Combining stated and revealed choice research to simulate the neighbor effect: The case of hybrid-electric vehicles

Jonn Axsen [a,*], Dean C. Mountain [b], Mark Jaccard [c]

[a] Institute of Transportation Studies, University of California at Davis, 2028 Academic Surge, One Shields Avenue, Davis, CA 95616, USA
[b] DeGroote School of Business, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4M4, Canada
[c] School of Resource and Environmental Management, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada

ABSTRACT

According to intuition and theories of diffusion, consumer preferences develop along with technological change. However, most economic models designed for policy simulation unrealistically assume static preferences. To improve the behavioral realism of an energy–economy policy model, this study investigates the "neighbor effect," where a new technology becomes more desirable as its adoption becomes more widespread in the market. We measure this effect as a change in aggregated willingness to pay under different levels of technology penetration. Focusing on hybrid-electric vehicles (HEVs), an online survey experiment collected stated preference (SP) data from 535 Canadian and 408 Californian vehicle owners under different hypothetical market conditions.
Revealed preference (RP) data was collected from the same respondents by eliciting the year, make and model of recent vehicle purchases from regions with different degrees of HEV popularity: Canada with 0.17% new market share, and California with 3.0% new market share. We compare choice models estimated from RP data only with three joint SP–RP estimation techniques, each assigning a different weight to the influence of SP and RP data in coefficient estimates. Statistically, models allowing more RP influence outperform SP influenced models. However, results suggest that because the RP data in this study is afflicted by multicollinearity, techniques that allow more SP influence in the beta estimates while maintaining RP data for calibrating vehicle

* Corresponding author at: Institute of Transportation Studies, University of California at Davis, 2028 Academic Surge, One Shields Avenue, Davis, CA 95616, USA. Tel.: +1 530 231 5007; fax: +1 530 752 6572.
 E-mail address: jaxsen@ucdavis.edu (J. Axsen).

class constraints produce more realistic estimates of willingness to pay. Furthermore, SP influenced coefficient estimates also translate to more realistic behavioral parameters for CIMS, allowing more sensitivity to policy simulations.

1. Introduction

The development and adoption of low and zero emissions technologies could curb emissions growth without substantial changes in human activity. This process of induced technological change is a lever that policymakers are increasingly looking to as a line of least resistance in reaching greenhouse gas (GHG) targets (Azar and Dowlatabadi, 1999). In road transportation, which accounts for 20% of Canada's GHG emissions (Environment Canada, 2004), the widespread adoption of hybrid-electric vehicles (HEVs) could reduce the GHG intensity of passenger vehicle travel (gCO_2/km). One key ingredient in the diffusion of such new technologies is the tendency for consumer preferences to change as the technology becomes more prevalent in the market—we refer to this tendency as the "neighbor effect," a term recently employed by Mau et al. (2008). This effect represents the tendency for the social costs of switching to a new technology to decrease as the adoption rate increases due to changes in social concerns, increased credibility, and learning from others with more experience (Yang and Allenby, 2003), as well as education, marketing and shifts in cultural norms (Norton et al., 1998).

One objective of this study is to estimate preference dynamics associated with the adoption of HEVs to improve the behavioral realism of CIMS, an energy–economy model which has been frequently applied to real-world policy research (e.g., Jaccard et al., 2003; Horne et al., 2005; Bataille et al., 2006). In addition to financial attributes, such as purchase price and fuel costs, CIMS represents the intangible costs associated with a new technology, such as consumer perceptions of quality, reliability, availability, social desirability or popularity. The neighbor effect is simulated as a decrease in intangible costs with increased market share. In this study, we seek to better represent the diffusion of new technologies at an economy-wide level. Because CIMS is inherently aggregate, we do not seek to individually identify the underlying components of preference change.

To investigate the preference dynamics for HEV adoption, discrete choice models are estimated with revealed (RP) and stated preference (SP) data, collected from two samples with different HEV market conditions (Canada and California). Discrete choice modeling has been frequently applied to low-emissions vehicles (e.g., Bunch et al., 1993; Ewing and Sarigollu, 2000; Potoglou and Kanaroglou, 2007). Because SP and RP data can have complementary strengths, a growing body of literature demonstrates the potential for "joint" modeling techniques (e.g., Swait et al., 1994; Brownstone et al., 2000; Train, 2003; Hensher et al., 2005). While RP data can realistically account for income and supply constraints, SP data can better represent hypothetical market conditions and eliminate problems of multicollinearity. We investigate joint modeling techniques applying different weights to SP and RP data. We are concerned not just with the statistical performance of such techniques, but also with subsequent implications for behavioral parameter estimates and policy simulations in CIMS.

The principle objectives of this study are to explore: (1) the role and value of the neighbor effect, (2) the use of SP and RP choice research to improve the behavioral realism of an energy–economy model, CIMS, (3) the value of integrating joint SP and RP modeling procedures to explain HEV vehicle choice and the influential tradeoff coefficients related to capital, fuel and performance, and (4) the implications of including such preference dynamics in policy simulations. Section 2 outlines CIMS and the multinomial logit (MNL) choice models used in this study. Section 3 details our data collection via a web-based survey, while Section 4 presents the results of SP, RP and joint choice model estimation. Section 5 illustrates the translation of choice model coefficients into behavioral parameters for CIMS. Section 6 demonstrates policy simulations with the behaviorally improved CIMS model, and Section 7 concludes.

2. The models

2.1. The energy–economy model: CIMS

CIMS is an energy–economy model that simulates the costs and GHG effects of a given abatement policy over a series of 5-year periods. CIMS is known as a "hybrid" model (Bohringer, 1998; Jaccard et al., 2003) because it combines the technological detail of "bottom-up" models with the behavioral realism and macro-economic equilibrium effects of "top-down" models (Jaffe and Stavins, 1994; Grubb et al., 2002). CIMS is technologically explicit in that it details thousands of technologies that compete for market share in a given node. The passenger vehicle node consists of technologies powered by gasoline and alternative fuels—the present study focuses only on conventional and hybrid-electric gasoline technologies. For a given period, market share is determined using a function that considers both the financial costs and intangible costs of each technology, j. The lifecycle cost (LCC) of j is:

$$LCC_j = [(CC_j + i_j) \times CRF + MC_j + EC_j] \tag{1}$$

where CC_j is the initial capital cost, i_j is the perceived intangible cost, MC_j is the annual maintenance costs, and EC_j is the annual energy cost of j. The cost recovery factor, CRF, is used to annualize upfront costs (CC_j and i_j), calculated as:

$$CRF = \frac{r}{1 - (1+r)^{-n_j}} \tag{2}$$

where r is the perceived discount rate of the decision maker, and n_j is the lifespan of technology j. The market share of technology j, MS_j, relative to a technology set K is determined by comparing LCCs as follows:

$$MS_j = \frac{LCC_j^{-\nu}}{\sum_{k=1}^{K}\{LCC_j^{-\nu}\}} \tag{3}$$

The ν parameter is a measure of market heterogeneity, representing how different consumers experience or perceive different LCCs across the economy. A high ν parameter indicates that consumers have relatively uniform preferences; a technology with a lower LCC will capture almost the entire market. In contrast, a low ν indicates the market shares for a technology set will be distributed relatively evenly, even if LCC values are substantially different. Thus, the ν parameter also dictates the sensitivity of the model to cost changes, where a higher ν estimate indicates higher sensitivity. More detailed discussions of the ν parameter can be found in Rivers and Jaccard (2005).

The purpose of this study is to improve the behavioral realism of CIMS with three parameters: r, ν, and of primary interest, i_j. CIMS has the capability to represent neighbor effects with the i_j parameter, following the declining intangible cost function.

$$i_j(t) = i_{\text{fixed}} + \frac{i_{\text{Var}}(0)}{1 + A\, e^{k \times MS_{t-1}}} \tag{4}$$

where $i_j(t)$ is the intangible cost of technology j at time t, i_{fixed} is the portion of initial intangible cost that is static, $i_{\text{Var}}(0)$ is the variable portion of intangible cost in time period zero, MS_{t-1} is the market share of the technology in the previous simulation period ($t-1$), and the A and k parameters represent the curve and rate of change of the intangible cost in response to increases in technology market share. We estimate this intangible cost curve and the ν and r parameters using empirically derived discrete choice models.

