Survey Evidence on the Importance of Fuel Availability to the Choice of Alternative Fuels and Vehicles

DAVID L. GREENE

Limited fuel availability is a critical factor in the marketability of new fuels. A survey of US households is used to estimate the value of fuel availability and its influence on choice of fuel for a fuel-flexible vehicle and the choice of a dedicated-fuel engine for a vehicle. The marginal value of availability decreases as the percent of stations offering a new fuel increases. For fuel-flexible vehicles the cost of lack of availability decreases from US $0.35/gallon at 1% to US $0.02/gallon when 50% of stations offer the fuel.

La disponibilité limité du carburant est un facteur critique pour la mise sur le marché de nouveaux carburants. On utilise une enquête réalisée dans les foyers américains pour estimer la valeur du carburant disponible et son influence quant au choix du carburant dans le cas d'un véhicule pouvant utiliser plusieurs sortes de carburants et quant au choix d'un type précis de carburant pour un véhicule donné. La valeur marginale de la disponibilité décroit en rapport avec l'accroissement du pourcentage des stations offrant un nouveau carburant. Dans le cas de véhicules pouvant accommoder plusieurs carburants, les coûts de l'absence de disponibilité décroîent de US$0.35/gallon à US$0.02 lorsque la disponibilité du carburant dans les stations services croît de 1% à 50%.

David L. Greene is Senior Research Staff Member at the Center for Transportation Analysis, Oak Ridge National Laboratory, Oak Ridge, Tennessee

1. Introduction

The effect of limited fuel availability on the demand for alternative fuels and vehicles is a critical factor in the transition to alternative fuels. Because petroleum fuels have been so dominant for so long, the relationship between fuel availability and fuel demand has received little attention. This paper briefly reviews previous studies, then presents new results from a survey of US households. The new data address the: (1) choice of fuel for a fuel-flexible vehicle, and (2) choice of a dedicated alternative fuel engine for a vehicle. Fuel availability is described in terms of the percent of refueling stations offering the new fuel.

Fuel availability and fuel choice questions were asked in two separate nationwide random phone surveys of 1,000 US households (Opinion Research, 1996) conducted November 7-10, 1996, and December 5-8, 1996. One question concerned a hypothetical purchase of an alternative fuel engine option costing less than a gasoline engine but otherwise equivalent. Respondents were asked for the smallest percent of stations offering the new fuel that would make the engine an acceptable choice. For flex-fuel vehicles, respondents in the first survey were asked three questions trading-off decreasing fuel availability for increasing price advantage. In the second an "orthogonal" set of choices paired decreasing availability with decreasing price advantage. Respondents were asked what percent of the time they would buy the alternative fuel given each price-availability combination. The resulting data were used to estimate binomial logit models of fuel choice, in which the probability of choosing an alternative fuel or engine depends solely on its price advantage and availability.

2. Background

2.1 Early Studies of Diesel and Natural Gas Vehicles

The waxing and waning of the diesel car market in the US in the 1980s produced valuable information about how buyers of alternative fuel vehicles perceive fuel availability. Based on surveys of diesel car buyers, Sperling and Kitamura (1986) concluded that provided diesel fuel outlets were ubiquitous and predictably located, even if only one in ten stations offered diesel fuel it would be a relatively minor consideration in vehicle purchase decisions. They noted that as the number of diesel fuel outlets increased from about 3-5% of total fuel outlets in the mid-1970s to 10-15% by 1983, the percent of survey respondents expressing concern about fuel availability before the purchase of a diesel vehicle decreased from 60% to less than 20%.

A subsequent survey of diesel car owners in California was conducted in 1986 by Sperling and Kurani (1987) to determine the importance of limited fuel availability in households' vehicle purchase decisions. They found that degree of concern about fuel availability was not strongly related to either consumer attributes (including income) or vehicle characteristics. Vehicle owners had similar or less difficulty finding diesel fuel than they expected. This is an important finding for this study since it suggests that consumers' expectations about fuel availability are likely to correspond well with their actual experience. Overall, with diesel fuel availability in California ranging from 10 to 25% over the 1976-1985 period, 61% of respondents reported being "not concerned" about fuel availability prior to purchasing a diesel; 27% were somewhat concerned, and only 12% were very concerned. Based on this evidence, Sperling and Kurani (1987) "hypothesized" that for dual fuel vehicles that could also use gasoline, at least 10%, and probably more, of all refueling stations would have to offer a new fuel before vehicles could successfully penetrate the household vehicle market. For dedicated AFVs they judged that at least 15% availability would be required.

Some evidence on the value of fuel availability to bi-fuel vehicle owners is provided by Greene's (1989a) analysis of a Canadian survey of CNG vehicle owners. In 1988, Energy Mines and Resources Canada conducted a survey by mail of 2,100 participants in its Natural Gas Vehicles Program and 731 responded. Nearly all respondents (80%) were continuing CNG users, suggesting some bias in the sample towards those favorable to CNG vehicles. The program provided subsidies to NGV purchasers. Natural gas also enjoyed a price advantage of 70¢ (1989 $)/gallon of gasoline equivalent. At the time, there were 115 public CNG refueling stations operating in Canada, but the location of stations relative to the respondents was not known. Inferences about the value of fuel availability were based on correlations between a respondent's reporting refueling problems (14% reported such difficulty) and: (1) his overall satisfaction with the CNG vehicle (vehicle choice), and (2) the frequency with which CNG was purchased (fuel choice.) For the vehicle satisfaction model the statistical results implied a penalty of -$0.99/gallon for those reporting refueling problems, and -$0.08 for a weighted average of participants. In the fuel choice model, the values inferred for refueling were -$1.66/gallon for those reporting problems, -$0.23/gallon for the sample as a whole (all in 1989 $). Those reporting refueling problems are clearly a biased sample for whom it is a greater than average problem. But the entire sample is biased in the other direction, since it includes predominantly those who were satisfied with their NG vehicles. Thus, Greene's (1989a) estimates might best be considered a pair of upper and lower bound estimates for the value of fuel availability rather than point estimates.

Survey evidence from the experience with CNG and LPG bi-fuel vehicles in New Zealand lends further support to the inference that concerns over fuel availability become minor when station availability reaches 10 to 20% (Kurani, 1992).

From zero in 1979, the number of stations on the North Island of New Zealand increased to 359 by 1987, 14% of total refueling outlets. Surveys of the general public in 1980 and 1981 indicated that fuel availability was a major concern, and inhibited conversions to CNG. During these years CNG outlets comprised 1.8% and 2.4% of total refueling stations. By 1988, CNG and LPG vehicle owners reported having had little concern about fuel availability prior to converting their vehicles and somewhat less difficulty in actually finding refueling outlets after conversion. Seventy-one percent reported that finding fuel was "not at all difficult" and only 1.4% reported that it was "very difficult." This same perception of CNG availability was shared by owners of gasoline-only vehicles. The increase from 2 to 14% of stations apparently dispelled most drivers' concerns about fuel availability for CNG bi-fuel vehicles in New Zealand.

2.2 California Alternative Fuel Vehicle and Fuel Choice Studies

Several important analyses of AFV preferences have been produced as part of research in support of California's clean fuel vehicles initiatives. An initial pilot study (Bunch et al., 1993) was followed by a full scale 1993 survey of households in most of urbanized California that has provided a rich source of data for analyzing consumers' preferences for AFVs (Golob et al., 1995). In the California Pilot study the range of fuel availability respondents were asked to consider was 10 to 100%. As a result, the inferences from this study cannot be expected to shed much light on the value of availability in the critical 0-15% range.

