Analytical Problems and Suggestions in the Analysis of Behavioral Economic Demand Curves

Short title: Suggestions on behavioral economic demand curve analysis

Jihnhee Yu*[a], Liu Liu[a], R. Lorraine Collins[b], Paula C. Vincent[b], Leonard H. Epstein[c]

[a] Department of Biostatistics, University at Buffalo, State University of New York, NY 14214, USA.

[b] Department of Community Health and Health Behavior, University at Buffalo, State University of New York, NY 14214, USA.

[c] Departments of Pediatrics and Social and Preventive Medicine, University at Buffalo, State University of New York, NY 14214, USA.

*Correspondence to: Jihnhee Yu, Department of Biostatistics, University at Buffalo, State University of New York, NY 14214, USA.

E-mail: jinheeyu@buffalo.edu

Abstract

Behavioral economic demand curves are innovative and well-thought approaches in psychosocial research to characterize the relationships between consumption of a substance and its price, shedding a light on substance addiction. Based on the linearity between price and elasticity, they are commonly established using a linear model through log-transformed values. Since first introduced (Hursh et. al, 1988), they have become popular tools to study reinforcing efficacy of substances; however, analytical techniques used in data analyses have not been changed much.

In this article, we particularly address following analytic issues that can influence interpretation of demand curves. First, we demonstrate that log transformation with different added values for handling zeros alters the analytical results substantially and thus dramatically changes derived values such as elasticity, $P_{\max}$, and $O_{\max}$. In addition, demand curves are often analyzed using an over-parameterized model, i.e., individual fitting is performed for each participant, giving rise to a lack of assessment of the variability among individuals.

To provide analytical suggestions in the analysis of behavioral economic demand curves, we apply a nonlinear mixed effects model, and we propose analytical formulas for the relevant standard errors of derived values such as $P_{\max}$, $O_{\max}$, and elasticity, which are not available in the current literature. We demonstrate that the proposed model greatly stabilizes the derived values regardless of using different added increments and provides substantially smaller standard errors and better inference than the over-parameterized model. We illustrate the analytical procedure using data from a relative reinforcement efficacy study of marijuana purchasing.

Keywords: Elasticity; Pmax; Purchasing task study; Omax; Variance estimation

1. Introduction

Behavioral economic demand curves describe the changing behavior of consumption of a drug, addictive substance, or other consumable commodity as a function of the cost represented by price [1-2]. They are expressed as a decreasing slope of the demand with increasing price to assess the reinforcing value of commodities. Hursh et al. [3] first proposed the linear elasticity equation for behavioral economic demand curves. It has been widely used for studies ranging from the purchase of single substance (e.g., [4]) to consumption of concurrently available drugs to compare relative reinforcements (e.g., [5]); however, analytical techniques used in data analyses have not been changed much. This article mainly focuses on the behavioral economic demand curve analysis in the context of a purchasing task study that evaluates the reinforcing efficacy of a study substance. Reinforcing efficacy of a substance is commonly assessed by the degree of efforts that individuals make to gain access to the substance [2].

In a purchasing task study [2, 4], typically, subjective demand (consumption) of a substance is observed at different hypothetical price points for each individual. Suppose that there are n participants in a study and consumption of substances is observed at k different price points. The linear elasticity equation [3] to describe the behavioral economic demand is commonly presented in the logarithmic coordinates as

$$\log C_{ij} = \log L + b\log(p_j) - ap_j + \varepsilon_{ij},\ i = 1,...,n,\ j = 1,...,k, \tag{1}$$

where p_j are j-th price and C_{ij} are corresponding amount of consumption for the i-th participant. In a typical linear regression, the residuals ε_{ij} are considered to have the independent and identical normal distribution with a constant variance, σ^2. Equation (1) is derived based on the first order linear relationship between elasticity and price [3], where elasticity is defined as $\partial \log C_{ij} / \partial \log p_j$. The parameter L is the amount of consumption at unit price and is referred to as derived-intensity [4]. The parameter a is the major parameter to derive the decreasing consumption when price increases. The parameter b is typically a small negative value and controls initial dip of the model, that is, the greater absolute value of b gives rise to the steeper decrease of initial consumption. The expenditure is defined as $C_{ij} p_j$. Price values p_j are considered to be fixed. In laboratory animal studies, price can be expressed as a function of both the response requirement and decreasing reward of the studied substance. For example, an increasing effort to obtain a study substance (e.g., pressing a lever) and a decreasing dose level of that substance for each effort give rise to an increase in price [6].