2.2. Choice models combining stated and revealed preference data

Discrete choice models quantify consumer trade-offs among product attributes (Train, 2003). Previous research has demonstrated the use of multinomial logit (MNL) models to empirically estimate behavioral parameters for CIMS (Rivers and Jaccard, 2005; Horne et al., 2005; Mau et al.,

2008). The MNL is based on random utility theory, assuming that a portion of the utility derived by an individual is unobservable. An individual's utility is broken into two components:

$$U_j = V_j + \varepsilon_j \tag{5}$$

where U_j, the utility of choice j, is the sum of V_j, observable or "representative" utility, and ε_j, unobservable utility. ε_j is treated as a random parameter with a mean of zero, following a Weibull distribution. This distribution simplifies the model, allowing estimation without simulation. Observable utility, V_j, is represented as:

$$V_j = X_j\beta + \alpha_j \tag{6}$$

where X_j is a vector of the attributes of choice j, β is a vector of coefficients weighting each of those attributes, and α_j is the alternative-specific constant, which represents the observable utility of each choice not captured by attributes specified in the model.

Similar to the market share function in CIMS, MNL models can estimate the probability of option j being chosen from choice set k, using Eq. (7). This probability can be equated with market share if the representative utility function is estimated from a large enough sample size, depicted as follows:

$$MS_j = \frac{e^{V_j}}{\sum_{k=1}^{K} e^{V_k}} \tag{7}$$

where MS_j is the estimated market share of choice j, which compares the observable utility of choice j to the observable utilities across the choice set K.[1]

Choice models can be estimated from SP data, RP data, or a combination thereof. Brownstone et al. (2000) combined SP and RP data to model vehicle preferences in California, concluding that the resulting joint MNL model was more robust than either the SP or the RP models alone. However, combining data sources is a complex process and requires judgment on the part of the researcher. Louviere et al. (2000) identify two main techniques, "pooled" and "sequential" estimation, that differ in how model coefficients are combined from the SP and RP data sources. Consider the two types of coefficients that make up the utility function in Eq. (6): the attribute coefficients (β) represent trade-offs among technology attributes; and the alternative specific constant terms (α) represent utility not captured in the specified β vector, and are responsible for calibrating RP models to fit observed market shares. The data "pooling" approach to joint estimation combines SP and RP data to estimate the β vector from both sources (β^{joint}) – where RP and SP influence can be weighted differently – while the α is estimated from the RP data only (α^{RP}). In contrast, the "sequential" approach estimates separate SP and RP models, then discards α^{SP} and the β^{RP} vector. The β^{SP} vector and α^{RP} are placed in a composite utility function, where α^{RP} is recalibrated to fit the real-life market shares represented in the RP data. Both methods of joint estimation have been successfully applied in various studies. Deciding between these methods depends on one's confidence in the SP and RP data; the "sequential" technique may be preferred if RP data is vulnerable to multicollinearity (Swait et al., 1994), while the "pooling" technique may be more appropriate if β vector estimates could benefit from both SP and RP influence.

In both techniques, researchers must account for different scale in observable utility (β and α estimates) relative to unobservable utility (ε_j) when combining coefficients from different models. Scale differences can arise because SP models hold constant all non-specified attributes while RP models cannot, often resulting in a larger ε_j variance for RP models. Train (2003) describes how to introduce a scale parameter, λ, to the utility function of both models, where the RP scale factor, λ^{RP}, is normalized to zero. Following the "sequential" technique, the β^{SP} vector extracted from the SP model would have to be adjusted by a scale factor, λ^{SP}. The resulting composite utility function would be:

$$V_j = X_j\left(\frac{\beta^{\text{SP}}}{\lambda^{\text{SP}}}\right) + \alpha_j^{\text{RP}} \tag{8}$$

---

[1] Eq. (7) typically estimates the probability of an individual choosing a particular technology. However, our interest is in aggregate preferences, so Eq. (7) is set to represent the economy using averaged attribute levels, without demographic variables (except income, which is also fixed to the sample average).

With this specification, λ^{SP} reflects the variance of unobserved factors in SP situation relative to RP situations. In the present study, scale parameters were estimated using an "artificial tree structure" nested logit technique, detailed in full by Louviere et al. (2000) and Hensher et al. (2005).

2.3. Translating choice model coefficients to an energy–economy model

Discrete choice model coefficients can be translated to the three behavioral parameters in CIMS: r, i, and v. Train (1985) explains how an implicit discount rate, r, can be estimated according to the ratio between β coefficients in a logit model:

$$\frac{\beta_{CC}}{\beta_{FC}} = \frac{r}{(1 - (1 + r)^{-n})} \tag{9}$$

where β_{CC} is the capital cost coefficient, β_{FC} is the coefficient for annual fuel costs, and n is the technology lifespan.[2] In short, this ratio represents the tradeoff between capital and operating costs—an estimate of the time value of money. A product with infinite product life has a discount rate equal to β_{CC}/β_{FC}.

The intangible costs of HEVs are estimated as the sum of ratios between each non-monetary coefficient and constant (Alberini et al., 2007), and the capital cost coefficient:

$$i_j = \sum_{n=1}^{N} \left(\frac{\beta_{nj}}{\beta_{CC}} \times X_{nj} \right) \tag{10}$$

where N is number of non-monetary attributes, β_{nj} is the coefficient for attribute n of technology j, X_n is the value for the non-monetary attribute n of technology j, and β_{CC} is the capital cost coefficient. In this study, we estimate i_j values using two non-monetary attributes: vehicle horsepower, β_{HP}, and the HEV constant, α_{HEV}. The declining A and k parameters of the intangible cost function, Eq. (4), can be fitted to i parameter estimates from several different HEV market share scenarios using the Solver algorithm in the Excel spreadsheet software package.

The third behavioral parameter, v, cannot be estimated directly from choice model coefficients. Instead, for several choice sets, Solver is used to select a v that consistently equates the market share predictions of the CIMS algorithm, Eq. (3), and the choice model, Eq. (7), such that:

$$\frac{e^{V_i}}{\sum_{j=1}^{J} e^{V_j}} = \frac{(LCC_j)^{-v}}{\sum_{k=1}^{K} (LCC_k)^{-v}} \tag{11}$$

This procedure can only be performed after all the terms of the *LCC* have been estimated. While there may be concern that the two sides of Eq. (11) are from substantively different models, this procedure serves to approximate the policy sensitivity of the CIMS market share algorithm to that of the empirically estimated choice model, and has been applied successfully in several previous studies (Horne et al., 2005; Rivers and Jaccard, 2005; Mau et al., 2008).

3. The data

SP and RP data were collected with a web-based survey conducted in March 2006. Survey samples were drawn from online panels constructed and maintained by market research companies in Canada and California. Because HEVs accounted for 0.17% (Autonews, 2006) and 3.0% (R.L. Polk and Co., 2006) of 2005 new vehicle market share in Canada and California, respectively, we expected to observe different RP intangible cost estimates between these regions. In both samples, respondents were screened to assure they (1) owned a new vehicle purchased in 2002 or later (when HEVs were reasonably available in the North American market), (2) were age 19 or older, (3) drove three times or more per week, and (4) resided in an urban centre. HEV owners were purposely overrepresented using

---

[2] Train's (1985) original formula specified operating costs instead of fuel costs. Hausman (1979) makes a similar inference using a qualitative choice model describing the purchase of room air conditioners.

a choice-based sampling strategy.[3] In a panel-based survey, it is difficult to estimate a response rate that is comparable to other survey modes.[4] However, we can report the ratio of completed surveys to invited respondents: 14% for the Canada sample and 2% for the California sample.[5] The latter ratio is smaller due to a less targeted sampling strategy, i.e., more invitations were sent to ineligible individuals. The final breakdown of respondents was 535 Canadians and 408 Californians, including 51 and 50 HEV owners, respectively. Each respondent completed 18 SP choice sets, yielding 9630 SP choice observations in the Canada sample and 7344 in the California sample.

The Internet was chosen as a survey medium due to its many benefits: flexibility of survey design, control over response environment, automation of data entry, and low cost per respondent. However, the Internet has three main drawbacks. First is the tendency for low response rates relative to mail and phone survey methods (Couper, 2002). Secondly, and more importantly, is vulnerability to non-response bias, where differences between respondents and non-respondents are correlated with survey variables. However, there is not a clear relationship between response rate and non-response bias (Groves, 2007; Champ and Welsh, 2007); in other words, a higher response rate does not necessarily make a sample more representative of the target population. However, Couper et al. (2007) note that web survey respondents tend to be disproportionately younger and of higher socioeconomic status than non-respondents. A third limitation of web-based surveys is coverage error, where as of 2006, Internet penetration in Canada was 67.5% of the population (Internet World Stats, 2007).

