Econometric Consequences of Combining Hedonic Model and Conjoint Analysis for Environmental Valuation

Arianto A. Patunru
University of Illinois

This paper addresses some important econometric issues when hedonic model and conjoint analysis are combined in valuing changes in environmental quality. Contrary to traditional hedonic price model which uses continuous price function, we follow the recent development that applies discrete-choice model. Conjoint analysis based on choice set is added to deal with econometric weaknesses encountered in hedonic model. We use random utility model as the framework for combining data. First, multinomial logit model is elaborated. However, studies have found that this model suffers from strong parameter restriction. We then study the plausibility of using nonparametric approach. 2002.

1. Introduction

This paper addresses some important econometric issues that may appear when two data sources, namely hedonic data and conjoint data, are combined in order to make a sound environmental valuation. In the second section we briefly overview methods commonly used in valuing the change in environmental quality. We then focus our attention to hedonic approach and conjoint analysis, the methods that can be combined together to obtain better estimates in environmental valuation. The third section summarizes the development of random utility model and multinomial logit model. We also discuss the property of independence from irrelevant alternatives. The fourth section addresses the data combination issues, including taste and scale heterogeneity, as well as the estimation function. The fifth section lays out possible implementation of the model. In the sixth section we briefly review the development in nonparametric approach to response model. Last section lists some possible further extensions.

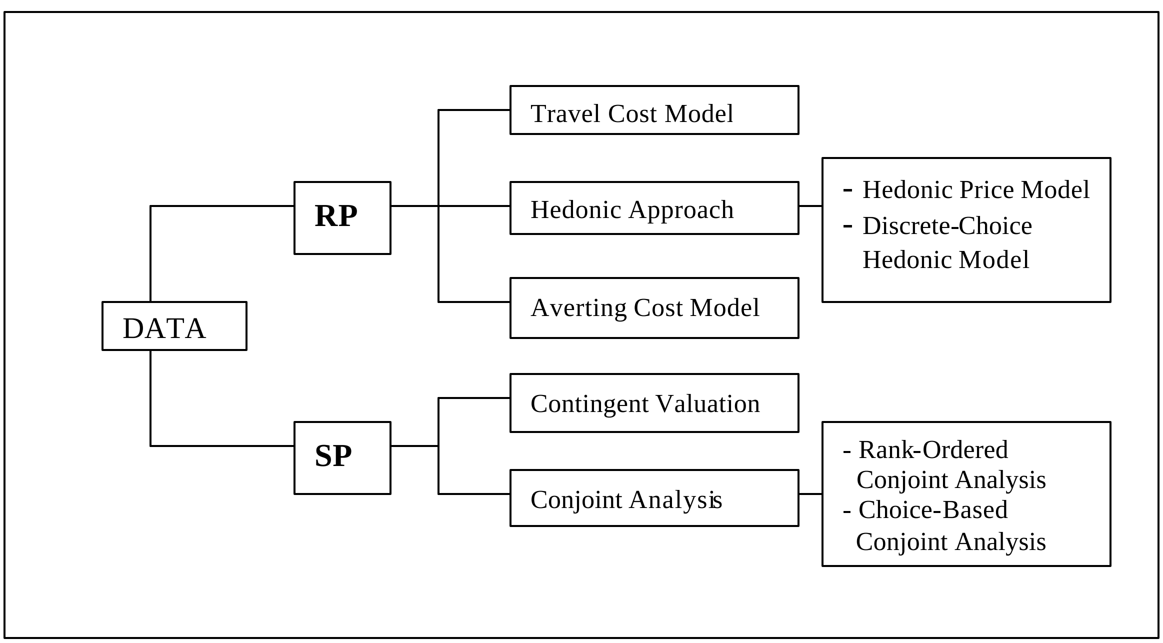
2. Environmental Valuation Method

Methods for valuing environmental quality can be divided into two groups, namely revealed preference (RP) methods and stated preference (SP) methods (see **Figure 1**). The first approach is determined by observing actual choices by individuals and using the information to infer the relation between money and environmental goods. Included in this category are travel cost method (analysis based on the cost of traveling to destination -- usually a recreational site, as well as the time spent on the traveling), averting cost model (model that analyzes expenditures on environmental quality improvement), and hedonic approach. This third approach refers to statistical techniques to measure how one variable affects another variable indirectly. For example, a hedonic pricing model could help determine how changes in air quality affect property values.

The second method, on the other hand, is determined by asking individuals how much worth an environmental good is. This is done for example by investigating how much they would be willing to pay for a reduction in pollution to some particular level. Contingent valuation usually uses yes-no type of questions in its survey format. This is more convenient for respondents; however it may provide insufficient information for the researcher. Conjoint analysis, in contrast, has more variations, i.e. it asks respondent's preference over one or more sets of choices.

Hedonic analysis has been used for environmental valuation extensively since its development by Griliches (1971) and Rosen (1974). One of the weaknesses of traditional hedonic techniques (i.e. "hedonic-price model") is that it uses continuous function to relate the price of a good to its attributes. This has been found problematic when we are dealing with valuing the effect of environmental change into a market good which is perceived as discrete by consumer in their preference function. In fact, consumer chooses *one* house, for example, that provides the highest utility from all the houses in the feasible choice set. Therefore, it is more appropriate to apply the discrete-choice hedonic model that overcomes this problem. To get a clearer picture, in subsection 2.1 we provide an example of traditional hedonic approach.

Figure 1. Methods for Environmental Valuation



Unlike hedonic analysis, conjoint analysis is a relatively new technique borrowed heavily from marketing literature (e.g. Green and Srinivasan, 1990). This analysis refers to methods of decomposition that estimates the structure of a consumer's preference given her evaluation of a set of alternatives that are prespecified in terms of levels of different attributes. We explain this method briefly in subsection 2.2.

2.1. Hedonic Analysis

We briefly formalize the hedonic price method as follows. Suppose an individual's utility is a function of housing, environmental quality, her characteristics, and any other goods. In addition, she manages to allocate her budget to buy housing and other goods. That is,

$$U_n = U_n(Z_j, Q_j, X_n, C_n) \tag{1}$$

$$Y_n = X_n + P_j(Z_n, Q_n) \tag{2}$$

where Z is housing characteristics, Q is an index of environmental quality, X is a composite commodity, C is individual characteristics, Y is income, and P is housing price.