Multinomial logit (MNL) models of vehicle and fuel choice estimated using the survey data permit inferences about the dollar value of availability.[1] In the vehicle choice model, perceived availability is represented as a quadratic function of the fraction of stations offering a new fuel. The marginal value of availability, the willingness to pay for a given increment in availability (say, 0.1), is therefore not a constant, but rather a linear function of availability. Three models were estimated by Bunch et al. (1993), but all imply virtually identical values for availability. The total willingness to pay for 100% versus 0% availability is $9,000 to $10,000 in vehicle purchase price equivalent, depending on which model formulation is used (1991 dollars assumed since the survey was conducted in 1991).

In the MNL fuel choice model, perceived availability is represented simply by the fraction of stations offering the fuel and so its marginal value is constant. Because Bunch et al. (1993) specified fuel costs in terms of cents/mile, the implied value of availability depends on the price of fuel and fuel economy. With gasoline at $1.20/gallon, the owner of a 28 MPG automobile would be willing to pay $1.68 to $1.77/gallon for full versus no fuel availability, or 1.7 to 1.8¢/gallon for each 0.01 increase. Drivers over 55 are apparently willing to pay more, $2.32/ gallon.

The California Pilot Study results imply a substantial willingness to pay for availability well beyond the 10 to 20% threshold identified by Sperling and others as the point at which motorists become unconcerned about availability. The fuel choice model implies that motorists value the last 10% of stations just as much as the first 10%, not an intuitive result. Even in the vehicle choice model with declining marginal willingness to pay, more than 60% of the $9,900 total willingness to pay is for the last 80% of availability.

The full-scale, 1993 California survey included two stated-preference experiments concerning choice of AFVs. Among the fourteen attributes used to characterize the alternative choices was "Service Station Availability," equal to the number of alternative fuel stations/gasoline station (e.g., "one methanol station for every ten gasoline stations") (Brownstone, 1995). This representation once more implies a constant marginal utility for a unit change in availability. Three service station availability measures were used, one for battery-powered electric vehicles, one for dedicated CNG vehicles, and another for both dual-fuel CNG and methanol flexible-fuel vehicles. Because the formulation of the choice models are rather complex, the inferred value of service station availability varies across income and demographic groups for each of the three availability measures. The resulting estimates of the full value of 100% versus 0% fuel availability are summarized in Table 1.

The California survey estimates of the value of fuel availability have four important features.

1/ Calculation of implied values of availability in the MNL model is described in Section 3, below.

First, the values are large, often as large as the total value of the vehicle. Second, the estimates are generally not statistically significant. Only the service station availability coefficient for dedicated fuel CNG vehicles in the one-vehicle household model is statistically significant at the 0.1 level, based on t-statistics. Third, fuel availability values for the battery electric and dedicated CNG vehicles are consistently larger than those for the flexible- and bi-fuel vehicles. Fourth, the range of values is quite wide: from -$36,000 to +$30,000 for one-vehicle households and from -$35,000 to +$36,000 for two-vehicle households.

The inference that fuel availability has less value for vehicles that are also able to use gasoline accords with intuition. Very large values of fuel availability may also seem appropriate at first but, on reflection, raise questions. Even if there are no stations offering methanol, why should an FFV which is otherwise essentially identical to a conventional vehicle and can function perfectly well on gasoline be worth $9,000 to $36,000 less to a potential buyer than if methanol were fully available? And is it reasonable to consider even a dedicated CNG vehicle, which at worst could be retrofitted to run on gasoline for several thousand dollars, essentially worthless if no CNG stations are available?

2.3 Other Indirect Evidence

Indirect evidence on the value of fuel availability for AFVs is provided by Segal's (1995) analysis of consumer survey data using conjoint analysis (Louviere, 1988). A random sample of 2,400 residential electricity customers of the Pacific Gas & Electric company were mailed a questionnaire asking them to rank descriptions of conventional, dedicated compressed natural gas (CNG), and battery electric vehicles. Based on the results of Segal's analysis one can infer that the value of 100% service station availability of CNG refueling versus only home refueling is $3,050, in vehicle purchase price equivalent. If home refueling were not an option, the value would certainly be greater than $3,050, so this must be considered a lower bound. While this value seems reasonable, at least one of the others implied by the model estimates does not. The implied value of refueling time is $14/minute, or $840/hour. Other constraints on refueling were also perceived as onerous by survey respondents. The coefficient estimates imply that not being allowed to refuel between 2 pm and 9 pm is worth a cost penalty of $3,300.

2.4 Summary

The early surveys of diesel and natural gas vehicle owners produced several important observations about fuel availability and vehicle and fuel choice. First, the early studies suggest that concern over fuel availability, and by implication its value to motorists, drops rapidly from very high levels when only 1-2% of stations offer a new fuel, to become a minor issue when 10-20% of stations offer the alternative fuel. Second, the surveys indicate that motorists' expectations about the difficulty of finding fuel were generally in accord with the actual difficulty they experienced. Third, fuel availability perceptions do not seem to vary greatly across income classes, gender, age, or any other socio-economic factor. These last two points suggest that the fuel availability problem is well understood by motorists and that they share a common understanding of it. These findings are encouraging, because they suggest that reasonably accurate inferences can be obtained from stated preference surveys in which respondents evaluate hypothetical fuel availability situations.

The comprehensive stated preference surveys conducted in California provide some additional information about the value of availability. Unfortunately, they shed little light on how availability will affect fuel and vehicle choice during the critical early stages of an alternative fuel market when retail outlets are likely to be scarce. In part, this is because the surveys did not focus on the critical range from 0 to 20% availability. But it is also because the linear functional form most often chosen to represent availability cannot represent a decline in marginal willingness to pay between 0% and 20%. In addition, availability was only one of many factors consumers were asked to consider simultaneously. It may be that more reliable and consistent responses could be obtained in a survey focused solely on the availability issue.

Table 1: Value of Service Station Availability Based on the California Study

Single Vehicle Households		**Vehicle Type**		
		Availability Coefficients		
Household Category	Net Capital Cost Coefficients	Electric 0.56	Ded. CNG 1.004	FFV Bi-fuel 0.2995
Inc. < $30K, child < 21	-0.00003290	$17,508	$30,517	$9,103
Inc. < $30K, no child < 21	-0.00006952	$8,285	$14,442	$4,308
$30 < > $75K, no child < 21	-0.00003950	$14,582	$25,418	$7,582
Inc. > $75K, child < 21	-0.00005253	$10,965	$19,113	$5,702
Inc. < $75K, no child < 21	0.00002766	($20,824)	($36,298)	($10,828)
Two-Vehicle Households		**Vehicle Type**		
		Availability Coefficients		
Household Category	Net Capital Cost Coefficients	Electric 0.5846	Ded. CNG 0.7408	FFV Bi-fuel 0.6312
Inc. < $30K, child < 21	-0.00007060	$8,280	$10,493	$8,941
Inc. < $30K, no child < 21	-0.00002882	$20,285	$25,704	$21,901
Inc. > $30K, lux veh, child < 21	0.00002205	($26,512)	($33,596)	($28,626)
Inc. > $30K, lux, no child < 21	0.00002118	$27,602)	($34,976)	($29,802)
Inc. > $30K, no lux, child < 21	-0.00001741	$33,578	$42,550	$36,255
Inc. > $30K, no lux, no child < 21	-0.00004112	$14,217	$18,106	$15,350

Source: Brownstone, 1995, Tables 3 & 4.

