Titre: Title:	Which variables influence electric vehicle adoption?
Auteurs: Authors:	Hamed Naseri, Owen Waygood, Zachary Patterson, & Bobin Wang
Date:	2024
Type:	Article de revue / Article
Référence: Citation:	Naseri, H., Waygood, O., Patterson, Z., & Wang, B. (2024). Which variables influence electric vehicle adoption? Transportation, 38 pages. https://doi.org/10.1007/s11116-024-10525-1

Document publié chez l'éditeur officiel
Document issued by the official publisher

Titre de la revue: Journal Title:	Transportation
Maison d'édition: Publisher:	Springer Nature
URL officiel: Official URL:	https://doi.org/10.1007/s11116-024-10525-1
Mention légale: Legal notice:	This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s11116-024-10525-1

Which variables influence electric vehicle adoption?

Hamed Naseri[1*], E.O.D. Waygood[2], Zachary Patterson[3], Bobin Wang[4]

[1]Department of Civil, Geological, and Mining Engineering, Polytechnique Montréal, Montréal, Canada, hamed.naseri@polymtl.ca

[2]Department of Civil, Geological, and Mining Engineering, Polytechnique Montréal, Montréal, Canada, owen.waygood@polymtl.ca
[3]Concordia Institute for Information Systems Engineering, Concordia University, Montréal, Canada, zachary.patterson@concordia.ca
[4]Department of Mechanical Engineering, Université Laval, Québec, QC, G1V 0A6, Canada, bobin.wang@gmc.ulaval.ca

* Corresponding Author: Hamed Naseri

Which variables influence electric vehicle adoption?

Abstract

Understanding the factors that will influence people's preferences for Electric Vehicles (EVs) over Internal Combustion Engine Vehicles (ICEVs) is crucial. A discrete choice experiment was designed and administered as an online survey resulting in 1,077 completed questionnaires. This study examined the influence of over 83 variables on preferences for EVs. As well, previous studies have used tailpipe emissions only to present GHG information, but in this study lifecycle GHG emissions of vehicles are presented. Five ensemble learning techniques and two interpretation techniques were employed to investigate individual decisions regarding selecting between EVs and ICEVs. The results demonstrate that when lifecycle emissions are presented, financial impacts are the principal influences on predicting preference for an EV over ICEV. Following the financial impacts are existing preferences for EVs and attitudes related to climate change. How the emissions are presented was the 12th and 9th most influential factor for BEVs and PHEVs respectively.

Keywords: Electric vehicle; EXtreme Gradient Boosting; Partial Dependence Plot; SHapley Additive exPlanations.

1. Introduction

The transportation sector contributes to nearly 24% of CO_2 emissions worldwide (Austmann 2021). Electric Vehicles (EVs) are being promoted to help reduce CO_2 emissions in the transportation sector. Many governments have applied new policies and financial incentives to promote EVs as well as develop advanced industrial technologies for EVs (Wu et al. 2021). As such, EVs have been subsidized (Li and Wang 2023; Lu et al. 2022), the charging station locations of EVs have been optimized (Zhang et al. 2019), fast-charging protocols have been developed (Makeen et al. 2022), the charging infrastructure of EVs have been improved (Unterluggauer et al. 2022), taxes have been imposed on ICEV users (Bonilla et al. 2022).

Although considerable effort has been taken to promote EVs, the share of EVs in the vehicle market is quite limited. The number of vehicles around the world is estimated to be 1.446 billion units (Abid et al. 2022), while the number of EVs around the world reached just forty million in 2023 (Ritchie 2024). Hence, the share of EVs is still tiny (approximately 1% of international vehicle stock), and promoting EVs needs to receive more attention (Austmann 2021).

The parameters impacting the willingness to pay for EVs should be determined to increase the share of EVs in the vehicle markets. Different variables can potentially influence EV preferences, which need to be investigated to promote EVs. Considering many variables in the modeling increases the problem's complexity, and a powerful method should be applied. In this case, machine learning methods can be helpful. However, powerful machine techniques are mainly black-box, and their results cannot be used for policy-relevant issues (Naseri et al. 2022a). Therefore, they should be synchronized with interpretation techniques to address this issue.

The objectives of this study are as follows:

- Developing an accurate model to predict individuals' engine choices (between EVs and gasoline vehicles).
- Comparing the performance of different ensemble learning methods to determine which technique can predict individuals' engine choices with the maximum accuracy.
- Testing the influence of many variables on the intention to buy EVs, including those that have not been tested.
- Detecting the most efficient vehicle label to maximize EV preferences.
- By synchronizing interpretation techniques with the most accurate ensemble learning method, we aim to provide results that are directly applicable to policy-making purposes, emphasizing the practical relevance of our research.

2. Background

Many studies have investigated the influence of various parameters on willingness to buy EVs. Table 1 summarizes some of those studies. As can be seen, the influence of socio-demographic variables, vehicles' properties, individual's attitudes and perceptions on the willingness to buy EVs have been generally tested.

Table 1. A summary of studies that investigated EV preferences.

Reference	Region	Tested variables	Significant variables	Method
Hidrue et al. (2011)	United States	Socio-demographic variables and vehicle attributes	Age, education, fuel price, vehicle size, level of concern about the environment, the possibility of installing an	Latent class random utility model

			electrical outlet at home, and annual driving mileage	
Ramos-Real et al. (2018)	Spain	Socio-demographic variables, car ownership, mobility characteristics, level of familiarity with technologies, and the level of awareness about the environment	Mobility patterns, environmental awareness, income, and openness to new technologies	Logit and probit models
Ling et al. (2021)	China	Customers' attitudes and socio-demographic variables	Gender, household income, purchase budget, inclination level to ICEV, and previously driving EVs	Multilevel logistic regression
Lashari et al. (2021)	South Korea	Socio-demographic variables, customers' perceptions and attitudes	Economic and environmental perceptions, house type, gender, age, and home location	A regression tree and a binary logistic regression
Simsekoglu and Nayum (2019)	Norway	Psychological factors and information	Perceived behavioral control and subjective norm	Principal component analysis and hierarchical multiple regression
Tunçel (2022)	Turkey	Customers' attitudes	Attitudes toward EVs	Structural Equation Modelling

Vafaei-Zadeh et al. (2022)	Malaysia	Individuals’ attitudes	Subjective norms, environmental self-image, and perceived behavioral control	Partial Least Square path modelling

The influence of GHG information presentation on EV preferences has been also tested (e.g., (Wang et al. 2023; Wang et al. 2021)). Considering 16 variables, an individual’s Climate Change Stage of Change (CC-SoC) was the most influential factor, followed by the price ratio between EVs and ICEVs (Naseri et al. 2023). CC-SoC is a measure capturing attitudes and behavior with respect to personal climate emissions, which was developed based on the Transtheoretical Model (Naseri et al. 2021c). How the GHG emissions were presented (or framings) was the 6th most influential factor. Using discrete choice analysis, previous research (Wang et al. 2023) estimated the willingness-to-pay for emissions by different framings. That research demonstrated how the WTP could reach over $C 600/tonne with different framings. As such, consideration of that influence is important. However, in that previous research, tailpipe emissions were used, which gives an unfair and unrealistic advantage to EVs, considering that climate change is a global, not regional, problem.

Although the impact of many variables on the intention to buy EVs has been examined, there are still many variables that might be influential on sustainable behavior, but their effects on EV purchase intention need to be tested. As such, social media usage (i.e., type of social media, frequency of usage, and reason for usage) was significant in sustainable purchases (Ali et al. 2022; Pop et al. 2020; Tazeen and Mullick 2023). As another example, morality can impact sustainable behavior (Lasarov et al. 2022). Previous experiences can improve individuals’ perceived behavioral control (Liu et al. 2020a). GHG information presentation can significantly increase the willingness to pay for EVs (Daziano et al. 2021). Descriptive norms can lead to behavioral changes (Ryoo et al. 2017). Moreover, public transit users and cyclists are more likely to buy EVs when incentives are given (Rudolph 2016). However, such variables need to receive more attention when promoting EVs.

To determine which variables impact the willingness to purchase EVs, many techniques have been applied, such as Maximum-likelihood Estimation (Li et al. 2013), Regression (Brase 2019), Discrete Choice Models (Wang et al., 2021), Structural Equation Modelling (Krishnan and Koshy 2021), and machine learning (Sobiech-Grabka et al. 2022). Recently, machine learning techniques have been widely used in different transportation problems because their prediction accuracy is generally more than conventional methods of travel behavior (Pham et al. 2022; Salas et al. 2022). Machine learning techniques can detect non-linear relations between the independent and dependent variables (Dong et al. 2022). Moreover, these methods are efficient for modeling large-scale datasets (Naseri et al. 2021a).

Hence, machine learning techniques have been recently applied to investigate which factors influence the intention to buy EVs. For instance, Dixit and Singh (2022) applied a Logistic Regression to predict who will buy EVs in India and spot which parameters significantly impact buying an EV. Socio-demographic variables and vehicle characteristics were considered in the developed model. The results suggested that gender, level of concern about the environment, age, income, vehicle cost, vehicle performance, EV range, and running cost were significant parameters on the intention to buy EVs. Higueras-Castillo et al. (2021) used Least Square Support Vector Machines to determine the factors impacting the intention to buy EVs in Spain, considering sociodemographics, vehicle properties, and incentives. The results indicated that electric range, incentives, and vehicle reliability were top variables in the intention to buy EVs.

Sobiech-Grabka et al. (2022) compared the performance of different machine learning techniques, including Random Forest, classification and regression trees, K-nearest neighbor, and Support Vector Machine, to predict individuals' willingness to buy EVs. Random Forest obtained the highest prediction accuracy, and its results showed that attitudes toward EVs and ICEVs had the most significant relative influence on the intention to buy EVs. Similarly, the performance of some machine learning methods (e.g., Gradient Boosting Machine, Random Forest, Extreme Gradient Boosting Machine, and Generalized Linear Model) on EV adoption prediction was tested (Bas et al. 2021b). Gradient Boosting could reach a higher prediction accuracy. A feature importance analysis was performed using the most accurate technique. Agreement with the role of EVs in the mobility system, household income, attitude toward the reliability of EVs, and attitude toward climate change were the variables with the most substantial relative influence on the intention to buy EVs.

Using the same survey, the performance of different machine learning techniques (i.e. Artificial Neural Networks, Support Vector Machines, Gradient Boosting Models, Deep Neural Networks, Extremely Randomized Forests, and Distributed Random Forests) in predicting individual engine preferences (EV and ICEV) were compared based on prediction accuracy (Bas et al. 2021a). Support Vector Machine slightly outperformed other machine learning methods in accurately predicting EV buyers. Since Support Vector Machine could not present the most important variables, Random Forest was applied to detect the variables' relative influence on the intention to buy EVs. The five most influential variables on the intention to buy EVs were the respondents' county and the preference of the following vehicle engine type, EV reduced tax, purchase price, and electric range, respectively.

As can be perceived from the literature, a limited number of variables (i.e., socio-demographic variables, customers' attitudes, climate change beliefs and attitudes, and vehicle attributes) are generally applied to develop models predicting EV buyers. On the other hand, the influence of important variables, such as social media usage, GHG information presentation, previous or planned future actions to reduce GHG emission, and opinions about other sustainable transportation modes, on the intention to buy EVs have not been extensively tested in previous studies. In terms of modeling, machine learning methods can apply many variables in the modeling process. However, powerful machine techniques are mainly black-box, and their results cannot be used for policy-relevant issues. To address this issue, ensemble machine learning methods are generally used to rank variables based on feature importance (Sun et al. 2022). However, these methods could not demonstrate the relation between these variables (e.g., directly, indirectly, linearly, quadratically, etc.) and the dependent variable (i.e., the intention to buy EVs).

To address these issues, this investigation attempts to develop a framework to predict who is more likely to choose EVs over ICEVs and determine which parameters strongly affect the intention to buy EVs. Hence, a Discrete Choice Experiment (DCE) is generated to implement an online survey. A set of new labels is designed to investigate the influence of lifecycle GHG emissions information on the intention to buy EVs and identify the optimal vehicle labels to promote EVs. Building on previous studies, many new variables are tested based on their influence on the intention to buy EVs or "green consumerism", including questions on social media usage, morality, previous or planned future actions to reduce GHG emission, descriptive norms, and opinions about other sustainable transportation modes. As such, a technique to handle such a large number of potential influences is required. Different machine learning techniques are employed, and the most accurate machine learning technique on the EV preference prediction problem will be determined. Then, SHapley Additive exPlanation (SHAP) will be used to recognize the priority of variables according to their influence on the EV preference. Finally, Partial Dependence Plot (PDP) is used to illustrate the relation between top variables and the intention to buy EVs.

2. Methodology

This study aims to determine the most influential variables in the intention to buy EVs and investigate which groups of people are more likely to buy EVs in Canada. The methodology flowchart of this study is shown in Figure 1. As can be seen, this study includes three sections: data collection, ensemble learning process, and analyzing variables. In the data collection process, a DCE and new vehicle labels are designed, and they are applied to implement an online survey. Consequently, five ensemble learning methods are applied for the modeling, they are tuned, and the most accurate one is detected. Subsequently, the most accurate classifier and SHapley Additive exPlanations (SHAP) are used to rank variables based on their influence on the intention to buy EVs. Ultimately, Partial Dependence Plot (PDP) is utilized to investigate the influence direction of top-ranked variables on EV purchase likelihood.

In this section, designing DEC and the data collection are initially explained. Then, the applied machine learning techniques are presented. Consequently, the techniques used for analyzing the influence of variables on the intention to buy EVs (i.e., SHAP and PDP) are presented.

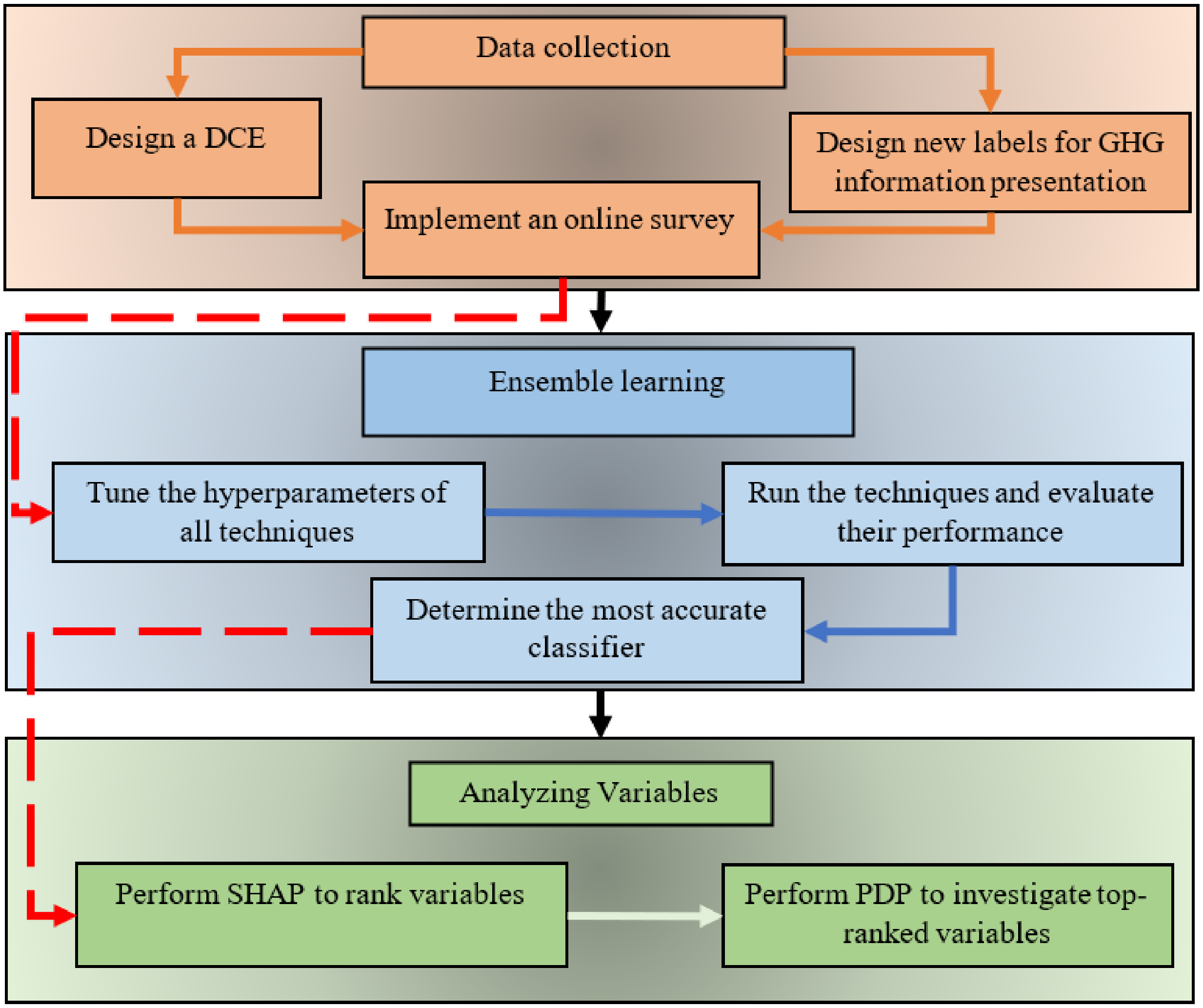


Figure 1. The methodology flowchart

2.1. Collecting data

A DCE was designed to administrate an online survey and collect the dataset. The survey was implemented in the Fall of 2022, and it included participants from six Canadian provinces: British Columbia (BC), Alberta (AB), Saskatchewan (SK), Manitoba (MB), Ontario (ON), and Quebec (QC). Since this was a survey about car purchases, all participants had a driving license and were over 18 years old (the legal driving age in Canada). One of the considered provinces (Quebec) is predominantly francophone, so the survey was offered in English and French. The survey included some trap questions to detect those who did not pay attention to all questions during the survey, and they were excluded from the final dataset. Ultimately, there were 1077 participants in the dataset. In this part, the DCE design is first presented. This is followed by the presentation of variables collected in the survey.