The first order condition of maximizing the above utility subject to the budget constraint yields:

$$\frac{\partial U_n}{\partial Z_j} \Big/ \frac{\partial U_n}{\partial X_n} = \frac{\partial P_j}{\partial X_n} \text{ and } \frac{\partial U_n}{\partial Q_j} \Big/ \frac{\partial U_n}{\partial X_n} = \frac{\partial P_j}{\partial Q_j} \tag{3}$$

Therefore, the individual will equate the marginal utility the market good (housing) and the marginal utility of environmental good to their associated marginal price.

Our interest is how to analyze the effect of a change in the environmental quality on the housing price. From this, we can infer how much the individual value the house given a certain level of environmental quality. Therefore, we can compute her willingness-to-pay for, say, an improvement in environmental quality (e.g. cleaner air). The steps involved are:

1. Estimate the inverse demand function: $P_j = P_j(Z_j, Q_j)$.
2. Derive the willingness-to-pay as the change of housing price due to change in environmental quality, i.e. $W = \partial P_j / \partial Q_j$.
3. Estimate marginal willingness-to-pay function using the estimated W regressed on income and the environmental measure: $\hat{W} = W(M_n, Q_j)$
4. Calculate total willingness-to-pay due to change in environmental quality over individuals:

$$TW = \sum_{n} \int_{Q_{j1}}^{Q_{j0}} W(M_n, Q_j) dQ$$

Implicit in the model above, we in fact employ the following assumptions:

1. Housing price is strongly affected by surrounding environmental quality.
2. Individual decision to purchase a house is a function of her expected exposure to environmental change.
3. Individual is willing to pay more for better environmental quality.
4. Individual takes into account all housing characteristics.
5. There are sufficient variations in environmental quality.
6. Market is in equilibrium.

One classic example to illustrate the application of the above framework is the study by Harrison and Rubinfeld (1978). They applied this technique on the data of Boston Metropolitan Area with 506 observations. We replicate their study using the same data (see **Appendix**). We find that despite high significance of estimation, the model is likely to suffer identification problem. This is because the willingness-to-pay is endogenous and must be estimated instead of being observed. Furthermore, the marginal price is likely to include errors. This is possible because the true hedonic price function (that is, the "inverse demand function") is unknown. The errors in measuring marginal prices are likely to be correlated with the endogenous variables in the price function as well as the income and taste variables. To deal with these problems, researchers suggest the use of discrete-choice hedonic model, as long as the preference function is structured carefully. In this case, it is assumed that consumer chooses **one** good *x* that gives her the highest utility out of all *x*'s in a universal **choice set**, where utility is a function of the good's attributes (see Cropper et al, 1993).

2.2. Conjoint Analysis

Conjoint analysis has been used in marketing field to measure consumer acceptance of products with multi attributes. Such analysis allows the inference of implicit weights for each attributes. This method has also been brought to environmental economics recently (e.g. Lareau and Rae, 1986, Roe et al, 1996). It requires a field survey in which surveyor presents respondents with a number of commodity descriptions, which vary according to the attributes described.

In rank-ordered conjoint analysis, consumers are asked to rank the desirability each commodity. Then, since price of one of the attributes is included in the description, the implicit price can be derived. Since this ranking approach does not resemble the actual behavior of consumers, it is not suggested for environmental valuation; in which the purchase of a single good affected by environmental condition (e.g. house) is of interest. On the other hand, the other variant of conjoint analysis, namely choice-based conjoint analysis asks consumer to choose one single good from a set of alternatives. For example, it analyzes consumer choice of one recreation site from a set of site alternatives (e.g. Adamowics et al, 1997). With an appropriate design of survey form, this method will be able to avoid inconsistency that might otherwise appear in rank-ordered mechanism.

The use of choice-based conjoint analysis can be seen as in example below. Suppose we want to measure the economic effect of cleaning up contaminated sediments in a harbor area. This is of interest since the contamination may affect the local economy in two ways. First, the cost of freight will increase (bigger ship/boats can not reach the harbor; smaller transportation means might be needed). Secondly, the fish both for consumption and for recreational purposes are badly affected. The study might focus on how the redevelopment of the harbor area could benefit its surrounding. In lights of hedonic approach, one way to measure (i.e. "put monetary value" on) environmental improvement is to assess the effects on residential property value in that area[1]. Say, in the survey we include the attributes listed in **Table 1** below[2].

Table 1. Example of Attributes and Levels Included in Conjoint Analysis

Attribute	Levels	Attribute	Levels
Pollution (change in pollution around harbor)	100 % less, 50 % less, 25 % less, no change, 25 % more	Class size (compared to current avg. amount of students per class)	Many less than average, a few less than average, about average, a few more than average, many more than average
House price (compared to current payment)	10 % less, no change, 10 % more, 15 % more, 20 % more	Public areas around harbor (park, beach, etc)	20 % less, no change, 25 % more, 50 % more, 100 % more
House lot size (compared to current situation)	20 % smaller, 10 % smaller, no change, 25 % larger, 50 % larger	Number of retail shops around harbor	No change, 15 more shops, 25 more shops, 50 more shops, 100 more shops
House size (compared to current situation)	20 % smaller, 10 % smaller, no change, 25 % larger, 30 % larger	Boat docking capacity	20 % less, 10 % less, no change, 20 % more, 50 % more
House age	20 % newer, 10 % newer, no change, 10 % older, 20 % older	…	…

[1] In Lancasterian view, a house can be seen as a bundle of characteristics – those it embodies physically (including its price), as well as those of neighborhood (e.g. schooling characterictics and environmental condition).

[2] Attributes and corresponding levels should be based on field observation (e.g. focus group discussion), as well as efficiency consideration (incl. task complexity, ambivalence, protest/fatigue effect, etc). For designing attribute-level mix, see e.g. Kuhfeld et al, 1994. They suggest a general approach for design selection, that is, by maximizing the determinant of Fisher information matrix consisting of the explanatory variables (attributes used).