To test for non-response and coverage bias in our samples, we assessed the representativeness of demographic distributions. The Canada sample was satisfactorily consistent with the geographic, gender, and age distributions of 2001 Canada Census data, as well as car/truck ownership split reported by Environment Canada (2004) data. Respondent income and education were slightly higher than census average, but both biases are consistent with findings that new car buying households tend to have higher income and education than the general population (Axsen and Kurani, 2008). Also, the demographic and attitudinal characteristics of the California sample were highly consistent with the Canada sample, justifying the extrapolation of results from the California data to represent a future version of the Canada market.

3.1. SP data collection

We designed an experiment to collect SP data under different market share conditions, randomly dividing each sample into three treatment groups. We used a common market research technique to create hypothetical scenarios using multimedia stimuli to forecast consumer responses to new technologies (Urban et al., 1996, 1997; Hoeffler, 2003). All respondents received information about HEVs, but details differed according to three penetration scenarios: (1) current HEV market share (0.17% in Canada, 3.0% in California), (2) moderate market share (10%), and (3) high market share (50%). We refer to these scenarios as MS1, MS2 and MS3, respectively.

To simulate consumer learning and word-of-mouth effects, each scenario included hypothetical information (based on real-world language) from three different source categories: a newspaper article, a manufacturer brochure, and personal testimonials (Table 1).[6] Lower penetration scenarios communicated more uncertainty about HEV technology from individuals with less direct HEV

---

[3] Choice-based sampling is used for research on technologies with low incidence in the general population; had we not oversampled HEV owners, our Canada sample would not have included a statistically significant number.

[4] A large number of invitations are emailed out to a panel maintained by the hired company, and once the sample quota is reached, further access to the web survey is blocked. Thus, we cannot truly estimate the percentage of contacted individuals that would have completed the survey, so a response rate is not directly comparable with that of a mail or telephone survey.

[5] Of the 3973 invitations that were sent out in the Canada sample, 1230 (31%) respondents logged into the survey, 538 of which did not meet initial screening criteria, and 137 qualifying respondents did not finish. Of the (approximately) 22,000 invitations sent out in the California sample, 2676 (12%) respondents logged into the survey, 1706 of which did not meet initial screening criteria, and 532 qualifying respondents did not finish.

[6] In reality it would take time for peoples' references or perceptions to change as learning and observations of actual market shares occur. And while our SP approach cannot uncover the complete picture of the time path of this learning process, our hypothetical scenarios allow us to gain important insight on these effects. Indeed, as described later, through these scenarios we do pick up statistically the significance of the effects of increased HEV penetration and of correspondingly more reliable information.

Table 1
Descriptions of HEV information packages.

Information	"Current" HEV share (MS1)	"Moderate" HEV share (MS2)	"High" HEV share (MS3)
HEV market share	0.17% (Canada), 3.0% (California)	10% (both regions)	50% (both regions)
Newspaper article	HEV sales, costs and benefits	Same as MS1 (adjusted to reflect HEV sales)	Same as MS1 (adjusted to reflect HEV sales)
Advertising brochures	Target: innovators Emphasis: cutting edge technology	Target: early adopters Emphasis: fuel savings and environmental benefits	Target: early majority Emphasis: financial savings
Testimonials	Sources: 1 friend, 1 stranger Little direct HEV experience High uncertainty	Sources: 2 friends, 1 stranger More HEV experience Less uncertainty	Sources: 2 friends, 1 family member More HEV experience Least uncertainty
Model availability for SP choice set	Limited: small car small SUV	Limited: all cars small SUV minivan small pickup	Unlimited

Table 2
Attribute levels in SP experiment (3^7 factorial design).

	Gasoline vehicle (GAS)	Hybrid-electric vehicle (HEV)
Capital cost (*CC*) (CAN $)	(1) 100% user [a] *CC*, (2) 90% user *CC* and (3) 75% user *CC*	(1) 110% user *CC*, (2) 120% user *CC* and (3) 150% user *CC*
Subsidy (*SUB*)[b] (CAN $)	None	(1) None, (2) 5% of HEV *CC* and (3) 10% of HEV *CC*
Performance (*HP*)[c] (Horsepower)	User *HP*	(1) 70% user *HP*, (2) 85% user *HP* and (3) 115% user *HP*
Fuel efficiency (*FE*)[d] (L/100 km)	(1) 80% (2) 100% (3) 120%	(1) 50%, (2) 75% and (3) 90%
Fuel price (*FP*)[d] (CAN $ per liter)	(1) 50% (2) 100% (3) 150%	
Fuel cost (*FC*)[e] (CAN $ per week)	(User *FC*) × (GAS *FE*) × *FP*	(User *FC*) × (HEV *FE*) × *FP*

[a] "User" indicates either the self-reported value entered by the respondent or the value automatically drawn from the vehicle database (based on the respondent's vehicle model).
[b] Described as a tax rebate received within 6 months of purchase.
[c] Not all respondents understand measurements of horsepower; we presented this attribute in terms of percentage change in performance (e.g., "15% better performance than your current vehicle"), as well as a corresponding horsepower quantity.
[d] Attribute not shown to respondent; only used to calculate weekly fuel cost and to estimate "pollution" level (see Table 3).
[e] The product of the users weekly fuel cost, the relative fuel efficiency of the vehicle, and the relative fuel price. For example, if the respondent normally spends $20 per week on fuel, and the presented vehicle has 80% fuel efficiency at a 150% fuel price, the presented attribute would be $20 × 0.80 × 1.50 = $24 per week.

experience. We assumed that advertising would tailor to different stages of product diffusion (following Rogers, 2003), following the innovator, early adopter and early/late majority stages summarized in Table 1. We also manipulated the availability of HEV models (or variety) because short-term limitations could be an important explanatory factor in HEV popularity.

Following the experimental treatment, respondents completed a series of binary choices, each presenting an HEV and conventional vehicle with varying attributes. Table 2 depicts the experimental design and Table 3 illustrates one hypothetical choice set. We specified 7 attributes with three levels, yielding a 3^7 factorial design that was simplified into a "main-effects only" orthogonal fractional factorial design of 18 choice sets.[7] The 7 attributes subject to the factorial design were: capital cost of the conventional vehicle, capital cost of the hybrid vehicle, subsidy provided for the hybrid vehicle, horsepower of the hybrid vehicle, conventional vehicle fuel efficiency, hybrid vehicle fuel efficiency, and gasoline price. While capital cost, subsidy and performance attributes were presented directly in

---

[7] An orthogonal factorial design has zero or negligible correlation among attributes, thus avoiding collinearity problems in the specification of choice models. We did not require the ability to model interactions, where resulting coefficients could not be easily translated to CIMS.

Table 3
Example SP choice set.

	Medium SUV gasoline vehicle	Small SUV hybrid-electric vehicle
Purchase price	$27,500	$35,000
Fuel cost per week	$35	$26
Pollution[a]	Same as current vehicle	25% less than current vehicle
Subsidy on purchase price	No subsidy	$3,500
Car performance	Same as current vehicle (150 HP)	15% better than current vehicle (172 HP)

[a] "Pollution" was not included as an independent attribute in the experimental design; a percentage increase or reduction was portrayed relative to their current vehicle, proportional to the *FE* attribute in Table 2.

the respondent choice sets, fuel efficiency and fuel price were aggregated into a single attribute (as well as a single coefficient): weekly fuel cost. The purpose of this aggregation was to allow for variation in fuel cost derived from technological advancements (for one vehicle) or fuel price (across vehicles), yielding a "cognitively sensible" range of fuel costs across choice sets (Hensher et al., 2005; Hensher, 2007).[8] All 18 choice sets were presented to each respondent—previous studies (e.g., Mau et al., 2008; Hensher, 2007), as well as careful pre-testing and retrospective analysis, indicated that this number of choices did not present an unreasonable respondent burden.[9]

Four key attributes were included in respondent choice sets: capital cost, fuel cost, performance and subsidy.[10] The first two attributes are essential to calculate discount rate in Eq. (9), while vehicle power (as a proxy for general performance) has been highlighted as an influential intangible attribute (Canadian Auto Agency, 2003; Horne et al., 2005). Although the HEV subsidy was explained as a reduction in upfront capital cost, we model it as a separate coefficient due to previous findings suggesting that a dollar of subsidy may be valued differently than a dollar of capital cost (Mau et al., 2008).[11] Together, all four attributes represented the key differences between HEVs and conventional gasoline vehicles in 2005, when available HEVs were generally more expensive (by 15–30%), more fuel efficient (by 20–40%), less powerful (by 15–25%) and more eligible for subsidies ($500–3000) than comparable conventional vehicles. In the choice sets, attribute levels were customized for each respondent, specified as percentage values that "pivoted" around the attribute of the respondent's current vehicle. This technique is common in SP literature (e.g., Mau et al., 2008; Hensher, 2007; Hensher et al., 2005) and helps to frame the choice scenario according to the respondent's actual experience.

