Evaluating popular sentiment of electric vehicle owners in the United States with real-time data from mobile platforms

Omar Isaac Asensio[1,*], Kevin Alvarez[2], Arielle Dror[3], Emerson Wenzel[4], and Catharina Hollauer[5]

[1]School of Public Policy and Institute for Data Engineering & Science, Georgia Institute of Technology, Atlanta, GA 30332, USA, ORCID 0000-0003-2143-5022
[2]North Carolina State University, Raleigh, NC, 27695, USA
[3]Smith College, Northampton, MA 01063, USA
[4]Tufts University, Medford, MA 02155, USA
[5]Georgia Institute of Technology, H. Milton Stewart School of Industrial and Systems Engineering, Atlanta, GA 30332, USA
[*]correspondence to: asensio@gatech.edu

ABSTRACT

By displacing gasoline and diesel fuels, electric cars and fleets offer significant public health benefits by reducing emissions from the transportation sector. However, public confidence in the reliability of charging infrastructure remains a fundamental barrier to adoption. Using large-scale social data and machine learning based on 12,720 U.S. electric vehicle charging stations, we provide national evidence on how well the existing charging infrastructure is serving the needs of the expanding population of EV drivers in 651 core-based statistical areas in the United States. Contrary to predictions, we find that stations at private charging locations do not outperform public charging locations provided by government. We also find evidence of higher negative sentiment in the dense urban centers, where issues of charge rage and congestion may be the most prominent. Overall, 40% of drivers using mobility apps have faced negative experiences at EV charging stations, a problem that needs to be fixed as the market expands.

Global investment in electric vehicle charging infrastructure is estimated to reach $80 billion USD by 2025[1]. In the United States, this investment growth marks an expected transition in policy support at the federal level to more aggressive actions at the state and local level. The transportation sector is now the dominant source of CO_2 emissions in the United States[2]. By displacing gasoline and diesel fuels, vehicle electrification strategies have captured the attention of policy makers and analysts due to the large expected public health benefits associated with reduced air pollution and tailpipe emissions[3–5]. However, while current EV infrastructure policies have focused on increasing the quantity of charging stations to meet future growth[6], not much attention has been paid to the quality of charging services, particularly at the consumer level. Service reliability is a key risk in the public provision of EV charging services, and hence a critical barrier to large-scale technology adoption.

Some scholars contend that the private sector, under the right incentives, can more effectively deliver public fast charging services as needed. Other scholars argue that large public investments in fast charging infrastructure could crowd out private investments and lead to wasteful spending on charging locations that would have been built anyway, e.g. what economists refer to as inframarginal participation. Still, other scholars argue that public charging serves a public good, particularly if sufficient incentives do not exist for private entrepreneurs and organizations to invest locally. This debate on public versus private provision of environmental public goods and services has a long tradition in economics[7,8] and public management[9,10], with mixed empirical evidence on whether decentralized local provision is most effective.

Subjective perceptions about the quality and reliability of public charging infrastructure are critical to building range confidence among existing EV owners[11,12]. More importantly, popular sentiment about EV charging station experiences could be even more critical to potential buyers in the electric vehicle purchase decision, particularly for consumers in under-served communities.

A major challenge to evaluating whether the current EV charging infrastructure is meeting the needs of the public is in access to available monitoring data. This is because EV mobility data is largely user-generated and is often owned by private entities[13,14]. For example, in the United States, charging transaction records are typically managed by tens of thousands of individual station hosts—each with the ability to independently set prices and charging policies (subject to State rules)—with no central repository or reporting requirements across network providers. As a result, given these high monitoring costs, national evidence on the quality of service provision in EV infrastructure has been scant.

In this article, we analyze evidence of electric vehicle charging station experiences in both public and private spaces, and at major points of interest. We use machine intelligence to automatically classify user reviews in 651 core-based statistical areas in the United States. In doing so, we demonstrate the potential to use machine learning (ML) to substantially reduce data aggregation costs by automatically classifying unstructured user reviews into positive and negative station experiences as an indicator of performance. Based on market data from 2011-2015, we show how a convolutional neural network trained on large-scale social data learns domain-specific terms and in effect, approaches the accuracy of human experts for sentiment classification. We then use this data to evaluate popular sentiment and test hypotheses about service provision on a national scale.

We discuss performance in the context of prediction policy problems[15] related to EV infrastructure. We further discuss directions for the use of machine learning tools in the analysis of government service delivery in near-real time and with dynamically growing datasets.

Mobility Data

Mobile applications are changing the scale and techniques by which user data can be aggregated[16,17]. Digital platforms in mobile phones enable users to search, locate, and pay for transportation services in real time. Given the rise in smart phone use for transportation services, it is possible to analyze—subject to the necessary privacy protections—mobility decisions for large populations with digital infrastructure[18,19]. In the context of electric vehicles, charging station locator apps help lower information and transaction costs. Users can search for available EV charging stations, pay for charging sessions, and interact with other users by uploading station photos and writing station reviews for the EV community.

Here we analyze unstructured consumer reviews at 12,720 US charging station locations as provided by a popular EV charge station locator app. The data consists of 127,257 reviews from 29,532 EV drivers during the period from 2011 to 2015. This includes data aggregated from 10 major EV charging networks in the US.

Given the dynamically growing data size, it would be too costly for researchers or government analysts to hand classify these reviews for performance assessment. For example, at a rate of 100 reviews per hour, it would take a human expert about 32 work weeks to analyze reviews by hand. As a solution to this problem, we deploy a machine learning algorithm to automatically process unstructured reviews by taking advantage of social data in a digital platform and natural language processing. This approach allows us to reduce processing times for research evaluation to just minutes of computation. In the next section, we give a brief overview of innovations in natural language processing used to compare human ratings versus automated machine ratings for performance assessment.

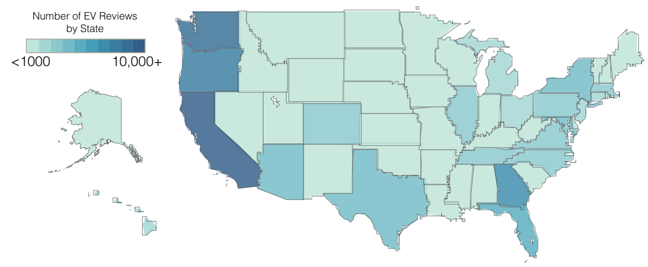

Figure 1. **U.S. Map of Active Charge Station Reviews.** This map shows the counts of electric vehicle charging station reviews per state from 2011 to 2015.

Natural Language Processing

Sentiment Classification Tasks

Sentiment classification is a classic problem in natural language processing, in which the overall polarity (positive/negative) of a body of text is predicted using a machine learning model. There have been a variety of approaches to this task over time. For example, in early work, Pang and Lee (2002) use bag-of-*N*-grams models with various algorithms (e.g. support vector machines (SVM), Naïve Bayes, and maximum entropy models). Scholars have argued that these simplified bag-of-*N*-grams models, which convert a recurring sequence of *N* items from sample text into word vectors, lose much of the information in the text that is useful for classification tasks[20]. For example, word order is lost, especially when using only unigrams, and so unstructured text that could have different meanings would receive the same vector space representation as long as they use the same words. In our domain with EV charging station reviews, this means that the two reviews: "no this charger is good", and "this charger is no good" would have identical representations. This is clearly a problem for sentiment classification. Most critically, the bag-of-*N*-grams approach fails to capture word semantics, and therefore cannot generalize across semantically-related, but distinct sequences of words[21] such as "it is working" or "it is functional." Another known issue with the bag-of-*N*-grams approach is that it often leads to high-dimensional document representations that tend to generalize poorly[22]. We therefore seek to implement more recent approaches that can capture word meanings and allow for lower dimensional document representations that can be effective for consumer charging reviews in EV transportation and mobility.