In order to capture respondents' preference based on the attributes above, they are asked to choose one combination of levels of attributes (one "profile") that they prefer most, out of a set of several combination. **Table 2** below is an example of such choice set. Respondents are asked to pick one of the three options[3]. It is obvious that there can be too many different combinations of attributes and levels. It is tempting to include possible combinations of levels of attributes, since this will ensure no collinearity (which is common in hedonic method). However, it will be very impractical and time consuming. For example the number of all potential combinations (complete factorial) in the above scenario is 5^8 x 3. Fortunately, one can apply fractional factorial design (see Louviere et al, 2000) to pin this down into smaller number, without exposing the data to significant multicollinearity, i.e. to maintain orthogonality[4]. Using factorial software, we can reduce the number of profiles significantly, that is to exclude aliases[5]. Furthermore, it is common in practice to rotate the profiles and attributes order to minimize "order effect" – a bias due to respondent's boredom of choice set layout. We should also consider task complexity and/or ambiguity resulted from the design[6]. That is, we need to assure reasonable task for the respondent, number of alternatives per choice set, number of choice set per survey group, the length of the survey, the number of questions (there are other questions in addition to choice sets), as well as wording, phrasing, and directions. Failure to address this issue may lead to cognitive overload. The data obtained from the survey are then used to estimate the willingness to pay for environmental improvement.

Table 2. Example of Survey Form

	Home A	**Home B**	**Home C**
Pollution	100 % less	No change	No change
Price	10 % more	No change	No change
Lot size	No change	30 % larger	No change
House size	50 % larger	No change	No change
House age	10 % older	20 % newer	No change
Class size	A few more students	A few more students	No change
Public areas	No change	20 % less	No change
Retail shops	15 more shops	25 % more	No change
Boat slips	20 % fewer	10 % fewer	No change

[3] Column C represents a status-quo, to accommodate respondents who are reluctant to changes offered so as to enhance task realism and, in marketing studies, to help estimate market penetration.
[4] For example, only include the main effects and outcast the interaction effects of attributes. Note that in creating orthogonal factorial design, the manipulation could be very cumbersome. Fortunately, there are some statistic software that can do this task (e.g. SPSS Conjoint 11.0)
[5] Applying ORTHOPLAN program in SPSS to our example above gives us 64 different profiles that are orthogonal one another. In practice we may further exclude some unacceptable profiles, or otherwise add some holdout cases to balance the survey. This will sacrifice full orthogonality in favor of efficiency. See e.g. Kuhfeld et al, 1994.
[6] Task complexity and ambiguity of conjoint design can be measured using "confusion index", entropy, ambivalence index, etc. (see e.g. Kordas, 2000).

3. Combining Hedonic and Conjoint Approach

We have briefly discussed that both hedonic method and conjoint analysis can be used to value the economic effect of environmental quality. However, each method has its own strengths and weaknesses (see **Table 3**) These strengths and weaknesses can be attributed to their nature of data used (i.e. revealed preference and stated preference, as described above).

Since both have strengths and weaknesses, it is tempting to use them together. The idea is to combine both approaches as to generate a joint model that enhances the strengths and diminishes the weaknesses. Here, we will use discrete-choice hedonic model to capture the retrospective aspects of data on consumer behavior (esp. we measure the consumer's willingness-to-pay for environmental improvement, by examining her historical behavior: how she selected, for example, housing location that provides best combination of attributes). On the other hand choice-based conjoint analysis is the attempt to mimic this revealed data with the consumer's stated preference (i.e. prospective data). In other words, the consumer is asked to identify her choice from hypothetical set of the good's (e.g. housing) characteristics.

Table 3. Discrete-Choice Hedonic Model and Choice-Based Conjoint Analysis

	RP: Hedonic Approach	**SP: Conjoint Analysis**
Strengths	- Based on observed behavior. - High reliability and face validity.	- Clear identification of the parameters to consider. - Orthogonal attribute data avoid collinearity. - Adequate number of observation for all attributes including uncommon ones. - Prespecification of alternatives within each choice sets.
Weaknesses	- Errors from omission of variables. - Collinearity between explanatory variables. - Inability to capture effectively preferences for uncommon attributes. - May involve arbitrary specification of feasible choice sets considered by consumer.	Hypothetical nature of the questions and choices

The combination of both approaches will provide an econometric model with more robust estimates and better identification of attributes, welfare measurements which are less prone to information and hypothetical biases, and more realistic measures of price responsiveness, since they are drawn from actual financial settings (Earnhart, 2001). As for our example of conjoint structure above, we can create a data structure for hedonic model consisting of the same attributes, as well as some uncommon attributes. In addition, individual demographic characteristics data are also needed (for conjoint analysis, this is usually conducted through supplemental telephone survey). However, the issues of what model to use as well as econometric consequences of combining the data become crucial. In the subsequent subsections we address these aspects respectively.

3.1. Random Utility Model and Multinomial Logit Model

The RUM is a model that decomposes overall utility into two components: deterministic- and random component. The latter may due to unobserved attributes, unobserved taste variations, measurement errors and imperfect information, and instrumental variables (Ben-Akiva and Lerman, 1985).

Suppose a representative agent[7] chooses one alternative from a set $C=\{1,\ldots,J\}$:

U_j^* = utility of j-th choice.

$$U_j = 1 \text{ if } U_j^* = \max\{U_1^*,...,U_J^*\} \quad (4)$$

= 0 otherwise.

We can decompose the utility into two parts:

$$U_j^* = V(s, z_j) + e_j \quad (5)$$

where V is the observed utility function ("subutility function"), $s \in S$ is socioeconomic characteristics of the agent, $z_j \in Z_j$ is array of observed attributes of alternative j, and e is unobserved attributes of the utility (i.e., an unobservable random term). If e_j is IID-EV1 (extreme value type one), that is:

$$F(e_j) = \exp(-\exp(-l(e_j - h)))$$

$$f(e_j) = l \exp(-l(e_j - h)) \cdot \exp(-\exp(-l(e_j - h))) \quad (6)$$

where l is a positive scale factor and h is location parameter, then the probability of alternative j being chosen, assuming h is zero, is:

$$\Pr_j = \Pr(V_j + e_j \geq V_i + e_i; \forall i \in C) \quad \text{where } C \text{ is the choice set.} \quad (7)$$

This is equivalent to (see McFadden, 1974):

$$\Pr_j \equiv \Pr(j \mid s, z_j) \equiv \Pr(U_j = 1 \mid s, z_j) = \frac{\exp(lV_j)}{\sum_{j \in C} \exp(lV_{j'})} \quad (8)$$

where l is scale factor.