3.2. RP data collection

RP data was elicited by asking respondents to provide details about their primary vehicle.[12] Rather than rely purely on respondent recall, we constructed a comprehensive database of vehicle make and models for each target year, allowing respondents to choose the year, make and model of

---

[8] Although this aggregation of attributes did slightly compromise the orthogonality of the original design, any observed correlation is small (<0.05) and we maintain that the 18 choice sets perform well in terms of attribute balancing, as discussed by Johnson et al. (2007).

[9] However, considerations of respondent burden and overall efficiency should be tested on a case by case basis (DeShazo and Fermo, 2002; Arentze et al., 2003; Scarpa and Rose, 2008). Prior to the implementation of this study, we conducted careful pre-tests to assure our design did not present an inappropriate burden to respondents. Also, the online survey instrument recorded respondent completion times—retrospectively confirming that the choice exercise was not overly burdensome.

[10] Fuel efficiency was also represented as a proportionally equal attribute called "pollution"—the intention was not to estimate a pollution coefficient, but instead to merely present to the respondent the portion of weekly fuel cost that corresponded with negative emissions (rather than fuel prices).

[11] While some may argue that consumers should be indifferent to the format of a price reduction, there is a growing literature (e.g., Eckel and Grossman, 2003; Davis et al., 2005; Davis and Millner, 2005) suggesting that various but differing forms of price reduction, including subsidies are valued differently. By incorporating the subsidy separately, rather than subtracting it from income, we are letting the data speak for themselves regarding whether the subsidy is perceived differently than just being subtracted form capital costs and income.

[12] The "primary" vehicle is the car or truck that the respondent used most routinely, which was purchased new in 2002 or later.

Table 4
Example RP choice set (reportedly owned vehicle and 11 randomly selected non-chosen alternatives, 2005 model year).

Class	Constant	Make and model	Price	Horsepower	Liters/100 km	Weekly fuel cost[b]
Subcompact car[a]	Small car	Volkswagon Beetle	$23,210	115	8.9	$30.00
Compact car	Small car	Ford Focus	$17,376	110	8.3	$27.98
Midsize car	Large car	Nissan Altima	$29,098	240	10.4	$35.06
Large car	Large car	Lexus LS 430	$82,800	290	11.7	$39.44
Small SUV	SUV	Toyota RAV 4	$25,933	148	10.0	$33.71
Midsize SUV	SUV	GMC Envoy	$40,619	270	14.1	$47.53
Large SUV	SUV	Cadillac Escalade	$66,047	345	16.3	$54.94
Minivan	Van	Dodge Caravan	$31,072	215	10.9	$36.74
Large van	Van	Ford E150	$34,043	225	15.4	$51.91
Small pickup	Truck	Mazda B2300	$17,478	140	9.3	$31.35
Large pickup	Truck	Chevrolet Avalanche	$40,991	285	16.6	$55.96
HEV	HEV	Toyota Prius	$27,140	110	4.8	$16.18

[a] For this example, the hypothetical respondent is assumed to own this vehicle.
[b] Weekly fuel cost for each non-chosen alternative = (user fuel cost) × (alternative fuel economy)/(user fuel economy).

their recently purchased vehicle with drop-down menus (manual entry was permitted for unlisted models).[13] We were thus able to collect RP data for three of the same attributes as the SP design: capital cost, weekly fuel cost and vehicle power. The database was also used to select the non-chosen alternatives that are required to model RP choice sets for each respondent. We followed a technique similar to that used by Brownstone et al. (2000) to randomly select a subset of available vehicles (because it is computationally prohibitive to include all possible vehicles), one from each of the 11 vehicle class categories not chosen by the respondent. Table 4 illustrates a hypothetical example for a respondent owning a Volkswagon Beetle (classified as a gasoline subcompact car). To complete the choice set, 11 other vehicles were selected from our database: 10 other gasoline vehicles, and 1 HEV. This method ensured that each choice set contained a significant degree of attribute variation, approximating the actual breadth of vehicle choices available to each respondent at the time of purchase.

4. Choice model results

4.1. Stated preference models

A multinomial logit (MNL) model estimated from all three Canada market share scenarios is presented under the heading "SP Canada" in Table 5. SP data from the 51 HEV owners were downweighted to reduce their influence according to actual market share (0.17%). All coefficient estimates were significant at a 99% confidence level, and of the expected sign.[14] The capital cost variable contributed the best fit of the SP data when divided by the log of household income (as in Brownstone et al., 2000; Brownstone and Train, 1999) to account for income effects.[15] This specification represents income as a deflator, where capital cost is less important for higher income households.[16] The bottom five rows of Table 5 portray the monetized value of several coefficient estimates, representing the value of an attribute unit in terms of purchase price, holding household income at the sample average of $78,000. For example, the average respondent

---

[13] We updated a 2003 Natural Resources Canada fuel efficiency database provided to include newer models. Retail price, fuel efficiency and horsepower were detailed for nearly 1500 vehicle (300 per model year). Vehicle class followed the 11-class system used by Greene et al. (2005).
[14] This model assumes a linear relationship between the attribute levels and perceived utility (except the alternative specific constants). We performed a test of linearity, and although the model appears to work well with these assumptions, we found that fuel cost was better specified as a quadratic variable. However, such a specification could not be translated to CIMS, and was thus not incorporated in this study.
[15] Commonly, MNL choice models will specify several demographic variables, in addition to choice attributes. However, it is difficult to translate demographic coefficients into CIMS behavioral parameters, which represent an entire economy. The only demographic variable we include is household income, which is set to the sample average of $78,000 for the purpose of translation to CIMS.
[16] Moreover, including income in this non-linear fashion is consistent with the non-homotheticiy of demand for consumer purchases (e.g., Banks et al., 1997; Denton et al., 2006).

Table 5
SP and RP choice model estimates.

Variable	SP Canada		RP Canada		RP California	
	Coeff.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat
SP CC/log(income)	−0.00717	−30.7	−0.000205	−7.2	−0.000735	−15.1
Fuel cost	−0.0342	−20.6	−0.0364	−4.1	−0.0337	−3.9
Horsepower	0.0117	15.2	0.000165	0.1	0.0218	12.2
SP subsidy	0.000103	6.7				
SP HEV constant	0.296	5.6				
SP HEV × MS2	0.174	3.0				
SP HEV × MS3	0.222	3.8				
RP large car constant			−1.240	−7.3	−1.118	−6.2
RP SUV constant			−1.133	−8.5	−0.919	−5.9
RP van constant			−0.683	−3.5	−2.681	−8.9
RP truck constant			−2.178	−9.2	−2.566	−11.2
RP HEV constant			−6.429	−5.6	−1.805	−4.7
Observations[a]	9630		542		408	
Max log-likelihood	−5191.6		−887.1		−600.5	
LL ratio[b]	0.147		0.329		0.408	
Fuel cost value[c]	−$233.01		−$867.54		−$224.08	
Horsepower value[c]	$79.68		$3.94		$145.12	
HEV value MS1[c]	$2,041.69		−$153,204.68		−$12,008.60	
HEV value MS2[c,d]	$3,201.84					
HEV value MS3[c,e]	$3,534.60					

[a] Some completed surveys removed due to missing income data and/or poor data.
[b] Ratio of log-likelihood of full model to log-likelihood of model with no coefficients.
[c] Ratio of variable to capital cost coefficient, assuming income = $78,000.
[d] Value of SP HEV constant + MS2 interaction.
[e] Value of SP HEV constant + MS3 interaction.

in this model would pay an extra $79.68 in vehicle purchase price for one extra unit of horsepower.