3. Theory and Estimation Methodology

The influences of fuel availability on the choice of dedicated AFV and on choice of fuel for a multi-fuel vehicle are represented here using a random utility, binomial logit (BNL) choice framework. The BNL random utility model expresses the value, U, of an option, i, as a function of its attributes (in this case price, P, and fuel availability $g(\sigma)$ where σ is the fraction of retail outlets offering a fuel and g is the "perceived availability" function), plus a random error ε_{ij} that varies across individuals, $j=1$ to N, and options, $i=1,2$. The constant A represents all other attributes besides P and $g(\sigma)$.[2]

2/ There may well be attributes of the respondents that interact with P and $g(\sigma)$ and influence their choices of fuel. Another way of putting it is that A, B, and C are in fact respondent-specific (i.e., Ai, Bi, Ci, each having a distribution across the population). In this case, let A, B, and C be the population means. The error term, ij, then contains the individual deviations from the population means (e.g., (Bi - B) P). The coefficient deviations from their respective means are then random variables, with mean zero by construction. As long as the deviations are independent of the included variables P and $g(\sigma)$, a reasonable assumption for a random survey, then the error term will still be uncorrelated with the variables P and $g(\sigma)$ and will also have mean zero. Thus, the above assumptions accepted, obtaining unbiased estimates of the population average parameters B and C does not require taking account of individual respondent's characteristics.

$$U_{1j} = A_1 + BP_1 + Cg(\sigma_1) + \varepsilon_{1j} \tag{1}$$

The probability, p_1, of choosing option 1 over option 2 is precisely the probability that $U_{1j} > U_{2j}$.

$$\begin{aligned}\text{Prob}(U_1 > U_2) = \ & \text{Prob}(A_1 + BP_1 \\ & + Cg(\sigma_1) + \varepsilon_1 > A_2 \\ & + BP_2 + Cg(1) + \varepsilon_2)\end{aligned} \tag{2}$$

When the error term ε has a Type 1 Extreme Value Distribution, p_1 is given by the following formula (e.g., Train, 1986).

$$P_1 = \frac{e^{U_1}}{e^{U_1} + e^{U_2}} = \frac{1}{1 + e^{U_2 - U_1}} \tag{3}$$

This function is especially convenient for estimation purposes. Note that the log of the odds in favor of purchasing option 2 is a linear function of the difference in their attributes.

$$\begin{aligned} ln\left[\frac{P_2}{P_1}\right] &= U_2 - U_1 \\ &= (A_2 - A_1) + B(P_2 - P_1) \\ &\quad + C[g(1) - g(\sigma)] \end{aligned} \tag{4}$$

A very important consideration is the functional form of $g(\sigma)$. The existing literature generally uses the simplest possible form namely, $g(\sigma)=\sigma$. This implies a constant value for each fractional improvement in station availability. Consumers would value the first 5% of stations exactly the same as the last 5%. Intuition suggests that the first few stations should be much more valuable than the last few, that is, the value of fuel availability should decrease as availability increases. One alternative is the exponential function ($g(\sigma)=e^{b\sigma}$). If the parameter $b<0$, then the *"cost"* of fuel availability will *decline* from a maximum of C at $\sigma=0$ to a minimum of Ce^b at full availability ($\sigma=1$). An alternative to the exponential function is a power function (σ^b). If $0<b<1$, the power function represents a *"value"* of fuel availability and will *increase* from 0 at $\sigma=0$ to a maximum of C at $\sigma=1$. Other possible forms for a value function include linear, which is a special case of the power function, namely $b=1$, and log (σ).

The function $g(\sigma)$ can be considered a way of transforming the availability measure (σ=fraction of stations offering a fuel) into a variable with a constant value to the consumer for each unit change (i.e., a constant marginal utility).[3] For example, the exponential form of $g(\sigma)$ implies that for each unit change in availability the cost of lack of availability will decrease by e^b. If $b=-2.88$, each 0.1 change in the fraction of stations offering the alternative fuel will reduced the cost of availability by 25%.

3/ For the exponential, power, or logarithmic function, the marginal utility of σ is a function of σ. The marginal value of station availability in dollars (given by the negative of the ratio of the marginal utility of σ to the marginal utility of fuel price) may decrease as σ increases. For the power function,

$$Cg(\sigma) = C\sigma^b; \quad \frac{dU}{d\sigma} = bC\sigma^{b-1}$$

$$-\frac{dU/d\sigma}{dU/dP} = -\frac{bC\sigma^{b-1}}{B}$$

For the exponential availability function the marginal cost of (lack of) availability may decrease as availability increases, provided that $b < 0$.

3.1 Choice of Dedicated Fuel Vehicle

In the new survey described below, a single question was asked to determine the trade-off between price and fuel availability in the choice of a dedicated fuel AFV:

> "Suppose you were buying a new car and could buy an optional engine that required a new fuel just as good as gasoline, and cost the same as gasoline. The optional engine costs $500 less, but the fuel is NOT available at all stations. What is the *smallest percent of stations* offering the new fuel that would make the new engine an acceptable choice?"

In the context of the model described above, this question asks for the level of fuel availability that exactly equates the price advantage with the fuel availability disadvantage (that makes the inequality in equation (2) an equality). Rearranging terms from equation (2) we interpret the value of σ given as the value that satisfies the following:

$$\begin{aligned} \varepsilon_{1j} - \varepsilon_{2j} &= (A_2 - A_1) + B(P_2 - P_1) \\ &\quad + C(g(1) - g(\sigma_j)) \end{aligned} \tag{5}$$

By the way the question was asked, we should have $A_2=A_1$, in other words, all characteristics of the vehicles other than price and fuel availability are identical. Also, if we assume that $g(\sigma)=e^{b\alpha}$, then for gasoline, $g(1)=e^b$. Making these changes, summing both sides of the equation across all respondents, and assuming that $\Sigma_j(\varepsilon_{1j}-\varepsilon_{2j})=0$ (please see footnote 2), we get the following condition for the ratio B/C,

$$\left(\frac{B}{C}\right) = \frac{m_g - e^b}{(P_2 - P_1)} \tag{6}$$

where m_g is the sample mean of the fuel availability measures, $g(\sigma_j)$. In the binomial logit random utility model, the ratio -C/B is the value (in dollars) of a one-unit change in the variable $g(\sigma)$.[4] Equation (6) says that this ratio depends on the

4/ While the value of a one-unit change in $g(\sigma)$ is constant, the value of a one-unit change in σ, is not.

price difference assumed and the sample mean of the fuel availability function. For the question asked, $(P_2 - P_1) = \$500$ for all respondents. The sample mean of $g(\sigma)$ depends on its functional form and the value of b as well as on the responses of those surveyed. The estimation method is described in the Appendix.

3.2 Choice of Alternative Fuel for Multi-Fuel Vehicle

In each of the two surveys, three questions were posed to elicit consumers' willingness to trade-off fuel availability for fuel price. In the first survey, the three questions were:

> " Suppose your car could use gasoline or a new fuel that worked just as well as gasoline. If the new fuel cost 25 (10/5) cents LESS per gallon but was sold at just one in 50 (20/5) stations, what percent of the time would you buy this new fuel?"

The fuel choice posed in the survey is once again a binomial (two alternative) choice problem. Consumers have been asked to estimate the percent of the time they would buy an alternative fuel, which corresponds to the expected relative frequency, or the probability, p_{1j} in equation (3), that the j^{th} respondent would buy the alternative fuel in the case of the t^{th} question. Well established statistical methods are available for estimating the parameters of this model (e.g., Greene, 1993, chap. 21).