---

[7] Here, we suppress difference in individuals. Later in the application section, we allow for the individuals to vary.

In most cases dealing with a single data set, the scale factor l is usually assumed equal to one (e.g. Cropper et al, 1993). However, in the present context where we are going to combine two data sets, the role of the scale parameter becomes crucial. Even though we cannot identify them separately for each data set, we fortunately can identify their ratio. We will discuss this later.

3.2. Independence from Irrelevant Alternatives

Before addressing the role of scale parameter, we need to consider the important property associated with using probabilistic discrete choice or choice-based models. This is the independence from irrelevant alternatives (IIA) property. This property says that the ratio of the probabilities of choosing one alternative over another is unaffected by the presence or absence of any additional alternatives in the choice set. That is,

$$\frac{P(i \mid \tilde{C}_n)}{P(j \mid \tilde{C}_n)} = \frac{P(i \mid C_n)}{P(j \mid C_n)} \qquad i, j \notin \tilde{C}_n \subset C \tag{9}$$

where C is choice set.

However, this may not be true in certain circumstances. A famous counterexample is the "Red Bus – Blue Bus" case coined by Luce (see Ben-Akiva and Lerman, 1985). Suppose that we have two modes of transportation, namely car and bus. Suppose further that a consumer's preferences to these means of transportation are equal. That is, Pr(car) = ½ and Pr(bus) = ½. Next, if another bus with different characteristic (e.g. the old bus is red, the additional bus is blue) is added, in order for IIA property to hold, the probability of choosing car, red bus, and blue bus are each 1/3. That is, Pr(car) = 1/3, Pr(red bus) = 1/3, and Pr(blue bus) = 1/3. However this seems unrealistic since the consumer is most likely to treat the two buses as a single alternative, hence Pr(car) = ½, Pr(red bus) = ¼, and Pr(blue bus) = ¼.

In regards with the "Red Bus- Blue Bus" violation, one is suggested to perform test for the IIA property. Fortunately, the random utility-multinomial logit model, thanks to IID-EV1 assumption, has the property of IIA. However, not all data set follows IID-EV1. This leads to departure from multinomial logit model, in order to deal with the issue of IIA violation. (For example, one can apply nested logit model (e.g. Chattopadhay, 2000) that is less restrictive and uses McFadden's generalized extreme value, rather than EV1). Fortunately, models that include socioeconomic attributes in an appropriate fashion may generate reasonable estimates since the deterministic component in the utility function should account for population heterogeneity (Earnhart, 2001; Ben-Akiva and Lerman, 1985). In addition, a good orthogonal design in conjoint analysis is likely to eliminate this potential problem.

Test for IIA property

The test for IIA property is conducted on each data set, or between data set. The idea is to test whether the estimated parameter does not change when a variation in the choice set is introduced. There are various types of test for this property. Here we use the Hausman-McFadden test (Hausman and McFadden, 1984; Louviere et al, 2000). First, we need to estimate the choice model with all alternatives. Then the specification under the alternative hypothesis (IIA violated), i.e. the model with a smaller set of choices, is estimated with a restricted set of alternatives and the same attributes. Finally, the test statistic is:

$$y = [\hat{b}_r - \hat{b}_u]'[V_u - V_r]^{-1}[\hat{b}_r - \hat{b}_u] \approx c^2 \qquad (10)$$

where subscript u and r denote unrestricted and restricted set (or, in line with our interest, can be applied to test the IIA property when combining hedonic data set and conjoint data set); degrees of freedom equal to the rank of $[V_u - V_r]$. It is important to note that, when calculating this test we have to be careful in making the choice sets compatible in order to get a non-singular matrix of $[V_u - V_r]$.[8]

The above test, however, is mainly for checking the IIA property. As mentioned in McFadden (1984), this test may fail because of misspecification rather than IIA. Therefore, it is often suggested to perform other tests (Wills, 1987). Such tests include the classical LR (likelihood ratio) test, Wald test, and Rao's Score test. For example, Score (LM) test is given by:

$$z = \left[\frac{\partial L(\hat{b}_r)}{\partial b}\right]' [V_u + V_r] \left[\frac{\partial L(\hat{b}_r)}{\partial b}\right] \approx c_k^2 \qquad (11)$$

where L is the log-likelihood

4. Combining Data

Suppose we have two data sets generated from hedonic approach and conjoint design, respectively. Since both assume random utility maximization, we then have two utility functions (see Louviere et al, 2000) as follow[9]:

---

[8] This is Hausman-McFadden test. Other tests include McFadden-Train-Tye test, Small-Hsiao test, and Horowitz test. Fry and Harris (1996) use a Monte Carlo simulation study to investigate the size and power properties of these tests. They find that majority of these tests have poor size and power properties in small samples. They then suggested another test based upon Dogit model (Gaudry, 1980). We find that Dogit model is similar to nested logit model (a generalization of McFadden's multinomial logit), which can be tested using e.g. score test (see below).

[9] We suppress the superscript (*) on *U*, for simpler notation.

$$U_j^h = V_j^h + e_j^h \qquad \forall j \in C^h \text{ and } U_j^c = V_j^c + e_j^c \qquad \forall j \in C^c \tag{12}$$

where the deterministic components V's can be written as

$$V_j^h = a_j^h + b^h X_j^h + wZ_j \text{ and } \quad V_j^c = a_j^c + b^c X_j^c + dW_j \tag{13}$$

where h indicates hedonic data and c conjoint data, a is alternative-specific constant, X is attributes common to hedonic- and conjoint sets, Z is attributes unique to hedonic set, and W is attributes unique to conjoint set.

Again, assuming e_j^h and e_j^c are IID-EV1, we have

$$\Pr_j^h = \frac{\exp(l^h V_j^h)}{\sum_{j \in C^h} \exp(l^h V_{j'}^h)} \text{ and } \Pr_j^c = \frac{\exp(l^c V_j^c)}{\sum_{j' \in C^c} \exp(l^c V_{j'}^c)} \tag{14}$$

We are interested to combine the two data sets in order to get better estimation. However, we can not compare the parameter estimates directly, since they are linked with the scale factors (in both models we can see that the V's are multiplied by l 's). By the same reason, we can not identify the scale factors individually. As a conclusion, we can not determine whether the observed difference is due to scale factor, true parameter, or both. We will address these issues respectively.