Unique HEV constants were estimated for each market share experimental group by adding two dummy α interaction terms for MS2 and MS3 scenarios. In all three market share scenarios, respondents assigned a positive value to HEVs independent of the other specified variables. Both market share interaction terms are positive, indicating that the value of an HEV increased from $2042 to $3202 ($\alpha_{HEV} + \alpha_{HEV} \times$ MS2) and to $3533 ($\alpha_{HEV} + \alpha_{HEV} \times$ MS3) in the 10% and 50% market share scenarios, respectively. Thus, the experimental treatment induced a statistically significant change in HEV preferences among scenario groups in the expected direction; that is, HEVs become more favorable with increased market penetration. However, we cannot conclusively infer the underlying drivers of this effect—that is, whether technological uncertainty, consumer influence, model availability or other factors were individually or collectively important.

The monetized coefficients in Table 5 are comparable to previous SP studies of alternative fuel vehicles. Valuation of fuel cost is close to Mau et al.'s (2008) estimates, and valuation of vehicle power is similar to that found by Ewing and Sarigollu (2000). The positive valuation of the HEV ranges between $2000 and $3500 among market share scenarios, which is within the $2000–5000 range of willingness to pay found by Potoglou and Kanaroglou (2007). However, like many SP choice models, when the attributes of real HEVs and comparable conventional vehicles are entered into Eq. (7), the resulting market share predictions of low-emissions technology (>30%) are unrealistically optimistic relative to the actual Canadian HEV market in 2005 (0.17%)—likely a result of hypothetical or social desirability bias. Our SP model may also have excluded other important factors, such as supply constraints indicated by the long waiting lists experienced by some HEV consumers (often 6 months or longer).[17]

---

[17] We also specified an MNL model using California SP data following the same methodology. Results were nearly identical to the Canada SP model, particularly for the monetized coefficients.

4.2. Revealed preference models

Table 5 also portrays MNL models estimated from Canada and California RP data.[18] Again, data in both samples were weighted to offset the intentional overrepresentation of HEV owners (choice-based sampling).[19] As anticipated, the RP vehicle attributes in this study were highly correlated, with correlation coefficients ranging from 0.32 (between capital cost and fuel efficiency) to 0.72 (between fuel efficiency and horsepower). Such multicollinearity can result in insignificant and/or counter-intuitive coefficient estimates. Despite this vulnerability, both RP models have significant Chi-square values, and most coefficient estimates are significant at a 99% confidence level. However, the Canada horsepower coefficient has a relatively low *t*-value and low value, likely resulting from observed correlations. Multicollinearity also appears to have affected the scale of coefficients; the monetized RP attribute values in Table 5 are alarmingly different from SP estimates. For example, the Canada monetized fuel cost coefficient of $867.54 is almost 4 times the magnitude of the corresponding SP estimates.

Five vehicle class constants were specified relative to the base class, "small car" (set to zero): large car, SUV, van, truck and HEV. The HEV constant is interpreted relative to small cars, where the HEV is the least desirable class in the Canada sample, indicating that on average, consumers would require a payment of $153,204.68 to equate an HEV with a conventional vehicle. Although this value may seem unrealistically large, the value is an economy-wide average representing a technology with very little market exposure. A large segment of the population may not have had any awareness of HEV technology, or lacked access to an HEV dealer; such a segment would have nearly infinite "intangible" costs. In contrast, the monetized California HEV constant was more comparable to other vehicle class constants, consistent with the higher rate of HEV penetration (3.0%). Overall, we have low confidence in these RP models. The observed drawbacks of the RP data illustrate the potential value of a jointly estimated choice model that could eliminate multicollinearity issues while including a more realistic vehicle class specification.

4.3. Joint models

We compare three alternative techniques to jointly estimate choice models from SP and RP data. Each model is fit to the Canada sample data and depicted in Table 6. Models were specified as follows:

Joint model 1: A "pooling" technique assigning equal weighting to the SP and RP data, estimating joint β coefficients for capital cost, fuel cost and horsepower, and unique α constants for the RP data (large car, SUV, van, truck and HEV). Each respondent yielded 1 RP observation and 18 SP observations; the SP data was weighted by 0.056 to limit influence to be equivalent with the RP data.

Joint model 2: A "pooling" technique similar to model 1, but without "corrective" weights—allowing the SP data to have 18 times more influence in jointly estimated β coefficients than the RP data.

Joint model 3: A "sequential" technique, directly extracting β coefficients from the SP model (Table 5) and calibrating the α constants to fit the RP data.

For each estimated joint model, we followed the "artificial tree structure" technique noted at the end of Section 2.2 and found no significant difference between RP and SP scale parameters (i.e., $\lambda^{SP} = \lambda^{RP} = 1$); Hensher et al. (2005) suggest this finding is common when RP and SP data is collected

---

[18] Another possibility would be to use a nested MNL procedure for the RP data, representing hierarchical decision making (e.g. choice among classes, then choice within class). We acknowledge the value of this possibility. However, we opted for the MNL procedure in this study for simplicity, and also to avoid assuming a particular nesting structure. For instance, Heffner et al. (2007) find the HEV buyers often compare vehicles from vastly different classes.

[19] All survey respondents purchased new vehicles between 2002 and 2006, so it was not entirely accurate to apply the 2005 HEV weighting to all respondents. However, the weighting procedure used in this study only allows weighting by choice, not vehicle year. Thus, we chose 2005 as a base year because the best market data was available for this year, and this year also coincides with the first simulation period in CIMS.

Table 6
Joint choice model estimates (Canada).

Variable	Joint model 1: "Pooled" weighted		Joint model 2: "Pooled"		Joint model 3: "Sequential"	
	Coeff.	t-Stat	Coeff.	t-Stat	Coeff.	t-Stat
SP *CC*/log(income)					−0.00072	n/a[a]
SP fuel cost					−0.0341	n/a[a]
SP horsepower					0.0117	n/a[a]
Joint *CC*/log(inc.)	−0.00028	−10.1	−0.00057	−28.1		
Joint fuel cost	−0.0341	−5.6	−0.0322	−16.2		
Joint horsepower	0.0033	2.5	0.0105	16.9		
RP large car constant	−1.330	−7.8	−1.449	−8.0	−1.379	−7.2
RP sport utility constant	−1.205	−9.2	−1.282	−8.9	−1.224	−7.7
RP van constant	−0.873	−4.9	−1.166	−7.6	−1.043	−7.2
RP truck constant	−2.371	−10.2	−2.791	−12.3	−2.833	−12.1
RP HEV constant	−6.124	−5.6	−5.531	−5.2	−5.437	−5.1
Observations[b]	10,162		10,162		532	
Max log-likelihood	−1223.5		−6689.9		−1003.2	
LL ratio[c]	0.497		0.555		0.241	
Fuel cost value[d]	−$594.98		−$278.65		−$233.01	
Horsepower value[d]	$57.60		$90.77		$79.68	
HEV value[d]	−$106,974.51		$47,885.94		$37,115.77	
California values						
Fuel cost value[d]	−$225.21		−$232.40		−$228.11	
Horsepower value[d]	−$139.15		−$108.36		−$86.07	
HEV value[d]	−$12,399.06		−$13,122.33		−$11,663.87	

[a] t-Stat not applicable because variables fixed (from SP model).
[b] Some completed surveys removed due to missing income data and/or poor data.
[c] Ratio of log-likelihood of full model to log-likelihood of model with no coefficients.
[d] Ratio of variable to capital cost, assuming income = $78,000.

from the same sample. Thus, we exclude further discussion of scale parameters in this study, but note the importance of exploring this issue in each application. Model results in Table 6 (and subsequent tables) are arranged according to degree of RP influence over β coefficients (left to right corresponds with highest RP influence to lowest). Each model's coefficient estimates are highly significant. Vehicle class constants vary among models, but rank order remains the same. Monetized attribute values generally correspond with the degree of RP influence; joint models allowing more RP data influence yield estimates closer to the RP-only model, while joint models allowing more SP data influence yield estimates closer to the SP-only model. The same joint estimation techniques were

Table 7
Comparing models fitted to RP data.