2.1.1. Discrete Choice Experiment

A DCE was designed to investigate the influence of different variables on the intention to buy EVs. Respondents needed to complete different choice tasks (hypothetical situations) in the designed DCE. The attributes of two choices (an EV and an ICEV) were presented in each choice task. Each participant needed to select an EV or an ICEV in different situations based on their preferences. In the DCE, the underlying experiment and precise model were identical to those of Wang et al. (2021). A total of 12 different choice tasks were required to be done by each respondent in the DCE. Six of these tasks were selecting between a Battery Electric Vehicle (BEV) and an ICEV. In the other six tasks, the participants needed to select between a Plug-in Hybrid Electric Vehicle (PHEV) and an ICEV.

In each choice task, the properties of vehicles (e.g., purchase price, GHG emissions, fuel cost) were changed. Since 1077 participants remained in the final dataset and each did six choice tasks for each EV type (i.e., BEV and PHEV), the number of data observations for each EV type was 1,077×6=6,462. The behavior of BEV-oriented and PHEV-oriented groups is different in Canada (Ferguson et al. 2018). Therefore, the influence of variables on the intention to buy each EV type (i.e., BEV and PHEV) was analyzed using two different models. That is, a separate model was developed to determine the impact of different variables on the intention to buy BEVs.

Another model was generated to investigate the influence of variables on the intention to buy PHEVs.

The choice attributes were the purchase price, yearly fuel or electricity cost, life cycle CO_2 emissions, and EV range. The attributes were changed in each choice task, and the participants needed to choose between an EV and an ICEV based on their preference. Fuel or electricity costs were calculated based on an annual driven distance of 20,000 kilometers since this value is generally considered by Natural Resources Canada (NRCan) to design vehicle labels. DCE was performed in Ngene using a D-efficient design. DCE was designed based on the attributes' levels shown in Table 2.

Table 2. The levels of attributes used in DCE.

Options	Purchase price (CAD)	Monthly fuel or electricity cost (CAD)	GHG emissions (kg CO_2/month)	EV range (km)
ICEV	[22000; 26000; 30000]	[900; 1420; 1720]	[210; 270; 306]	——
PHEV	[30000; 38000; 48000]	[600; 840]	[168; 186]	[40; 85; 203]
BEV	[40000; 48000; 56000]	[444; 568]	[132; 150]	[240; 423; 600]

The attributes shown in Table 2 were used in the vehicle labels. Eight labels were applied to present GHG information, indicating in Figure 2. The first label (Figure 2 (a)) is currently applied in Canada, called NRCan label. NRCan label presents the tailpipe emission of vehicles. Nonetheless, it has been proven that the tailpipe emission cannot accurately reflect overall emissions because many other processes (e.g., vehicle production, maintenance, operation, recycling, and battery production) generate GHG emissions (Kosai et al. 2022). Hence, the life cycle GHG emissions of vehicles were calculated according to De Souza et al. (2018), and life cycle emissions were shown in other labels (Figure 2 (b) to (h)).

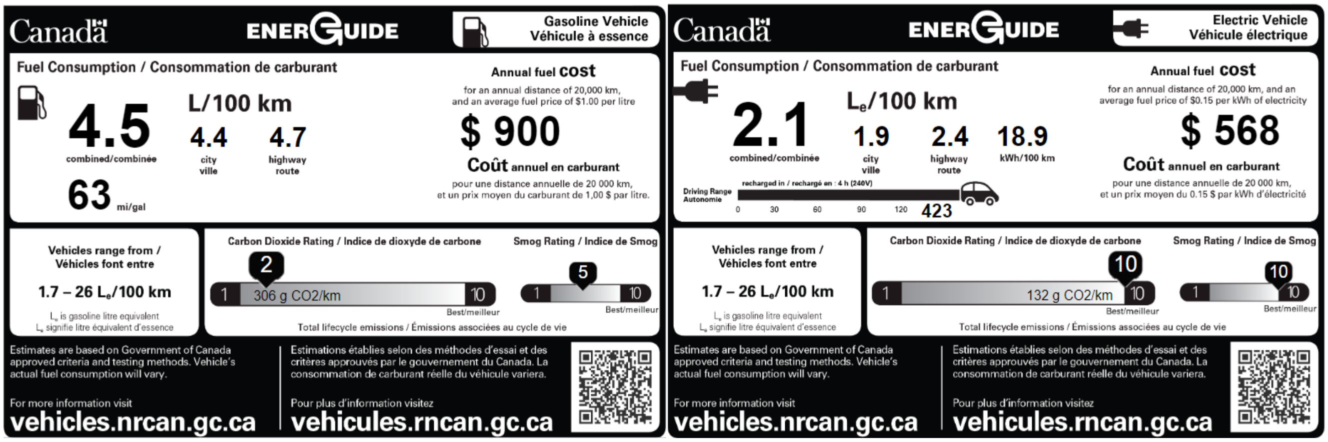
The NRCan label has a 1 (worst) to 10 (best) scale indicating the carbon dioxide rating of vehicles. The carbon dioxide rating ranged from 6 to 9 for ICEVs in NRCan. For instance, it is 7 in Figure 2 (a). This value is ten for both BEVs and PHEVs since NRCan considers only tailpipe emissions. The carbon dioxide rating was redesigned for the other treatments that present life cycle emissions (e.g., life cycle NRCan). The carbon dioxide rating for life cycle emissions was estimated to be between 2 and 6 for ICEVs based on their life cycle emissions. For PHEVs, a range from 7 to 8 was estimated and for BEVs, a range from 9 to 10 was found based on their life cycle emissions. The carbon dioxide rating values were redesigned due to two reasons. First, the carbon dioxide rating difference between BEVs and PHEVs is significantly higher when life cycle emission is considered than the tailpipe emission. Second, it is possible to examine the influence of the carbon dioxide rating range on the individual's decisions.

The second treatment (NRCan label with life cycle emissions; Figure 2 (b)) applies the frame of NRCan but presents the life cycle GHG emissions rather than the tailpipe emissions. The third treatment (emojis; Figure 2 (c)) was the best label for maximizing the willingness to pay for EVs in the previous study in Canada (Xun et al. 2022). To test whether the location and predominance of the GHG scale has an impact on choice, the label shown in Figure 2 (d) was developed. An important psychological distancing occurs with respect to climate change where people think of the problem as being in the future or affecting other regions of the world (Fielding et al. 2014). Researchers have found that framing climate change as something happening now in your neighborhood and affecting people like you can encourage climate action (McDonald et al.

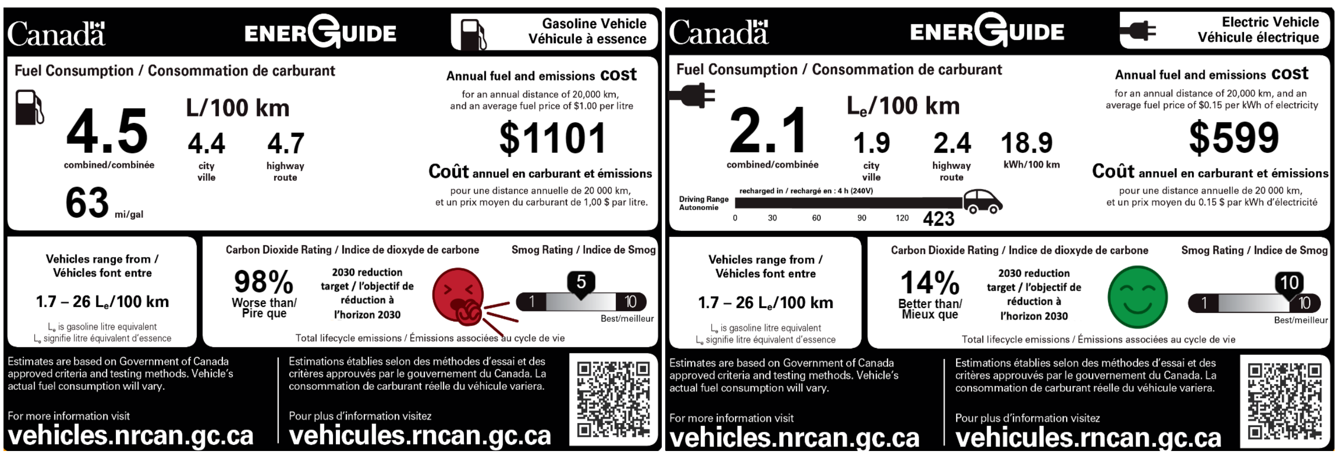
2015). Hence, treatments five to seven (Figure 2 (e, f, g)) were designed according to Canada's current consequences of climate change. Current consequences of climate change in Canada include widespread wildfires and devastating floods (Government of Canada 2021; Wang et al. 2020). Finally, the last treatment was designed to incorporate patriotism (Canadian maple leaf) and purity (the health of the leaf is increasing poor with higher emissions) based on research from social psychology relating to moral foundations and climate change (Feinberg and Willer 2013).

The goal of the government of Canada at the time of study was to reduce annual GHG emissions by 40% from 2005 by 2030 (Government-of-Canada 2021). Therefore, the values shown in treatments three and eight were designed regarding the 40% GHG reduction goal. Previous research in Canada had used a 30% reduction (e.g., Wang et al., 2021).

(a) Treatment 1: NRCan (tailpipe emission)



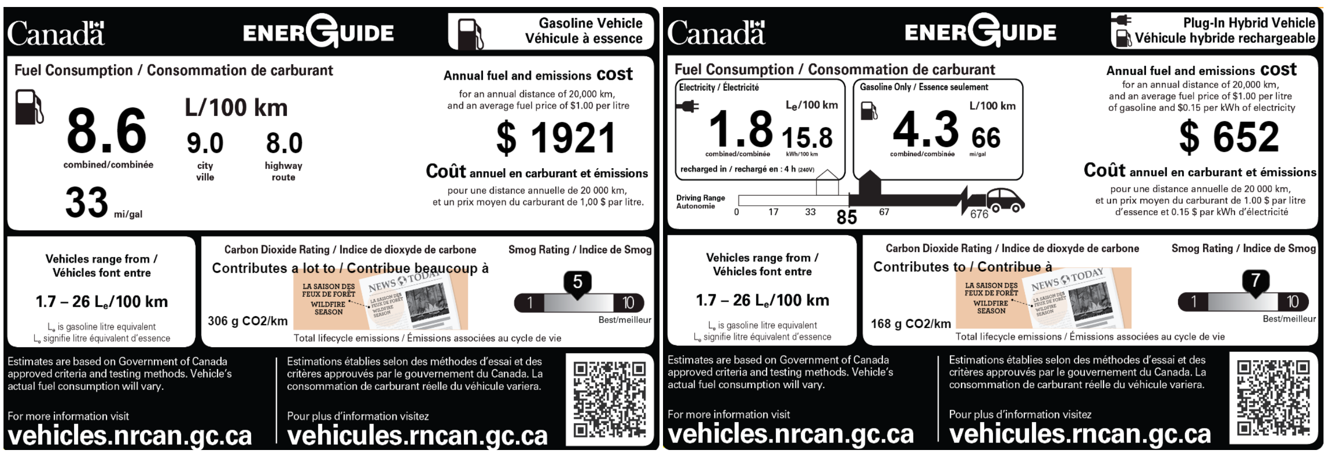
(b) Treatment 2: NRCan (life cycle emissions)



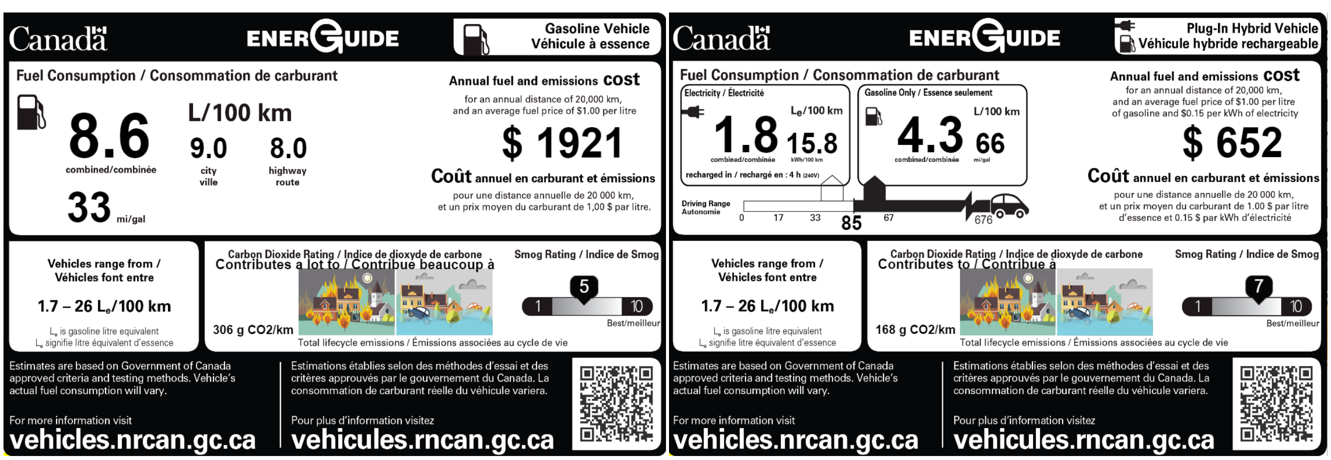
(c) Treatment 3: emojis

(d) Treatment 4: different layout of label

(e) Treatment 5: flood



(f) Treatment6: fire



(g) Treatment7: disaster

(h) Treatment 8: leaves

Figure 2. The applied labels for presenting GHG information.

2.1.2. Variables Tested

As mentioned, conventional models to predict who will buy EVs have generally considered a few variables, such as socio-demographic variables, customers' attitudes, climate change beliefs and attitudes, and vehicle attributes. However, there are many possible influences on such choices. As such, this study added variables related to social media usage, morality (Haidt and Joseph 2008), how GHG emissions information is presented (Wang et al. 2023), previous or planned future actions to reduce GHG emission, descriptive norms (Smith et al. 2012), and opinions about other sustainable transportation modes. Those variables were added to the conventional models to develop an accurate prediction model and detect the top variables influencing the intention to buy EVs. The list of variables applied in the modeling is presented in Appendix (Table A1). In the final model, 83 variables were employed. Regarding the vehicle properties, EV range and different ratios, including the purchase price of EV to ICEV, the fuel cost of EV to ICEV, and the emission of EV to ICEV, were used in the modeling.

Select socio-demographic and climate change beliefs and attitudes of respondents are indicated in Table 3. Respondents for the survey were sought based on relative population sizes of the provinces, gender splits, and age groups. As can be seen in Table 3, after cleaning the data of respondents who failed the "trap" questions, the most common attributes of the retained respondents were: males (51.4%), bachelor's degree holders (38%), European origins (51.4%), Ontario residents (43.6%), vehicle owners (92%), and full-time workers (51.8%). Moreover, 41% of respondents were very or extremely worried about climate change, while 30.7% did not worry or were slightly worried.

Table 3. Select socio-demographic and level of concern about climate change of respondents.

Variable	Frequency (%)	Variable	Frequency (%)
Gender		Province and language	
Male	51.4	British Colombia	15.6
Female	48.1	Alberta	12.8
Other	0.5	Saskatchewan	3.3
Education attainment		Manitoba	4
No education	0.2	Ontario	43.6
Elementary school	0	Quebec (English)	3.2
Less than high school equivalent	1.3	QC (French)	17.5
High school diploma or equivalent	15.3	Worried about climate change	
Registered Apprenticeship or other trades certificate or diploma	6.1	Not at all worried	9
College, CEGEP or other non-university certificate or diploma	21.6	Slightly worried	21.7

University bachelor's degree	38	Somewhat worried	28.3
Degree in dentistry, medicine, optometry or veterinary medicine	1.3	Very worried	25.9
Master's degree	13.8	Extremely worried	15.1
Doctoral degree	2.4	Currently own a vehicle	
Ethnicity		Yes	92
North American Aboriginal	1.4	No	8
Other North American	19.2	Employment	
European	51.4	Full-time	51.8
Latin, Central and	1.3	Part-time	9
South American	0.9	Homemaker	1.5
African	0.8	Full-time student	3.5
Asian	17.5	Retired	29.4
Oceania	0.4	Not currently employed, but looking for work	2.2
Prefer not to answer	7.1	Other	2.6

2.2. Modeling process

After preparing the datasets, each data set (BEV and PHEV) was split into two sub-datasets, including testing data and training data. That is, 20% of each data was randomly considered testing data, and the other 80% was considered training data. Training data was used to train the prediction models, and testing data was used to evaluate their prediction power and detect the most accurate classification technique. Five ensemble techniques were applied for modeling: Light Gradient Boosting Machine (LGBM), Random Forest (RF), eXtreme Gradient Boosting (XGB), Adaptive Boosting (AB), and Categorical Boosting (CB).

First, the hyperparameters of these techniques were optimized using Grid Search and K-fold cross-validation for each dataset (K considered to be five (Naseri et al. 2022b)). Consequently, all ensemble learning techniques were run using their optimal hyperparameter values for each dataset separately. The performance of these techniques on the BEV and PHEV datasets was compared, and the most accurate classifier on each dataset (BEV and PHEV) was detected. Then, the most accurate technique was used to determine which variables have the strongest relative influence on the intention to buy BEVs and PHEVs. The applied ensemble learning techniques are briefly introduced in the Appendix (Section A1).

2.3. Analyzing variables

One of the main goals of this study is to realize which variables are the most influential parameters to predict EV preference and determine how top-ranked variables impact the intention to buy EVs. However, it is difficult to interpret the results of machine learning techniques, since they are often black-box. In this regard, SHAP was employed to prioritize variables based on their relative influence on vehicle engine preferences.

SHAP is a technique to interpret the results of machine learning methods, introduced by Lundberg and Lee (2017). SHAP outperformed other interpretation techniques in terms of computational performance and consistency with human intuition (Lundberg and Lee 2017). This method applies game theory and local explanations to predict the relative influence of each variable on the response variable (Naseri et al. 2024). An additive variable attribution technique is used by SHAP, in which input variables are linearly added to an output model. Then, the relative influence of input variables on the response variable is determined by analyzing their marginal change (Mangalathu et al. 2020). SHAP needs to use machine learning methods to calculate the relative influences. Hence, the most accurate classifier on each dataset was synced with SHAP to calculate the top-ranked variables, affecting the intention to buy BEVs and PHEVs.

After detecting the top-ranked variables on the EV preferences, these variables should be analyzed for policy purposes in the transportation sector. In this regard, PDP was used to represent the influence direction of top variables on the EV purchase likelihood. PDPs can illustrate the nonlinear and complex relationships between the input and output variables (Alnahit et al. 2022). PDP generates new data samples by replacing a variable with different feasible values. Subsequently, a machine learning model is used to predict the response variable of generated data samples. Finally, an average value over the predicted response variables is used to calculate the partial dependence of that variable over its marginal distribution (Inglis et al. 2022). Therefore, the top-ranked variables detected by SHAP were investigated using PDP and the most accurate classification technique on both data sets (BEV and PHEV).