Important quantities that can be derived from Model (1) include $P_{\max}$: the price to achieve the maximum expenditure, $O_{\max}$: the maximum expenditure, and e_p:elasticity at price p as defined by

$$P_{\max} = (1+b)/a, \tag{2}$$

$$O_{\max} = L P_{\max}{}^{1+b} e^{-(1+b)}, \tag{3}$$

and

$$\mathrm{e}_p = b - ap. \tag{4}$$

Often, elasticity is summarized in the average price. Let e indicate the average elasticity, i.e., $\mathrm{e} = b - a\overline{p}$ where $\overline{p} = \sum_{j=1}^{k} p_j / k$. Elasticity (4) becomes -1 at $P_{\max}$ and the demand moves from inelastic (> -1) to elastic (< -1) when a price value passes $P_{\max}$. Thus, a large value of $P_{\max}$ shows that the elastic demand of a substance is achieved only at a high price, indicating strong reinforcement of a substance of interest [6].

Although fitting Model (1) seems straightforward, there are a few analytical issues to be considered in common data analysis practice. The first issue is in taking the logarithm of data points. There are a few different ways of carrying out the log transformation to handle zero values in consumption and price, and we observe that the estimated parameter values based on Model (1) change dramatically by different log transformation strategies, resulting in unreliable parameter estimations. No literature specifically addresses the potential implication of using different log transformation strategies. In this article, we will provide a detailed account of the effects of the different log transformation strategies on model fitting. The second issue is the usage of over-parameterized models where they are fitted for each participant, giving rise to inefficient inferences. The third issue is the lack of consideration of a clustering effect in multiple observations from the same individuals.

To address these issues, in this article, we propose to use a nonlinear mixed effects model [7-8] to fit behavioral economic demand curves as an alternative to conventional data analysis practices. This article has the following structure. In Section 2, we will illustrate the analytical problems in conventional data analyses using data from a purchasing task study. In Section 3, we propose a nonlinear mixed effects model for behavioral economic demand curves, and analytical formulas of the variance estimations for the derived values, $P_{\max}$, $O_{\max}$ and e. We will also examine the large sample properties of the variance estimations through a Monte Carlo study. In Section 4, we will demonstrate the application of the proposed model to behavioral economic data. Section 5 is devoted to additional discussion.

2. Analytical problems illustrated

Throughout this article, the analytical problems and an actual data analysis will be described based on data obtained from a relative reinforcement efficacy study of marijuana purchasing carried out at University at Buffalo (henceforth, referred as to the marijuana purchasing data) [9].

In the study, participants completed a computerized simulation task in which they were asked how many average-sized high grade marijuana joints they would smoke for 4 hours at various prices. The price per joint was: 0, 0.1, 0.25, 0.5, 1, 2, 4, 5, 7.5, 10, 15, 20, 30, 40, 80 and 160 in U.S. dollars. The data that we use in this article are from 59 participants. By fitting a behavioral economic demand curve to simulated demand at hypothetical prices, the study aims to assess the reinforcing efficacy of marijuana in the study population. Note that the study design is adapted from the purchasing task study by MacKillop and Murphy [4] to assess demand curves for drugs.