Indicator	Region	RP-only	Joint model 1: "Pooled" weighted	Joint model 2: "Pooled"	Joint model 3: "Sequential"
Log-likelihood	Canada	−887.1	−890.9	−950.7	−1003.2
	California	−600.5	−600.9	−611.6	−631.6
Chi-square[a]	Canada	420.0	412.4	292.6	187.7
	California	641.4	640.4	619.0	579.1
Correct predictions[b]	Canada	44.7%	44.3%	41.8%	41.5%
	California	48.0%	47.3%	46.3%	44.6%
Submodel predictions[c]	Canada	47.6%	46.6%	42.9%	42.5%
	California	50.7%	49.5%	49.5%	47.1%

[a] Compared to base model with constants only. All Chi-square values significant at 99% confidence level.
[b] Within RP data only.
[c] Average accuracy over 10 iterations of submodel re-estimations; in each iteration the model was estimated with 90% of sample (randomly selected) and used to predict choices of the remaining 10%.

Table 8
Comparing joint model market share predictions (Canada, income = $78,000).

Type	Actual[a]	RP-only	Joint model 1: "Pooled" weighted	Joint model 2: "Pooled"	Joint model 3: "Sequential"
Small car	55.48%	61.12%	60.94%	62.56%	64.15%
Large car	8.50%	10.82%	11.49%	12.64%	12.42%
SUV	16.00%	9.16%	9.54%	9.26%	8.13%
Van	15.39%	17.10%	16.31%	14.48%	14.60%
Truck	4.45%	1.71%	1.62%	0.96%	0.62%
HEV	0.17%[b]	0.10%	0.10%	0.09%	0.08%

[a] Calculated from Canada RP sample ($n = 542$).
[b] HEV market share was weighted to correct for oversample of HEV owners, and match 2005 HEV sales in Canada.

applied to the California data, and the monetized attribute values are depicted in the bottom three rows of Table 6.

Table 7 compares the statistical performance of the three joint models and the RP-only model, all fitted to the Canada and California RP data. Models allowing more influence from RP data are statistically superior on all four indicators. The maximum log-likelihood and Chi-square values improve with greater RP influence (from right to left). We also present the percentage of correct choice predictions within the model, as well as the accuracy of models re-estimated with 90% of the data in predicting the remaining 10% of choices that were randomly removed (averaged over 10 iterations). In both cases, results are consistent; models with more RP influence perform more accurately. However, as found by Swait et al. (1994), differences are not substantial. Table 8 compares the market share predictions of these four models using attribute levels from the six vehicles specified in Table 9 (median values for each vehicle class in the Canada RP data). Resulting predictions are compared with actual percentages observed in the Canada RP data. Predictions are fairly consistent across the models, which generally correspond with actual market shares aside from some noticeable differences (e.g., underestimating SUV and truck market share). HEV market share predictions in all models are fairly consistent and appropriately low.

In summary, we find that models with greater RP influence perform slightly better according to statistical measures, but monetized coefficients estimated with greater SP influence are more realistic and consistent with empirical research. Interestingly, the market share predictions of the four models do not vary substantially. Overall, we favor the more realistic coefficient estimates of models 2 and 3; these estimates are of primary concern for the main objective of this study: translating coefficients into dynamic behavioral parameters for CIMS.

5. Adding behavioral realism to an energy–economy model

Using the joint models and RP-only model specified above, we estimated the three main behavioral parameters of CIMS depicted in Table 10. First, Eq. (9) was used to estimate the private discount rate, r, using capital cost and fuel cost coefficients estimated from each joint choice models.[20] The "sequential" model estimate of 21.3% is most consistent with previous SP choice studies (Horne et al., 2005; Mau et al., 2008), and approximates the middle of the range of automotive discount rates reviewed by Train (1985). With added RP influence, r estimates decline to unrealistically low values due to increasingly large fuel cost coefficient estimates.

The intangible costs of HEVs were estimated from the non-monetary coefficients of the joint choice models, using Eq. (10).[21] For example, the initial i_{HEV} in joint model 3 ($MS_{t-1} = 0$) is $i_{fixed} + i_{var} = \$38,469$, which is the sum of reduced HEV horsepower relative to the average small car in Table 9 $(127 - 110) \times \$79.68 = \$1355$, plus the HEV constant value ($37,115). Fig. 1 plots this estimate according to HEV market share (Canada = 0.17%) along with the i_{HEV} estimate from the California "sequential" model ($13,127) to represent the 3.0% HEV market share scenario. The 10% and

---

[20] The technological lifespan, n, was assumed to be 16 years, as is specified for all new vehicles in CIMS.
[21] Because intangible cost is a *relative* value, it must be calculated relative to a competing technology. We used the "small car" class as a reference to compare constant and horsepower estimates.

Table 9

Attributes levels for choice models (median values from Canada RP survey data).

Vehicle type	Purchase price	Horsepower	L/100 km	Weekly fuel cost[a]
Small car	$22,000	127	8.1	$25.45
Large car	$28,000	200	10.3	$32.37
SUV	$31,000	200	11.6	$36.45
Van	$28,750	200	10.9	$34.25
Truck	$38,500	238	14.5	$45.41
HEV	$32,000	110	4.7	$14.77

[a] Weekly average assumes $0.86 per liter of gasoline, driving 19,000 km per year.

Table 10

Behavioral parameter estimates for CIMS model.

Type	RP-only	Joint model 1: "Pooled" weighted	Joint model 2: "Pooled"	Joint model 3: "Sequential"
r parameter	−0.5%	4.2%	17.2%	21.3%
Small car i_{fixed} parameter[a]	$0	$0	$0	$0
Large car i_{fixed} parameter[a]	$29,284	$19,031	$5,918	$3,595
HEV i_{fixed} parameter	−$2,180	−$2,180	−$2,180	−$2,180
HEV i_{var} parameter	$155,552	$110,134	$51,609	$40,649
HEV A parameter	1.487	0.951	0.313	0.237
HEV k parameter	57.484	58.486	61.990	61.819
v parameter	3.9	4.1	5.3	6.1

[a] Only static intangible cost parameters (i_{fixed}) used for "small car" and "large car" technologies.

50% points in Fig. 1 are estimated from the Canada SP-only models (Table 5) from the MS2 and MS3 scenarios, respectively. Notice that in these SP scenarios, the HEV is assumed to have an overall intangible "benefit" relative to a conventional small car. In each model, the A and k parameters were estimated to fit the four i points, yielding the parameter estimates in Table 10. This same procedure was completed for all four models, where only the first two points (Canada and California) varied. Each intangible cost curve followed a similar pattern as seen in Fig. 1; consumers perceive HEV intangible costs to be very high when HEVs first enter the market, but costs decline substantially as they diffuse and stabilize once HEVs reach 10% market share.

The last behavioral parameter, v, was estimated using Eq. (11) to represent market heterogeneity in the CIMS model. Using the vehicle attribute levels in Table 9, we used the Solver function in Excel to find a v value that minimized the differences between vehicle market share predictions in the choice model, Eq. (7), and the CIMS function, Eq. (3). The optimal v values ranged from 3.9 to 6.1 among models. Previous vehicle choice studies estimated similarly low v parameters for the vehicle choice node, ranging from 2.4 to 5.2 (Horne et al., 2005; Mau et al., 2008).

Fig. 2 depicts HEV market share forecasts in CIMS using parameter estimates form each modeling approach in Table 10.[22] We use the technology parameter values in Table 11 to compete three technologies in the passenger car node. All curves start realistically low in 2005, and follow the s-curve penetration pattern typical of new technologies. Parameters estimated from choice models with more RP influence penetrate faster; joint model 3, the "sequential" model, follows the slowest penetration forecast. Although it is difficult to comment on the "accuracy" of market share forecasts, we can draw an important observation: the RP-only model and joint model 1 predict highly optimistic penetration scenarios, whereas parameters translated from joint models 2 and 3 yield lower, and likely more realistic, penetration scenarios. We next explore the sensitivity of these models to policy intervention.