If all consumers evaluated the price-availability trade-off in the same way (i.e., had the same utility function), then one should expect the same response from all consumers. But all consumers do not have the same utility function. First, it is assumed that each individual's utility function contains a random component, μ_{jt}, independent of the values of price or availability. Second, it is assumed that because of differences in geography, travel patterns, income, and other factors, each respondent has his own way of valuing fuel price and fuel availability. In terms of the linear utility function, these differences can be expressed as differences (A_j, B_j, C_j) from the mean parameters for the population as a whole (A, B, C).

$$\begin{aligned} U_{1jt} &= (A_1 + A_{1j}) + (B + B_j)P_{1t} \\ &\quad + (C + C_j)\, g(\sigma_{1t}) + \mu_{jt} \\ &= A_1 + BP_{1t} + Cg\,(\sigma_{1t}) + \varepsilon_{jt} \end{aligned} \tag{7}$$

where

$$\varepsilon_{jt} = A_{1j} + B_j P_{1t} + C_j\, g(\sigma_{1t}) + \mu_{jt}$$

Note that the differences will have means of zero and can be considered to vary randomly across the population just as μ does. Assuming that the perception of availability, $g(\sigma)$ also varied across individuals would not change the basic structure of equation (7) but only make the error term, ε, somewhat more complicated. All of the individual random components are collected in the single error term, ε, which, because of the assumed independence of A_j, B_j, C_j and μ_{jt} from the explanatory variables, will also be uncorrelated with P and $g(\sigma)$. If ε has the type I extreme value distribution, then the probability that an individual chosen at random from the population will prefer the alternative fuel is the binomial logit function of the average utilities of the two fuels shown above in equation (3).[5]

The surveys provide random samples of individuals who have revealed their choice probabilities, p_{1i}, at the same time revealing their ε_{jt}. Each individual, however, answered three fuel choice questions. Although it may be reasonable to assume independence of error terms across individuals (in particular, to assume independence across j of the A_j, B_j, and C_j), it is certainly not reasonable to assume independence across questions for the same individual. Indeed, equation (7) implies a correlation based on the commonality of the A_j, B_j, and C_j parameters for the same individual. This implies that statistical methods used to estimate these parameters using individual responses must recognize the existence of respondent-specific components.

Because the answers to the fuel choice question are percents, when divided by 100 they are continuous on the interval (0,1). As a result, the binomial logit equation can be transformed into a linear equation by taking the log of the ratio of probabilities (odds ratio). Parameters of the linear equation can be estimated by least squares methods

5/ This assumption was made for convenience and was not tested.

for panel data (Greene, 1993, chapter 16.4). The estimating form of the model is obtained from equation (4) by substituting $P_2 = (1-P_1)$ to obtain the following:

$$ln\left[\frac{1}{P_{1jt}}-1\right] = U_{2jt} - U_{1jt} \tag{8}$$
$$= (A_2 - A_1) + B(P_{2t} - P_{1t})$$
$$+ C[g(1) - g(\sigma_{1jt})] + \varepsilon_{2jt} - \varepsilon_{1jt}$$

If respondents correctly interpreted and accepted the premises of the questions, it follows that $(A_2 - A_1) = 0$, since the question states that there are no differences between the two fuels other than price and availability. As a result, there should be no statistically significant constant term in the regression. The presence of a statistically significant constant term could indicate either lack of fit for a particular model or rejection by some respondents of the premises of the question.

4. Survey Results

Fuel availability and fuel choice questions were asked in two, separate nationwide surveys administered by CARAVAN® Opinion Research Corporation under contract to the National Renewable Energy Laboratory[6] during the periods November 7-10, and December 5-8, 1996. Each time, one thousand adults were surveyed by telephone based on a national probability sample. The design for the fuel choice questions asked in both surveys is illustrated in Table 2. Note that the central question (10¢ price advantage and 5% availability) was asked of both groups to determine whether the two groups' responses were consistent. The question concerning dedicated fuel vehicle (engine) choice was asked only on the first survey.

Table 2: Design of Fuel Choice Questions

Availability / Price Advantage	\$0.25	\$0.10	\$0.05
1/50 Stations	Survey 1		Survey 2
1/20 Stations		Survey 1&2	
1/5 Stations	Survey 2		Survey 1

4.1 Choice of Dedicated Fuel Vehicle

Asked for the smallest percent of stations offering a new fuel that would make a vehicle that cost \$500 less but required a new fuel an acceptable choice, almost one-fifth (18.6%) of those answering indicated that less than 25% of all stations need offer the new fuel. Of the 1,004 persons surveyed, 252 answered "don't know" or "none." The response "none" is problematic, because it is not crystal clear whether respondents really meant that zero stations would be acceptable or if they simply chose not to respond. Since it seems implausible that so many respondents would choose the new engine if no stations offered the fuel, it was assumed that "none" meant "no response." The mean response was 42.6%, and the median 50%. The most common response was exactly 50%, given by 212 individuals. Only 10% of the respondents felt that less than 10% availability would be adequate. Only 8% felt that more than 80% of all stations needed to offer the new fuel.

The mean response for males was 40%, for females 44%. Responses differed little by region, race, or household size. Those over 55 years of age required, on average, 35% of stations to carry the new fuel, other age groups asked for more than 40%, on average. The acceptable level of fuel availability increased by 6% from the lowest (39%) to the highest (45%) of five income groups.

4.2 Choice of Fuel for Bi-Fuel Vehicles

Responses to the three questions asked in the November survey about the effects of price and fuel availability on alternative fuel purchases were disturbingly similar. About one-fifth of the respondents indicated that they would never buy the alternative fuel under any of the three conditions. Here, the response "none" was interpreted as "zero" percent of the time, since this was a plausible response for this question. Twenty to 30%, on the other hand, indicated they would buy the new fuel more than half of the time. The most common response was 50%, and the average purchase frequency ranged from 32.8 to 37.7% for the three options.

As Figure 1 shows, there are only minor differences among the distributions of responses to the three questions posed in the November survey. As a result, it was not entirely clear whether:

6/ Special thanks are due to Philip Patterson of the US Department of Energy and to Vince Schaper of the National Renewable Energy Laboratory for securing the survey data.

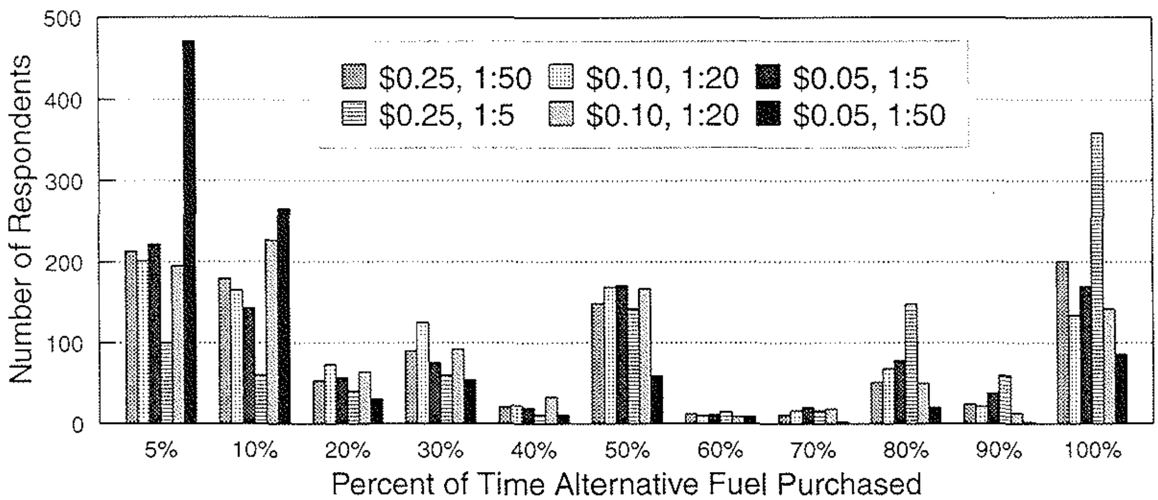

Figure 1: Alternative Fuel Purchase Frequency Distributions for the November and December 1996 Surveys

1) respondents found all three choices approximately equally attractive; or 2) respondents were unable or unwilling to understand the question, or effectively comprehend the differences among the alternatives, and therefore simply gave the same response to all questions. If the latter were true, it would call into question the use of the stated preference methodology for this subject. A closer examination of the data provided some reasons to think the respondents did understand the question and did respond thoughtfully. First, only 5 to 10% of the respondents answered "don't know" to the question. If the questions had baffled respondents, a higher percentage of "don't knows" would be expected. Second, responses differed systematically, albeit slightly, by age, gender, income, and education level. Those aged 25-44, females, those with a college degree and the highest income group, preferred high availability to low cost. Respondents without a high school diploma much preferred low cost.