Convolutional Neural Networks

Recently, different types of neural networks have seen some success in sentiment classification tasks[21,23–25]. However, these algorithms need to be adapted and optimized for specific domains before they can be useful. For this study, we use a convolutional neural network (CNN) and build on a model architecture similar to that proposed by Kim (2014). We choose this approach as CNN-based classifiers have been

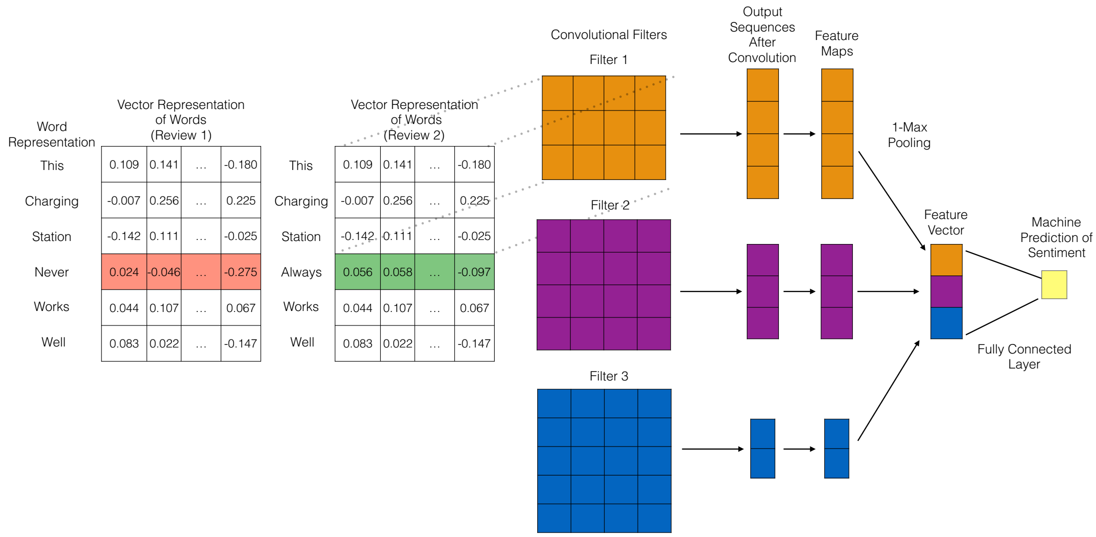

Figure 2. Model Architecture for the Convolutional Neural Network. The first stage shown depicts the matrix representation of the review text. Each row is a vector representation of a word, which captures information about word similarity. The second stage is the convolutional layer for feature extraction. These convolutional filters learn what words to look for in the reviews. There can be several convolutional filters per filter size such as filter heights of 3, 4 or 5. Next, we represent the output sequences after convolution. The next stage represents the feature map, which is the result of applying an activation function to the outputs of convolution. Finally, we apply max-pooling to capture the most important feature from each feature map, and concatenate these together. This feature vector is then input into a fully connected layer for classification.

shown to achieve state-of-the-art results for sentence level classification of short user generated texts, so we evaluate these advances for possible implementation in our domain. We compare the performance of our CNN classifier to two commonly used models, namely SVM and logistic regression, for which we use the bag-of-*N*-grams approach as baseline.

Convolutional neural networks first gained popularity in computer vision and have recently been demonstrated to be effective in several natural language processing tasks[26–29]. Briefly, each review text is first converted into a sequence of tokens, where each word is replaced with a number, such as the index of the word in the total vocabulary set. These tokenized sequences are padded with null tokens to all have the same length, and are then converted into a matrix (see Figure 2). Each row in the matrix is a vector representation of a word and is normally referred to as a word embedding. Similar words are closer together in the vector space than dissimilar words[21]. In our implementation, we use pre-trained *word2vec* word embeddings, which have been trained on approximately 100 billion words and phrases from Google news[20]. To capture domain-specific semantics, word embeddings are updated as the model is trained. A key innovation of the model architecture is that it flexibly allows for both unsupervised neural language learning from pre-trained word vectors, while also allowing for supervised learning of domain-specific terms through back propagation[23]. A summary of the model architecture is provided in Figure 2.

Results

Machine Classification

Using a convolutional neural network, we classify electric vehicle charging station experiences over a four year period from 2011 to 2015. We ask: how well do the machine intelligence predictions agree with human predictions? We know from our Cohen's κ=.84 achieved when building our training set that inter-rater agreement between human experts is high, but it is not perfect. As such, we note that binary sentiment classification in this domain is difficult, even for human experts. With this in mind, it is encouraging that the CNN classifier achieved a sentiment prediction accuracy of 84.1% when compared to human labels (Table 1). To further demonstrate the efficiency of our classifier, we also report precision and recall measures of 0.87 and 0.83, respectively. These results indicate an improvement over other commonly used baseline models for classification (e.g. SVM or logistic regression, as shown in Table 1). The use of neural language models to extract sentence level features have recently been shown in the analysis of short user reviews and texts[23,25]. Here we demonstrate state-of-art performance for classifying EV charging station reviews in the context of transportation and mobility.

In our series of experiments, we find that the convolutional neural network successfully identifies domain-specific patterns of natural language. For instance, a commonly used term that may be recognized by subject matter experts, but not necessarily by the general population, is the notion of

Table 1. Model Performance

Model	Accuracy (Percent)	Precision	Recall
CNN	**84.1**	**0.87**	**0.83**
SVM	76.5	0.78	0.79
LR	78.5	0.79	0.82

Comparison of our convolutional neural network (**CNN**) against baseline methods. For this comparison, we use support vector machine (**SVM**) and logistic regression (**LR**). For both of these baseline models, we use a bag-of-N-grams document representation with identical features.

Table 2. Descriptive Statistics, Public and Private

	Public	Private	Total
Positive	12,237	55,327	67,564
Negative	10,376	47,061	57,437
Total	22,613	102,388	125,001

Counts of machine classified reviews of binary sentiment by public and private ownership. 2,256 reviews were submitted in locations where it was impossible to discern whether it was public or private.

"ICE-ing." To be "iced" or "ICE'd" is an informal term that refers to cases in which an internal combustion engine vehicle is parked in a space normally reserved for EV drivers. ICE-ing is a common source of charge rage—the feeling drivers get when they are unable to find a charger—and often reflects negative sentiment as its represents a violation of a community norm. For example: *"Came here on a Sunday around 11:30am and every spot was ICEd,"* or *"I was iced by a blue Dodge Journey."* For non-experts, these reviews might lead to ambiguious classifications due to insufficient domain knowledge otherwise common to EV drivers. That artificial intelligence can detect ICE-ing in this context, and reach the accuracy of human experts, albeit in a matter of minutes of computation, is notable. With this illustrative example, we show how machine learning can be deployed to detect natural language associated with complex behavioral norms such as charging etiquette and other informal rules among a community of users. Such capabilities could also substantially reduce infrastructure evaluation costs and help equip utility managers and station operators with rapid response capabilities to improve service times. We suggest future research to explore further uses of machine intelligence to identify behavioral mechanisms related to charge rage, congestion and other station failures.

In the next section, we use our best prediction model to test common assumptions about charging behavior in public and private spaces and at key points of interest.

Sentiment Analysis

We find evidence of significant EV charging station use in all major US geographic areas (Figure 1). One could expect reviews data to reflect mainly positive sentiment, if the supporting EV infrastructure is working very well; or mainly negative sentiment, if station reviews are primarily a repository of complaints. Past research in marketing and psychology studies have shown that people tend to weigh negative information more than positive information during evaluation, which could suggest a negativity bias in the incidence of user reviews[30,31]. We know that consumers are more likely to share negative information across a platform when community ties are relatively weak, and both positive/negative messages when community ties are strong[32]. Given the high level of engagement that we see among the community of EV users, which has been described as a 'community of enthusiasts', it is not surprising that we find mixed valence in the informational exchanges across the platform. From our best performing model, we find that 68,876 reviews have a positive sentiment and 58,381 have a negative sentiment for a total of 127,257 classified reviews. In order to compare the incidence of negative sentiment for econometric analyses, we created a negativity index of conditional probabilities across stations, where 0 means all reviews at a given station location are positive, and 1 means all reviews at a given station location are negative. A higher predicted sentiment probability (closer to 1) would therefore not be desirable.