3. Results and discussions

This study aims to detect the most accurate ensemble learning technique to predict who is more likely to prefer BEVs and PHEVs over ICEVs and identify the top variables impacting EV preference. Hence, five ensemble learning methods were employed to detect the most accurate classifier. Consequently, SHAP was applied to rank variables according to their impact on EV preferences. Ultimately, PDP was used to illustrate the relationship between top-ranked variables and the EV purchase likelihood. In this part, the outcomes of ensemble learning techniques are initially presented. Afterward, the results of SHAP are presented. Finally, PDPs for top-ranked variables are illustrated.

3.1. Accuracy of ensemble learning methods

The training data is used for training and tuning the models. Following this, the testing data accuracy of all classification techniques for both datasets is calculated, shown in Figure 3. The results for the BEV and then PHEV are discussed. As can be seen, XGB is the most accurate classifier to predict individual choices between BEV and ICEV, followed by LGBM, RF, CB, and AB. The testing data accuracy of XGB is 0.31%, 0.46%, 2%, and 7% higher than LGBM, RF, CB, and AB. Similarly, XGB reaches the maximum prediction accuracy to predict individuals' decisions when the choices are PHEVs and ICEVs. The testing data accuracy of XGB on the PHEV dataset is 0.2%, 0.5%, 2.7%, and 9.2% more than LGBM, CB, RF, and AB, respectively. Hence, XGB is the most accurate classifier to predict individual choices on both datasets. Accordingly, XGB is applied in the following analyses (SHAP and PDP) since it outperforms other ensemble learning techniques when comparing the prediction accuracy.

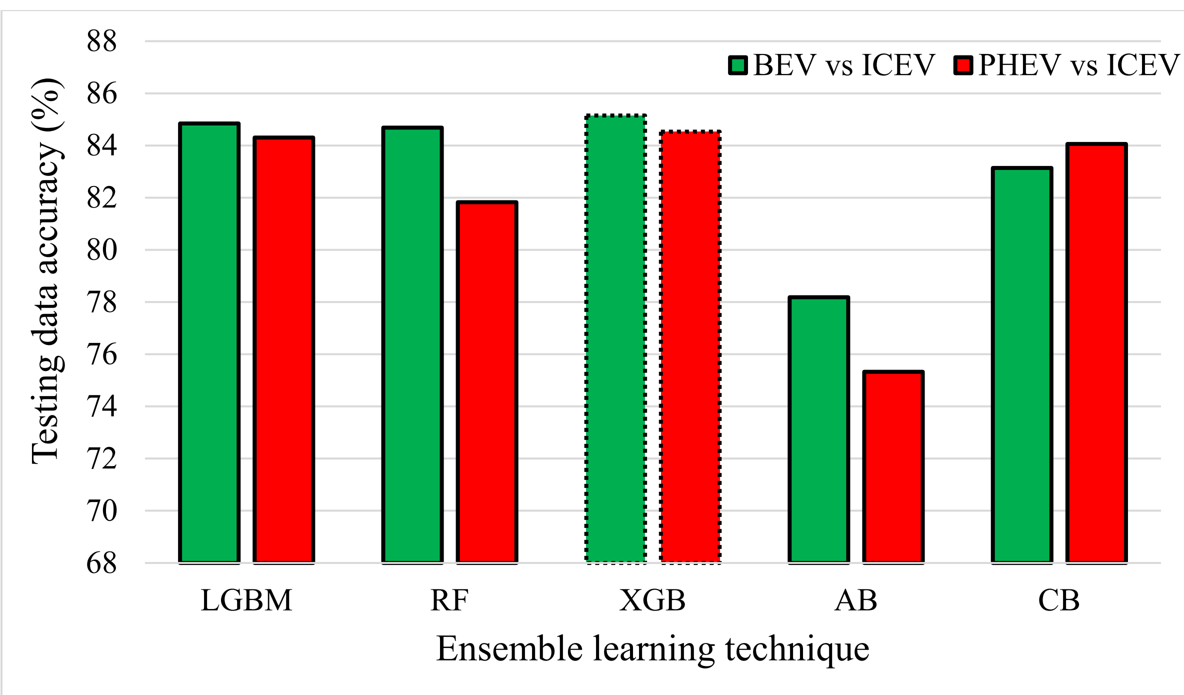


Figure 3. Testing data accuracy of machine learning techniques on both BEV and PHEV datasets.

3.2. The relative influence of variables on the EV preferences

As discussed, SHAP and XGB are used to prioritize variables regarding their impact on the intention to buy EVs. The SHAP values for 20 top-ranked variables on the BEV dataset are indicated in Figure 4. As can be seen, the price ratio of BEV to ICEV is the most influential variable on the intention to buy BEVs, with a SHAP value of 0.924. The interest in buying an EV if a significant discount by the government were to be given is the second top variable on the BEV purchase intention. The third variable is the ratio of the energy cost for an EV (electricity cost) to energy cost of an ICEV (fuel cost). Hence, financial factors are the most influential variables on the intention to buy BEVs. Experience or intention on using EVs, the preferred year to eliminate new ICEV sales in Canada, and the level of carbon pollution tax support are the following variables. Interestingly, the ranking of treatment (GHG information presentation) is 11 out of 83 variables. From the social media usage variables, the frequency of Instagram usage is the only variable available in the top 20.

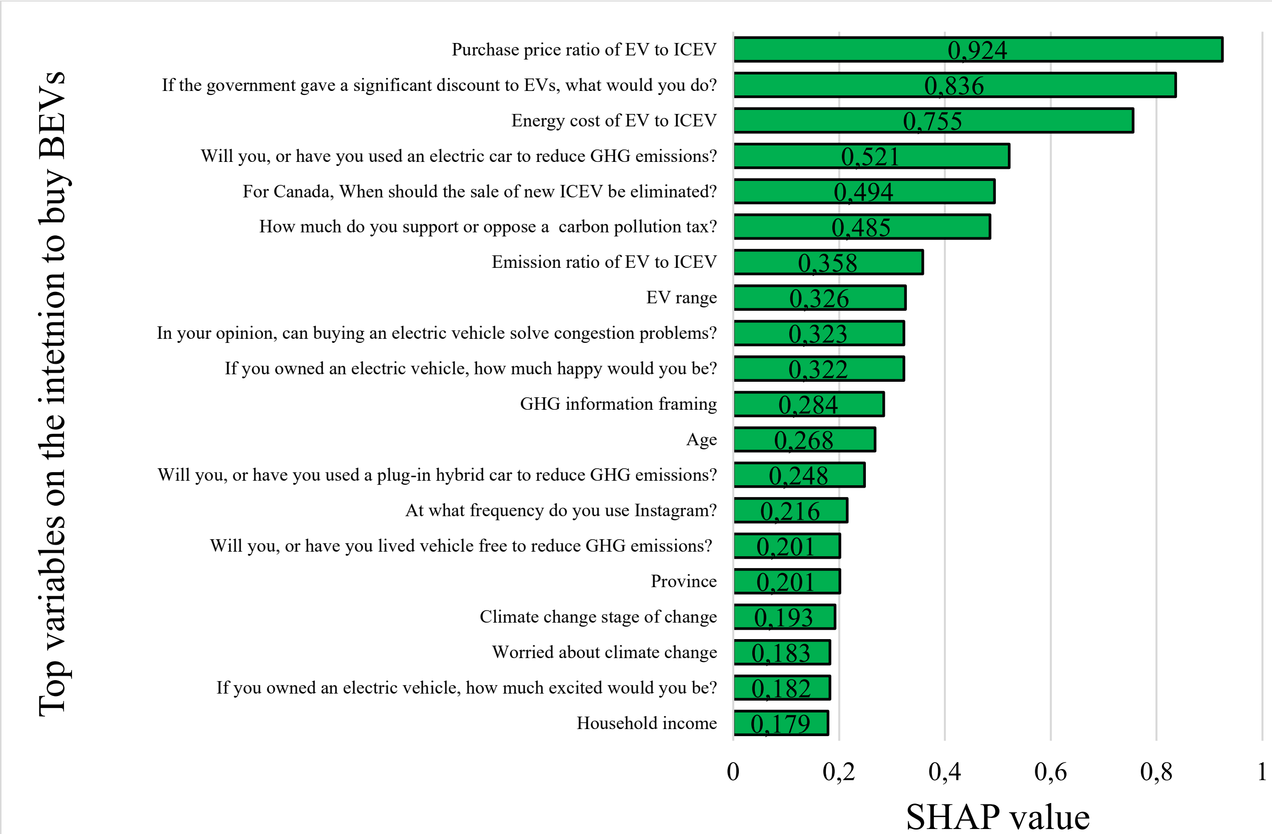


Figure 4. SHAP values of top variables on the intention to buy BEVs.

The SHAP values of top variables on the intention to buy PHEVs are illustrated in Figure 5. Like BEV, the price ratio of BEV to ICEV, actions given providing a significant discount by the government, and the electricity cost of EV to the fuel cost of ICEV are the top three variables impacting PHEV preferences. However, the difference in SHAP values of first-ranked (price ratio of BEV to ICEV) and other variables in PHEV is significantly higher than that of BEV. Therefore, it can be postulated that the intention to buy PHEV is more influenced by EV to ICEV purchase price than BEV. The experience in using a PHEV, the level of happiness when owning an EV, and the preferred year to eliminate new ICEV sales in Canada are third to six parameters affecting the PHEV preference.

The ranking of GHG information presentation (i.e., treatment) is nine. Hence, the influence of GHG information presentation on PHEV preference can be more than BEV preference since the ranking of this variable on the intention to buy BEV is 11 (higher than 9). EV range has a higher ranking on the BEV preference than PHEV preference since its ranking is 8 and 10 on the BEV and PHEV datasets, in the order mentioned. In both EV engine types, age is the most influential socio-demographic variable (with a ranking of 7 on PHEV and 12 on BEV).

The ranking of household income on the PHEV preference is 12 while the ranking of this variable is 20 on the intention to buy BEV. The relative influence of household income is less than other financial factors since having a high income is a necessary condition to pay an additional amount to buy EVs, while it is not a sufficient condition. That is, individuals with higher income may also prefer to choose cheaper vehicles (ICEVs). The variables related to attitudes toward climate change are not within the top 20 variables on the intention to buy PHEV. On the other hand, the ranking of climate change stage of change and level of worried about climate change on BEV preference is 17 and 18. The ranking of these two environmental variables on the PHEV preference is 21 and 26, respectively. Therefore, climate change beliefs and attitudes can more influence the intention to buy BEV than PHEV.

In the previous study (Naseri et al. 2023), the relative influence of socio-demographic variables and vehicle attributes on the EV purchase likelihood is evaluated considering both BEV and PHEV as EVs. In that study, the tailpipe emissions were used. The results showed that CC-SoC, the price ratio of EV to ICEV, province, age, education, and GHG information framing were top variables on EV preferences. Hence, their results are generally in line with the results of this study. The major difference between that investigation and the current study is the relative influence of climate change stage of change. Climate change stage of change was the first-ranked variable in Naseri et al. (2023) study, while here, its rank is 17. This suggests that if lifecycle emissions are presented, an individual's climate change attitudes and behavior are less important as EVs are products with no emissions associated to them.

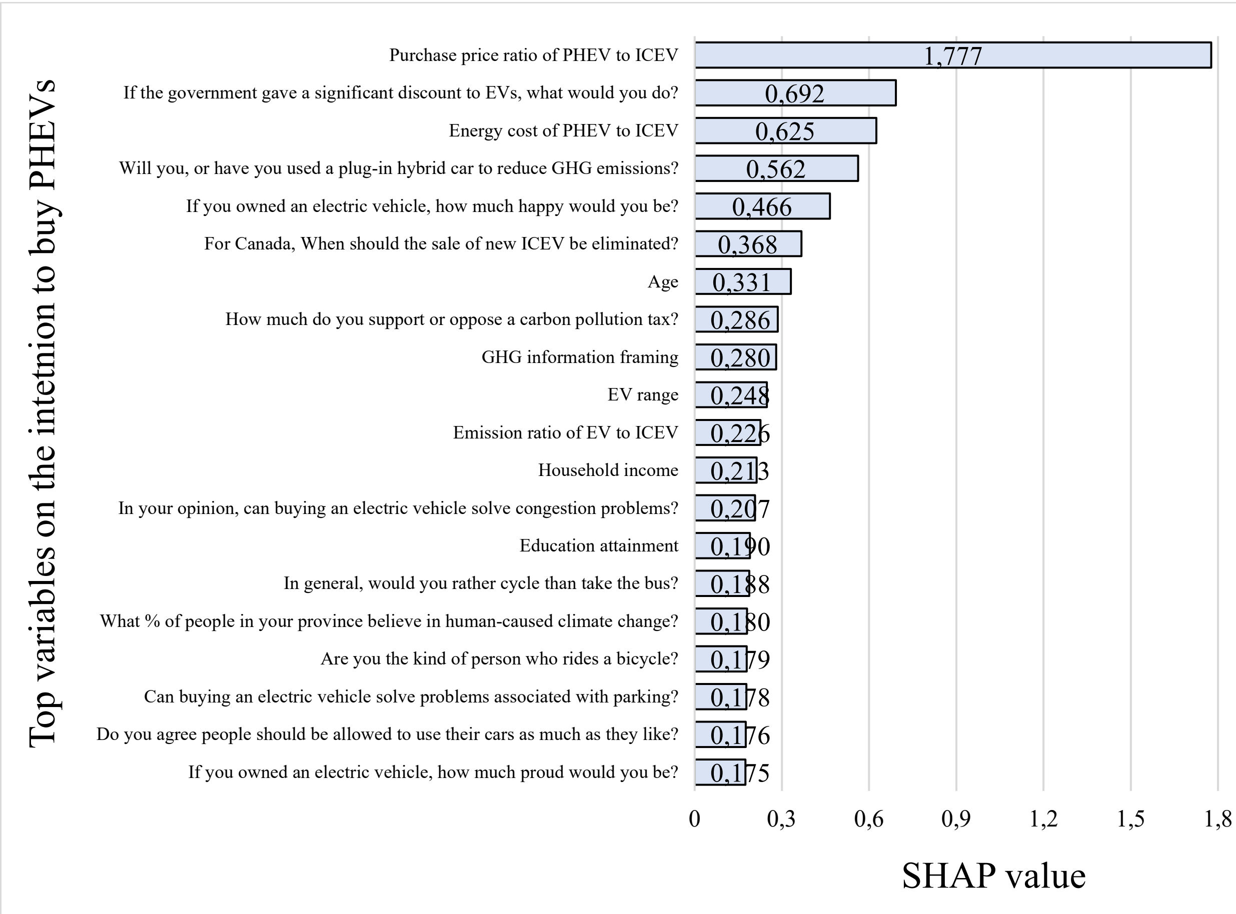


Figure 5. SHAP values of top variables on the intention to buy PHEVs.

To check the robustness of the results, SHAP is also synchronized with other ensemble learning techniques, and SHAP values of variables are calculated using LGBM, RF, CB, and AB for both models (i.e., BEV and PHEV). Table 4 presents the ranking of XGB's top ten variables when other ensemble learning methods are used to calculate SHAP values. As shown, the ranking of variables in LGBM is quite similar to that of XGB (the most accurate model), and it can be related to the similar prediction accuracy of XGB and LGBM. On the other hand, the ranking of variables in AB (the model with minimum accuracy) is different from that of XGB in many cases. Hence, these results prove the importance of using the most accurate technique.

Table 4. The ranking of top variables using SHAP and ensemble learning techniques.

	Ranking of top variables on the intention to buy BEVs				
	XGB	LGBM	RF	CB	AB
Purchase price ratio of EV to ICEV	1	1	1	1	1
If the government gave a significant discount to EVs, what would you do?	2	2	2	2	17
Energy cost of EV to ICEV	3	3	4	3	3
Will you, or have you used an electric car to reduce GHG emissions?	4	4	3	4	23
For Canada, when should the sale of new ICEV be eliminated?	5	6	6	15	44
How much do you support or oppose a carbon pollution tax?	6	5	5	5	2
Emission ratio of EV to ICEV	7	7	30	17	41
EV range	8	8	13	9	14
In your opinion, can buying an electric vehicle solve congestion problems?	9	10	14	6	5
If you owned an electric vehicle, how much would you be happy?	10	12	8	7	34
	Ranking of top variables on the intention to buy PHEVs				
	XGB	LGBM	RF	CB	AB
Purchase price ratio of PHEV to ICEV	1	1	1	1	1
If the government gave a significant discount to EVs, what would you do?	2	2	2	4	6
Energy cost of PHEV to ICEV	3	3	4	2	2
Will you, or have you used a PHEV car to reduce GHG emissions?	4	5	5	3	21
If you owned an electric vehicle, how much happy would you be?	5	6	3	5	9
For Canada, when should the sale of new ICEV be eliminated?	6	4	7	6	15
Age	7	7	15	19	11
How much do you support or oppose a carbon pollution tax?	8	8	13	7	4
GHG information framing	9	14	25	26	41
EV range	10	11	11	8	33

3.3. Influence direction of top ranked variables on the intention to buy EVs

Although the variables are ranked based on their relative influence on the intention to buy BEVs and PHEVs, it is not clear how these variables impact vehicle engine choices. To address this issue, the effects of top-ranked variables (determined in the previous section) on EV preferences are scrutinized using PDP. In this section, the influence direction of EV to ICEV purchase price ratio, action given providing a significant discount to EVs by the government, EV to ICEV fuel cost ratio, experience or intention to use EVs, preferred year to eliminate the sales of ICEVs, the level of support in carbon pollution tax, and age on EV preferences is investigated using PDP. Similarly, PDP is applied to investigate the impact of GHG information presentation on the EV purchase likelihood and identify the optimal vehicle label to promote EVs.