---

[22] All simulations assumed base gasoline price predictions used by CIMS as of 2006 (decreasing from $0.87 L^{-1} in 2005 to $0.67 L^{-1} in 2020).
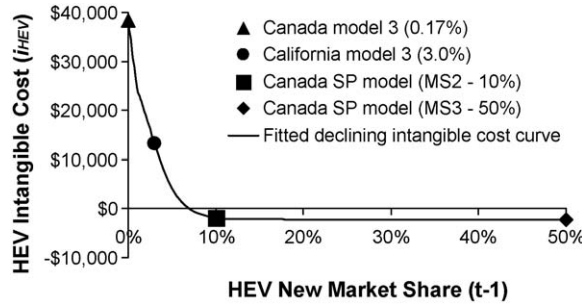
Fig. 1. Intangible cost and function estimates ("sequential" estimates shown for illustration).
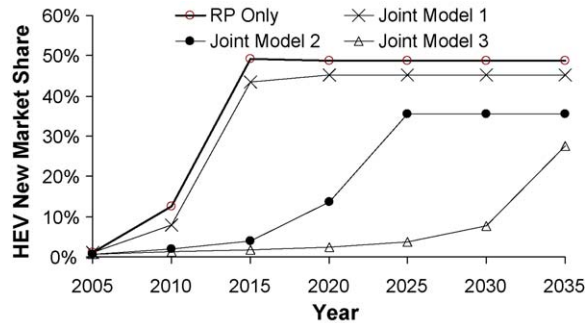

Fig. 2. New HEV market share forecasts for Canada (business as usual).

Table 11
Technology parameter estimates for CIMS model.

Vehicle type	CC_{2005}	CC_{2020}	n	MC	EC_{2005}[b]	EC_{2020}[c]
Small car	$22,000	$22,000	16	$1751	$1336	$1026
Large car	$28,000	$28,000	16	$1751	$1699	$1305
HEV	$32,000	$29,500[a]	16	$1751	$775	$596

[a] HEV production costs set to exogenously reduce by 9% per year from 2005 to 2015.
[b] Canada energy outlook estimate of $25.05/GJ of gasoline for 2005 (NRCAN, 2006).
[c] Canada energy outlook estimate of $19.24/GJ of gasoline for 2020 (NRCAN, 2006).

6. Selected policy simulations

As the final stage of this study, we used the behavioral parameters in Table 10 and technological parameters in Table 11 to simulate three climate change policies in the transportation sector of Canada: (1) a $3000 subsidy to HEV capital cost, (2) a feebate scheme simultaneously subsidizing lower emissions vehicles and taxing higher emissions vehicles at a rate of $1000 per liter per 100 km relative to average efficiency, and (3) a $100 per tonne carbon tax corresponding to an increase in gasoline price of about $0.24 per liter. These simulations are intended to illustrate the capabilities of the CIMS model with empirically based, dynamic behavioral parameters—not as a formal policy analysis.

We limit our analysis to the car competition node, excluding potential higher level effects (e.g., switching from truck to cars and from vehicles to transit). Table 12 summarizes HEV new market share forecasts, fleet average GHG intensity (grams of CO_2 per km), and reductions in GHG intensity relative to 2005 and relative to business as usual (BAU) in the same forecast year. Again, the RP-only parameter estimates produce more optimistic HEV new market share predictions than the SP-informed models,

Table 12
Policy simulations for CIMS model: Canada.

Simulation run 2005 GHG intensity (gCO$_2$/km)	RP-only (195.1)		Joint model 1 (195.5)		Joint model 2 (196.1)		Joint model 3 (196.0)	
Simulation year	2015	2030	2015	2030	2015	2030	2015	2030
(1) BAU								
HEV new market share %	49.1	48.8	43.5	45.2	4.1	35.5	1.8	7.8
New fleet average GHG intensity (gCO$_2$/km)	153.2	153.5	158.3	156.9	193.1	165.6	194.9	189.7
% intensity reduction from 2005	21.4	21.3	19.0	19.7	1.5	15.6	0.6	3.2
(2) $3000 subsidy[a]								
HEV new market share %	53.6	53.3	50.1	50.9	6.9	45.4	3.0	43.8
New fleet average GHG intensity (gCO$_2$/km)	149.3	149.6	152.6	151.8	190.6	156.9	193.8	158.2
% intensity reduction from 2005	23.4	23.3	21.9	22.3	2.8	20.0	1.1	19.3
% intensity reduction from BAU	2.5	2.5	3.6	3.2	1.3	5.2	0.5	16.6
(3) $1000 feebate[b]								
HEV new market share %	57.9	57.6	56.6	56.6	17.8	55.8	7.2	56.6
New fleet average GHG intensity (gCO$_2$/km)	145.4	145.7	146.7	146.6	179.9	147.2	188.5	146.4
% intensity reduction from 2005	25.4	25.3	25.0	25.0	8.3	24.9	3.8	25.3
% intensity reduction from BAU	5.1	5.1	7.4	6.5	6.8	11.1	3.3	22.8
(4) $100/tonne carbon tax[c]								
HEV new market share %	52.1	51.8	47.6	48.1	6.1	38.3	2.5	34.9
New fleet average GHG intensity (gCO$_2$/km)	150.8	151.1	154.9	154.5	191.4	163.2	194.3	166.1
% intensity reduction from 2005	22.7	22.5	20.7	21.0	2.4	16.8	0.9	15.3
% intensity reduction from BAU	1.6	1.6	2.1	1.5	0.9	1.4	0.3	12.4

[a] Subsidy implemented as $3000 reduction in HEV capital cost, beginning in 2010.
[b] Feebate implemented as $1000 per L/100 km, using the "small car" fuel use as a pivot point (8.1 L/km). The "large car" was assigned a capital cost increase of $2200, while the HEV capital cost was reduced by $5600, beginning in 2010.
[c] Carbon tax implemented as increase of per liter gasoline price by $0.23, beginning in 2010.

around 50% by 2030 in all policy simulations, as well as BAU—indicating low sensitivity to pricing mechanisms and policy simulation. In contrast, joint model 3 shows much higher sensitivity, where policy is predicted to reduce 2030 GHG intensity by 16.6–22.8% relative to BAU. Based on these simulations, joint models 2 and 3 again appear to yield more realistic behavioral patterns. We suspect RP parameter problems result from the observed distortions induced by multicollinearity.

7. Conclusion

This study investigated the use of SP and RP choice research to provide a sounder empirical basis for the behavioral realism of CIMS, an energy–economy model. We highlight four main conclusions from this exercise. First, the estimation of joint choice models has proven to be a useful technique for modeling vehicle choice. SP models provided realistic and empirically consistent tradeoff coefficients (β), such as capital cost, fuel cost and performance, but yielded overly optimistic HEV market share predictions. RP models were better calibrated to actual vehicle market shares, but β coefficients appeared to be unreliable due to multicollinearity. Joint modeling techniques were found to generally improve upon models using only SP or RP data.

Second, we found that the joint models allowing more influence from SP data (e.g., the "sequential" technique) performed best of the joint methods explored. Joint models allowing more RP influence over the β coefficients appeared to distort estimates towards the RP-only model estimates. In contrast, the sequential technique allowed for a more intelligent specialization of tasks; SP data was used solely for its strength, β coefficient estimates, while RP data was used for its strength, calibration to real-world market share with constants. Like Swait et al. (1994), we find that although RP-influenced coefficients perform better from a purely statistical perspective, we prefer the sequential technique due to other performance metrics: more realistic and empirically consistent β coefficient estimates and more realistic CIMS behavioral parameter estimates (e.g., discount rate), as well as the resulting translation into CIMS parameters that maintain sensitivity to policy. We also note that the unweighted

pooled technique, which allowed SP data to have 18 times more influence than RP data in β coefficient estimates, performed as well as the sequential method. Future research may explore the optimal level of blending the influence of SP and RP data.

Third, this study demonstrated the existence and potential importance of neighbor effects in simulating technical change. HEV willingness to pay estimates derived from the SP, RP and joint choice models were consistent with expectations; scenarios with higher HEV penetration yielded higher valuation of HEVs. Although we are fairly confident in intangible cost estimates from the joint models, the 10% and 50% market share were drawn only from SP data—which may be vulnerable to hypothetical bias. Despite this drawback, we feel that the overall shape of the intangible cost curve is useful, demonstrated by the empirically consistent s-curve market forecasts when propagated into CIMS.

Lastly, policy simulations illustrate that parameters translated from choice models allowing more SP influence resulted in HEV penetration scenarios that were more realistic and sensitive to policy simulation than RP dominated models. For our particular objectives, the sequential joint model or the pooled joint model allowing for a high influence of SP data yield parameter estimates that allow for more useful technology competitions in a hybrid energy–economy model. However, this recommendation may vary across data sets and applications.