Table 3. Descriptive Statistics, Urban/Rural Type

	Rural	Urban Cluster	Urban Center	Total
Positive	4,932	4,990	58,954	68,876
Negative	2,322	2,320	53,739	58,381
Total	7,254	7,310	112,693	127,257

Counts of machine classified reviews of binary sentiment by geographic area type as defined by U.S. Census designations.

The mean predicted sentiment across all station reviews in both urban and non-urban areas is 0.39. This means that we predict a negative experience in consumer charging reviews by EV drivers who use charging station locator apps nearly 40% of the time. While this number might not seem high at first, it is analogous to predicting a negative experience four out of ten times that a driver goes to a gas station to fill up a car and writes about the experience. We argue that a greater focus on the quality of the electric vehicle charging experience is needed.

Discussion

Public versus Private Stations

Theory predicts that under the right incentives, private charging stations should outperform those run by government entities[8]. However, in practice, it is unclear whether sufficient incentives exist for private station hosts to maintain a high level of service quality, especially in the reselling of electric

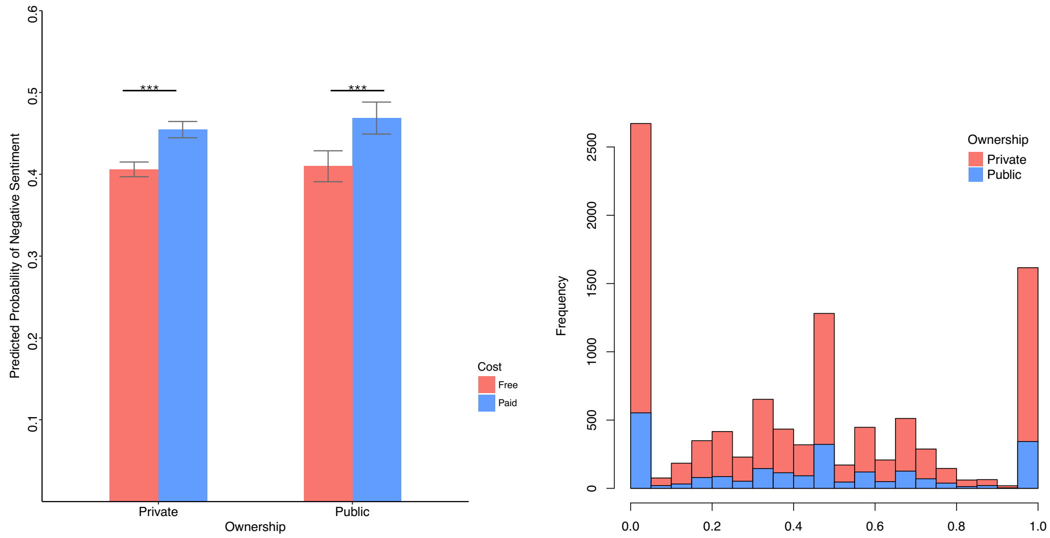

Figure 3. Predicted Probability of Negative Sentiment in Public and Private Spaces. Stations for pay have a higher probability of negative sentiment in both public and privately owned stations (p = 0.00). While we find differences in distribution (K-S test p-value p=.00), we find no statistically significant difference in mean sentiment between public and privately owned stations ($\mu_{\textbf{private}} = \mathbf{0.42}$, $\mu_{\textbf{public}} = \mathbf{0.44}$, $\mathbf{p = 0.14}$).

power, where capital cost recovery is often challenging and retail electricity prices are low. Here we test the hypothesis that private charging stations more effectively deliver charging services versus public stations provided by government. We consider a broad definition of public stations such as those that have been geolocated at points of interest (POI) that include government and municipal buildings, public libraries, rest areas, transit centers, public parks and visitor centers. We define private stations as those that have been geolocated at POIs that include hotels, retail/food establishments, shopping centers, healthcare facilities, workplaces and other non-residential locations. It is important to note that private charging locations do not imply that charging access is restricted to the public. This is because privately owned or managed EV stations are a key part of the publicly accessible charging infrastructure. Only about 1% of private charging locations that contain user check-ins and reviews on the popular mobile platform are considered restricted access.

In Table 2, we provide descriptive statistics for the raw counts of machine classified reviews at both public and private charging destinations. Contrary to expectation, we do not find a statistically significant difference in the mean predicted sentiment between public and private charging station locations (see Figure 3). We validated this finding by adjusting for factors driving selection to review, and other observable location characteristics. Observable location characteristics include the type and number of networks available (e.g. Chargepoint, Blink, SemaConnect, Aerovironment, EVgo, Tesla Supercharger, GE Wattstation, etc.), the type and number of connector plug technologies (e.g. J-1772, CHAdeMO, SAE Combo, Tesla supercharger, NEMA plug, etc.), and other driver-identifiable location attributes by point of interest. To mitigate possible unobserved heterogeneity, we also include the station rating as a proxy variable for unobserved quality attributes. Additionally, we considered a more narrow definition of public chargers with POI restricted to government-only stations in order to verify the result of statistical parity in consumer sentiment between public and private stations. Additional information is available in the Materials and Methods section.

We interpret this finding in two ways. First, our results indicate that private charging station locations do not outperform those that may be publicly owned or managed. Second, from a public choice perspective, our results provide some evidence that the private provisioning of EV charging services could be an alternative to large, publicly managed charging infrastructure. For example, one anonymous reviewer writes about the substitutability of a public charger for a private charger: *"Be careful if you plan on charging here, there are two cars that tend to bogart these chargers try the city hall chargers."* Evidence of statistical parity in consumer perceptions between public and private charging locations addresses a concern raised by the National Research Council on barriers to EV infrastructure growth[33]. We caution however, that our performance indicator captures popular sentiment from the standpoint of national consumer reviews, and not a power systems delivery perspective, which requires further investigation and integration with consumer data.

As shown in Figure 3, we find that paid charging stations consistently receive a higher proportion of negative reviews as compared to free stations. Not surprisingly, this result holds whether the station is in a public or private location.

This finding suggests users may have higher expectations for service reliability when paying for charging services. It is plausible that EV station location, whether public or private, may not be the dominant factor affecting service reliability. For example, stations located on public properties could have enjoyed the same (or perhaps even higher) level of operation and maintenance subscription services. In the next section, we use location microdata to investigate possible regional differences.

Urban versus Rural Areas

We compare the performance of stations in urban versus non-urban areas. For this analysis, we merge the geocoded station location data with geographical designations using standard U.S. Census definitions[34]. These include dense urban centers or urbanized areas (e.g. populations greater than 50,000), smaller urban clusters (populations between 2,500 and 50,000) and rural areas (populations less than 2,500). Based on the 2010 Census results, there are 486 urbanized areas and 3,087 urban clusters in the US. Here the differences in consumer sentiment are considerable. According to one view, EV drivers in areas with the lowest density of charging stations are more likely to experience issues of range anxiety, possibly leading users in these areas to publish more negative reviews. Therefore, from a supply side or resource availability hypothesis, areas with greater access to charging stations should garner the most positive reviews. Interestingly, we find the highest incidence of negative sentiment not in the rural areas or smaller urban clusters, but rather in the dense urban centers (see Figure 4). This is intriguing because approximately 89% of all user reviews are in urban centers, which reflects the state of the built infrastructure (Table 3).

After controlling for important station location and timing factors, we find that urban charging stations exhibit a statistically significant 12-15 percent increase in the negativity score as compared to non-urban charging locations (Table 6, Models IV-VI).