In order to combine the hedonic data and conjoint data, we have to make sure that model parameters are equal (while controlling for scale differences). Following Swait and Louviere (1993), we use the following test procedure.

We estimate

$$\Pr_j^h = \frac{\exp(l^h V_j^h)}{\sum_{j \in C^h} \exp(l^h V_{j'}^h)} \text{ and } \Pr_j^c = \frac{\exp(l^c V_j^c)}{\sum_{j' \in C^c} \exp(l^c V_{j'}^c)}$$

individually and obtain the estimates of $l^c a^c, l^c b^c, l^c w$ with log likelihood L^c from the first estimation and the estimates of $l^h a^h, l^h b^h, l^h w$ with log likelihood L^h from the second one.

Next, we estimate the two models simultaneously while imposing $b^h = b^c = b$ and normalizing $l^h = 1$, and obtain the estimates of a^h, a^c, b, w, d with log likelihood L^{hc}. Finally, we calculate test statistic $y = -2[(L^h + L^c) - L^{hc}] \approx c^{2}_{k-1}$. If our test result is in the acceptance region, then it implies that the hedonic- and conjoint data contain similar preference structures.

4.1. The Role of Scale Factor

Scale factor or precision parameter is an inverse function of the standard deviation of the unobserved effects s. This factor can vary between data sets, between alternatives or between time periods (Louviere et al, 2001). For EV1 distribution, we have

$$s^{2} = \frac{p^{2}}{6l^{2}} \quad \text{(for proof, see Ben-Akiva and Lerman, 1985).}$$

Therefore, the smaller the variance, the bigger the scale factor. This indicates that model parameters and the characteristics of the error terms are closely related. That is, *parameters estimated from two identical specifications using different data sources with different variances will likely be different in magnitude, even if the true parameters are identical.* The knowledge of scale factors helps us determine if we need to rescale a data set to be comparable with another. If we want to combine hedonic data set and conjoint data set and estimate the parameters jointly, we have to control for these scale factors. However, since we cannot identify each one of them (separately), we should normalize either one. If we, following usual practice, set the scale parameter of hedonic data equal to 1, than we have $s^{2}_{h} = \frac{p^{2}}{6}$. Thus, we are left to estimate $a^{h}, a^{c}, b, w, d, l^{c}$, where now l^{c} is the conjoint data's scale relative to hedonic data scale.

4.2. Estimation

The log likelihood of the pooled of hedonic data and conjoint data is:

$$\begin{aligned} l(a^{h}, a^{c}, b, w, d, l^{c}) = &\sum_{n} \sum_{\Pr_j \in C_n^h} U_{jn} \ln \Pr^{h}_{jn}(X^{h}_{jn}, Z_{jn} \mid a^{h}, b, w \\ &+ \sum_{n} \sum_{\Pr_j \in C_n^c} U_{jn} \ln \Pr^{c}_{jn}(X^{c}_{jn}, W_{jn} \mid a^{c}, b, s, l^{c}) \end{aligned} \qquad (15)$$

where $U_{jn} = 1$ if n chooses j.

$= 0$ otherwise.

McFadden (1984) shows under general conditions that maximum likelihood estimation of the multinomial logit provides estimators that are asymptotically efficient and normally distributed. A simple way to estimate the parameters in the joint log likelihood above is as follows:

Define the range of l^c.

1. For each l^c, create a pooled data matrix:

$$Q(l^c) = \begin{bmatrix} I^h & 0 & X^h & Z & 0 \\ 0 & l^c I^c & l^c X^c & 0 & l^c W \end{bmatrix}$$

2. Define joint log likelihood:

$$y\,(a^h, a^c, b, w, d \mid l^c) = \sum_n \sum_{j \in C_n} y_{jn} \ln \Pr_j [Q(l^c)]$$

3. Obtain the estimates of a^h, a^c, b, w, l^c from model solution that gives the maximum value of the log likelihood.

4.3. Taste Heterogeneity

In addition to scale heterogeneity discussed above, multinomial response models should also deal with taste heterogeneity. This refers to differences across decision-makers in their intrinsic preference for choice alternatives; that are unobserved by analyst. One way to address this is to apply unified mixed logit model proposed by Bhat and Castellar (2002). The model is "unified" in the sense that it addresses not only taste heterogeneity, but also scale difference, and even choice set characteristics. They assume a linear specification and use error decomposition as follows:

$$U_{njt} = a_n^{'} z_{njt} + q_n^{'} \left[(1 - d_{nt,h}) \times \left(\sum_{s=1}^{Tn} d_{ns,h} D_{njs} \right) \right] + e_{njt} \quad \text{and} \quad e_{njt} = e_{njt}^0 + m' z_{njt}^0 \quad (16)$$

where U_{njt} is individual's j utility from choosing choice j out from choice set t. Dummy d equals 1 when t-th choice set in stated preference model (i.e. the conjoint model) corresponds to her revealed preference choice (i.e. the hedonic model)[10]; and zero otherwise. Dummy D equals 1 when choice j is picked up by individual j given choice set t. The errors compose of two terms. The first one is assumed to follow IID-EV1 across individuals and choices, and also independently (but not identically) distributed across choice sets. It has a scale parameter of $l_{nt} = [(1 - d_{nt,h}) \times l] + d_{nt,h}$. The factor z^0 in the second term refers to a vector of unobserved utility components of the choices in any sets. Thus, this term leads to

[10] The subscript h is therefore due to "hedonic".

heteroscedasticity and correlation across observed data. Clearly this specification accommodates the scale difference between hedonic and conjoint data. Also it allows for taste heterogeneity across individuals, as well as differences in perception towards choice sets. One drawback of this model, however, is its highly parametrical assumptions.

5. Implementation

In line with our previous example on hedonic pricing model for valuing environmental change, we can apply the combined discrete choice hedonic approach and choice-based conjoint analysis for the same case. We can write the utility function as:

$$U^*_{jn} = V_{jn}(Z_j, Q_j, X_n, C_n; b) + e_{jn} \qquad (17)$$

where Z is a vector of observed housing characteristics, Q is an index of environmental quality, X is composite commodity, C is individual characteristics, and b is the vector of parameters to estimate. Subscripts j and n refer to housing and individual, respectively. If we assume that the errors are IID Type I-Extreme Value, that IIA property holds, and hedonic- and conjoint data are independent, then we can apply the estimation procedure explained above.