2.1 Problem in log transformation

Coefficients based on log-log coordinates in (1) provide the useful description regarding the elasticity. The log transformation is often considered as a valid data analytical step especially for handling positive data with skewed distributions [10]; however, when data include zeros, the log transformation is not possible unless certain small values are added to the data. In purchasing task studies, the range of prices typically includes zero (e.g., $0), and consumption eventually decreases to zero when price increases. In the available literature, it is commonly stated that some arbitrary low values (such as 0.01) are added to zero values (e.g., [4, 11]). The different added values (henceforth referred to as δ) may seriously affect analytical results since arbitrary small values may have huge differences in the log scale. For instance, with a log transformation using the natural log, the values of 10^{-2} and 10^{-5} are transformed to -4.60517 and -11.51293, thus the initial difference (i.e., 0.00999) becomes a sizable difference after the log transformation (6.90776). In addition, besides the matter of the magnitude of δ , we found in some literature that zero values for consumption are either replaced by small values or simply removed [11-12]. Thus, we consider the following strategies in handling zero values.

Strategy 1: For each participant, the first zero consumption value is replaced by a small value and subsequent consumption values are removed [11-12]. Zero prices are replaced by a small value.

Strategy 2: Zero values of both consumption and price are replaced by a small value.

Strategy 3: Small values are added to all values, i.e., $C_{ij} + \delta$ and $p_j + \delta$ for all i, j in (1).

Strategy 3 essentially does a parallel shift of the data points. The idea of Strategy 3 is to preserve the overall trend by adding the same small increment to all data points. We fit Model (1) to the marijuana purchasing data using the different log transformation strategies and magnitudes of δ (Figures 1 and 2). The least square method is used for curve fitting. Figure 1 shows the fitted curves with different strategies with $\delta = 0.01$ in log-log scale. The bullet points in Figure 1 are the log of the average consumption corresponding to each price. Strategies 2 and 3 provide similarly fitted curves while Strategy 1 seems quite different from the other two. Strategy 1 shows a slight increase of consumption initially then a sharp change from being inelastic to elastic; it does not explain the initial decrease in consumption well and exaggerates the elasticity of the later part in the price range. On the other hand, Strategies 2 and 3 provide models that show a pretty constant decrease of consumption throughout the price range. That is, these different strategies provide quite different interpretations of consumption behavior in terms of the elasticity. Overall, the fitted curves do not explain the trend well throughout the price range regardless of the different strategies.

Because Strategy 1 eliminates some available data points and is not the most efficient way to use the data points, we will exclude Strategy 1 in the following discussions.

Figure 2 shows the fitted curves with various values of δ based on Strategies 2 and 3. It clearly demonstrates that the different δ values lead to quite different estimated parameter values. All presented strategies show serious fitting issues. The fitted curve with $\delta = 0.5$ is provided to demonstrate the fact that, although such a large number is likely to interrupt the trend of the data more seriously in the log transformation by shifting the data points more than the smallest increment in the price range and thus should not be used, it provides better curve estimation than other δ's. The different fitting curves seriously alter the estimations of derived quantities such as (2), (3) and the average elasticity as shown in Table 1 as well. For example, the estimated $P_{\max}$ becomes more than double when δ is changed from 0.00001 to 0.01. One may argue that the log transformation can provide the variance stabilization making more homoscedastic variances throughout the price range; however, we found that some arbitrarily added δ to zero consumption may result in quite heteroscadastic variances after log transformation.

We conclude that linear model fitting using the log transformation and Model (1) fails to provide a reliable result.

2.2. Improved performance of the nonlinear model without log transformation

Alternatively to Model (1), we can use the equivalent nonlinear model

$$C_{ij} = l p_j^{\,b} e^{-a p_j} + \varepsilon_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, k. \quad (5)$$

Still we need to add some increments to fit the nonlinear model since $C_{ij} = 0$ is not possible for $p_j < \infty$; however, without log transformation, adding small increments to the data points barely affects the overall trend of the data. We note that the elasticity based on Model (5) has the exactly same interpretation as that with Model (1), although Model (5) is not expressed in log-log coordinates.

Fitted curves of (5) based on Strategies 2 and 3 are compared in Figure 3. Overall, the nonlinear approach follows the trend of the data much better than the linear approach. In the figure, Strategy 2 (left plot) and Strategy 3 (right plot) have little distinction. Figure 3 also shows the fitted curves based on various δ are quite similar.