Our finding that EV charging stations in dense urban centers significantly underperform those in smaller urban clusters or rural areas where population and station densities are lower, could be indicative of a broader range of service quality issues in the largest EV markets. For example, many users report a lack of functional stations upon arrival, as well as issues related to congestion or lack of availability: *"some person is just pulling plugs without any note; i'll review footage on my security cam."* or *"Both spots taken. One by a Volt that's finished charging... Seriously time for more EVSE stations."* In Table 4 and Table 5, we summarize the predicted (negative) sentiment probabilities for both free and paid charging stations in the 18 largest U.S. metro areas and top 20 U.S. states by number of reviews. Although user reviews exist in all 50 states, the dominant source of activity is in California with 54,684 reviews or 43% of all consumer reviews in the dataset. The Los Angeles metro area for example, is the largest CBSA for charging station reviews through 2015 with 22,878 or 18% of all reviews in the dataset. The mean predicted negative sentiment in Los Angeles ranges from 0.43 to 0.56, which means a given user is nearly just as likely to report a negative consumer charging experience as a positive one. This is higher than the estimated U.S. national average sentiment score that we report of 0.39. Beyond resource availability, our results suggest that service reliability is already a key factor impacting consumer sentiment in the largest EV markets. Next, we evaluate the results by points of interest.

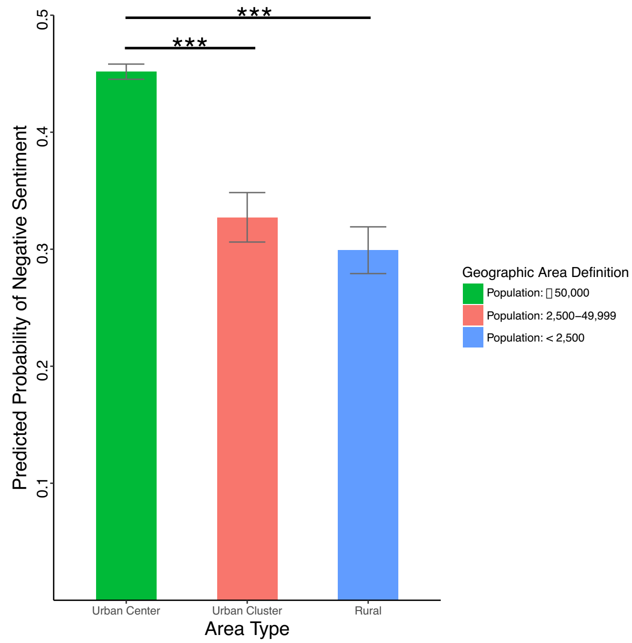

Figure 4. Predicted Probability of Negative Sentiment by Geographical Area. EV chargers in urban centers have a significantly higher probability of negative sentiment as compared to urban clusters and rural areas (p = 0.00).

Points of Interest

We summarize the results of our sentiment analysis by point of interest in Figure 5. The best performing private stations are at points of interest such as hotels/lodging destinations, restaurants and food establishments, and other service related POIs. This is to be expected as private station hosts in these locations often provide subsidized or complimentary EV charging services as a way to attract and cater to specific clientele. This suggests some incentive-based management practices. The highest performing POIs include parks and recreation as well as visitor centers, RV parks, and hotels/lodging. All of these POIs are associated with travel destinations and based on our reading of reviews at the locations, we believe that destination range anxiety could factor into positive reviews since drivers may be more willing to sacrifice some unsatisfactory conditions for the needed charge. Some of the worst performing POIs from the consumer reviews include car rental locations

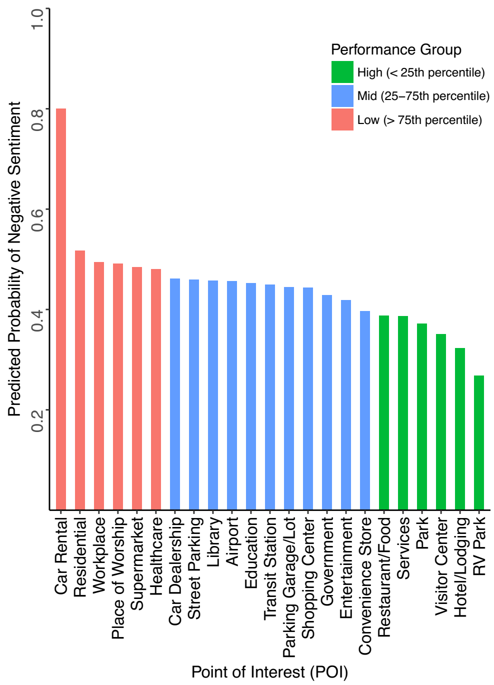

Figure 5. Predicted Probability of Negative Sentiment by Point of Interest

and car dealerships. This is consistent with recent evidence on car dealership practices at the point of sale, which have been documented as promulgating barriers to EV adoption[35]. Workplace and mixed use residential locations with ground floor retail establishments are also relatively low performing POIs. For example, many EV users at workplace and mixed use residential locations complain that EV stations can be difficult to access or that there is poor signage for public accessibility. We provide more detailed point estimates of POI location performance in Table 6.

Policy Implications

Large-scale social data from digital platforms can offer a number of benefits for research evaluation efforts, particularly in evaluating charging behavior in emerging EV infrastructure. We show that using computational tools such as neural network based language models, it is possible to develop more sophisticated performance indicators from unstructured data that offer the potential to update in near real-time. This is a major step forward from current practice that relies on indirect travel surveys or simulations, which can be costly and time-intensive to administer[36]. We argue that consumer reviews are an important input to learning about infrastructure challenges faced by current EV owners and should be prioritized when designing policies related to EV infrastructure access. This is particularly important in the design of "EV ready" or "EV capable" policies that require new buildings for example to install and maintain a certain number of EV charging stations using building codes and ordinances[37]. Such policies have grown in popularity in many cities such as Atlanta, Denver and Palo Alto, but largely without data or deliberation about service quality from existing EV drivers or other consumer groups.

Further, mobile apps can aggregate consumer data automatically at a large-scale, but independent station hosts and operators currently have little incentive to share data across network providers. Centralized reporting and secure data sharing across charging networks and utility jurisdictions would allow for more efficient resource planning decisions, particularly in resilience considerations between power systems delivery and emerging transportation infrastructure. One key exception to platform data sharing is the listing of EV charging stations maintained on an annual basis by the Alternative Fueling Station Locator tool hosted by the U.S. Department of Energy under the Clean Cities Program. While an invaluable tool, this digital repository of EV charging stations does not currently contain any real-time availability information, network status, user information or options for community engagement. We argue that policies to encourage greater information sharing as well as standardization in the quality of charging service delivery are necessary.

In this article, we show how machine intelligence can approach the accuracy of human experts for sentiment classification tasks, while showing promise for automatically learning domain specific terms in emerging EV infrastructure. Big data and machine learning techniques can automate the process of discovering new mobility patterns and detecting behavioral failures from consumer data, but they do not replace the need to keep a human-in-the-loop. It should be noted that, due to the classifier being trained by a human, the classifier is only as unbiased as the human rater. Not all consumer reviews can be relevant or actionable. Nevertheless, by expanding administrative records with real-time data from digital platforms, it is now possible to track station performance in both accessible and remote areas in ways that were not previously possible. Further, the use of machine learning tools as a pre-processing step for policy analyses can be helpful to determine quality requirements in both coverage and demand assessments related to transportation infrastructure[38]. This focus on big data and real-time mobility in digital platforms will become increasingly important over time, as driver incentives and other supply-driven policies designed to reduce externalities from electric vehicles do not typically address or affect driver behavior[39,40].

Finally, as EV infrastructure grows, we argue that it is not just the quantity of available charging stations that matter to consumers, but also the quality of the charging station experience. From our results, a key focus for quality improvement should be in the dense urban centers, where reports of ICE-ing, and a lack of available or functional stations are prominent and appear to drive negative consumer reviews. While paid stations consistently receive higher negative sentiment related to consumer expectations about the charging experience, community interactions also reveal emerging norms about charger etiquette and prosocial behavior primarily designed to help others in the community. However, further research is necessary to determine the most important mechanisms of user dissatisfaction in order to help resource allocation decisions or predict failures before they occur. This line of inquiry could pave the way towards the automatic detection of mobility decisions in near real-time, using the voice of the consumer as an input to gauging quality in charging service delivery. Over the next several years, we expect new investment of about \$80 billion USD in electric vehicle supply equipment[1]. From the perspective of consumer reviews, we argue that it is not enough to just invest money into increasing the quantity of available stations, it is also important to invest money into reliable infrastructure that actually works.