Analogues to the calculation of the total willingness-to-pay in section 2.1, for this discrete model, it is more convenient to calculate the compensating variation as a measure of welfare (see McConnell, 1995). This is formalized as:

$$CV = -\frac{1}{j}[\ln(\sum_{j\in K}\exp(V_{jn1})) - \ln(\sum_{j\in K}\exp(V_{jn0})] \qquad (18)$$

where j is the coefficient of housing price term ("marginal utility of income") and V_{jn1}, V_{jn0} are subsequent- and initial observed utility level, respectively.

6. Nonparametric Approach

There has been growing literature on drawbacks of using original McFadden's multinomial logit model. In particular, studies have shown that the model is very restrictive, since it requires parametric assumption on the observed utility function and predetermined distribution of the error term.

Manski (1985) argues that the assumptions used in most parametric (binary-) response model, namely mean independence and no perfect collinearity among regressors are not adequate to identify the parameter vector. In addition, Koenker (2000) shows that mean independence is not an appropriate approach for choice models (instead, he suggested median independence, and in general, quantile approach). Manski suggested a maximum score estimation technique that only assumes utility function linear in a finite dimensional parameter, while probability of choice *j* being chosen is generated from random terms that are IID (identically and independently distributed) conditional on (*s, z*). Kordas (2000) synthesizes Manski's maximum score estimation and Koenker's quantile approach and comes up with a *binary* regression quantile. Using a linear binary specification, Kordas proposed a maximum score estimator as follows:

$$\hat{b}(t) = \arg\max_{b} J^{-1} \sum_{j=1}^{J} [U_j - (1-t)] \mathrm{I}(x_j' b \geq 0) \qquad (19)$$

where $\hat{b}(t)$ is the estimated b at the t -th quantile, x_j represents (s, z) for *j*-th choice, and I is an indicator function – 1 if the inequality holds, 0 otherwise. The corresponding binary quantile regression is:

$$\hat{b}(t) = \arg\min_{b} J^{-1} \sum_{j=1}^{J} r_t (U_j - \mathrm{I}(x_j' b \geq 0) \qquad (20)$$

where $r_t(v) \equiv [t - \mathrm{I}(v < 0)] \cdot v$ is Koenker-Bassett check function[11].

Even though the above approaches are less restricted than the logit model, it still needs to assume a parametric function on the observed utility. Matzkin (1993) offers a fully nonparametric approach that does not require parametric specification of both utility function and the random terms. Rewrite (7):

$$\Pr(j \mid s, z) = \Pr(e_i - e_j \leq V_j(s, z_j) - V_i(s, z_i) \quad \forall i \in C,\ i \neq j \qquad (21)$$

the cumulative distribution of *e* conditional on (*s, z*) is therefore:

$$\Pr(j \mid s, z) = P_j(V^{(j)}(s, j) \mid (s, z) \equiv \mathrm{P} \qquad (22)$$

where $V^{(j)}(s, z_j)$ is the (*J*-1)-dimensional vector $\lfloor V_j(s, z_j) - V_i(s, z_i) \rfloor \ \forall i \in C,\ i \neq j$. Unlike Manski, Matzkin's approach only requires monotonicity, concavity, and homogenous of degree one of the utility function – conditions of which are commonly implied by economic theory[12].

Matzkin claimed that when the errors are only assumed to be IID (identically and independently distributed) conditional on (*s, z*), on can nonparametrically identify *V*.

[11] Horowitz (1992) offers smoothed version (i.e. continuous and differentiable) of such estimators that converges more rapidly and has a tractable asymptotic distribution theory.

[12] If both *V* and P are known, the model becomes parametric. If either *V* or P is known, the model is semiparametric, and if both *V* and P are unknown, the model is nonparametric.

Adapting Matzkin, let **V** and **P** be the set of all V and P, respectively. Let the probability measure of (s, z) be ***G***. Then the choice probabilities from a pair (V, P) $\in$ (**V** x **P**) is $\Pr(j \mid (s,z);(V,\mathrm{P})) \;\; \forall j = 1,...,J$.

Thus, $\Pr(j\}(s,z):(V,\mathrm{P})) = P_j(V^{(j)}(s,z) \mid (s,z))$.

Following Matzkin's definition for identification, a function V^* is identified in the set **V** when P* $\in$ **P** if for any $V \in$ **V** such that $V \neq V^*$ and all P$\in$ **P** there exists a set $D \subset (S \times Z)$ and an alternative $j \in C$ such that ***G***(D) > 0 and for all $(s,z) \in D$. That is,

$$\Pr(j \mid (s,z);(V,\mathrm{P})) \neq \Pr(j \mid (s,z);(V^*,\mathrm{P}^*)) \tag{23}$$

Definition (20) can also be seen as the following: a pair (V^*, P*) is identified in the set (**V**, **P**) if for all (V, P)$\in$ (**V**, **P**) such that $(V,\mathrm{P}) \neq (V^*,\mathrm{P}^*)$ there exists a set $D \subset (S \times Z)$ and an alternative $j \in C$ such that ***G***(D) > 0 and for all $(s,z) \in D$.

Furthermore, Matzkin proved that when J coordinate functions (of each function in **V**) are identical continuous function, the identification of V^* in **V** only requires that no two functions in V are monotone transformations of each other, conditional on s.

In order to define the maximum score estimation, suppose we have $w = (y, s, z)$ where $y = (y_1,...,y_J)$ and $y_j = 1$ if choice j is chosen and 0 otherwise. Next, let $y^* = (y_1^*,...,y_J^*)$ such that $y_j^* = V^{(j)}(s,z) - e^{(j)}$ where $e^{(j)}$ is a vector $[e_i - e_j] \; \forall i \in C$, $i \neq j$. Finally, we define a function f such that:

$$f(w,V) = \begin{cases} 1 \text{ if } \; y_j > 0 \text{ and } V^{(j)}(s,z) > 0 \\ \quad 0 \qquad\qquad \text{otherwise} \end{cases} \tag{24}$$

Then, the maximum score estimator is:

$$\hat{V}_N = \arg\max_{V \in \mathbf{V}} J^{-1} \sum_{j=1}^{J} f(w^j, V) \tag{25}$$

Matzkin proved that when identification definition (see above) holds and **V** is concave and compact, the maximum score estimator $\hat{V}$ is strongly consistent.

