59%
Model 5	77.64%	60.16%	92.44%	70.39%

Alternative Probability Thresholds

Table C8: Predictive Performance of XGBoost for Predicting Visits using Alternative Threshold Probability of 0.68

Model	Accuracy	Recall	Precision	F1 Score
Model 1	64.11%	56.47%	75.17%	62.39%
Model 2	75.29%	59.93%	88.05%	69.24%
Model 3	75.99%	60.44%	89.14%	69.97%
Model 4	77.81%	62.76%	91.79%	72.19%
Model 5	77.83%	63.57%	92.05%	72.86%

Table C9: Predictive Performance of XGBoost for Predicting Visits using Alternative Threshold Probability of 0.39

Model	Accuracy	Recall	Precision	F1 Score
Model 1	68.15%	56.47%	75.17%	62.39%
Model 2	75.99%	59.93%	88.05%	69.24%
Model 3	76.53%	60.44%	89.14%	69.97%
Model 4	78.82%	62.76%	91.79%	72.19%
Model 5	79.29%	63.57%	92.05%	72.86%

Alternative Metrics: Variable Importance and Area under the Curve

Figure C1: Variable Importance Plot for Model 5 Across Restaurant-Weeks

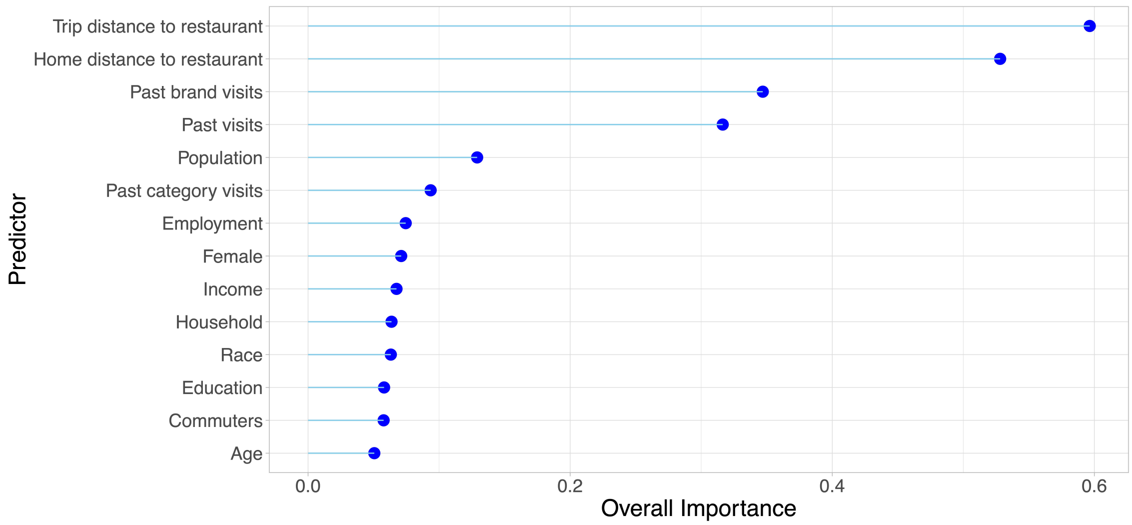
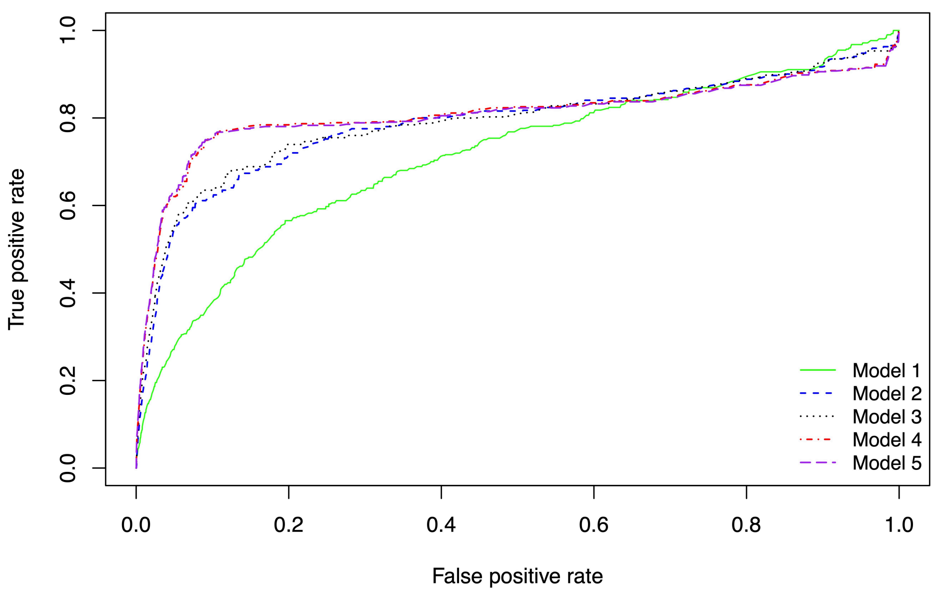


Figure C2: ROC Curves



Note: The area under the curve (AUC) metrics for Models 1-5 (described in Table 2) are 0.711, 0.788, 0.790, 0.809, and 0.807 respectively. Model 1 uses baseline demographics with no location information. Models 4 and 5 with geo-tracking information have a similar performance along the ROC curve and are generally superior to Models 1-3.

WEB APPENDIX D

CATEGORIZATION OF RESTAURANTS

Table D1: 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.	