3.3.1. EV to ICEV purchase price ratio

The impacts of BEV to ICEV and PHEV to ICEV purchase price ratio on the probability of choosing EVs over ICEVs are indicated in Figure 6 and Figure 7. As can be seen, there is a linear trend between the BEV to ICEV purchase price and the probability of buying a BEV. If the purchase price ratio increases from 1.33 to 1.53, the BEV purchase likelihood is reduced by nearly 5%. Moreover, by increasing the BEV to ICEV price ratio from 1.33 to 2.18, the BEV purchase likelihood is reduced by 16.6%. When the BEV to ICEV price ratio is at its largest (2.55), the BEV purchase likelihood is reduced by 18.9% compared to the price ration of 1.33. The results of this analysis are in line with the results of Naseri et al. (2023).

Similarly for PHEV, increasing the purchase price ratio (compared to ICEV) leads to a reduction in the probability of buying a PHEV. The probability of choosing a PHEV over an ICEV is maximum when the PHEV purchase price equals the ICEV. In other words, when the PHEV to ICEV purchase price ratio is one, the PHEV purchase likelihood is 80.3%. However, increasing the price ratio from 1 to 1.15, the probability of selecting PHEVs over ICEVs is reduced by 29.3% (from 80.3% to 51%). After this threshold (ratio of 1.15), the PHEV purchase likelihood is reduced at a lower rate. Increasing the PHEV to ICEV purchase price ratio from 1.15 to 1.46 causes a 12.2% reduction in the likelihood of selecting PHEVs over ICEVs. After this ratio (1.46), a significant drop in PHEV purchase likelihood is not seen as the reduction in likelihood is only 1.5% if the price ratio increases from 1.46 to 2.18.

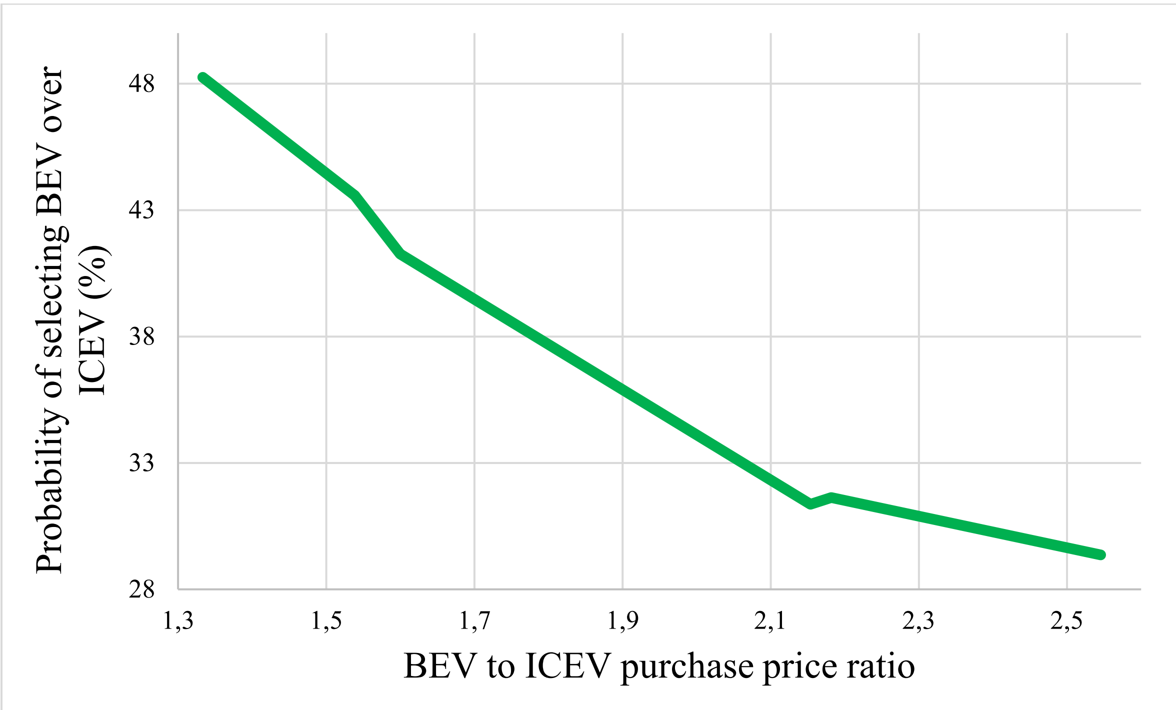


Figure 6. The impact of BEV to ICEV purchase price on the probability of choosing BEV over ICEV.

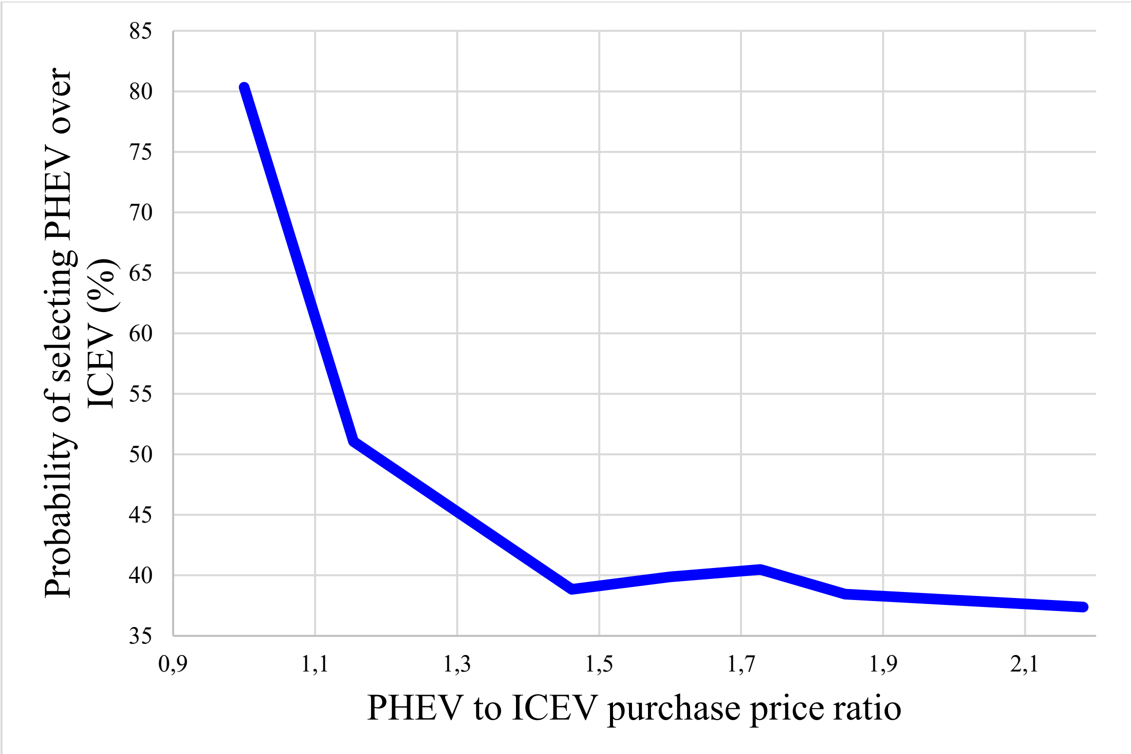


Figure 7. The impact of PHEV to ICEV purchase price on the probability of choosing PHEV over ICEV.

3.3.2. Intentions to buy an EV if a significant discount to EVs was given

The influence of purchase intentions of EVs if the government gave a significant discount on EVs for BEVs and PHEVs is indicated in Figure 8 and Figure 9. As shown, those who already own an EV are most likely to choose BEVs over ICEVs when a significant discount on EVs is provided by the government. Those who declared that "I don't currently have a car, but would buy an EV" are the second group with the highest probability of preferring BEVs over ICEVs; however, their BEV purchase likelihood is roughly 1% less than EV owners. The BEV purchase likelihood of those who intend to buy an EV as an additional vehicle and buy an EV to replace their current vehicles is 5% and 5.2% less than those who have already an EV. Those who stated that they would still not buy an EV are least likely to prefer BEVs over ICEVs, and the probability of selecting BEVs over ICEVs for this group is 20% less than for EV owners.

Likewise, those who already have an EV are the most likely to prefer PHEVs to ICEVs if a significant discount on EVs is offered by the government. Of note is that the difference between the current EV owners and those with no intention to buy an EV is much smaller for the PHEVs (0.7% versus 20%). The PHEV purchase likelihood of selecting PHEVs over ICEVs for those who do not have a car but would buy an EV is approximately equal to EV owners. Hence, providing a significant discount to BEVs appears to be much more important than for PHEVs.

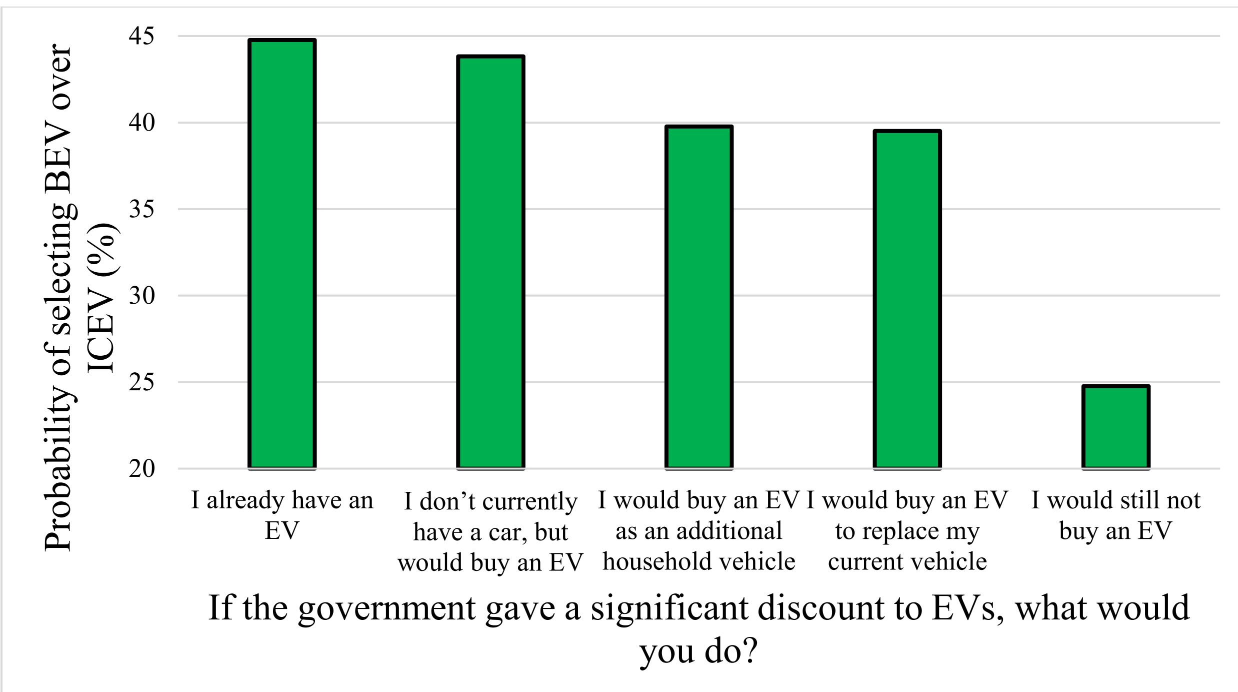


Figure 8. The impact of actions given a significant discount to EVs on the probability of choosing BEV over ICEV.

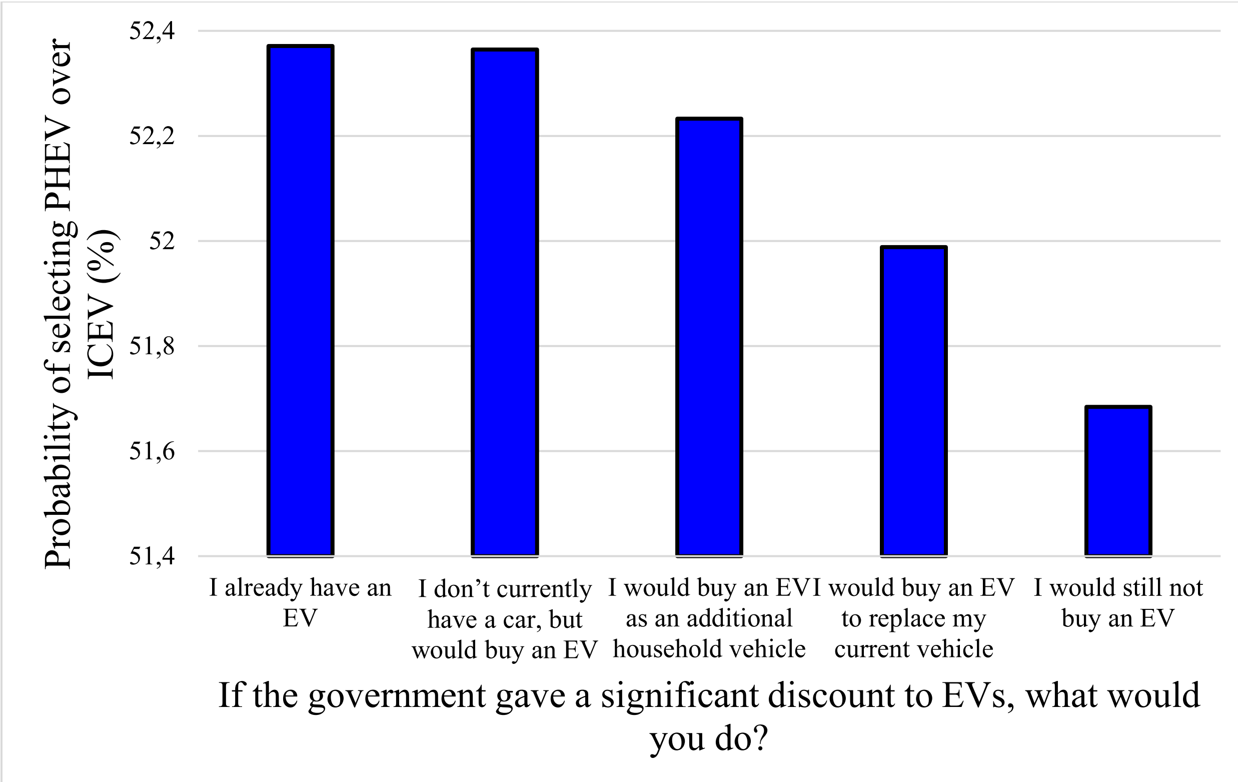


Figure 9. The impact of actions given a significant discount to EVs on the probability of choosing PHEV over ICEV.

3.3.3. EV to ICEV fuel cost ratio

The EV to ICEV fuel cost ratio is the third most influential variable on the intention to buy both BEVs and PHEVs. The influence of EV to ICEV electricity/fuel cost ratio on the probability of choosing BEV and PHEV over ICEV is indicated in Figure 10 and Figure 11. As indicated, by increasing the fuel cost ratio, the probability of choosing EVs over ICEVs is slightly reduced. However, if this ratio becomes higher than a threshold, the EV preference likelihood is sharply reduced. This threshold is 0.4 for BEV and 0.6 for PHEVs (e.g., this means that the energy costs of a BEV would be only 40% as compared to an ICEV). A lower ratio for BEV compared to PHEV is possible as PHEVs use gasoline (or similar) like with ICEVs. If the BEV electricity to ICEV fuel cost is increased from 0.25 to 0.4, the BEV purchase likelihood is reduced by only 1.3%. Nonetheless, increasing the BEV electricity to ICEV fuel cost from 0.4 to 0.63 results in a 10% drop in the probability of preferring BEVs over ICEVs. By increasing the PHEV to ICEV fuel cost from 0.35 to 0.6, the PHEV preference likelihood is reduced by 1.6%. Increasing PHEV to ICEV fuel cost from 0.6 to 0.93 leads to a 5.9% reduction in the probability of choosing PHEVs. Therefore, it is recommended to keep the EV to ICEV fuel cost ratio less than 0.4 for BEV and 0.6 for PHEVs to promote EVs.

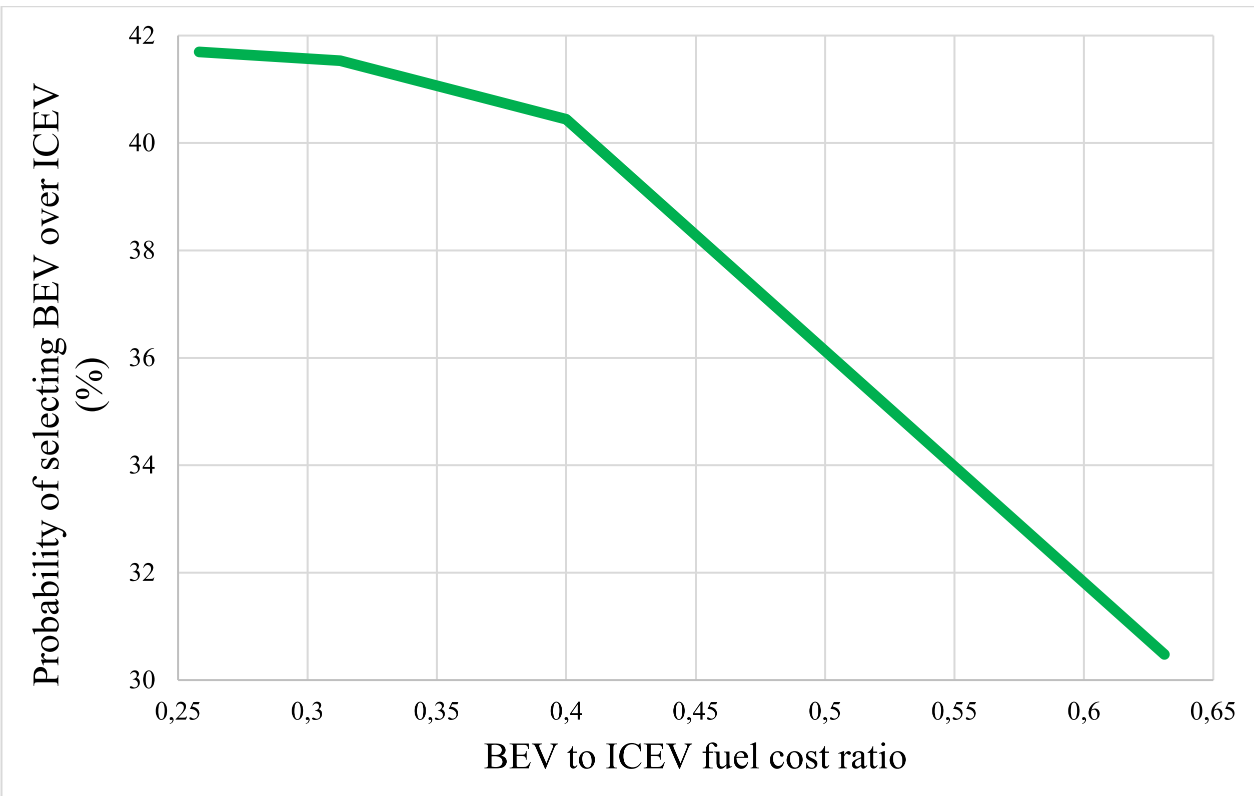


Figure 10. The impact of BEV to ICEV energy cost ratio on the probability of choosing BEV over ICEV.