The second round of survey data not only proved that respondents did understand the questions and respond thoughtfully, but revealed remarkable consistency between the two survey groups in their responses to the question both surveys had in common. Although respondents to the second survey gave very different answers to the two new questions not asked in the first survey, they gave nearly identical responses as the first group to the repeated question. Both surveys contained the case of a $0.10/gallon alternative fuel price advantage and 5% station availability. Two entirely different groups of 1,000 respondents gave nearly identical responses to this question. The average response to this question in the November survey was 32.8%. The average in the December survey was 31.7%. According to the survey design criteria, a difference of four percentage points is required for statistical significance at the 95% level. Grouping the data into five reasonable intervals, shown in Figure 2, a χ^2 goodness of fit test does not reject the hypothesis that the two distributions are the same at the 0.05 significance level.

5. Inferences on the Value of Fuel Availability

Data collected in the November and December, 1996 surveys permit inferences about the way consumers believe they will trade-off fuel price for fuel availability, in other words, to estimate the value to consumers of fuel availability. First, responses to the question about the percent of stations that would have to offer a new fuel to make the choice of an alternative fuel engine acceptable were used to fit the binomial logit equation to the cumulative frequency distribution of responses. Next, the questions concerning the choice of fuel for a bi-fuel vehicle are used to statistically estimate the parameters of the binomial choice model, using a random effects model technique (Greene,

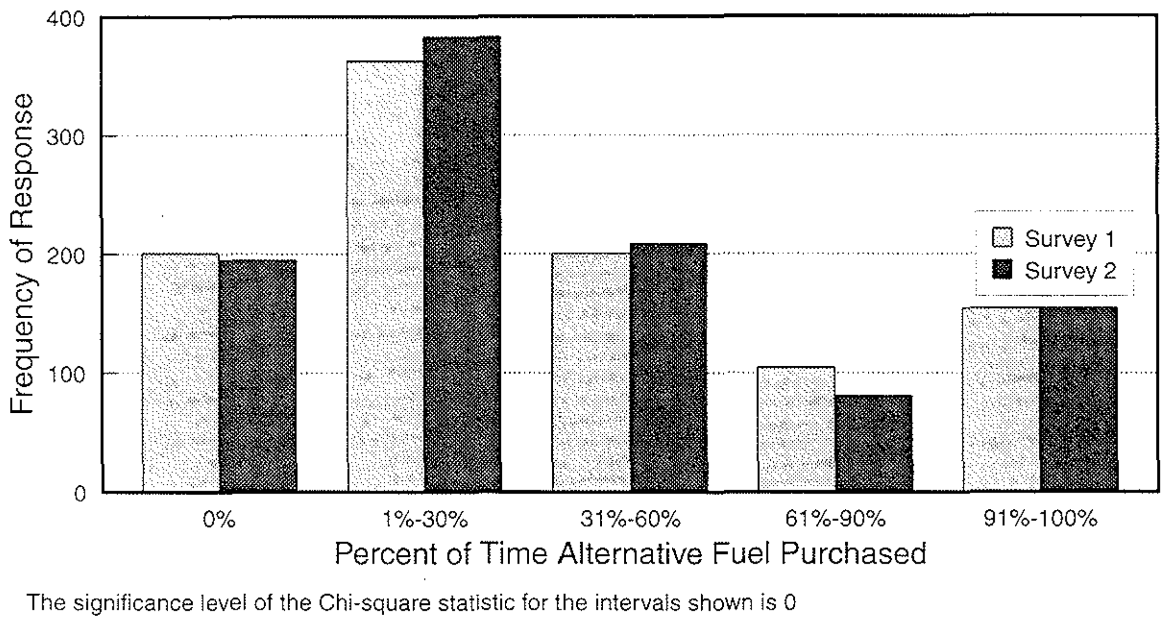

Figure 2: Comparison of Responses to the Fuel Purchase Question Asked in Both Surveys

1992, chap. 29).

5.1 Results for Dedicated AFV Fuel Availability

The parameter estimates that best fit the cumulative distribution of survey responses imply a total value of fuel availability for the dedicated alternative fuel engine of about \$1,300, new car purchase price equivalent.[7] This is equivalent to about \$0.38/ gasoline equivalent gallon (based on discounted lifetime fuel consumption). An alternative estimate of the full value of fuel availability can be derived from the mean response using equation (6). This turns out to be just slightly higher, about \$1,550. Thirteen to fifteen hundred dollars is much smaller than the estimates of \$9,000 and up derived from the California surveys. However, in this survey the question implied that only the car's engine, not the entire vehicle would be useless if no fuel were available. There is no mention of differences in storage tanks, transmissions, etc. It would be reasonable, therefore, for respondents to assume that the engine could be retrofitted or replaced if no fuel were available, thereby placing an upper limit on the buyer's potential losses.

The estimated parameters imply a high sensitivity to availability for the choice alternative fuel engine. For example, the model predicts that if the engines were priced equally, less than 1% of respondents would consider the new engine acceptable if less than 10% of stations offered fuel for it. At 1% availability the willingness to pay for another percentage point increase (to 2%) is \$34. This decreases to about \$20 at 20% availability and by the time 50% of stations carry the fuel, the marginal value is about \$9/percentage point.

Using the exponential availability function, the best fit of the binomial logit model to the survey frequency data was obtained when $b=-2.03$, $C=-6.92$, and $B=-0.00566$. These values provide a reasonable fit to the observed cumulative distribution, as shown in Figure 3. For the frequency intervals shown in Figure 3, the χ^2 goodness of fit measure is 36.6, as opposed to a value of 205.4 under the null assumption that all intervals are equally probable. The fit of the model to the relative frequency distribution is less exact, as shown in Figure 4 and confirmed by the χ^2 value of 157.5. This may be largely due to the tendency of

7/ The exponential form of the fuel availability function was used in deriving these estimates. The power-function form was also fitted and produced very similar results.

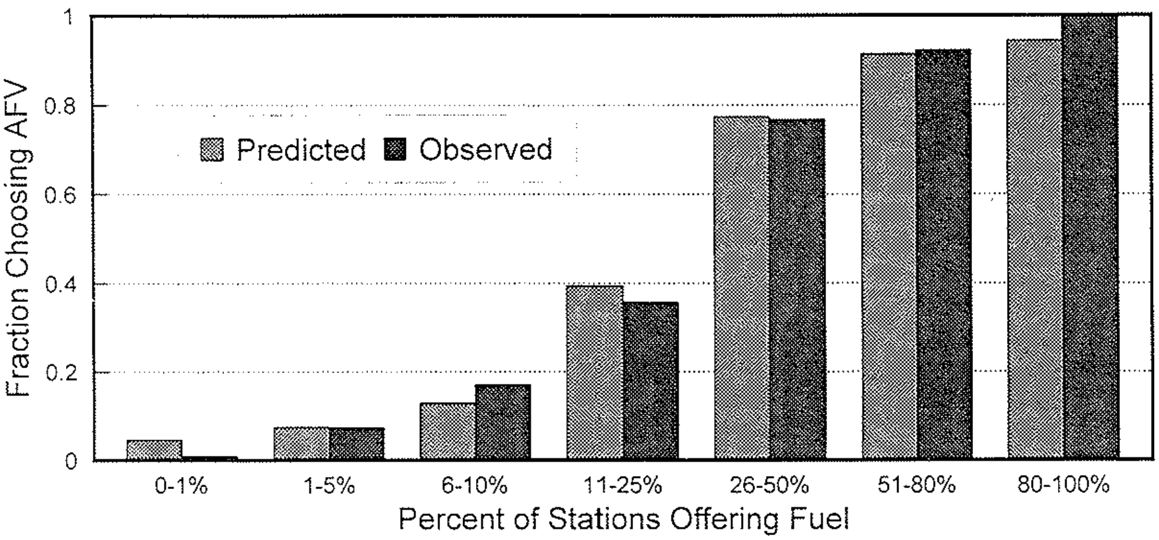

Figure 3: Cumulative Frequency Fit of Exponential Fuel Availability Model of Alternative Fuel Engine Choice