Materials and Methods

Training Data

In any supervised ML classification task, it is necessary to obtain ground truth labels. To generate these labels, two human raters served as experts for sentiment classification. Each human rater independently coded a set of reviews. These reviews are approximately balanced in polarity. In total, there are 8,953 hand-classified reviews in our training set. We achieved the best inter-rater reliability ($\kappa = 0.84, SE = 0.7$)[41] by treating this task as a two-class problem, meaning reviews reflect a binary sentiment (positive/negative). Neutral classifications are found to be very difficult for human rater tasks in this domain.

Selection of Hyper-parameters

We used various strategies to select our hyper-parameters. Building on prior literature, we selected 1-max pooling, dropout regularization[42] with a rate of .6, and a ReLU activation function in our convolutional layer, as these hyper-parameters have been shown to improve accuracy[43]. In particular, the dropout technique was implemented to prevent over-fitting[42]. In our implementation, we confirm that an L2 constraint had no discernible performance improvement and therefore we do not include it in our model[43]. Other hyper-parameters include a batch size of 128; learning rate of .001; filter heights of 3, 4, and 5; 100 filters for each filter height. Filter widths are 300, which are set to the dimensionality of the word embeddings. Our unit of analysis for each review is at the station level.

Measuring Outcomes of Interest

For a given charging station i and review period *year*, we define the *Negativity Score* as the count of negative reviews as a fraction of the total count of reviews:

$$\text{NegativityScore}_{i,year} = \frac{\text{Count of negative reviews}_{i,year}}{\text{Total count of reviews}_{i,year}} \quad (1)$$

By construction, the share of negative reviews at a charging station is normalized to lie in the unit interval $[0,1]$. Boundary observations of the dependent variable at 0 indicates that all reviews at a charging station are positive, and boundary observations at 1 indicates that all reviews at a charging station are negative. A higher negativity score is undesirable. We also group the charging stations by location group, g (i.e. there can be more than 1 station ID at a given location) and the year of the review to provide a rate of users leaving a review relative to the amount of use of the class of station. Users can check-in to the platform and leave a review which enters into the *Count of reviews*, or check-in without leaving a review, which enters into the count of other *Check-ins*. We define the *review rate* as

$$\text{ReviewRate}_{i,g,year} = \frac{\text{Count of reviews}_{\text{i,g,year}}}{\text{Count of reviews}_{i,g,year} + \text{Check-ins}_{\text{i,g,year}}} \quad (2)$$

Fractional response models

We used the outputs of the CNN classifier as a pre-processing step for econometric analysis of consumer sentiment. Given the limitations of linear estimation methods such as OLS for bounded dependent variables, we implemented a fractional response model (FRM) for the probability share data based on the quasi-maximum likelihood (QMLE) estimator[44,45]. We present some elements of FRM models as developed by Papke and Woolridge[44,45] and then apply it to machine learning outputs in mobility data. In the standard FRM setup, we are interested in the conditional expectation of the fractional response variable $y_{i,t}$ on a group-specific vector of explanatory variables $\mathbf{x}_{i,t}$ such as,

$$E(y_{i,t}|\mathbf{x}_{i,t}) = G(\mathbf{x}_{i,t}\theta), \quad i = 1, \ldots, N \quad (3)$$

where $G(\cdot)$ is a non-linear function such as the cdf that satisfies $0 \leq G(\cdot) \leq 1$, the fractional dependent variable is defined only on $0 \leq y_{i,t} \leq 1$, and θ is a parameter vector of interest. Observations at the extremes of the outcome distributions are estimated directly using the Bernoulli log-likelihood function, given by

$$l_{i,t}(\theta) \equiv y_{i,t} \log[G(\mathbf{x}_{i,t}\theta)] + (1 - y_{i,t}) \log[1 - G(\mathbf{x}_{i,t}\theta)] \quad (4)$$

In our dataset, some charging station reviews may be classified as all negative or all positive at a given location. Given the presence of boundary observations at 0 and 1, the pooled Bernoulli QMLE of θ does not require dichotimizations of the dependent variable and is computed as

$$\hat{\theta} = \arg\max_{b} \sum_{i=1}^{N} l_{i,t}(\theta) \quad (5)$$

We note that this approach overcomes 3 important limitations found in comparable methods. First, we account for the bounded nature of the data and do not assume a linear conditional mean or constant linear effects, which requires stronger assumptions for estimation. Second, commonly used log-odds methods are not well-defined for boundary values 0 and 1 present in the data, and often require ad-hoc adjustments such as arbitrarily chosen constants. Third, methods based on two-limit Tobit models may be appropriate for censored data with boundary observations at both limits, but its application to fractional data that is not defined outside the boundary limits is hard to justify. For a more detailed review of fractional regression models, see Ramalho et al.[46]

Selection Effects of Providing A Review

The decision to provide a review is a voluntary one. It conditions the interpretation of information developed by analyzing a sample of reviews. Charging station activity outside the digital platform is inherently unobservable. To address the selection effects, we attempt to explain the likelihood of giving a review as a function of characteristics of the charging location and timing. In Eq. (2), we normalize the empirical review counts by total platform engagements including user check-ins without reviews. In this way, we are able to adjust our estimates of the importance of explanatory variables on the empirical review rate by a measure of total charging station usage beyond review activity. For example, during the period of study, there are 276,744 total user check-ins on the platform, of which 127,257 contain reviews.

Our main estimating equation relates the review rate and negativity score as a function of one or more of the explanatory variables. This includes point of interest (POI) information, geographical area such as urban, suburban or rural, the type and count of networks available, the type and count of station connectors available, and our designation as public stations based on station geolocation. Due to data limitations, we could not adjust for car type of the driver as that information is voluntary, so we had a biased subsample. Additionally, we also tested specifications that included the proprietary station quality rating [numeric score 1-10] as a proxy for possible unobserved heterogeneity. We estimate the following general equation,

$$\begin{aligned} Outcomes_{i,g,year} &= \alpha_{i,year} + Public_{i,g} + POI_{i,g} \\ &+ Networks_{i,g,year} + Connectors_{i,g} + Rating_i \end{aligned} \tag{6}$$

In Table 6, we report the main results. The main drivers of the review rate include geographical area (whether urban or non-urban) and point of interest location information. We also find a statistically significant effect of the type and number of station connectors available, and the type of charging network such as the service provider, but not the number of networks available at a station location, which can range from 1 to 3 networks at a location ID. This suggests choice in charging network service provider is not yet a significant factor. Given our main interest in the public provision of charging services, we confirmed our finding of no significant effect of public locations (or more narrowly defined government only locations) on the review rate. This result is robust to our proxy for unobserved quality attributes as measured by the station quality rating provided to us by the platform provider (Table 6, Models II-III). Overall, for factors driving the selection to review, location matters; as does the network type, connector technology and other quality related factors. In Table 6, we do not show the point estimates for individual networks or plug types, but these results are available upon request from the authors.

In Table 6, we condition on all observable characteristics from our aggregate selection analysis to then compute the average partial effects for factors driving the negativity score. The analysis reveals that urban chargers accoubnt for a statistically significant 12-15 percent increase in the negativity score, as compared to non-urban charging stations (Table 6, Models IV-VI). Similarly, we also confirmed our finding of no statistically significant effect of private versus public stations, which is robust to both a broad and narrow definition of public stations, and unobserved quality factors.