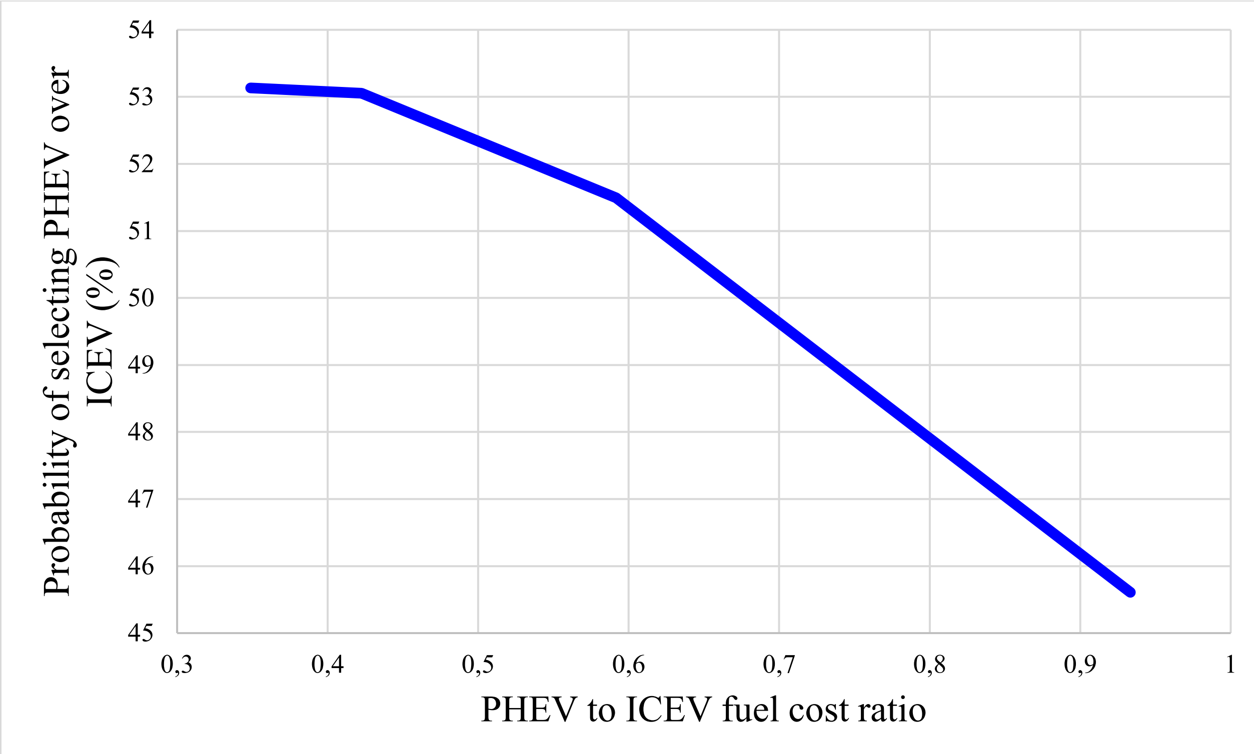


Figure 11. The impact of PHEV to ICEV fuel cost ratio on the probability of choosing PHEV over ICEV.

3.3.4. Current use or intention to use an EV

The question "Will you, or have you used an electric car to reduce GHG emissions?" is the fourth variable impacting the intention to buy BEVs. In the same way, the question "Will you, or have you used a plug-in hybrid car to reduce GHG emissions?" is the fourth influential variable on the PHEV purchase likelihood. The influence of the intention to drive BEVs and PHEVs to reduce GHG emissions on the probability of choosing BEV and PHEV over ICEV is displayed in Figure 12 and Figure 13. As can be seen, experience in using an EV to reduce GHG emissions maximizes the probability of preferring EVs over ICEVs. Moreover, those who intend to use EVs to reduce GHG emissions (mentioned "will do") behave roughly the same as those who already do. Intention is often seen in behavioral models as a key explanation of actual behavior (e.g., Ajzen (1991)).

In the BEV dataset, the probability of selecting BEVs over ICEVs for those who already used EVs and those who will use EVs to reduce their GHG emissions is 2.8% and 2.7% more than those who are not interested in using an EV to reduce their GHG emissions. In the PHEV dataset, the experience or intention to use PHEVs to reduce GHG emissions has a higher influence on the EV purchase likelihood. In other words, individuals who already used PHEVs and will use PHEVs to reduce GHG emissions are 5.6% and 5.4% more likely to select PHEVs over ICEVs.

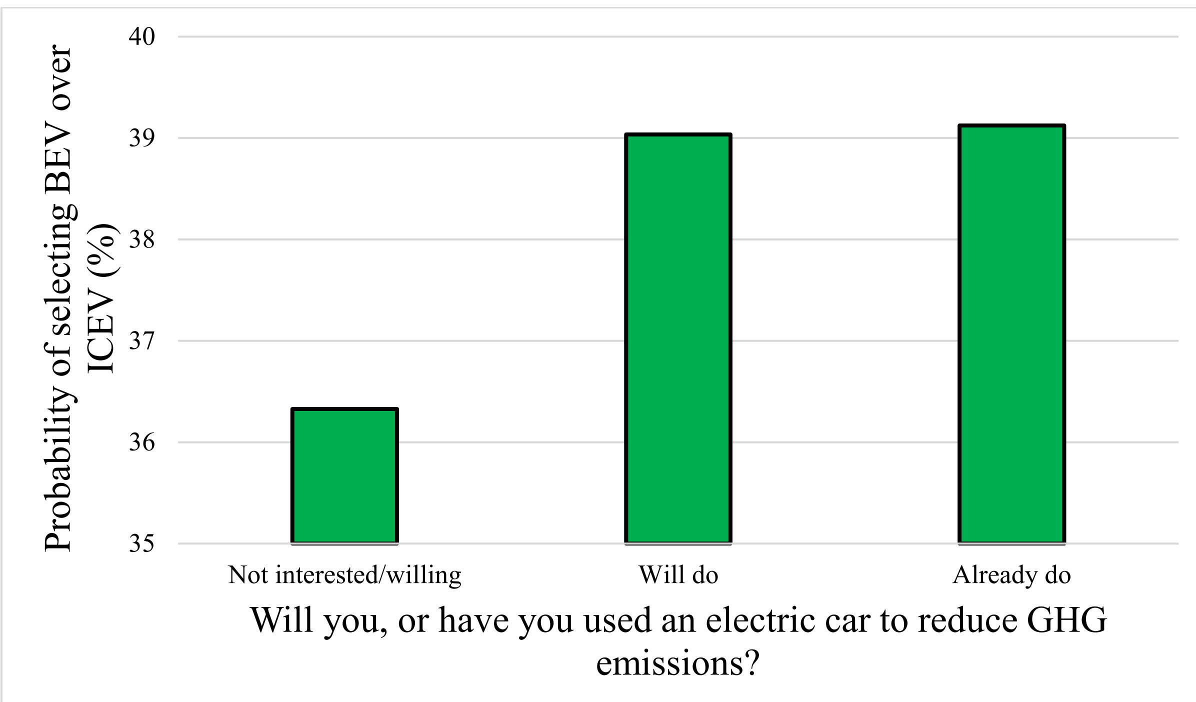


Figure 12. The impact of experience or intention to use EVs on the probability of choosing BEV over ICEV.

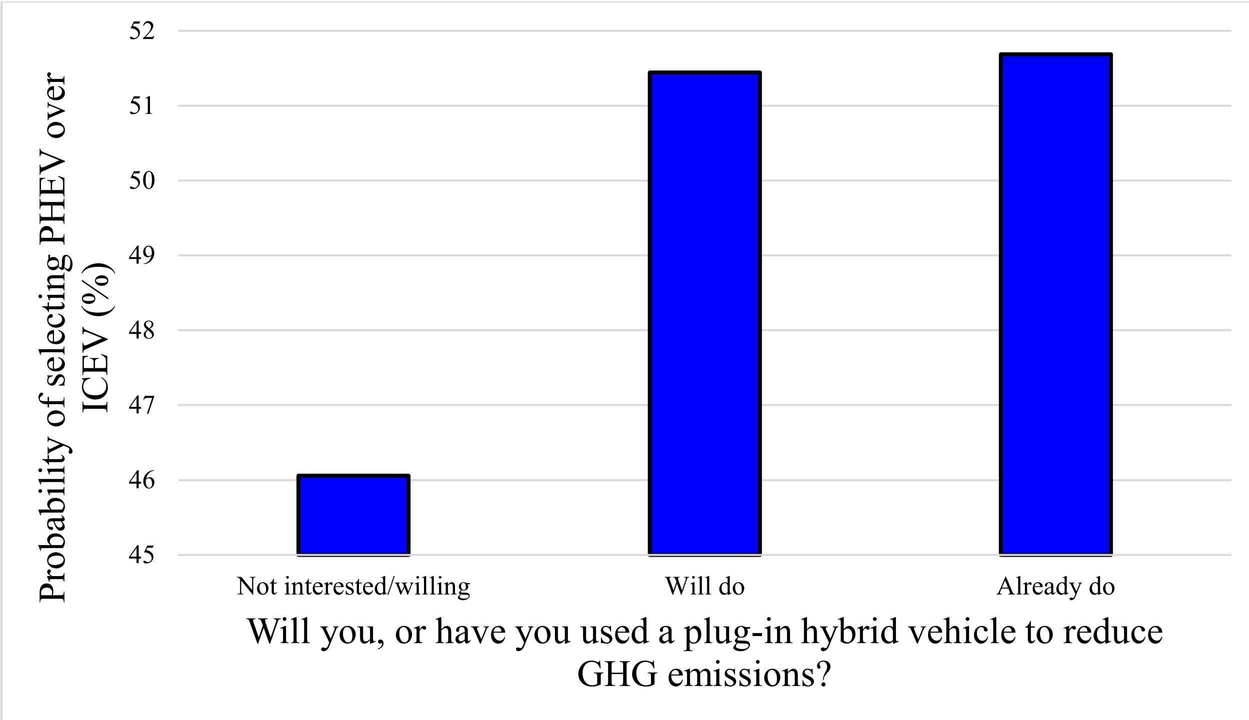


Figure 13. The impact of experience or intention to use PHEVs on the probability of choosing PHEV over ICEV.

3.3.5. Preferred year to eliminate the sales of ICEVs

The respondent's preference for the year when the government of Canada should eliminate the sale of ICEVs is the fifth and sixth most influential variable on the BEV and PHEV preference, respectively. The influence of the preferred year to eliminate the sales of ICEVs on the probability of choosing BEV and PHEV over ICEV is illustrated in Figure 14 and Figure 15. The participants who prefer that the sale of new ICEVs should be eliminated immediately have the highest intention to buy EVs. Those who prefer that the government forbids ICEV sales at a later date (or never) are less likely to select EVs over ICEVs.

The respondents who do not want the government to ever forbid the sale of ICEVs are least likely to prefer EVs to ICEVs. The BEV purchase likelihood of those who prefer that ICEV sales should be eliminated immediately, in 2025, 2030, 2035, 2040, and 2050 is 4.6%, 4.6%, 4.2%, 3.3%, 2.2%, and 1.5% more than those who believe that ICEV sales should never be stopped. The probability of choosing PHEVs over ICEVs for those who prefer that ICEV sales should be eliminated immediately, in 2025, 2030, 2035, 2040, and 2050 is 2.7%, 2.6%, 2.5%, 1.5%, 0.1%, and 0.1% higher than those never want to see the sales of ICEV stopped. Thus, wanting the government to forbid the sales of ICEVs sooner is an influential consideration for EV preference.

Surprisingly, the probability of selecting a BEV over an ICEV is only 40% among respondents favoring an immediate elimination of ICEV sales, which is relatively low. This can be related to other strong variables such as vehicle attributes (e.g., purchase price and GHG emission ratios). That is, although some individuals believe that new ICEV sales should be immediately eliminated, in nearly 60% of choice selections, they select ICEVs over BEVs in the cases that the BEV to ICEV purchase price ratio is very high or emission ratio of BEV to ICEV is not low.

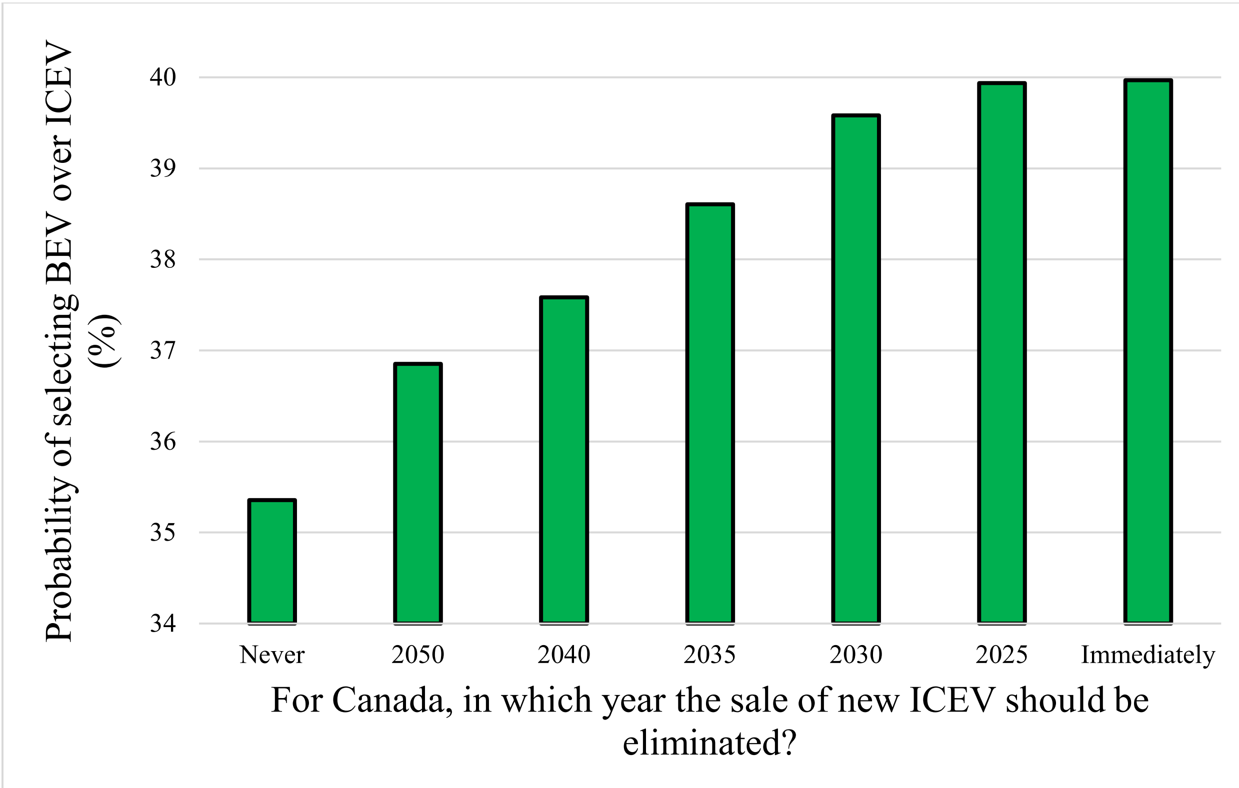


Figure 14. The impact of the preferred year to eliminate the sales of ICEVs on the probability of choosing BEV over ICEV.

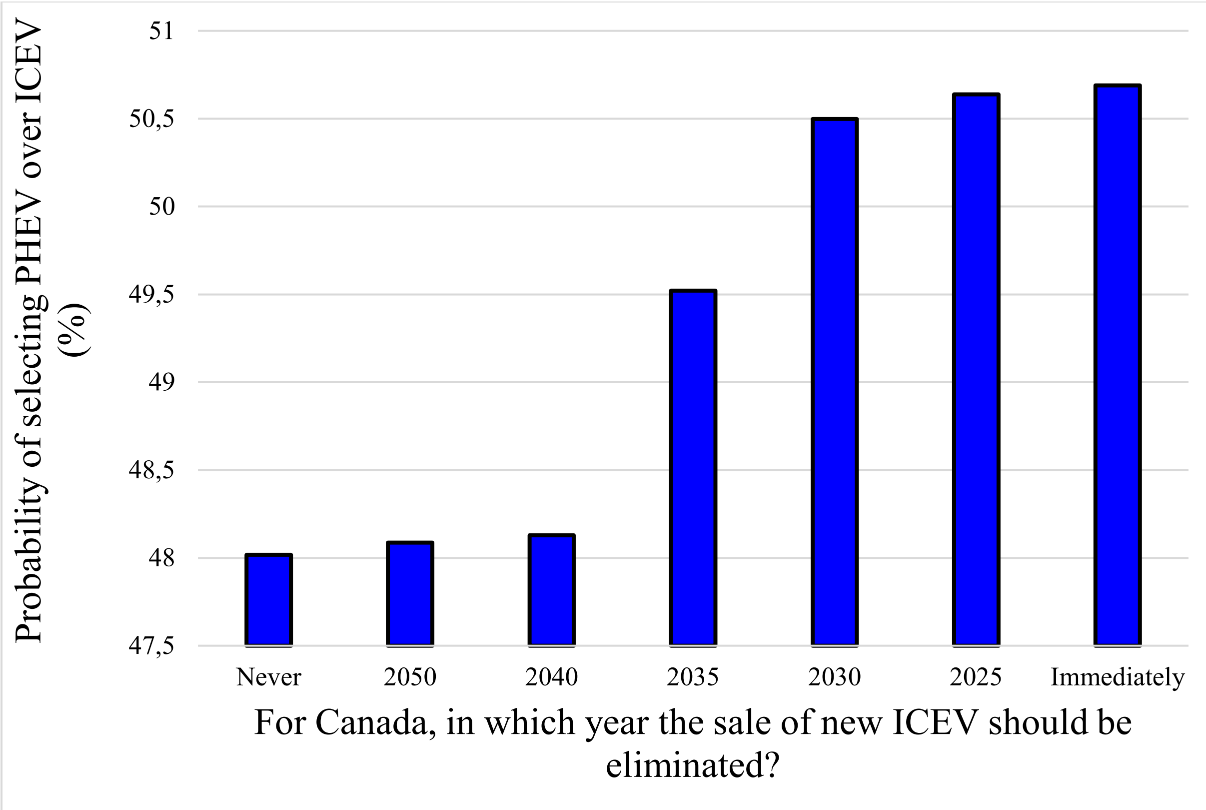


Figure 15. The impact of the preferred year to eliminate the sales of ICEVs on the probability of choosing PHEV over ICEV.