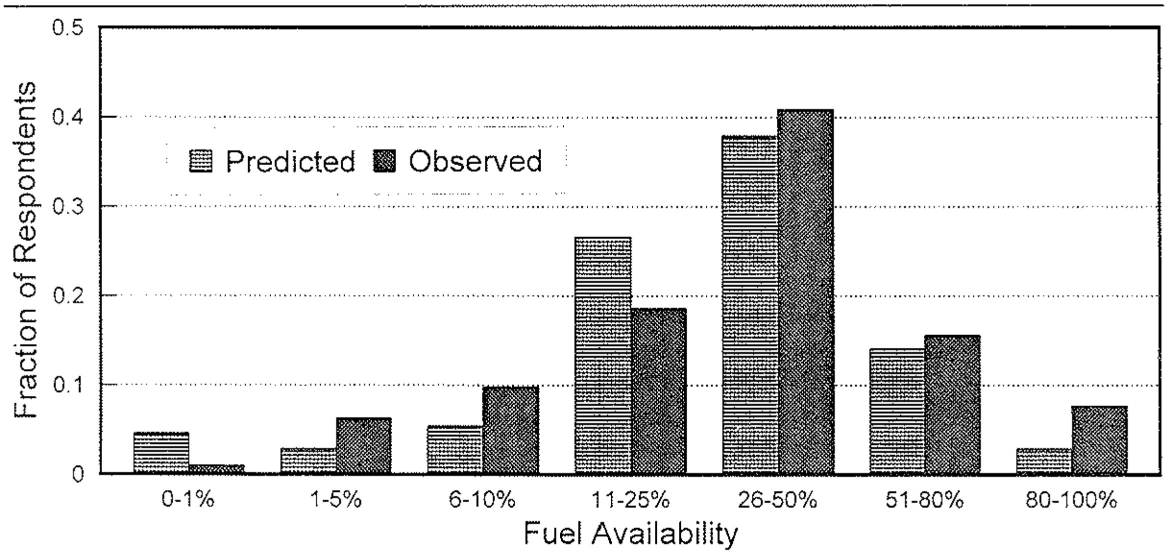

Figure 4: Relative Frequency Fit of Exponential Fuel Availability Model of Alternative Fuel Engine Choice

respondents to choose round numbers, such as 50%, or 10%, thereby producing a "lumpy" rather than a smooth frequency distribution.

The alternative fuel engine choice models also have important implications for the price-sensitivity of choice of alternative fuel vehicles. There is a sizable econometric literature on motor vehicle choice, but essentially all of this literature concerns the choice among different makes and models of vehicles, rather than the choice of fuel options for the same make and model of vehicle. The alternative fuel engine choice question specifically states that the engines apply to the same vehicle and are identical in all respects except for their price and the availability of fuel they require. In other words, the unobserved attributes of the choices should be minimal. This implies that they are very close substitutes, which implies that the choice between them should be highly price elastic.

Price elasticities in the logit model are not constant, but vary with market share and price. Specifically, the price elasticity of engine choice is $\beta = B\ (1-p_i)P$, where B is the coefficient of price difference in the binomial logit model, p_i is the probability of choice (or market share) of alternative i, and P is price. If we take P to be the price difference between the two engines, then at P = \$500, $p_i=0.01$, and $B=-0.00525$ the price elasticity of engine choice is -2.6. But this is the elasticity with respect to price *difference*. The corresponding elasticity with respect to total vehicle price would be much higher. For example, at a price of \$20,000 and 1% market share, the price elasticity of choice of the alternative fuel engine would be about -100. Even at 50% market share the elasticity with respect to full vehicle price would be -50. These elasticities are about an order of magnitude greater than those found in the literature for revealed preference choices among makes and models, and about five times as large as Greene (1986) found in his study of diesel engine choice (again, based on revealed preference data). Of course, these elasticity estimates apply to choice situations in which there are many unobserved attributes. These new results suggest that choice among very similar alternative fuel options may be highly price elastic.

5.2 Results for Fuel Choice

The sample design for the choice of fuel for a bifuel vehicle (Table 2) specifies three levels of price and three levels of fuel availability, all within the range 0 to 20%. Thus, while the data adequately describe the level and rate of change of the cost of fuel availability within this critical interval, the sample design does not readily allow discrimination among alternative functional forms. Parameters of the binomial logit fuel choice model were estimated using the pooled individual observations from the two surveys, by means of random effects model regression analysis, a technique especially suited to panel data. LIMDEP version 6.0 software, specifically the REGRESS procedure with options for panel data was used to estimate the binomial logit equations (Greene, 1992, chap. 29). The exponential and power function models were estimated by an iterative search over values of b, so as to maximize the adjusted R^2 of the random effects regression model. The results are summarized in Table 3.

The first issue is whether the random effects model for panel data is the correct model. The Breusch-Pagan Lagrange Multiplier statistic tests whether there are actually significant panel effects or whether the assumption of random errors across all observations (as in the classical regression model) is sufficient. The Breusch-Pagan test statistics decisively reject the hypothesis of no panel effects, for all the alternative models. The relative importance of respondent-specific effects can be seen in comparing the adjusted R^2 values for the random effects (REM) and least squares dummy variable (LSDV) fixed effects models. The fixed effects statistically explain about five times as much variation in the data as the price and availability variables do. Clearly, there are strong respondent-specific effects and the error terms for an individual respondent are related. The Hausman chi-square statistics test whether the best representation of the panel effects is by a set of respondent-specific constants (fixed effects) or by correlation among the error terms (random effects). Large values of the Hausman statistic favor the fixed effects model. The data once again clearly and decisively support the random effects model, consistent with the assumptions made in formulating equation (8).

All four of the alternative availability functions fit the data about equally well. The range of adjusted R^2 from the worst to the best fitting model is 0.0011, or about 1%. It is not surprising that the functional form could not be clearly identified, since the questions referred to only three availability levels. One feature that does discriminate among the functional forms is the presence or absence of a constant term in the regression equation. As the questions were framed, there should be no constant terms, since the fuels are purported to be identical save for price and availability. Therefore, a significant constant term either indicates lack of fit, or is a result of consistent misinterpretation of the question on the part of the respondents. The exponential model is the only functional form that did not contain a statistically significant constant, and therefore should be preferred over the other functional forms.

Total and marginal availability costs curves for the exponential model are shown in Figure 5. The

Table 3: Results of Random Effects Model Estimation of Fuel Availability Choice Models

	Exponential	Power	Logarithmic	Linear
Constant	-0.0503*	-3.0365	-2.1782	-6.5505
	(0.1045)	(0.1991)	(0.1075)	(0.3280)
C (Availability)	-3.2651	6.4819	0.9114	9.4217
(std. error)	(0.1212)	(0.2406)	(0.0330)	(0.3544)
b (exponent)	-5.35	0.620	n.a.	n.a.
B (Price)	-9.1451	-9.1563	-10.447	-8.6539
(std. error)	(0.3392)	(0.3392)	(0.3442)	(0.3391)
Group Effects Test	1742.5	1742.8	1772.7	1729.3
Breusch-Pagan LM	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(signif. level)				
Fixed v. Random	0.0001*	0.0001*	0.0001*	0.0001*
Hausman's chi-square	(0.9999)	(0.9999)	(0.9999)	(0.9999)
(signif. level)				
REM R-squared	0.0907	0.0907	0.0896	0.0905
LSDV Adj. R-sq.	0.5790	0.5791	0.5834	0.5770
Value of 100% v. 1%				
Per gallon	$0.34	$0.67	$0.40	$1.08
Purchase Price Equiv.	$1,067	$2,115	$1,274	$3,417

* Not statistically significant.