We compare the derived values $P_{\max}$, $O_{\max}$, and e in Table 2. The derived values are more stable with various δ compared to the dramatic changes in Table 1. It is notable that the estimated values in Table 2 are quite different from those in Table 1, suggesting that the study conclusion can be misled by the linear approach combined with the log transformation.

These results clearly demonstrate that curve fitting without log transformation produces much more robust parameter estimations and better fitting to the data.

We note that, although Model (5) gives improved curve fitting compared to Model (1), it fits a single curve for the entire participants, thus it does not consider the variability among different individuals. Figure 3 also shows that there is a lack of fit toward the end of the price range.

2.3. Over-parameterized model

For behavioral economic demand analyses, it is the customary analytical practice that a separate curve is fitted to each participant based on either Model (1) (e.g. [4, 13]) or Model (5) (e.g., [3, 14]). Based on Model (5), one can express a model allowing individual fitting as

$$C_{ij} = l_i p_j^{b_i} e^{-a_i p_j} + \varepsilon_{ij},\ i = 1, \ldots, n,\ j = 1, \ldots, k. \tag{6}$$

Compared to Model (5), Model (6) considerably increases the number of parameters to be estimated, giving rise to an over-parameterized model. After individual model fitting, the overall model coefficients and the derived values (2), (3) and (4) are estimated by simply averaging the individually obtained parameter estimates, and the corresponding standard errors are subsequently obtained. Note that, in the literature on behavioral economic demand curve analysis, the parameter estimations and standard errors are commonly presented; the corresponding statistical tests rarely accompany those articles. Conceptually, fitting a different model for every participant can be considered as an opposite model building scheme to Model (5). While Model (6) can incorporate variability from different individuals in parameter estimation, it does not take into account deviations of individuals from average behavior based on the demand curve. Thus, the model does not directly reflect the variability between and within individuals, which may present a problem in generalizing the results to a population with similar characteristics [15].

When the over-parameterized model (6) is applied to the marijuana purchasing data, the estimated residual standard error σ is 2.547, much smaller than 8.595, the estimated σ based on Model (5); however, Model (6) does not provide better fitting of the data than the fitted curve based on Model (5) as shown later in this article (Section 4). This is because of the practice of averaging individually estimated parameters in an ad-hoc manner rather than fitting a model based on the entire data. Consequently, Model (6) is seriously affected by unusual data values, and oftentimes estimate $P_{\max}$ and $O_{\max}$ that are not in the valid boundary. For example, in the marijuana purchasing data, the derived $P_{\max}$ for one participant is estimated as -17,333.1 and the corresponding $O_{\max}$ does not exist. Another participant shows an estimated $P_{\max}$ of 1269.96. Such estimations either inflate the standard errors or do not contribute the parameter estimation by eliminating them. In Section 3, using the Monte Carlo study, we will show that the over-parameterized model produces larger standard errors in general and lower coverage rates for confidence intervals compared to a model fitting strategy based on the entire data.

3. Model proposed

3.1 A nonlinear mixed effects model

We demonstrated that parameter estimation by the nonlinear model is much more robust than that by the linear approach through the log-transformed data. However, the nonlinear model assumes a single line for all individuals, and thus it may not explain the variability between individuals giving rise to increasing the residual standard error. The over-parameterized model reduces the residual standard error considerably comparing to the nonlinear model; however, the parameter estimation and their variance estimation depend on the ad-hoc approach resulting in the loss of data points and problem with generalizability.