In Table 7, we provide supplementary regression results comparing the performance of the FRM approach versus a standard OLS estimator. While we find the estimates to be qualitatively similar, we see that FRM generates more conservative estimates compared with OLS, which over-estimates the magnitude of the effects as expected.

Comments on Defining Public and Private Stations

In order to determine whether or not the chargers on public properties were also publicly owned and managed, we contacted a random sample of 170 public EV charging locations, stratified by network (Blink, ChargePoint, etc.). We then attempted to contact each location through a combination of email and phone calls to ask the following questions, “Are the charging stations at this property owned by the organization?” and “Are the charging stations at this property managed by the organization?”. We also contacted several major EV charging networks directly (e.g. SemaConnect, ChargePoint, GreenLots, and Blink network) to determine whether or not they operate/maintain charging units on behalf of their customers. For example, we found that while GreenLots network manages all of their stations on behalf of stations hosts, station owners from the other three major networks we contacted can decide whether or not they want to enter into a contract/warranty for servicing. Overall, we found four possibilities regarding station ownership and maintenance on public properties:

1. Stations are both owned and managed by public entities (such as those in Colton, California).
2. Stations are owned by public entities, but managed by private EV charging networks (such as the one at the Anaheim Intermodal Transit Center in Anaheim, California).
3. Stations are owned by public entities, but managed by a local contractor (such as the station at Roswell City Hall in Roswell, GA).
4. Stations are neither owned nor managed by public entities (such as the station at the Minnesota Department of Natural Resources in St. Paul, Minnesota).

After contacting 170 stations, we were able obtain answers to our management question at 32 locations. Of these 32 locations, 10 were managed by the public entity and 22 were either managed by an EV charging network or managed by a private company. We were also able to get answers to our ownership question at 23 locations. Of these 23 locations, the stations at 14 locations were owned by the public entity, and the stations at 9 locations were not. We believe that the management structure can potentially be an important driver of proper functioning of EV chargers and hence, the consumer experience. However, the managerial aspects of public versus private operation, while it is out of the scope of the current paper, we highlight as important differences for future research.

Study limitations

While we demonstrate gains using machine learning in this domain, there remain key areas for technical improvement. First, it may be necessary to increase the size of the training data to achieve higher convergence between human and machine classifications, especially in dynamically growing social datasets where topic categories may be broad. For reference, we calculated an alternative agreement score between the human predictions and machine predictions by treating the machine as a separate rater. The resulting $\kappa = .68$ suggests additional optimization could be necessary to increase reliability scores. However, due to computational complexity, it may be difficult to fully optimize all hyper-parameters to reach a global optimum. Future work can explore optimal filter sizes. For example, one approach proposed by Zhang and Wallace is to conduct "a line search over the single filter region size to find the 'best' single region size"[43]. This could be a promising approach to further improve accuracy in subpopulations or in training sets with different types of human raters. We leave this as future work for topic modeling.

Another limitation of our analysis is that while we are able to quantitatively evaluate sentiment from consumer reviews, additional information is needed to identify the psychological basis for negative charging experiences. It would be useful to develop topic classifications and accompanying training data with ground truth labels that describe the various sources of negative consumer experience. This might allow for deeper identification of mechanisms and algorithmic classification for policy analysis.

Code Availability

Computer code and algorithm replication materials have been deposited on Github under a general public license GPL-3.0-or-later at DOI: 10.5281/zenodo.1419830.

Data Availability

All data generated or analysed during this study are not publicly available due to privacy restrictions from the platform provider. These data are however available from the corresponding author upon reasonable request and with the permission of the platform provider.

References

1. Navigant. Market Data: EV Market Forecasts: Global Forecasts for Light Duty Plug-In Hybrid and Battery EV Sales and Populations: 2017-2016 (2017).
2. EPA. Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2016 (2018). EPA Document No. 430-R-18-003.
3. Michalek, J. J. *et al.* Valuation of plug-in vehicle life-cycle air emissions and oil displacement benefits. *Proc. Natl. Acad. Sci.* **108**, 16554–16558, DOI: 10.1073/pnas.1104473108 (2011).
4. Tessum, C. W., Hill, J. D. & Marshall, J. D. Life cycle air quality impacts of conventional and alternative light-duty transportation in the united states. *Proc. Natl. Acad. Sci.* **111**, 18490–18495, DOI: 10.1073/pnas.1406853111 (2014).
5. Holland, S. P., Mansur, E. T., Muller, N. Z. & Yates, A. J. Are there environmental benefits from driving electric vehicles? the importance of local factors. *Am. Econ. Rev.* **106**, 3700–3729, DOI: 10.1257/aer.20150897 (2016).
6. Li, S., Tong, L., Xing, J. & Zhou, Y. The market for electric vehicles: Indirect network effects and policy design. *J. Assoc. Environ. Resour. Econ.* **4**, 89–133, DOI: 10.1086/689702 (2017). https://doi.org/10.1086/689702.
7. Andreoni, J. Impure altruism and donations to public goods: A theory of warm-glow giving. *The Econ. J.* **100**, 464–477 (1990).
8. Kotchen, M. Green markets and private provision of public goods. *J. Polit. Econ.* **114**, 816–834 (2006).
9. Warner, M. & Amir, H. Managing markets for public service: The role of mixed public–private delivery of city services. *Public Adm. Rev.* **68**, 155–166, DOI: 10.1111/j.1540-6210.2007.00845.x (2008). https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6210.2007.00845.x.
10. Warner, M. & Hebdon, R. Local government restructuring: Privatization and its alternatives. *J. Policy Analysis Manag.* **20**, 315–336, DOI: 10.1002/pam.2027 (2001).
11. Carley, S., Krause, R. M., Lane, B. W. & Graham, J. D. Intent to purchase a plug-in electric vehicle: A survey of early impressions in large us cites. *Transp. Res. Part D: Transp. Environ.* **18**, 39–45 (2013).
12. Sovacool, B. K. & Hirsh, R. F. Beyond batteries: An examination of the benefits and barriers to plug-in hybrid electric vehicles (phevs) and a vehicle-to-grid (v2g) transition. *Energy Policy* **37**, 1095 – 1103, DOI: https://doi.org/10.1016/j.enpol.2008.10.005 (2009).
13. Asensio, O. I. & Walsh, S. E. Mobile apps for workplace charging: A big data field experiment in electric vehicles. *Acad. Manag. Glob. Proc. 2018* **Surrey**, 208, DOI: 10.5465/amgblproc.surrey.2018.0208.abs (2018).

14. Williams, B. & DeShazo, J. Pricing workplace charging: financial viability and fueling costs. *Transp. Res. Rec. J. Transp. Res. Board* 68–75 (2014).

15. Kleinberg, J., Ludwig, J., Mullainathan, S. & Obermeyer, Z. Prediction policy problems. *Am. Econ. Rev.* **105**, 491–95, DOI: 10.1257/aer.p20151023 (2015).

16. Asensio, O. I. & Delmas, M. A. Nonprice incentives and energy conservation. *Proc. Natl. Acad. Sci.* **112**, E510–E515, DOI: 10.1073/pnas.1401880112 (2015).

17. Asensio, O. I. & Delmas, M. A. The dynamics of behavior change: Evidence from energy conservation. *J. Econ. Behav. Organ.* **126**, 196 – 212, DOI: https://doi.org/10.1016/j.jebo.2016.03.012 (2016).

18. Alexander, L., Jiang, S., Murga, M. & González, M. C. Origin–destination trips by purpose and time of day inferred from mobile phone data. *Transp. Res. Part C: Emerg. Technol.* **58**, 240–250, DOI: 10.1016/j.trc.2015.02.018 (2015).

19. Gonzalez, M. C., Hidalgo, C. A. & Barabasi, A.-L. Understanding individual human mobility patterns. *Nature* **453** (2008).

20. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. & Dean, J. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 3111–3119 (2013).

21. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444, DOI: 10.1038/nature14539 (2015).

22. Le, Q. & Mikolov, T. Distributed representations of sentences and documents. In *International Conference on Machine Learning*, 1188–1196 (2014).