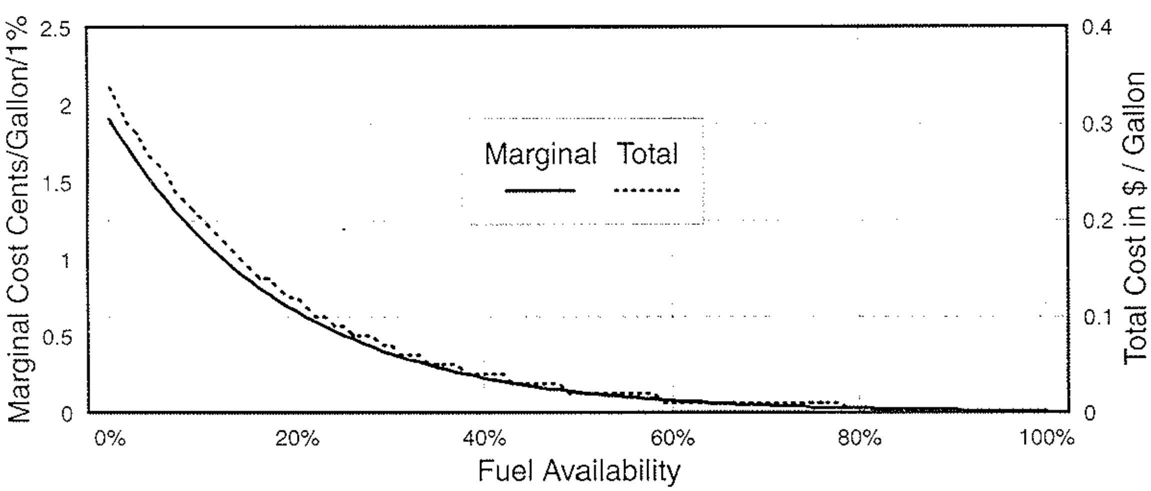

Figure 5: Value of Fuel Availability for Bi-Fuel Vehicles: Exponential Availability Function

total cost/gallon of fuel availability begins at about $0.35 at near zero and declines rapidly to about $0.10/gal. at 20% availability.

The sensitivity of alternative fuel market share to fuel availability ranges from inelastic to slightly elastic depending on both the level of availability and market share (Table 4). The alternative fuel's share is inelastic at small percentages, but large

Table 4: Elasticities of Alternative Fuel Choice Exponential Availability Model

	Availability Elasticities			Price Elasticities
Elasticity at	2%	10%	20%	Price = $1.30
1 %	0.31	1.01	1.19	-11.77
10 %	0.28	0.92	1.08	-10.70
50 %	0.16	0.51	0.60	-5.94
90 %	0.03	0.10	0.12	-1.19

relative changes in small percentages are actually small absolute changes. For example, a 2 percentage point change in availability from 2 to 4% is a 100% increase, and would produce about a one-third increase in market share in the exponential or power function models. Starting with a market share of 10%, a 10 percentage point increase in availability from 10 to 20% would again be a 100% increase, and would produce roughly a 100% increase in market share.

Fuel choice is highly price elastic. For alternative fuel market shares less than 20%, the price elasticity of fuel choice is about -10. With gasoline at $1.30/gallon (approximately the national average price when the survey was conducted) price elasticity in the exponential model ranges from nearly -11.8 at 1% market share to -9.5 at 20%, down to -5.9 at a 50% market share (Table 4).

These elasticities are somewhat lower than those found in econometric analyses of choice among gasoline grades or levels of service based on revealed preference data. For choice among premium and regular grade versions of leaded and unleaded fuel, Greene (1989b) estimated price elasticities of -15 to -20 at 50% market share, using data from 1982-1985. In an analyses of demand for full versus self-service gasoline, Phillips and Schutte (1988) found elasticities in the range of -35 to -40. In general, the more similar two options are with respect to attributes not represented in the choice model, the greater the price elasticity of choice. A plausible explanation for the lower elasticities estimated from the 1996 survey data is that there are significant unobserved attributes associated with fuel availability, namely the specifics which make the value of fuel availability to an individual differ from those of the group as a whole (see equation (7) for the mathematical representation of this effect). These would include such factors as annual mileage, local geographical effects including the density of refueling stations, vehicle range, the value of time, as well as personal preferences. Despite the fact that the fuel options are described as identical except for price and availability, there are sufficient respondent-specific factors affecting the value of availability to make the "unobserved" attributes of this choice more important than the unobserved attributes in choice of gasoline grade, for example.[8]

An interesting feature of the logit model is that the sensitivity of alternative fuel market share to availability depends on the price advantage of the alternative fuel (or, equivalently, its net advantages in other respects). Figure 6 illustrates that as the price advantage of the alternative fuel increases, its market penetration increases more rapidly with availability. With no price advantage, even a 60% availability gives only about a 20% market share. With a $0.10/gallon price advantage, a 20% market share is obtained at less than 25% availability. Given a $0.25/gallon price advantage, it takes less than 5% availability to attain a 20% market share. At 20% availability, an 80% market share is reached.

6. Conclusions

The 1996 stated preference survey data appears to provide a sound basis for understanding how consumers believe they will respond to the limited availability of alternative fuels. The fact that respondents understood the question well and responded thoughtfully is reflected both in the consistency of the results with basic economic principles and in the close correspondence between responses to the same question asked to two entirely different national samples. The results are also generally consistent with the conclusions of previously published studies, in that they indicate a steep decline in concern with availability between zero and 20%, and substantial importance within that range. Fuel availability appears to be a problem with which consumers are quite familiar, which they understand well, and about which they

8/ The presence of significant unobserved attributes does not imply that there should be a constant term in the model. The model is specified and estimated in such a way that these unobserved factors should have an average value of zero and should be uncorrelated with the explanatory variables. It is the variance of the unobserved attributes, rather than there mean, which affects the price elasticity of fuel choice.

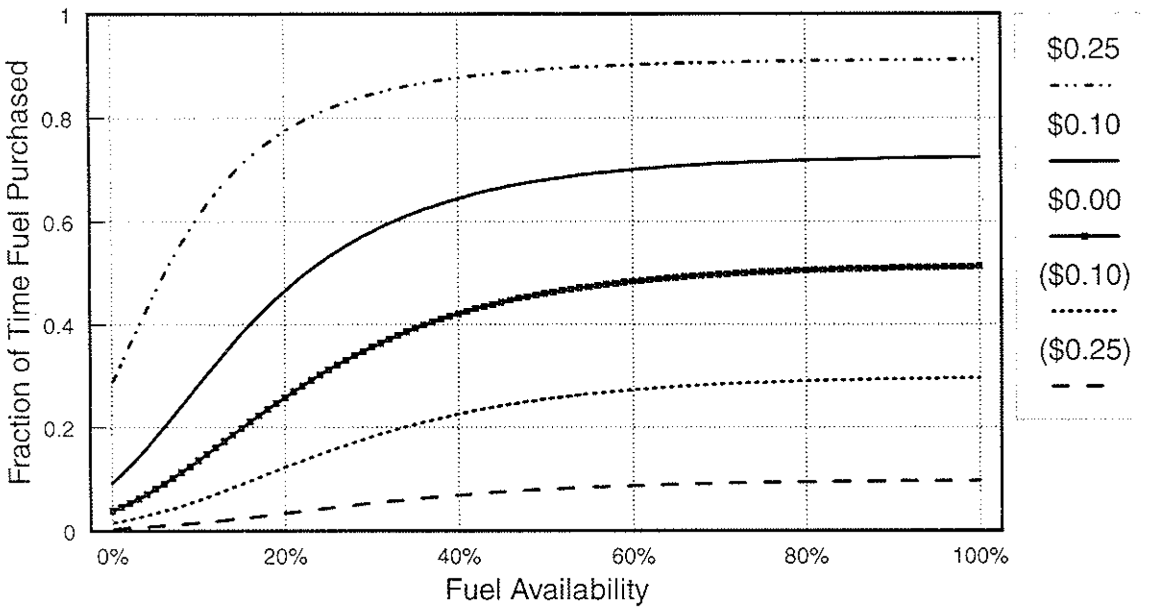

Figure 6: Alternative Fuel Market Share as a Function of Fuel Availability and Price Advantage

can give logical and consistent responses to a straightforward and simple question. All of this, together with the fact that there is no apparent reason for respondents to give biased responses to the questions asked, suggests that the results of this stated preference survey may be a good prediction of how consumers actually will behave when faced with limited fuel availability.