7. Future Direction

One possible flaw if multinomial logit model is used is related to the IIA property. In most empirical studies, the violation of IIA can be taken care by redesign the data orthogonality in conjoint analysis and/or by adding more related socioeconomic attributes so as to maintain the pattern of the main effects. However, this seems weak, theoretically. Some studies have addressed this issue by using models that are less restrictive then multinomial logit model, to allow for deviation from IIA property. This includes the use of nested logit model, mixed logit model, heteroscedasticity extreme value model, etc. Unfortunately so far these developments have not come into combining data sources. Instead they are applied to single data source. In line with our discussion, the main difficulty in using the more advanced models while attempting to combine data sources is how to deal with the scale parameter. Hensher and Greene (2002) have started to treat the scale parameter in more complex, nested logit model. However, their concern of scale parameter is motivated by the difference of data structure between upper nest and lower nest in the nested logit tree. If we want to combine two sources of data, things become more complicated, because we have to take into account of difference between parallel nests.

Maximum score estimation adapted as fully nonparametric approach seems very promising. It encompasses the above problem about IIA and moreover it answers skepticism on the strongly restricted multinomial logit model. However, there has been no application so far of this theoretical model. Also, it still needs to be developed in order to allow for data combination from different sources, such as hedonic and conjoint process.

Another issue is related to potential misclassification in the response variable. It is very possible that when we are trying to capture consumer's preference, we classify the data inappropriately. This off course will lead to a misleading result. A very recent attempt to address this issue is Ramalho (2002). She uses GMM technique to estimate her model and outlines a specification test to detect misclassification. However, her test is due to choice-based sample case. That is, sample which is obtained from endogenous stratification according to individual responses. In our study for environmental valuation, often times we have to deal with exogenous random sampling, in which a sequence of decision makers are drawn and their choice behaviors are observed. It is interesting to extend Ramalho's GMM model to the case of random sampling.

Reference

Adamowicz, W.,J. Louviere, M. Williams, 1994. Combining Revealed and Stated Prefe rence Methods for Valuing Environmental Amenities. *J. Environmental Economics and Management*. 26:271-92.

Adamowicz, W., J. Swait, P. Boxall, J. Louviere, M. Williams, 1997. Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *J. Environmental Economics and Management*. 32:65-84.

Ben-Akiva, M., S.R. Lerman, 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.

Bhat, C.R. and S. Castelar, 2002. A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area. *Transportation Research B*. 36(7):539-616.

Chattopadhyay, S., 2000. The Effectiveness of McFadden's Nested Logit Model in Valuing Amenity Improvement. *Regional Science and Urban Economics*. 30:23-43.

Cropper, M.L. L. Deck, N. Kishor, K.E. McConnell. 1993. Valuing Product Attributes Using Single Market Data: A Comparison of Hedonic and Discrete Choice Approaches. *Review of Economics and Statistics*. 75(2):225-32.

Earnhart, D., 2001. Combining Revealed and Stated Preference Methods to Value Environmental Amenities at Residential Locations. *Land Economics*. 77(1):12-29.

Fry, T.R.L. and M.N. Harris, 1996. A Monte Carlo Study of Tests for the Independence of Irrelevant Alternatives Property. *Transportation Research B*. 30(1):19-30.

Gaudry, M.J.I., 1980. Dogit and Logit Models for Travel Mode Choice in Montreal. *Canadian J. Economics*. 13:268-79.

Green, P.E., V. Srinivasan, 1990. Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *J. Marketing*. 54:3-19.

Griliches, Z. 1971. Introduction: Hedonic Prices Revisited. in Z. Griliches (ed), *Price Indexes and Quality Change*. Harvard University Press.

Harrison, D., D.L. Rubinfeld (1978), Hedonic Prices and the Demand for Clean Air. *J. Environmental Economics and Management*, 5:81-102.

Hausman, J., D. McFadden. 1984. Specification Tests for the Multinomial Logit Model. *Econometrica*. 52(5):1219-40.

Hensher, D.A, W.H. Greene. 2002. Specification and Estimation of the Nested Logit Model: Alternative Normalizations. *Transportation Research Part B*. 36:1-17.

Horowitz, J.L., 1992. A Smoothed Maximum Score Estimator for the Binary Response Model. *Econometrica*. 60(3):505-31.

Koenker, R., 2001. *Applied Econometrics*. Unpublished Lecture Note (Econ 472). Univ. of Illinois at Urbana-Champaign.

Koenker, R., 2002. *Econometrics I*. Unpublished Lecture Note (Econ 476). Univ. of Illinois at Urbana-Champaign.

Kordas, G. 2000. *Binary Regression Quantiles*. Unpublished Doctoral Dissertation. Univ. of Illinois at Urbana-Champaign.

Kuhfeld, W.E., R.D. Tobias, and M. Garrett, 1994. Efficient Experimental Design with Marketing Research Applications. *J. Marketing Research*, 31:545-57.

Lareau, T.J., D.A. Rae, 1986. Valuing WTP for Diesel Odor Reductions : An Application of Contingent Ranking Technique. *Southern Economic J*. 728-42.

Louviere, J.J., D.A. Hensher, J.D. Swait, 2000. *Stated Choice Methods: Analysis and Application*. Cambridge Univ. Press.

Maddala, G.S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge Univ. Press.

McConnell, K.E., 1995. Consumer Surplus from Discrete Choice Models. *J. Environmental Economics and Management*. 29:263-70.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. in Zarembka, P. (ed), *Frontiers of Econometrics*. New York: Academic Press.

McFadden, D. 1978. Modelling the Choice of Residential Location. in A. Karlqvist, L. Lundqvist, F. Snickars, J. Weibull (eds), *Spatial Interaction Theory and Planning Models*. 75-96. North Holland: Amsterdam.

McFadden, D. 1984. Econometric Analysis of Qualitative Response Model. in Z. Griliches and M.D. Intriligator (eds), *Handbook of Econometrics*. Vol II. Elsevier Scince Publishers BV.

Ramalho, E.A. 2002. Regression Models for Choice-Based Samples with Misclasification in the Response Variable. *J. Econometrics* 106:171-201.