23. Kim, Y. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746–1751 (Association for Computational Linguistics, Doha, Qatar, 2014).

24. Socher, R. *et al.* Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, 1631–1642 (2013).

25. dos Santos, C. & Gatti, M. Deep convolutional neural networks for sentiment analysis of short texts. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 69–78 (2014).

26. Yih, W.-t., Toutanova, K., Platt, J. C. & Meek, C. Learning discriminative projections for text similarity measures. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning*, 247–256 (Association for Computational Linguistics, 2011).

27. Shen, Y., He, X., Gao, J., Deng, L. & Mesnil, G. Learning semantic representations using convolutional neural networks for web search. In *Proceedings of the 23rd International Conference on World Wide Web*, 373–374 (ACM, 2014).

28. Kalchbrenner, N., Grefenstette, E. & Blunsom, P. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188* (2014).

29. Collobert, R. *et al.* Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **12**, 2493–2537 (2011).

30. Sen, S. & Lerman, D. Why are you telling me this? an examination into negative consumer reviews on the web. *J. Interact. Mark.* **21**, 76–94, DOI: 10.1002/dir.20090. https://onlinelibrary.wiley.com/doi/pdf/10.1002/dir.20090.

31. Mizerski, R. W. An attribution explanation of the disproportionate influence of unfavorable information. *J. Consumer Res.* **9**, 301–310 (1982).

32. Chu, S.-C. & Kim, Y. Determinants of consumer engagement in electronic word-of-mouth (ewom) in social networking sites. *Int. J. Advert.* **30**, 47–75, DOI: 10.2501/IJA-30-1-047-075 (2011). https://doi.org/10.2501/IJA-30-1-047-075.

33. Board, T. R. & Council, N. R. *Overcoming Barriers to Deployment of Plug-in Electric Vehicles* (The National Academies Press, Washington, DC, 2015).

34. US Census. 2010 Census Urban and Rural Classification and Urban Area Criteria (2010).

35. Gerardo Zarazua de Rubens, L. N. & Sovacool, B. K. Dismissive and deceptive car dealerships create barriers to electric vehicle adoption at the point of sale. *Nat. Energy* **3**, 501–507, DOI: 10.1038/s41560-018-0152-x (2018).

36. Rezvani, Z., Jansson, J. & Bodin, J. Advances in consumer electric vehicle adoption research: A review and research agenda. *Transp. Res. Part D: Transp. Environ.* **34**, 122 – 136, DOI: https://doi.org/10.1016/j.trd.2014.10.010 (2015).

37. DOE. Plug-In Electric Vehicle Deployment Policy Tools: Zoning, Codes, and Parking Ordinances (2015). Alternative Fuels Data Center, Technology Bulletin August 2015.

38. DOE. National Plug-In Electric Vehicle Infrastructure Analysis (2017). Office of Energy Efficiency and Renewable Energy, DOE/GO-102017-5040.

39. DeShazo, J. R. Improving incentives for clean vehicle purchases in the united states: Challenges and opportunities. *Rev. Environ. Econ. Policy* **10**, 149–165, DOI: 10.1093/reep/rev022 (2016).

40. DeShazo, J., Sheldon, T. L. & Carson, R. T. Designing policy incentives for cleaner technologies: Lessons from california's plug-in electric vehicle rebate program. *J. Environ. Econ. Manag.* **84**, 18 – 43, DOI: https://doi.org/10.1016/j.jeem.2017.01.002 (2017).

41. Cohen, J. A Coefficient of Agreement for Nominal Scales , A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **20**, 37–46, DOI: 10.1177/001316446002000104 (1960).

42. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *The J. Mach. Learn. Res.* **15**, 1929–1958 (2014).

43. Zhang, Y. & Wallace, B. A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. *arXiv:1510.03820 [cs]* (2015). ArXiv: 1510.03820.

44. Papke, L. E. & Wooldridge, J. M. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *J. Appl. Econom.* **11**, 619–632, DOI: 10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1.

45. Papke, L. E. & Wooldridge, J. M. Panel data methods for fractional response variables with an application to test pass rates. *J. Econom.* **145**, 121 – 133, DOI: https://doi.org/10.1016/j.jeconom.2008.05.009 (2008). The use of econometrics in informing public policy makers.

46. Ramalho, E. A., Ramalho, J. J. & Murteira, J. M. Alternative estimating and testing empirical strategies for fractional regression models. *J. Econ. Surv.* **25**, 19–68.

Acknowledgements

We thank the generous support of the Anthony and Jeanne Pritzker Family Foundation, the Sustainable LA Grand Challenge, and the Civic Data Science REU program at Georgia Tech (NSF Award No. IIS-1659757). We are grateful to Ellen Zegura and Chris Le Dantec for feedback and discussions. For valuable research assistance, we thank Mary Elizabeth Burke and Soobin Oh. Special thanks to Norman Hajjar.

Author contributions statement

O.I.A. directed the research and wrote the paper; A.D., E.W., and K.A. developed code, analyzed data and wrote the paper; K.A. and C.H. implemented algorithms and performed experiments. All authors reviewed the manuscript.

Additional information

Competing Interests

The authors declare no competing interests.

Table 4. Probability of Negative Sentiment for Top 20 States in the United States

State	Public			Private			No. of Reviews
	Free	Paid	p-value	Free	Paid	p-value	
California	0.44	0.47	0.10	0.48	0.50	0.11	54,684
Washington	0.46	0.44	0.70	0.38	0.43	0.06	7,830
Oregon	0.42	0.46	0.72	0.34	0.41	0.02	7,027
Georgia	0.46	0.44	0.83	0.38	0.48	0.00	6,623
Florida	0.36	0.41	0.39	0.46	0.44	0.54	4,420
Maryland	0.39	0.39	0.96	0.42	0.39	0.50	3,541
Arizona	0.44	0.47	0.90	0.36	0.50	0.00	3,365
New York	0.43	0.40	0.66	0.38	0.44	0.12	2,894
Texas	0.53	0.45	0.34	0.42	0.50	0.02	2,615
Virginia	0.48	0.40	0.46	0.38	0.46	0.05	2,588
Pennsylvania	0.42	0.55	0.30	0.40	0.45	0.24	2,536
North Carolina	0.40	0.45	0.64	0.39	0.50	0.06	2,181
Colorado	0.40	0.51	0.28	0.34	0.36	0.78	2,161
Illinois	0.53	0.56	0.76	0.48	0.44	0.33	2,105
Massachusetts	0.38	0.52	0.12	0.40	0.42	0.58	2,074
Tennessee	0.52	0.52	0.97	0.42	0.49	0.19	1,983
Michigan	0.34	0.35	0.94	0.39	0.28	0.03	1,488
Ohio	0.31	0.49	0.19	0.40	0.48	0.21	1,443
New Jersey	0.45	0.47	0.82	0.41	0.44	0.72	1,390
Hawaii	0.59	0.69	0.69	0.56	0.58	0.72	1,259

Results of t-tests for free and paid stations by public and private ownership in the top 20 states by numbers of reviews in the United States.