3.3.6. Level of support in carbon pollution tax

The level of support or opposition to the Canadian carbon pollution tax is the sixth and eighth influential variable on the probability of choosing BEVs and PHEVs over ICEVs, respectively. The influence of the level of carbon pollution tax support on the probability of choosing a BEV or a PHEV over ICEV is illustrated in Figure 16 and Figure 17. As can be seen, there is a direct correlation between the level of carbon pollution tax support and the probability of choosing EVs over ICEVs. Those who strongly support the carbon pollution tax in Canada are 4.7% and 2.8% more likely to prefer BEVs and PHEVs over ICEVs than those who oppose the carbon pollution tax. The influence of supporting is stronger for BEVs than PHEVs which likely reflects the fact that BEVs' lifecycle emissions are lower.

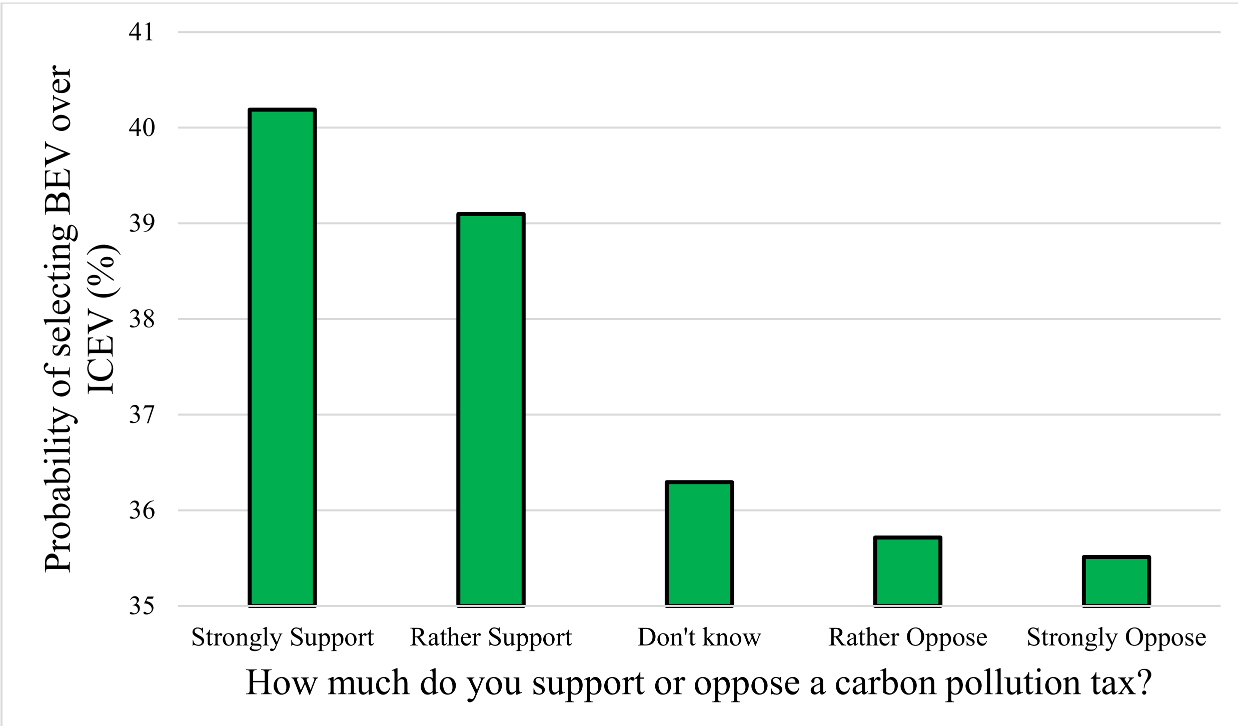


Figure 16. The impact of the level of carbon pollution tax support on the probability of choosing BEV over ICEV.

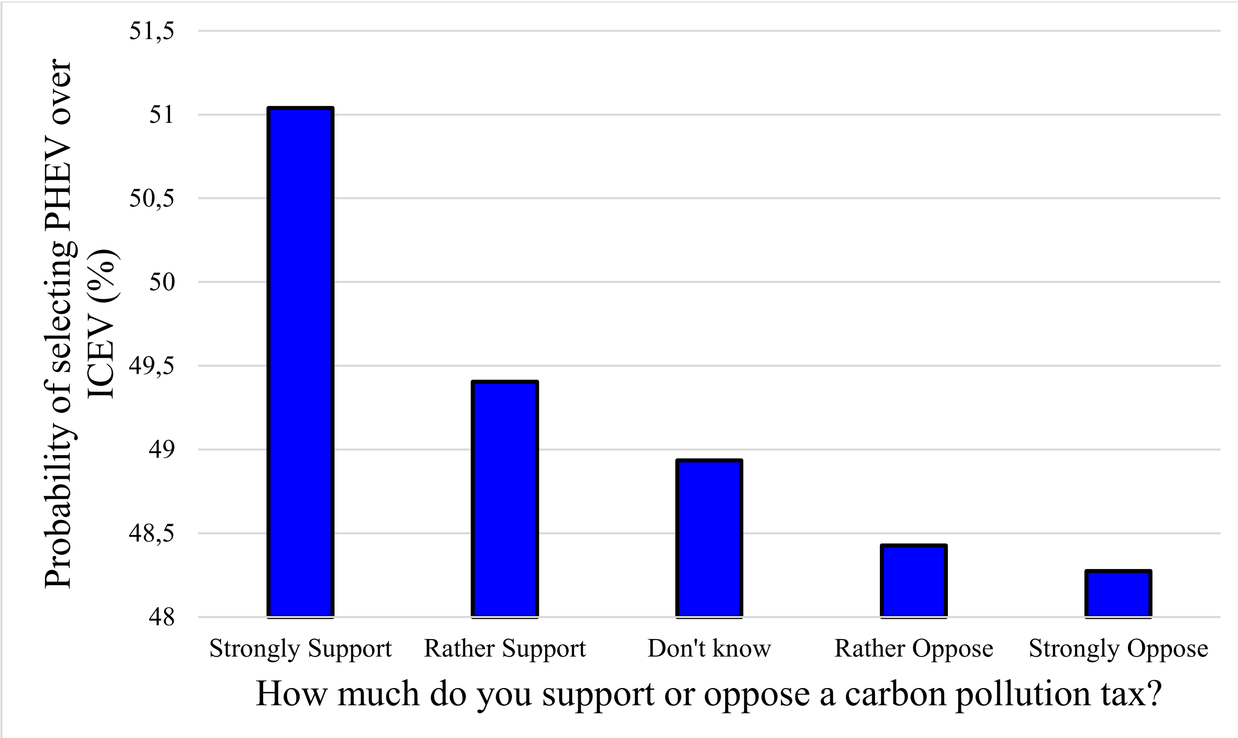


Figure 17. The impact of the level of carbon pollution tax support on the probability of choosing PHEV over ICEV.

3.3.7. GHG emissions information presentation

How GHG emissions information is presented (or framed) is an influential variable that can be changed much easier by the government than the other variables. Cost ratios would depend on either subsidies (tax money diverted to that) or additional charges (those who pollute must pay for this negative externality). Other influences above relate to attitudes. However, the government can establish standards on how GHG emissions information is framed so that individuals are better informed about how their choice impacts this critical problem. The ranking of this variable on the intention to buy BEVs and PHEVs is 11th and 9th respectively. The influence of framing on the probability of choosing BEV and PHEV over ICEV is illustrated in Figure 18 and Figure 19. As can be perceived, all of the framings developed and tested perform better than the current mock-up in Canada (NRCan) regarding EV purchase likelihood increment.

The Flood framing is the most efficient label to promote BEVs, followed by the framings Fire, Emoji, Leaves, General disaster, Layout change, NRCan-LC, and finally the current NRCan framing. Fire outperforms other framings in terms of maximizing the PHEV preference likelihood, followed by General disaster, Emoji, Leaves, Flood, Layout change, NRCan-LC, and then the NRCan. In both cases, this relates to psychological distancing (both geographic and time) and highlighting current local impacts. Replacing the NRCan with the Flood label can increase the probability of selecting BEVs and PHEVs over ICEVs by 2.4% and 1.9%, in the order mentioned. Further, the Fire label can increase the probability of selecting BEVs and PHEVs over ICEVs by 2.3% and 2.1% compared to the NRCan framing. Using the life cycle emissions with the NRCan framing (NRCan-LC) increased the likelihood to choose and EV by roughly 1%. In the design of the scale for the life cycle emissions, BEVs remained at 10/10, PHEVs were reduced from 10/10 to 8 or 9/10 and ICEVs were typically ranked lower than for the standard tailpipe emissions NRCan label. For example, an ICEV that was previously rated as 7/10 would be rated as 4/10. This larger difference between EVs and ICEVs likely explains why a greater preference was found for the NRCan-LC framing. However, it should be noted that when life cycle emissions were used, the influence of framing was reduced (Naseri et al., 2023), likely because the overall advantage of EVs over ICEVs was diminished.

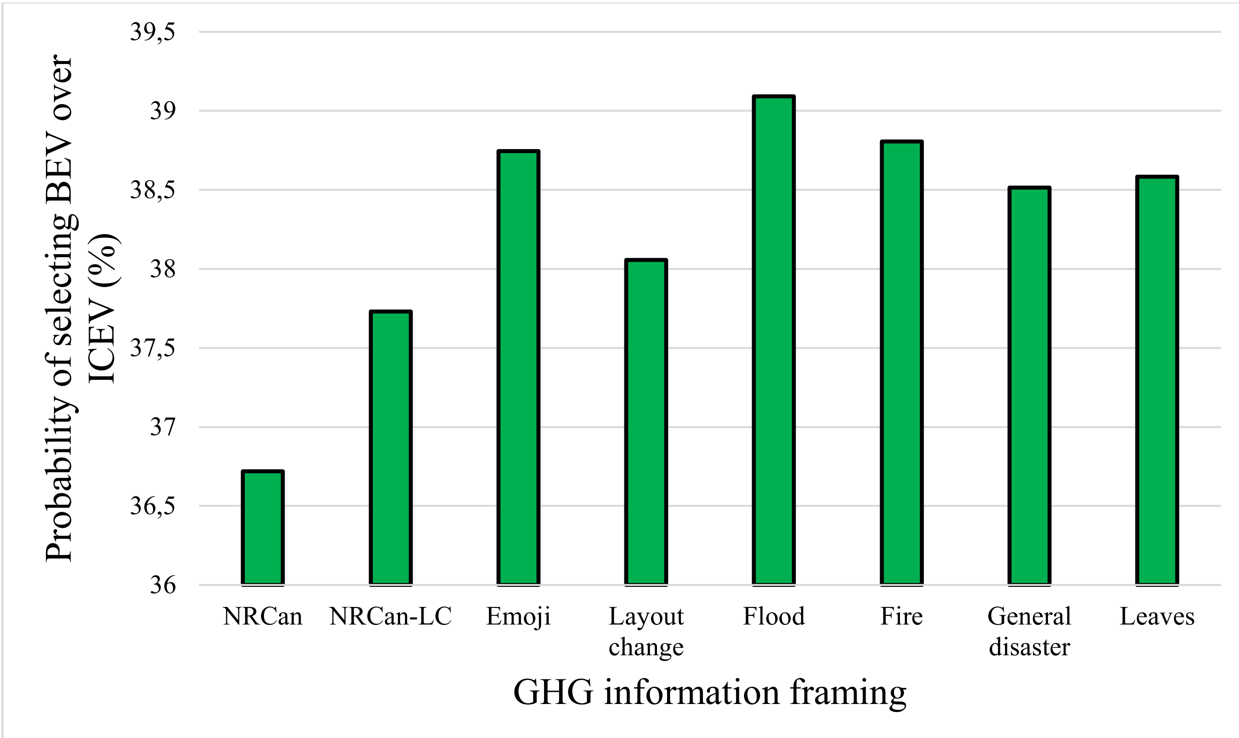


Figure 18. The impact of GHG information presentation on the probability of choosing BEV over ICEV.

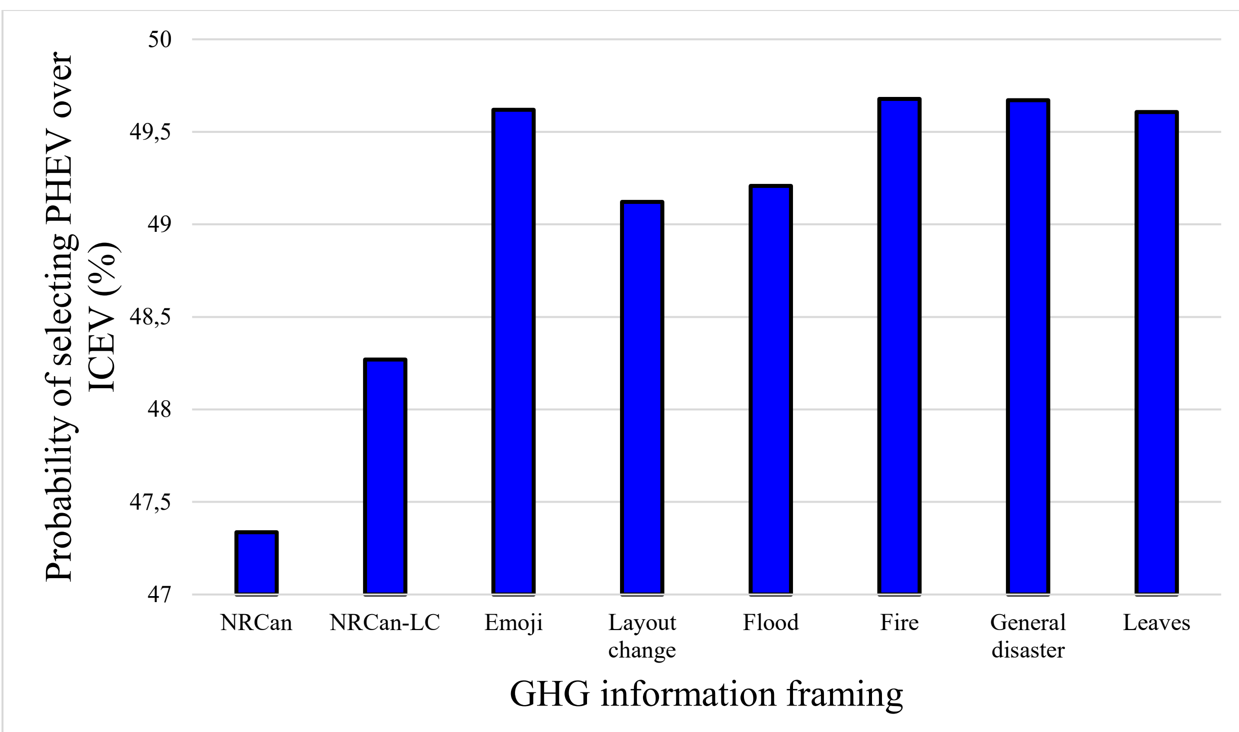


Figure 19. The impact of GHG information presentation on the probability of choosing PHEV over ICEV.

3.3.8. Age

Age was found to be the top-ranked variable among socio-demographic variables. The impacts of age on the probability of choosing EVs over ICEVs are indicated in Figure 20 and Figure 21. As shown, people aged 38 and 47 are most likely to prefer BEVs to ICEVs. For the other age groups, the difference in the probability of BEV preference is not significantly different. People younger than 45 years are more likely to choose PHEVs over ICEVs. After this threshold (age of 45), by increasing the age, the PHEV preference likelihood is slightly reduced. For those aged under 46 (except 36 and 37 years), the probability of preferring PHEVs to ICEVs is higher than 50%.

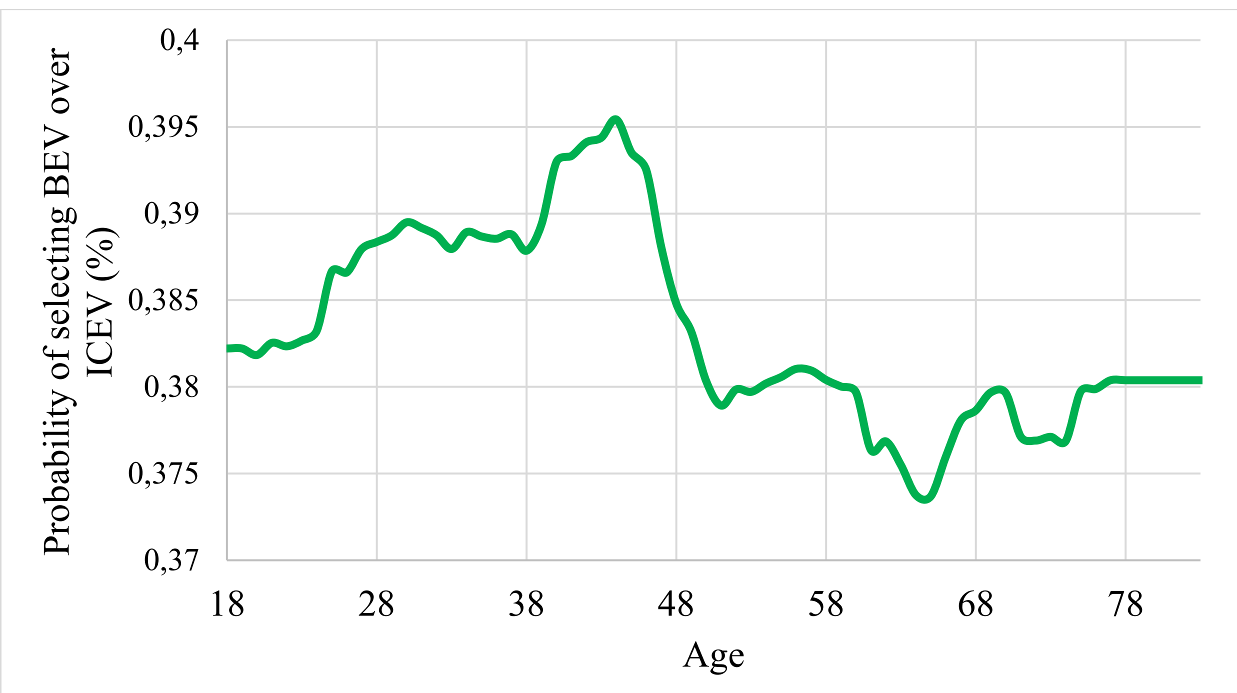


Figure 20. The impact of age on the probability of choosing BEV over ICEV.