Unlike previous studies, this analysis has focussed on the value of fuel availability to consumers during the initial phases of alternative fuel market penetration when availabilities will be low. As a result, it provides some useful information about the critical range between 0 and 20% market availability. While it was not possible to definitively discriminate among alternative functional forms for perceived availability over the range of data (2 to 20%), only the exponential availability function fits the data, is consistent with the premises of the question, and behaves reasonably outside the range of the data.

Consumer willingness to pay for (or, alternatively the cost of) availability for a bi-fuel vehicle begins at about $0.35/gallon, declines to about $0.10/gallon at 20% availability, and to $0.02/gallon at 50%. Marginal willingness to pay for availability decreases from 2¢/gallon/percentage point at near zero availability to about 0.5¢/gallon/percentage point at 25% availability. This low willingness to pay for greater availability is consistent with the qualitative findings of previous surveys that consumers are relatively unconcerned about availability at levels above 20%.

Willingness to pay for availability in the case of choice of a dedicated fuel engine appears to be in the range of $1,000 to $2,000, purchase price equivalent. The choice of alternative fuel engine is highly sensitive to price when all other vehicle characteristics are equal. For either functional form, the elasticity of choice of engine with respect to total vehicle purchase price is on the order of -100 for low fuel availabilities. This value may be construed as a limiting price elasticity for the choice of nearly identical alternative fuel technologies for a given make and model of vehicle. A good example might be the choice between an alcohol/gasoline fuel flexible engine or a conventional gasoline engine. Alternative fuel technologies that significantly change other vehicle attributes (e.g., compressed natural gas or battery electric) would have much lower price elasticities of vehicle choice.

7. References

Brownstone, D., D.S. Bunch, T.F. Golob, and W.

Ren (1995) 'A Transactions Choice Model for Forecasting Demand for Alternative-Fuel Vehicles,' University of California, Irvine (forthcoming, *Transportation Economics*, 1996).

Bunch, D.S., M. Bradley, T.F. Golob, R. Kitamura, and G.P. Occhiuzzo (1993) 'Demand for Clean-Fuel Personal Vehicles in California: A Discrete-Choice Stated Preference Study,' *Transportation Research*, 27A: 237-53.

Davis, S.C. and D.M. McFarlin (1996) *Transportation Energy Data Book*, Ed. 16, ORNL-6898 (Center for Transportation Analysis, Oak Ridge National Laboratory: Oak Ridge, Tennessee).

Golob, T.F., S. Kim, and W. Ren (1995) 'A Structural Model of Vehicle Use in Multi-Vehicle Households,' presented at the *Annual Meeting of the Transportation Research Board*, Washington, DC, January 23-28.

Greene, D.L. (1989a) 'Fuel choice for Dual-Fuel Vehicles: An Analysis of the Canadian natural Gas Vehicles Survey,' *SAE Technical Paper Series No. 892067* (Society for Automotive Engineers: Warrendale, PA).

Greene, D.L. (1989b) 'Motor Fuel Choice: An Econometric Analysis,' *Transportation Research A*, 23A:3:243-53.

Greene, D.L. (1986) 'The Market Share of Diesel Cars in the USA, 1979-83,' *Energy Economics*, 8:1:13-21.

Greene, W.H. (1993) *Econometric Analysis*, second edition (Prentice Hall: Englewood Cliffs, New Jersey).

Greene, W.H. (1992) *LIMDEP User's Manual and Reference Guide*, Version 6.0 (Econometric Software, Inc.: Bellport, NY).

Kurani, K.S. (1992) 'Application of a Behavioral Market Segmentation Theory to New Transportation Fuels in New Zealand,' Ph.D. Dissertation, Institute of Transportation Studies, University of California, Davis.

Louviere, J.J. (1988) *Analyzing Decision Making: Metric Conjoint Analysis* (Sage University Press: Newbury Park).

Opinion Research Corporation (1996) 'Caravan: Interest in New Fuel,' prepared for the National Renewable Energy Laboratory, December 5 (Opinion Research Corp.: Princeton, NJ).

Phillips, O.R. and D.P. Schutte (1988) 'Identifying Profitable Self-Service Markets: A Test in Gasoline Retailing,' *Applied Economics*, vol. 20, February, pp. 263-72.

Segal, R. (1995) 'Forecasting the Market for Electric Vehicles in California Using Conjoint Analysis,' *The Energy Journal*, 16:3:89-111.

Sperling, D. and R. Kitamura (1986) 'Refueling and new Fuels: An Exploratory Analysis,' *Transportation Research A*, 20A:1:15-23.

Sperling, D. and K.S. Kurani (1987) 'Refueling and the Vehicle Purchase Decision: The Diesel Car Case,' *SAE Technical Paper Series No. 870644*, Society of Automotive Engineers, Warrendale, PA.

Train, K. (1986) *Qualitative Choice Analysis*, (MIT Press: Cambridge, Massachusetts).

Appendix A: Method of Estimating the Engine Choice Parameters

If we assume that equation (6) holds, then two of the three parameters must be estimated. Either B or C can be derived from knowledge of the other and b using (6). This suggests that one search for values of b and C that give the best correspondence between the sample and the estimated binomial logit model, and use the ratio of B to C calculated via equation (6) to derive B. Alternatively, one can search for values of b, B, and C that give the best fit and use the ratio B/C as a check on the validity of the resulting estimates. An iterative search was used to determine the values of b, B and C that provide the best fit of cumulative frequencies predicted by the binomial logit model to the observed responses. For this purpose, responses were grouped into seven availability intervals as illustrated in Figure 3. At each interval's upper cutpoint, σ_j, the expected fraction of respondents who would prefer the alternative fuel option can be obtained by inserting σ_j in the binomial logit equation and, using current guesses for B, C, and b, calculating p_{1j} , as in equation (3). The predicted cumulative frequency is obtained by multiplying p_{1j} by the sample size, N. If O_j and F_j are the observed and predicted cumulative frequency counts in the j^{th} class, the parameters are chosen to minimize the χ^2 statistic. Iterations were stopped when changing parameter estimates by 0.1% produced no further decrease in χ^2.

$$\chi^2 = \sum_{j=1}^{N} \frac{(O_j - F_j)}{O_j} \tag{A.1}$$

Fitting to the cumulative frequencies does not give exactly the same results as fitting to relative frequencies, although the differences are not great. The cumulative distribution was preferred because of the "lumpiness" of the survey responses (e.g., respondents tend to give convenient, round percentages, such as 0%, 10%, or 50%). This creates pronounced spikes in the relative frequency distribution at those points (Figure 4). In addition, the small number of cumulative frequency classes also helps to smooth the distribution. The fit to relative frequency is clearly worst in the 26-50% category. This category, however, is no doubt inflated by the tendency of respondents to prefer responding with round numbers (50% was the modal response). Underestimation of this category is, therefore, not necessarily a bad thing.