Overall, this study reveals the vast complexity and uncertainty inherent in the investigation of preference dynamics. Current preferences are difficult to disaggregate, and future preference scenarios are highly speculative. However, preference dynamics are an important component of technological change, and further research efforts are required to ensure the design of effective and feasible climate policies that intend to induce technological change.

References

Alberini, A., Longo, A., Veronesi, M., 2007. Basic statistical models for stated choice studies. In: Kanninen, B. (Ed.), Valuing Environmental Amenities Using Stated Choice Studies. Springer, pp. 203–228.

Arentze, T., Borgers, A., Timmermans, H., DelMistro, R., 2003. Transport stated choice responses: effects of task complexity, presentation format and literacy. Transport Research Part E 39, 229–244.

Autonews, 2006. Autonews Datacenter. http://www.autonews.com.

Axsen, J., Kurani, K., 2008. The Early U.S. Market for PHEVs: Anticipating Consumer Awareness, Recharge Potential, Design Priorities, and Energy Impacts. Institute of Transportation Studies, University of California, Davis. Research Report UCD-ITS-RR-08-22.

Azar, C., Dowlatabadi, H., 1999. A review of technical change in assessment of climate policy. Annual Review of Energy and the Environment 24, 513–544.

Banks, J., Blundell, R., Lewbel, A., 1997. Quadratic angel curves and consumer demand. Review of Economics and Statistics 79, 527–539.

Bataille, C., Jaccard, M., Nyboer, J., Rivers, N., 2006. Towards general equilibrium in a technology-rich model with empirically estimated behavioural parameters. Energy Journal 2, 93–112.

Bohringer, C., 1998. The synthesis of bottom-up and top-down in energy policy modelling. Energy Economics 20, 234–248.

Brownstone, D., Bunch, D.S., Train, K., 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. Transportation Research Part B 34, 315–338.

Brownstone, D., Train, K., 1999. Forecasting new product penetration with flexible substitution patterns. Journal of Econometrics 89, 109–129.

Bunch, D., Bradley, M., Golob, T., Kitamura, R., Occhiuzzo, G., 1993. Demand for clean-fuel vehicles in California: a discrete-choice stated preference pilot project. Transportation Research Part A 27 (3), 237–253.

Canadian Auto Agency, February 2003. Autopinion: vehicle ownership survey results. Carguide.

Champ, P., Welsh, M., 2007. Survey methodologies for stated choice studies. In: Kanninen, B. (Ed.), Valuing Environmental Amenities Using Stated Choice Studies. Springer, pp. 21–42.

Couper, M., 2002. New technologies and survey data collection: challenges and opportunities. In: Invited paper presented at the International Conference on Improving Surveys, Copenhagen, August, p. 2002.

Couper, M., Kapteyn, A., Schonlau, M., Winter, J., 2007. Noncoverage and nonresponse in an Internet survey. Social Science Research 36, 131–148.

Davis, D., Millner, E., 2005. Rebates, matches, and consumer behavior. Southern Economic Journal 72, 410–421.
Davis, D., Millner, E., Reilly, R., 2005. Subsidy schemes and charitable contributions: a closer look. Experimental Economics 8, 85–106.
Denton, F., Mountain, D., Spencer, B., 2006. Age, retirement, and expenditure patterns: an econometric study of older households. Atlantic Economic Journal 34, 421–434.
DeShazo, J., Fermo, G., 2002. Designing choice sets for stated preference methods: the effects of complexity on choice consistency. Journal of Environmental Economics and Management 44, 123–143.
Eckel, C., Grossman, P., 2003. Rebate vs. matching: does how we subsidize charitable contributions matter? Journal of Public Economics 87, 681–701.
Environment Canada, 2004. Canada's greenhouse gas inventory 1990–2002.
Ewing, G., Sarigollu, E., 2000. Assessing consumer preferences for clean-fuel vehicles: a discrete choice experiment. Journal of Public Policy & Marketing 19 (1), 106.
Greene, D., Patterson, P., Singh, M., Li, J., 2005. Feebates, rebates and gas-guzzler taxes: a study of incentives for increased fuel economy. Energy Policy 33 (6), 757.
Groves, R., 2007. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly 70 (5), 646–675.
Grubb, M., Kohler, J., Anderson, D., 2002. Induced technical change in energy and environmental modelling: analytic approaches and policy implications. Annual Review of Energy and the Environment 27, 271–308.
Hausman, J., 1979. Individual discount rates and the purchase and utilization of energy-using durables. Bell Journal of Economics 10 (1), 33–54.
Heffner, R., Kurani, K., Turrentine, T., 2007. Symbolism in California's early market for hybrid electric vehicles. Transportation Research Part D 12, 396–413.
Hensher, D., 2007. Attribute processing in choice experiments and implications on willingness to pay. In: Kanninen, B. (Ed.), Valuing Environmental Amenities Using Stated Choice Studies. Springer, pp. 135–158.
Hensher, D., Rose, J., Greene, W., 2005. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge, UK.
Hoeffler, S., 2003. Measuring preferences for really new products. Journal of Marketing Research (JMR) XL, 406–420.
Horne, M., Jaccard, M., Tiedemann, K., 2005. Improving behavioural realism in hybrid energy-economy models using discrete choice studies of personal transportation decisions. Energy Economics 27 (1), 59.
Internet World Stats, 2007. Canada: Internet Usage and Broadband Usage Report, July 2007, http://www.internetworldstats.com/am/us.htm (accessed January 27, 2008).
Jaccard, M., Nyboer, J., Bataille, C., Sadownik, B., 2003. Modelling the cost of climate policy: distinguishing between alternative cost definitions and long-run cost dynamics. The Energy Journal 24 (1), 49.
Jaffe, A., Stavins, R., 1994. Energy-efficiency investments and public policy. The Energy Journal 15 (2), 43–65.
Johnson, R., Kanninen, B., Bingham, M., Ozdemir, S., 2007. Experimental design for stated choice studies. In: Kanninen, B. (Ed.), Valuing Environmental Amenities Using Stated Choice Studies. Springer, pp. 159–202.
Louviere, J., Hensher, D., Swait, J., 2000. Stated Choice Methods: Analysis and Application. Cambridge University Press, Cambridge, UK.
Mau, P., Eyzaguirre, J., Jaccard, M., Collins-Dodd, C., Tiedemann, K., 2008. 'The neighbour effect': simulating dynamics in consumer preferences for new vehicle technologies. Ecological Economics 68 (1), 504–516.
Natural Resources Canada, 2006. Fuel Consumption Guide 2006. http://vehicles.gc.ca.
Norton, B., Costanza, R., Bishop, R.C., 1998. The evolution of preferences: why 'sovereign' preferences may not lead to sustainable policies and what to do about it. Ecological Economics 24 (2–3), 193–211.
Potoglou, D., Kanaroglou, P., 2007. Household demand and willingness to pay for clean vehicles. Transportation Research Part D 12, 264–274.
R.L. Polk & Co., 2006. Hybrid Vehicle Registrations More than Double in 2005. Latest News. http://usa.polk.com.
Rivers, N., Jaccard, M., 2005. Combining top-down and bottom-up approaches to energy-economy modelling using discrete choice methods. The Energy Journal 26 (1), 83.
Rogers, E., 2003. Diffusion of Innovations, fifth ed. Free Press, New York.
Scarpa, R., Rose, J., 2008. Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. Agricultural and Resource Economics 52, 253–282.
Swait, J., Louviere, J., Williams, M., 1994. A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice. Transportation 21 (2), 135–152.
Train, K., 1985. Discount rates in consumers' energy-related decisions: a review of the literature. Energy 10 (12), 1243–1253.
Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, MA.
Urban, G.L., Hauser, J.R., Qualls, et al., 1997. Information acceleration: validation and lessons from the field. Journal of Marketing Research (JMR) 34 (1), 143.
Urban, G.L., Weinberg, B.D., Hauser, J.R., 1996. Pre-market forecasting of really-new products. Journal of Marketing 60 (1), 47.
Yang, S., Allenby, G.M., 2003. Modelling interdependent consumer preferences. Journal of Marketing Research (JMR) 40 (3), 282.