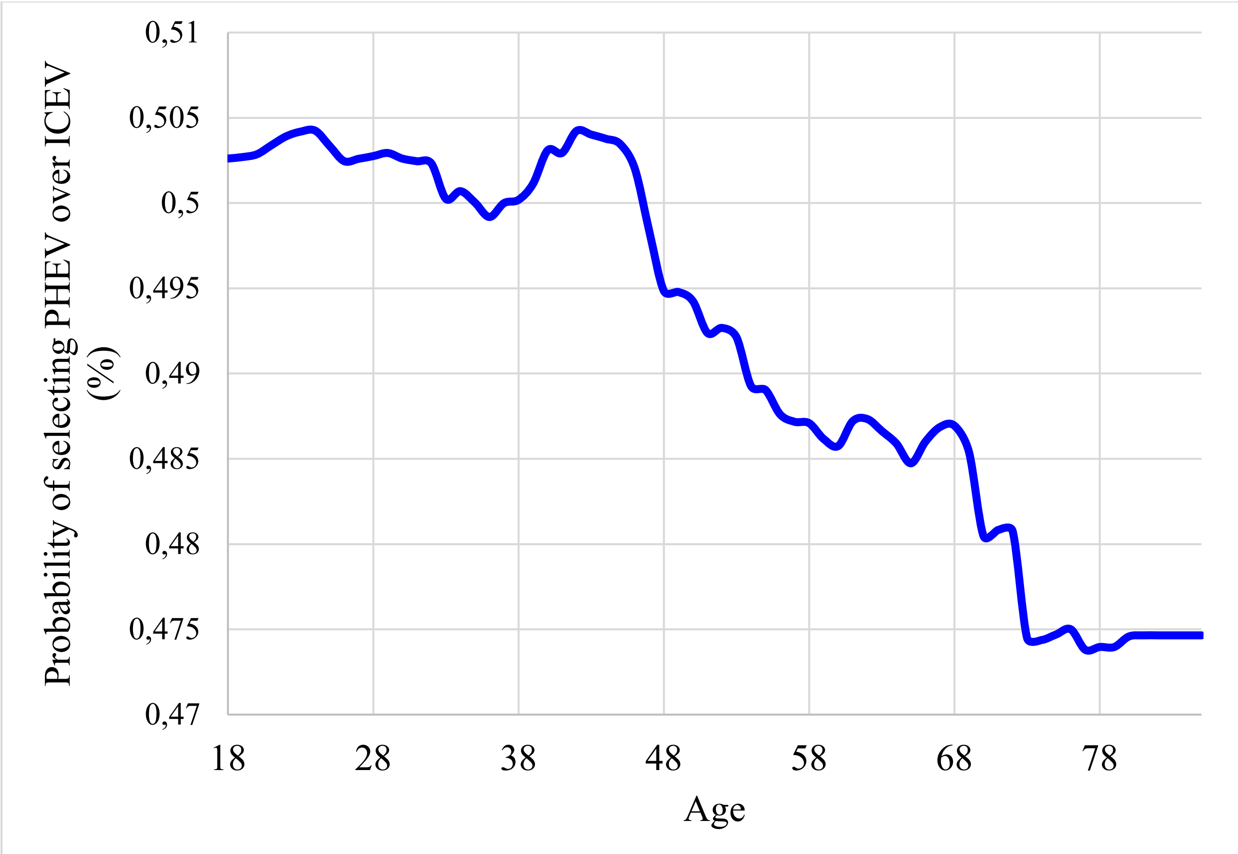


Figure 21. The impact of age on the probability of choosing PHEV over ICEV.

4. Conclusions

This study aimed to generate an accurate model to predict individuals' EV preferences and identify who might be more willing to buy battery electric vehicles and plug-in hybrid electric vehicles. Previous research has primarily used tailpipe emissions, but this study examined how presenting lifecycle emissions might change preferences. Further, as compared to previous studies, this research examined the influence of many variables related to socio-demographic variables, customers' attitudes, climate change beliefs and attitudes, vehicle attributes, social media usage, morality, GHG information presentation, previous or planned future actions to reduce GHG emissions, descriptive norms, and opinions about other sustainable transportation modes. All in all, 83 variables were applied in the modeling.

Five ensemble learning methods were used to develop prediction models for BEVs and PHEVs separately using 1077 completed questionnaires by Canadians with a driver's license. In both datasets, XGB was the most accurate classification technique, and it reached a testing data (20% of sample) accuracy of 85.2% and 84.5% for the BEV and PHEV datasets. Therefore, it is suggested that XGB is applied to predict EV preferences in future studies.

SHapley Additive exPlanations were employed to prioritize variables based on their influence on the intention to buy EVs. The results suggested that the price ratio of BEV to ICEV was the most influential variable on the intention to buy BEVs, followed by the change in vehicle ownership given a significant EV discount by the government, the electricity cost of BEV to the fuel cost of ICEV, current use or intention to use an EV, the preferred year to eliminate new ICEV sales in Canada, and the level of carbon pollution tax support are the next most influential variables.

For PHEVs, similar outcomes were found: price ratio of PHEV to ICEV, vehicle ownership change following a significant government discount on EVs, the electricity/fuel cost ratio, current use or intended use of a PHEV to reduce GHG emissions, the level of happiness if they owned an EV, and the preferred year to eliminate new ICEV sales in Canada. Therefore, the monetary variables were within the top three variables in promoting EVs when lifecycle emissions are used. In both EV engine types (BEV and PHEV), age was the most influential socio-demographic variable. For BEVs, it was the 12th most influential, while for PHEVs, it was the 7th.

Finally, the Partial Dependence Plot was used to investigate the influence direction of top-ranked variables on EV purchase likelihood. For both types of EVs, the principal influences on preference over ICEVs were economic when life cycle emissions were used. For BEVs, the preference changed with respect to the price ratio was 17%, for the intention to buy an EV given a significant subsidy from the government it was 20%, and for the energy cost ratio it was a 12% change. The overall intention to buy reached nearly 50/50 (48%) when BEVs had a price ratio of 1.33, but dropped to about 1/3rd (31%) when the price ratio was 2.15. Therefore, vehicle manufacturers should try to reduce the purchase price of EVs to maximize EV sales. The government of Canada can determine EV subsidies according to the relationship between EV and ICEV purchase price ratios and the likelihood of EV preference. For instance, if the government aims to increase the share of BEVs in the vehicle market to 48%, the subsidy should reduce the BEV to ICEV purchase price ratio to 1.33.

A 4% difference in preference could be seen between those who already own an EV or those who have no car currently with those who would use the subsidy to purchase an additional vehicle or replace a current one. Those who would not consider buying an EV even with a significant discount were much less likely to choose a BEV. This latter result is quite different from the influence on PHEV preference, which changed only by 0.7 between those who already own an EV and those who stated that they would not purchase one even with a significant subsidy. For PHEVs, it was primarily with the cost ratio that preference was decided as the preference changed from 80% at parity to 39% when the PHEV was 1.5 times or more expensive. There was a 7% change in preference when the energy cost ratio changed from 0.35 to 0.9.

The intention to purchase an EV to reduce climate change emissions was the fourth most influential variable for both types of EVs. The impact was a change of roughly 3% (versus no intention) and 5% for BEVs and PHEVs, respectively. Related to this, how the GHG emissions information is presented (or framed) influenced preferences. In this study, the most effective framings related to psychological distancing: floods and fires. The Fire label could increase the probability of selecting BEVs and PHEVs over ICEVs by 2.3% and 2.1% if it is replaced with the current label in Canada. The Flood label could increase the probability of selecting BEVs and PHEVs over ICEVs by 2.4% and 1.9%, respectively. Therefore, it is recommended that the government of Canada and vehicle manufacturers apply floods and fires labels to maximize EV preferences. This study indicated further evidence that how GHG emissions information is framed can have an influence on preferences. However, the overall influence was reduced in this study, which was likely related to using lifecycle emissions rather than tailpipe emissions.

The results of this study identified which groups are more likely to buy EVs. Therefore, policy makers and vehicle manufacturers can target those groups to promote EVs. For example, this information can be used to advertise EVs optimally. That is, EV advertisement and promotion offers can be sent to the groups with a higher probability of EV preference.

Various limitations should be mentioned. The sample was not fully representative of the Canadian population for one as not all the regions of Canada are present and differences between regions can be seen. Further, anyone who possesses a driver's license could participate, but differences with those with an intention to buy in the near future might be different. From an analysis perspective, only ensemble learning techniques to predict EV preferences were applied. Therefore, it is recommended that the performance of ensemble learning techniques is compared with other powerful machine learning techniques to determine the most accurate classification technique for the mentioned problem. Moreover, it is recommended that the total cost of ownership is considered in future surveys to better capture financial implications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI in scientific writing

The authors declare that generative AI was not applied during the preparation of this work.

Acknowledgments

This research was supported by (a): the the Fonds de recherches du Québec – Nature et Technologie (FRQNT) [grant number 2019-GS-261583]; (b): Trottier Energy Institute, Ph.D. Excellence Scholarship; (c): Social Sciences and Humanities Research Council [grant number 435-20200-1292] (d): Fonds de recherches du Québec – Nature et Technologie (FRQNT) [grant number 322727]; (e): Centre interdisciplinaire de recherche en operationalisation du développement durable (CIRODD)'s program SYNERGIE.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: H N, E-O-D W, Z P, B W; data collection: H N, E-O-D W, and B W; analysis and interpretation of results: H N, E-O-D W, Z P; draft manuscript preparation: H N. All authors reviewed the results and approved the final version of the manuscript.

References

Abid, M., Tabaa, M., Chakir, A., Hachimi, H.: Routing and charging of electric vehicles: Literature review. Energy Reports. 8, 556–578 (2022). https://doi.org/10.1016/j.egyr.2022.07.089

Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211 (1991). https://doi.org/10.1016/0749-5978(91)90020-T

Ali, M., Ullah, S., Ahmad, M.S., Cheok, M.Y., Alenezi, H.: Assessing the impact of green consumption behavior and green purchase intension among millennials toward sustainable environment. Environ. Sci. Pollut. Res. 30, 23335–23347 (2022). https://doi.org/10.1007/s11356-022-23811-1

Alnahit, A.O., Mishra, A.K., Khan, A.A.: Stream water quality prediction using boosted regression tree and random forest models. Stoch. Environ. Res. Risk Assess. 36, 2661–2680 (2022). https://doi.org/10.1007/s00477-021-02152-4

Austmann, L.M.: Drivers of the electric vehicle market: A systematic literature review of empirical studies. Financ. Res. Lett. 41, 101846 (2021). https://doi.org/10.1016/j.frl.2020.101846

Bas, J., Cirillo, C., Cherchi, E.: Classification of potential electric vehicle purchasers: A machine learning approach. Technol. Forecast. Soc. Change. 168, 120759 (2021)(a). https://doi.org/10.1016/j.techfore.2021.120759

Bas, J., Zou, Z., Cirillo, C.: An interpretable machine learning approach to understanding the impacts of attitudinal and ridesourcing factors on electric vehicle adoption. Transp. Lett. (2021)(b). https://doi.org/10.1080/19427867.2021.2009098

Bonilla, D., Arias Soberon, H., Galarza, O.U.: Electric vehicle deployment & fossil fuel tax revenue in Mexico to 2050. Energy Policy. 171, 113276 (2022). https://doi.org/10.1016/j.enpol.2022.113276

Brase, G.L.: What Would It Take to Get You into an Electric Car? Consumer Perceptions and Decision Making about Electric Vehicles. J. Psychol. Interdiscip. Appl. 153, 214–236 (2019). https://doi.org/10.1080/00223980.2018.1511515

Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794. ACM, New York, NY, USA (2016)

Das, P.K., Bhat, M.Y.: Global electric vehicle adoption: implementation and policy implications for India. Environ. Sci. Pollut. Res. 29, 40612–40622 (2022). https://doi.org/10.1007/s11356-021-18211-w

Daziano, R., Waygood, E.O.D., Patterson, Z., Feinberg, M., Wang, B.: Reframing greenhouse gas emissions information presentation on the Environmental Protection Agency's new-vehicle labels to increase willingness to pay. J. Clean. Prod. (2021). https://doi.org/10.1016/j.jclepro.2020.123669

Dhananjay, B., Sivaraman, J.: Analysis and classification of heart rate using CatBoost feature ranking model. Biomed. Signal Process. Control. 68, 102610 (2021). https://doi.org/10.1016/j.bspc.2021.102610

Dixit, S.K., Singh, A.K.: Predicting Electric Vehicle (EV) Buyers in India: A Machine Learning Approach. Rev. Socionetwork Strateg. 16, 221–238 (2022). https://doi.org/10.1007/s12626-022-00109-9

Dong, Y., Sun, Y., Waygood, O., Wang, B., Huang, P., Naseri, H.: Insight into the Nonlinear Effect of COVID-19 on Well-Being in China: Commuting, a Vital Ingredient. J. Transp. Heal. 25, 101424 (2022). https://doi.org/10.1016/j.jth.2022.101424

Feinberg, M., Willer, R.: The Moral Roots of Environmental Attitudes. Psychol. Sci. 24, 56–62

(2013). https://doi.org/10.1177/0956797612449177

Ferguson, M., Mohamed, M., Higgins, C.D., Abotalebi, E., Kanaroglou, P.: How open are Canadian households to electric vehicles? A national latent class choice analysis with willingness-to-pay and metropolitan characterization. Transp. Res. Part D Transp. Environ. (2018). https://doi.org/10.1016/j.trd.2017.12.006

Fielding, K.S., Hornsey, M.J., Swim, J.K.: Developing a social psychology of climate change, (2014)

Government-of-Canada: Government of Canada confirms ambitious new greenhouse gas emissions reduction target, available at https://www.canada.ca/en/environment-climate-change/news/2021/07/government-of-canada-confirms-ambitious-new-greenhouse-gas-emissions-reduction-target.html. (2021)

Government of Canada: Canada’s top 10 weather stories of 2021 - Canada.ca, available at: https://www.canada.ca/en/environment-climate-change/services/top-ten-weather-stories/2021.html, https://www.canada.ca/en/environment-climate-change/services/top-ten-weather-stories/2021.html

Haidt, J., Joseph, C.: The Moral Mind: How Five Sets of Innate Intuitions Guide the Development of Many Culture-Specific Virtues, and Perhaps Even Modules. In: The Innate Mind (2008)

Hidrue, M.K., Parsons, G.R., Kempton, W., Gardner, M.P.: Willingness to pay for electric vehicles and their attributes. Resour. Energy Econ. 33, 686–705 (2011). https://doi.org/10.1016/j.reseneeco.2011.02.002

Higueras-Castillo, E., Guillén, A., Herrera, L.J., Liébana-Cabanillas, F.: Adoption of electric vehicles: Which factors are really important? Int. J. Sustain. Transp. 15, 799–813 (2021). https://doi.org/10.1080/15568318.2020.1818330

Hu, G., Yin, C., Wan, M., Zhang, Y., Fang, Y.: Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier. Biosyst. Eng. (2020). https://doi.org/10.1016/j.biosystemseng.2020.03.021

Hussain, S., Mustafa, M.W., Jumani, T.A., Baloch, S.K., Alotaibi, H., Khan, I., Khan, A.: A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection. Energy Reports. 7, 4425–4436 (2021). https://doi.org/10.1016/j.egyr.2021.07.008

Inglis, A., Parnell, A., Hurley, C.B.: Visualizing Variable Importance and Variable Interaction Effects in Machine Learning Models. J. Comput. Graph. Stat. 31, 766–778 (2022). https://doi.org/10.1080/10618600.2021.2007935

Jeon, H., Seo, W., Park, E., Choi, S.: Hybrid machine learning approach for popularity prediction of newly released contents of online video streaming services. Technol. Forecast. Soc. Change. 161, 120303 (2020). https://doi.org/10.1016/j.techfore.2020.120303

Kosai, S., Zakaria, S., Che, H.S., Hasanuzzaman, M., Rahim, N.A., Tan, C., Ahmad, R.D.R., Abbas, A.R., Nakano, K., Yamasue, E., Woon, W.K., Amer, A.H.A.: Estimation of Greenhouse Gas Emissions of Petrol, Biodiesel and Battery Electric Vehicles in Malaysia Based on Life Cycle Approach. Sustain. 14, 5783 (2022). https://doi.org/10.3390/su14105783

Krishnan, V.V., Koshy, B.I.: Evaluating the factors influencing purchase intention of electric vehicles in households owning conventional vehicles. Case Stud. Transp. Policy. 9, 1122–1129 (2021). https://doi.org/10.1016/j.cstp.2021.05.013

Lasarov, W., Mai, R., Hoffmann, S.: The backfire effect of sustainable social cues. New evidence on social moral licensing. Ecol. Econ. 195, 107376 (2022). https://doi.org/10.1016/j.ecolecon.2022.107376

Lashari, Z.A., Ko, J., Jang, J.: Consumers' intention to purchase electric vehicles: Influences of user attitude and perception. Sustain. 13, 6778 (2021). https://doi.org/10.3390/su13126778

Li, K., Wang, L.: Optimal electric vehicle subsidy and pricing decisions with consideration of EV anxiety and EV preference in green and non-green consumers. Transp. Res. Part E Logist. Transp. Rev. 170, 103010 (2023). https://doi.org/10.1016/j.tre.2022.103010

Li, X., Clark, C.D., Jensen, K.L., Yen, S.T., English, B.C.: Consumer purchase intentions for flexible-fuel and hybrid-electric vehicles. Transp. Res. Part D Transp. Environ. 18, 9–15 (2013). https://doi.org/10.1016/j.trd.2012.08.001