Table 5. Probability of Negative Sentiment for 18 Core Based Statistical Areas in the United States

State	Public			Private			No. of Reviews
	Free	Paid	p-value	Free	Paid	p-value	
Los Angeles-Long Beach-Anaheim, CA	0.49	0.52	0.43	0.56	0.53	0.14	22,878
San Francisco-Oakland-Hayward, CA	0.44	0.45	0.83	0.55	0.51	0.23	8,951
Atlanta-Sandy Springs-Roswell, GA	0.46	0.55	0.31	0.46	0.49	0.44	5,442
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.50	0.42	0.39	0.45	0.44	0.85	3,452
Phoenix-Mesa-Scottsdale, AZ	0.12	0.45	0.19	0.42	0.53	0.01	2,863
New York-Newark-Jersey City, NY-NJ-PA	0.53	0.38	0.10	0.46	0.42	0.32	2,060
Chicago-Naperville-Elgin, IL-IN-WI	0.46	0.62	0.15	0.49	0.44	0.28	1,781
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.42	0.43	0.95	0.47	0.54	0.24	1,586
Boston-Cambridge-Newton, MA-NH	0.43	0.58	0.11	0.42	0.46	0.36	1,438
Dallas-Fort Worth-Arlington, TX	0.47	0.57	0.52	0.45	0.59	0.01	1,139
Nashville-Davidson-Murfreesboro-Franklin, TN	0.60	0.55	0.75	0.41	0.42	0.86	1,082
Denver-Aurora-Lakewood, CO	0.64	0.60	0.87	0.40	0.37	0.65	1,042
Detroit-Warren-Dearborn, MI	0.49	0.38	0.29	0.45	0.43	0.75	792
Minneapolis-St. Paul-Bloomington, MN-WI	0.33	0.60	0.02	0.39	0.33	0.38	658
Austin-Round Rock, TX	0.54	0.25	0.03	0.32	0.37	0.47	508
Hartford-West Hartford-East Hartford, CT	0.36	0.45	0.69	0.45	0.48	0.85	471
Kansas City, MO-KS	0.29	0.54	0.43	0.36	0.29	0.54	488
Chattanooga, TN-GA	0.22	0.42	0.37	0.51	0.65	0.41	292

Results of t-tests for free and paid stations by public and private ownership in 18 CBSA in the United States.

Table 6. Main Results

	Review Rate			Negativity Score		
	FRM	*FRM*	*FRM*	*FRM*	*FRM*	*FRM*
	(I)	(II)	(III)	(IV)	(V)	(VI)
Geographical Area						
Urban	-0.021**	-0.038***	-0.039***	0.149***	0.123***	0.122***
	(0.008)	(0.007)	(0.007)	(0.013)	(0.012)	(0.012)
Non-Urban	0.025**	0.016*	0.016	0.016	0.004	0.004
	(0.012)	(0.010)	(0.010)	(0.016)	(0.014)	(0.014)
Type of Location						
Public	-0.010	-0.012		0.010	0.008	
	(0.013)	(0.015)		(0.015)	(0.012)	
Government			-0.023			0.004
			(0.015)			(0.013)
Station Characteristics						
Number of Connectors	-0.082***	-0.074***	-0.074***	-0.011***	-0.005	-0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Number of Networks	-0.011	-0.012	-0.012	0.030*	0.020	0.020
	(0.014)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)
Quality Rating		-0.042***	-0.042***		-0.058***	-0.058***
		(0.002)	(0.002)		(0.002)	(0.002)
Points of Interest						
Residential	0.063*	0.019	0.014	0.076***	0.020	0.019
	(0.035)	(0.041)	(0.041)	(0.025)	(0.019)	(0.019)
Shopping	-0.107***	-0.101***	-0.105***	0.043***	0.048***	0.046***
	(0.011)	(0.011)	(0.011)	(0.014)	(0.012)	(0.012)
Restaurants	-0.017	-0.014	-0.018	-0.003	0.000	-0.002
	(0.014)	(0.013)	(0.012)	(0.017)	(0.014)	(0.015)
Healthcare	0.040**	0.026	0.022	0.022	0.008	0.006
	(0.017)	(0.016)	(0.016)	(0.018)	(0.016)	(0.016)
Hotel and Lodging	0.058***	0.047***	0.043***	-0.070***	-0.082***	-0.084***
	(0.012)	(0.012)	(0.012)	(0.015)	(0.013)	(0.013)
Workplace	0.027*	0.017	0.012	0.005	-0.007	-0.009
	(0.015)	(0.013)	(0.013)	(0.015)	(0.014)	(0.014)
Supermarket	-0.076***	-0.079***	-0.083***	0.083***	0.082***	0.081***
	(0.012)	(0.013)	(0.013)	(0.017)	(0.015)	(0.015)
Car Dealership	0.031***	0.019*	0.015	0.042***	0.032***	0.031**
	(0.010)	(0.010)	(0.010)	(0.013)	(0.012)	(0.012)
Education	0.055***	0.038**	0.025	0.032*	0.010	0.015
	(0.015)	(0.019)	(0.017)	(0.019)	(0.013)	(0.014)
Entertainment	0.007	0.011	0.007	0.025	0.028	0.026
	(0.017)	(0.016)	(0.016)	(0.020)	(0.019)	(0.019)
Convenience and Gas Station	-0.001	-0.005	-0.009	0.015	0.012	0.010
	(0.011)	(0.011)	(0.011)	(0.016)	(0.014)	(0.014)
Transit Station	-0.041***	-0.026	-0.043***	0.018	0.034**	0.040**
	(0.016)	(0.016)	(0.014)	(0.018)	(0.015)	(0.016)
RV Park	0.186***	0.143***	0.139***	-0.085***	-0.154***	-0.156***
	(0.021)	(0.017)	(0.017)	(0.029)	(0.035)	(0.035)
Outdoor	0.007	0.009	-0.007	-0.032*	-0.027	-0.021
	(0.027)	(0.027)	(0.026)	(0.019)	(0.018)	(0.019)
Airport	0.017	0.022	0.006	-0.006	0.003	0.010
	(0.018)	(0.018)	(0.016)	(0.040)	(0.036)	(0.036)
Services	0.005	0.013	0.008	-0.035	-0.023	-0.025

Table 6. Main Results

	Review Rate			Negative Score		
	FRM	*FRM*	*FRM*	*FRM*	*FRM*	*FRM*
	(I)	(II)	(III)	(IV)	(V)	(VI)
	(0.022)	(0.019)	(0.019)	(0.025)	(0.024)	(0.024)
Place of Worship	-0.051	-0.025	-0.029	0.018	0.055	0.053
	(0.081)	(0.063)	(0.063)	(0.026)	(0.038)	(0.038)
Shopping Center	0.012	0.024	0.020	-0.055***	-0.017	-0.018
	(0.085)	0.065)	0.065)	0.021)	0.022)	0.022)
Library	0.021	0.018	0.002	0.034	0.029	0.036
	(0.019)	(0.024)	(0.022)	(0.026)	(0.024)	(0.024)
Street Parking	-0.008	-0.016	-0.032	0.075***	0.061**	0.067**
	(0.025)	(0.024)	(0.022)	(0.029)	(0.029)	(0.030)
Visitor Center	-0.044	-0.029	-0.033	-0.013	0.002	0.000
	(0.028)	(0.026)	(0.026)	(0.028)	(0.026)	(0.026)
Car Rental	0.235***	0.175***	0.171***	0.303***	0.221**	0.219**
	(0.042)	(0.036)	(0.036)	(0.084)	(0.089)	(0.089)
Clustered SE	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	127,257	127,257	127,257	127,257	127,257	127,257
R^2	0.117	0.235	0.235	0.049	0.120	0.120

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7. Comparison of FRM and OLS Estimators

	Review Rate		Negative Score	
	FRM	*OLS*	*FRM*	*OLS*
	(I)	(II)	(III)	(IV)
Geographical Area				
Urban Area	-0.039***	0.078***	0.122***	0.254***
	(0.007)	(0.015)	(0.012)	(0.018)
Non-Urban Area	0.015	0.131***	0.004	0.129***
	(0.010)	(0.017)	(0.014)	(0.019)
Type of Location				
Public	0.031	0.032	0.028	0.014
	(0.042)	(0.058)	(0.031)	(0.036)
Location Attributes				
Number of Connectors	-0.074***	-0.050***	-0.005	0.002
	(0.004)	(0.004)	(0.004)	(0.005)
Number of Networks	-0.012	0.211***	0.020	0.276***
	(0.017)	(0.029)	(0.016)	(0.030)
Quality Rating	-0.042***	-0.033***	-0.058***	-0.039***
	(0.002)	(0.003)	(0.002)	(0.004)
POI controls	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes
Number of Observations	127,257	127,257	127,257	127,257
R^2	0.236	0.372	0.120	0.750

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$