Roe, B., K.J. Boyle, M.F. Teisl, 1996. Using Conjoint Analysis to Derive Estimates of Compensating Variation. *J. Environmental Economics and Management*. 31:145-59.

Rosen, S. 1974. Hedonic Price and Implicit Markets: Product Differentiation in Pure Competition. *J. Political Economy*. 82:34-55.

Swait, J., J. Louviere, 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *J. Marketing Research*. XXX(Aug): 305-14.

Wills, H. 1987. Specification Tests for the Logit Model. *J. Econometrics*. 34:263-74.

Appendix: An Example of Traditional Hedonic Analysis

Using their data, we replicate their study as follows[13]. Suppose the inverse demand function is:

$$\ln MEDV = a + b_1 RM^2 + b_2 AGE + b_3 \ln DIS + b_4 \ln RAD + b_5 TAX + b_6 PTRATIO$$
$$+ b_7 B + b_8 \ln LSTAT + b_9 CRIM + b_{10} ZN + b_{11} INDUS + b_{12} CHAS + b_{13} NOX^r + e$$

where[14] MEDV is median value of owner-occupied homes in dollar, RM is the average number of rooms per dwelling, AGE is proportion of owner-occupied unit built prior to 1940, DIS is weighted distances to five Boston employment centers, and RAD is index of accessibility to radial highways. TAX is full-value property-tax rate per \$10,000, PTRATIO is pupil-teacher ration by town, B is $1000(B_k - 0.63)^2$ where B_k is the proportion of blacks in town. LSTAT is percentage of lower status of the population, CRIM is per capita crime rate by town, ZN is proportion of residential land zoned for lots over 25,000 sq.ft. INDUS is proportion of non-retail business acres per town, CHAS is Charles River dummy variable (1 if tract bounds river; 0 otherwise), and NOX is nitric oxide concentration (parts per 100 million).

Using OLS estimation, we obtain:

Table A1. Estimation of Inverse Demand Function

lnmedv	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rmsq	.0063285	.0013123	4.82	0.000	.0037501	.0089068
age	.0000907	.0005263	0.17	0.863	-.0009433	.0011248
lndis	-.1912552	.0333924	-5.73	0.000	-.2568646	-.1256458
lnrad	.0957106	.0191342	5.00	0.000	.0581158	.1333054
tax	-.0004203	.0001227	-3.43	0.001	-.0006614	-.0001793
ptratio	-.0311224	.0050132	-6.21	0.000	-.0409724	-.0212725
b	.0003637	.0001031	3.53	0.000	.0001611	.0005663
lnlstat	-.3711573	.0250095	-14.84	0.000	-.420296	-.3220186
crim	-.0118644	.0012447	-9.53	0.000	-.01431	-.0094189
zn	.0000802	.0005056	0.16	0.874	-.0009133	.0010736
indus	.0002395	.0023636	0.10	0.919	-.0044045	.0048836
chas	.0913952	.0332023	2.75	0.006	.0261594	.1566309
noxsq	-.0063805	.0011314	-5.64	0.000	-.0086035	-.0041575
_cons	11.46553	.1544401	74.24	0.000	11.16209	11.76897

Adj. R-squared: 0.8008   F(13,492): 157.13

We can see that our variable of interest, *NOX* is significantly different from zero with the expected sign[15]. Next, we calculate the marginal willingness-to-pay according to:

$$W = \frac{\partial MEDV}{\partial(-NOX)} = 2 * (0.0064) * NOX * MEDV$$

The estimated value of W is then used as the dependent variable on the marginal willingness-to-pay function:

$$\ln W = a_0 + a \ \ln NOX + a_2 \ln INC + e \ ,$$

[13] For data, see http://lib.stat.cmu.edu/datasets/boston.

[14] In brackets are expected signs.

[15] For the value of r, Harrison and Rubinfeld performed a grid search and found that r = 2 fit best. We follow this conclusion.

estimation of which yields:

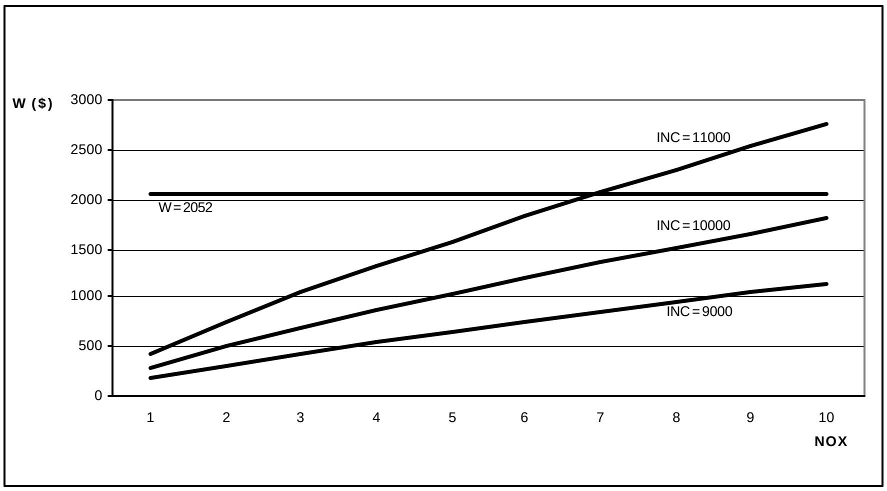
Table A2. Estimation of Marginal WTP Function

lnw	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lnnox	.8061104	.0273939	29.43	0.000	.7522898	.8599309
lninc	4.429781	.0653331	67.80	0.000	4.301421	4.55814
_cons	-35.15628	.6282828	-55.96	0.000	-36.39066	-33.9219

Adj. R-squared: 0.9014, F(2, 503): 2300.36

We can graph the relationship between the reduction in nitrogen oxide concentration and the marginal willingness-to-pay for some income level. This is shown in the following figure:

Figure A1. Marginal Willingness-to-Pay for NOX Redux



The graph suggests that the NOX concentration increasingly affects the marginal willingness-to-pay. In addition, the premium that high- income households are prepared to pay rises as the NOX level increases (although in a slightly decreasing rate).

Finally we can calculate the total willingness-to-pay as follows:

$$TW = \sum \int_{NOX1}^{NOX0} \exp(\hat{a}_0 + \hat{a}_1 \ln NOX + \hat{a}_2 \ln INC) dNOX = 99{,}187.87$$