Li, Y., Zou, C., Berecibar, M., Nanini-Maury, E., Chan, J.C.W., van den Bossche, P., Van Mierlo, J., Omar, N.: Random forest regression for online capacity estimation of lithium-ion batteries. Appl. Energy. 232, 197–210 (2018). https://doi.org/10.1016/j.apenergy.2018.09.182

Ling, Z., Cherry, C.R., Wen, Y.: Determining the factors that influence electric vehicle adoption: A stated preference survey study in Beijing, China. Sustain. 13, 11719 (2021). https://doi.org/10.3390/su132111719

Liu, A., Ma, E., Qu, H., Ryan, B.: Daily green behavior as an antecedent and a moderator for visitors' pro-environmental behaviors. J. Sustain. Tour. 28, 1390–1408 (2020)(a). https://doi.org/10.1080/09669582.2020.1741598

Liu, Y., Lyu, C., Khadka, A., Zhang, W., Liu, Z.: Spatio-Temporal Ensemble Method for Car-Hailing Demand Prediction. IEEE Trans. Intell. Transp. Syst. 21, 5328–5333 (2020)(b). https://doi.org/10.1109/TITS.2019.2948790

Lu, T., Yao, E., Jin, F., Yang, Y.: Analysis of incentive policies for electric vehicle adoptions after the abolishment of purchase subsidy policy. Energy. 239, 122136 (2022). https://doi.org/10.1016/j.energy.2021.122136

Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Advances in

Neural Information Processing Systems. pp. 4766–4775 (2017)

Makeen, P., Ghali, H.A., Memon, S.: A Review of Various Fast Charging Power and Thermal Protocols for Electric Vehicles Represented by Lithium-Ion Battery Systems. Futur. Transp. 2, 281–299 (2022). https://doi.org/10.3390/futuretransp2010015

Mangalathu, S., Hwang, S.H., Jeon, J.S.: Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. Eng. Struct. 219, 110927 (2020). https://doi.org/10.1016/j.engstruct.2020.110927

McDonald, R.I., Chai, H.Y., Newell, B.R.: Personal experience and the "psychological distance" of climate change: An integrative review, (2015)

Naseri, H., Aliakbari, A., Javadian, M.A., Aliakbari, A., Waygood, E.O.D.: A novel technique for multi-objective sustainable decisions for pavement maintenance and rehabilitation. Case Stud. Constr. Mater. 20, e03037 (2024). https://doi.org/10.1016/J.CSCM.2024.E03037

Naseri, H., Ehsani, M., Golroo, A., Moghadas Nejad, F.: Sustainable pavement maintenance and rehabilitation planning using differential evolutionary programming and coyote optimisation algorithm. Int. J. Pavement Eng. 0, 1–18 (2021)(a). https://doi.org/10.1080/10298436.2021.1873331

Naseri, H., Jahanbakhsh, H., Foomajd, A., Galustanian, N., Karimi, M.M., Waygood, E.O.D.: A newly developed hybrid method on pavement maintenance and rehabilitation optimization applying Whale Optimization Algorithm and random forest regression. https://doi.org/10.1080/10298436.2022.2147672. 1–13 (2022)(a). https://doi.org/10.1080/10298436.2022.2147672

Naseri, H., Shokoohi, M., Jahanbakhsh, H., Golroo, A., Gandomi, A.H.: Evolutionary and swarm intelligence algorithms on pavement maintenance and rehabilitation planning. Int. J. Pavement

Eng. 1–15 (2021)(b). https://doi.org/10.1080/10298436.2021.1969019

Naseri, H., Waygood, E.O.D., Wang, B., Patterson, Z.: Application of Machine Learning to Child Mode Choice with a Novel Technique to Optimize Hyperparameters. Int. J. Environ. Res. Public Health. 19, 16844 (2022)(b). https://doi.org/10.3390/ijerph192416844

Naseri, H., Waygood, E.O.D., Wang, B., Patterson, Z.: Interpretable Machine Learning Approach to Predicting Electric Vehicle Buying Decisions. Transp. Res. Rec. J. Transp. Res. Board. 036119812311695 (2023). https://doi.org/10.1177/03611981231169533

Naseri, H., Waygood, E.O.D., Wang, B., Patterson, Z., Daziano, R.A.: A Novel Feature Selection Technique to Better Predict Climate Change Stage of Change. Sustainability. 14, 40 (2021)(c). https://doi.org/10.3390/su14010040

Pham, H., Jiang, X., Zhang, C.: Causality and Advanced Models in Trip Mode Prediction: Interest in Choosing Swissmetro. OSF Prepr. (2022). https://doi.org/10.31219/OSF.IO/M4W38

Pop, R.A., Saplacan, Z., Alt, M.A.: Social media goes green-the impact of social media on green cosmetics purchase motivation and intention. Inf. 11, 447 (2020). https://doi.org/10.3390/INFO11090447

Prabhavathi, C., Mahesh, A., Ajay, S.V., Mythili, R.: Malware Prediction Using XGBOOST and CATBOOST. J. Eng. Sci. 13, 620–626 (2022)

Ramos-Real, F.J., Ramírez-Díaz, A., Marrero, G.A., Perez, Y.: Willingness to pay for electric vehicles in island regions: The case of Tenerife (Canary Islands), (2018)

Ritchie, H.: Tracking global data on electric vehicles - Our World in Data. (2024)

Rudolph, C.: How may incentives for electric cars affect purchase decisions? Transp. Policy. 52, 113–120 (2016). https://doi.org/10.1016/j.tranpol.2016.07.014

Ryoo, Y., Hyun, N.K., Sung, Y.: The Effect of Descriptive Norms and Construal Level on

Consumers' Sustainable Behaviors. J. Advert. 46, 536–549 (2017). https://doi.org/10.1080/00913367.2017.1396514

Salas, P., De la Fuente, R., Astroza, S., Carrasco, J.A.: A systematic comparative evaluation of machine learning classifiers and discrete choice models for travel mode choice in the presence of response heterogeneity. Expert Syst. Appl. 193, 116253 (2022). https://doi.org/10.1016/j.eswa.2021.116253

Simsekoglu, Ö., Nayum, A.: Predictors of intention to buy a battery electric vehicle among conventional car drivers. Transp. Res. Part F Traffic Psychol. Behav. 60, 1–10 (2019). https://doi.org/10.1016/j.trf.2018.10.001

Smith, J.R., Louis, W.R., Terry, D.J., Greenaway, K.H., Clarke, M.R., Cheng, X.: Congruent or conflicted? The impact of injunctive and descriptive norms on environmental intentions. J. Environ. Psychol. 32, 353–361 (2012). https://doi.org/10.1016/j.jenvp.2012.06.001

Sobiech-Grabka, K., Stankowska, A., Jerzak, K.: Determinants of Electric Cars Purchase Intention in Poland: Personal Attitudes v. Economic Arguments. Energies. 15, 3078 (2022). https://doi.org/10.3390/en15093078

De Souza, L.L.P., Lora, E.E.S., Palacio, J.C.E., Rocha, M.H., Renó, M.L.G., Venturini, O.J.: Comparative environmental life cycle assessment of conventional vehicles with different fuel options, plug-in hybrid and electric vehicles for a sustainable transportation system in Brazil. J. Clean. Prod. (2018). https://doi.org/10.1016/j.jclepro.2018.08.236

Sun, Y., Dong, Y., D. Waygood, E.O., Naseri, H., Jiang, Y., Chen, Y.: Machine-Learning Approaches to Identify Travel Modes Using Smartphone-Assisted Survey and Map Application Programming Interface. Transp. Res. Rec. J. Transp. Res. Board. 036119812211064 (2022). https://doi.org/10.1177/03611981221106483

Tazeen, F., Mullick, N.H.: The Impact of Social Media Platforms 'Facebook and Instagram' in Influencing Purchasing Behaviour of Green Products. Vision. (2023). https://doi.org/10.1177/09722629221133960

Tunçel, N.: Intention to purchase electric vehicles: Evidence from an emerging market. Res. Transp. Bus. Manag. 43, 100764 (2022). https://doi.org/10.1016/j.rtbm.2021.100764

Unterluggauer, T., Rich, J., Andersen, P.B., Hashemi, S.: Electric vehicle charging infrastructure planning for integrated transportation and power distribution networks: A review, (2022)

Vafaei-Zadeh, A., Wong, T.K., Hanifah, H., Teoh, A.P., Nawaser, K.: Modelling electric vehicle purchase intention among generation Y consumers in Malaysia. Res. Transp. Bus. Manag. 43, 100784 (2022). https://doi.org/10.1016/j.rtbm.2022.100784

Wang, B., Waygood, E.O.D., Daziano, R.A., Patterson, Z., Feinberg, M.: Does hedonic framing improve people's willingness-to-pay for vehicle greenhouse gas emissions? Transp. Res. Part D Transp. Environ. 98, 102973 (2021). https://doi.org/10.1016/j.trd.2021.102973

Wang, B., Waygood, E.O.D., Ji, X., Naseri, H., Loiselle, A.L., Daziano, R.A., Patterson, Z., Feinberg, M.: How to effectively communicate about greenhouse gas emissions with different populations. Environ. Sci. Policy. 147, 29–43 (2023). https://doi.org/10.1016/J.ENVSCI.2023.05.015

Wang, J.A., Sulla-Menashe, D., Woodcock, C.E., Sonnentag, O., Keeling, R.F., Friedl, M.A.: Extensive land cover change across Arctic–Boreal Northwestern North America from disturbance and climate forcing. Glob. Chang. Biol. 26, 807–822 (2020). https://doi.org/10.1111/gcb.14804

Wu, Y.A., Ng, A.W., Yu, Z., Huang, J., Meng, K., Dong, Z.Y.: A review of evolutionary policy incentives for sustainable development of electric vehicles in China: Strategic implications. Energy Policy. 148, 111983 (2021). https://doi.org/10.1016/j.enpol.2020.111983

Xun, J., Waygood, E.O.D., Wang, B., Naseri, H., Loiselle, L.A., Patterson, Zachary Feinberg, M.: Exploring the Effects of New Framing Techniques for Greenhouse Gas Emissions. In: Transportation Research Board 101st Annual Meeting. , Washington DC, United States (2022)

Zhang, H., Tang, L., Yang, C., Lan, S.: Locating electric vehicle charging stations with service capacity using the improved whale optimization algorithm. Adv. Eng. Informatics. 41, 100901 (2019). https://doi.org/10.1016/j.aei.2019.02.006

Zhou, Z., Wang, M., Huang, J., Lin, S., Lv, Z.: Blockchain in Big Data Security for Intelligent Transportation With 6G. IEEE Trans. Intell. Transp. Syst. (2021). https://doi.org/10.1109/TITS.2021.3107011

Zhu, S., Zhu, F.: Cycling comfort evaluation with instrumented probe bicycle. Transp. Res. Part A Policy Pract. 129, 217–231 (2019). https://doi.org/10.1016/j.tra.2019.08.009

Appendix

Table A1. The list of variables applied in the modeling.

Socio-demographic	Actions to reduce GHG emissions
Age	Will you, or have you upgraded light bulbs to reduce GHG emissions?
Household income	Will you, or have hanged dry clothes to reduce GHG emissions?
Education attainment	Will you, or have you taken short shower rather than bath to reduce GHG emissions?
Gender	Will you, or have you recycle to reduce GHG emissions?
Province	Will you, or have you washed clothes in cold water to reduce GHG emissions?
Ethnicity	Will you, or have you used a plug-in hybrid car to reduce GHG emissions?
Language	Will you, or have you eaten a plant based diet to reduce GHG emissions?
Vehicle ownership	Will you, or have you used an electric car to reduce GHG emissions?
Employment status	Will you, or have you used an electric-assist bicycle to reduce GHG emissions?
How many people are in your household including you?	Will you, or have you bought green energy to reduce GHG emissions?

Political spectrum	Will you, or have you avoided one long-haul flight to reduce GHG emissions?
How many people have a driver’s license in your household?	Will you, or have you lived vehicle free to reduce GHG emissions?
Do you have plans to own or lease a vehicle in the next year ?	Will you, or have you had one fewer child to reduce GHG emissions?
Do you have any children aged under 6 years old?	Will you, or have you vote for parties that support strong climate action to reduce GHG emissions?
Do you have any children aged between 6 and 12 years old?	Will you, or have you volunteered or contributed to groups that support strong climate action to reduce GHG emissions?
Do you have any children aged between 13 and 18 years old?	Descriptive norm
Do you have any children aged over 18 years old?	In your opinion, what percentage of people support government action to reduce emissions causing climate change?
Do you have any children?	In your opinion, what percentage of people in your province believe in human-caused climate change?
Have you driven a personal car/truck/SUV/etc. in the past 12 months?	Social media usage
Vehicle Attributes	At what frequency do you use Facebook?
Purchase price ratio of EV to ICEV	At what frequency do you use Instagram?
Electricity cost of EV to fuel cost of ICEV	At what frequency do you use Twitter?
EV range	At what frequency do you use TikTok?

Emission ratio of EV to ICEV	At what frequency do you use Snapchat?
Climate change beliefs and attitudes	Do you use social media to find and share information (news, opinions, etc.)
Climate change stage of change	Do you use social media to socialise with friends and family
Worried about climate change	Do you use social media for entertainment (games, videos, etc.)
Do you think human activity is a major cause of climate change?	Do you know the reason for your social media usage?
How much do you support or oppose a carbon pollution tax?	To what extent do you trust the information that you see on social media sites
Do you support government action on climate change?	To what extent do you consult your social media when making purchase decisions
Mobility restriction	Opinion about other sustainable transportation modes
Can you walk for more than a short time?	Are you the kind of person who rides a bicycle?
Can you use a bicycle?	Do you feel you should cycle more to keep fit?
Can you use public transport?	Do you find cycling stressful?
Attitude toward EVs	Do you think cycling can be the quickest way to get around?
The energy for electric vehicles costs less per kilometer (vs gasoline), so effectively the cost of driving would be less. If you bought an electric vehicle, you would drive less, more, or the same?	Do you like travelling by bicycle?

For Canada, in which year the sale of new internal combustion vehicles should be eliminated?	Are you the kind of person to take the bus?
If the government gave a significant discount to electric vehicles, what would you do?	In general, would you rather cycle than take the bus?
In your opinion, can buying an electric vehicle solve problems associated with climate change emissions?	Are you the kind of person that likes to walk a lot?
In your opinion, can buying an electric vehicle solve problems associated with congestion?	Do you feel you should walk more to keep fit?
In your opinion, can buying an electric vehicle solve problems associated with air quality?	Do you like traveling by walking?
In your opinion, can buying an electric vehicle solve problems associated with parking?	Do you feel a moral obligation to reduce your emissions of greenhouse gases?
In your opinion, can buying an electric vehicle solve problems associated with traffic safety?	Do you agree people should be allowed to use their cars as much as they like?
If you owned an electric vehicle, how much would you be happy?	GHG framing
If you owned an electric vehicle, how much would you be proud?	Treatment
If you owned an electric vehicle, how much would you be excited?	

A1. Ensemble learning methods

The applied ensemble learning methods are briefly explained in this section.

A1.1 Light Gradient Boosting Machine (LGBM)

Light Gradient Boosting Machine (LGBM) is a powerful and fast machine learning algorithm that can be used to model classification and regression problems and capture the importance of independent variables on the response variable. Microsoft developed LGBM, which is an updated version of conventional boosting methods. Since LGBM is an ensemble learning method, it merges some weak learners to generate a robust classifier (Liu et al. 2020b). A parallel learning process is utilized by LGBM, which can considerably reduce computational costs and memory usage. LGBM uses a leaf-wise leaf growth strategy, limiting the depth growth when splitting the data samples (Naseri et al. 2022b). The same layer of leaves can be simultaneously split using the mentioned leaf growth strategy. Accordingly, multi-threaded optimization can be implemented in LGBM. As a result, the model's complexity is automatically controlled, and the model is less likely to be overfitted (Zhou et al. 2021).

A1.2 Random Forest (RF)

Random Forest (RF) is a prediction method, which can be used for regression, classification, and outlier detection. This model is efficient for modeling both large-scale and small-scale datasets (Naseri et al. 2022a). RF generates different decision trees with different forms. Then, it aggregates the generated decision trees with a bagging technique to develop a powerful estimation model. Subsequently, all decision tree algorithms are executed, and they predict the response variable individually (Naseri et al. 2021b). RF predicts the response variable using all the values predicted by decision trees and a majority voting process (Li et al. 2018). RF cedes intelligibility and swallows computation power since it combines skill enhancement and decision tree models (Sun et al., 2022).

2.2.3. eXtreme Gradient Boosting (XGB)

XGB is a robust ensemble learning technique extensively used for prediction purposes. XGB uses parallel learning and is a fast prediction method (Jeon et al. 2020). This method can accurately model complex prediction problems (Chen and Guestrin 2016). Like other ensemble learning methods, XGB generates a given number of decision tree algorithms and combines them to form an ensemble prediction model. XGB iteratively increases the prediction accuracy using an optimization framework. The purpose of the optimization framework is to find the best structure for decision trees by minimizing their corresponding prediction error in each iteration. XGB uses first-order and second-order gradients to optimize the mentioned structures. Moreover, the objective function of the applied optimization problem contains a regularized term, which can reduce the likelihood of overfitting (Zhu and Zhu 2019).

2.2.4. Adaptive Boosting (AB)

Adaptive Boosting (AB) is an iterative ensemble learning technique. This technique applies a boosting process to combine weak learners. AB generate a given number of weak learners and assign them equal weights. Then, the weak learners are implemented, and their prediction error is evaluated. Consequently, the weights of weak learners are updated based on their prediction error in each iteration (Naseri et al. 2021c). That is, higher weights are assigned to learners with fewer prediction errors. During different iterations, the structure of weak learners and their weights are optimized to obtain an accurate prediction model (Hu et al. 2020).

2.2.5. Categorical Boosting (CB)

Categorical Boosting (CB) is an ensemble learning method, solving classification and regression prediction problems using gradient boosting. CB applies Bayesian estimators to reduce the probability of overfitting (Dhananjay and Sivaraman 2021). CB replaces categorical variables with binary variables in all iterations, reducing the computational cost. This method provides efficient target-based statistics for modeling categorical input variables, which reduces the runtime (Hussain et al. 2021). One of the main differences between CB with robust ensemble learning techniques (e.g., XGB) is its symmetric (balanced) structure. As part of the splitting procedure, CB uses balanced tree structures, keeping the same depth of the tree across all nodes (Prabhavathi et al. 2022).