A nonlinear mixed effects model [15] can provide some middle ground between Models (5) and (6). The nonlinear mixed effects model or the hierarchical nonlinear model has been widely applied in diverse research areas to analyze data with repeated measurements [8]. Using a random coefficient, the nonlinear mixed effects models allow individual variability and incorporate clustering effect into the model, similarly to multilevel linear models. To analyze the economic demand curve analysis, we propose to use a nonlinear mixed effects model as

$$C_{ij} = l_i p_j^{b_i} e^{-a_i p_j} + \varepsilon_{ij},\ i = 1,\ldots,n,\ j = 1,\ldots,k, \tag{7}$$

where

$$\begin{pmatrix} l_i \\ a_i \\ b_i \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{pmatrix} = \boldsymbol{\beta} + \mathbf{b}_i,\ \ \mathbf{b}_i \sim N(\mathbf{0}, \Psi),\ \varepsilon_{ij} \sim N(0, \sigma^2 f(p_j)),$$

and

$$\Psi = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}.$$

In Model (7), $\boldsymbol{\beta}$ indicates the fixed effect and $\mathbf{b}_i$ indicates the random effect that allows the variability of parameters for each individual. This additional variability defines the coefficients unique to each individual so that it reflects the cluster effect within an individual. In actual model building, the necessity of these random effects can be examined visually (see Section 4). The variance-covariance matrix of random effects Ψ shows the general positive definite structure. The residuals ε_{ij} are independent and have heteroscedastic variances as a function of the price, reflecting the fact that the consumption variability tends to be large initially then decreases with increasing prices.

Based on the fitted model, the variances of the derived values need to be obtained. Let $\hat{\beta}_i$ indicate the estimates of the fixed effects $\beta_i, i = 1,2,3$, $Var(\hat{\beta}_i) = \sigma^2_{\hat{\beta}_i}$ and $Cov(\hat{\beta}_i, \hat{\beta}_j) = \sigma_{\hat{\beta}_i \hat{\beta}_j}$. Also, $\hat{P}_{\max}$ and $\hat{O}_{\max}$ are the estimates of $P_{\max}$ and $O_{\max}$, respectively, where β_i are replaced by $\hat{\beta}_i$. We can obtain the following result.

Proposition 1: The variances for $\hat{P}_{\max}$ and $\hat{O}_{\max}$ are approximated by

$$Var(\hat{P}_{\max}) \simeq -2\hat{\beta}_2^{-3}(1+\hat{\beta}_3)\sigma_{\hat{\beta}_2\hat{\beta}_3} + \hat{\beta}_2^{-4}(1+\hat{\beta}_3)^2\sigma^2_{\hat{\beta}_2} + \hat{\beta}_2^{-2}\sigma^2_{\hat{\beta}_3}, \text{ and} \tag{8}$$

$$\begin{aligned} Var(\hat{O}_{\max}) \simeq \hat{\beta}_2^{-4}(1+\hat{\beta}_3)^2 \left(\frac{1+\hat{\beta}_3}{\hat{\beta}_2}\right)^{2\hat{\beta}_2} e^{-2(1+\hat{\beta}_3)} \Big\{ -2\hat{\beta}_2(1+\hat{\beta}_3)\sigma_{\hat{\beta}_1\hat{\beta}_2}\hat{\beta}_1 + \hat{\beta}_2^{\,2}\sigma^2_{\hat{\beta}_1} + (1+\hat{\beta}_3)^2\hat{\beta}_1^{\,2}\sigma^2_{\hat{\beta}_2} \\ + 2\hat{\beta}_1\hat{\beta}_2(\hat{\beta}_2\sigma_{\hat{\beta}_1\hat{\beta}_3} - (1+\hat{\beta}_3)\sigma_{\hat{\beta}_2\hat{\beta}_3}\hat{\beta}_1)\log\left(\frac{1+\hat{\beta}_3}{\hat{\beta}_2}\right) + a^2\hat{\beta}_1^{\,2}\sigma^2_{\hat{\beta}_3}\log\left(\frac{1+\hat{\beta}_3}{\hat{\beta}_2}\right)^2 \Big\}. \end{aligned} \tag{9}$$

Since $P_{\max}$ and $O_{\max}$ in (2) and (3) are at least twice differentiable, the result above is the direct application of the multivariate Taylor's expansion. Note that, in Proposition 1, a specific underlying distribution is not required. Using the properties of the variance calculation, the variance of the estimated average elasticity $\hat{e}$ replacing β_i by $\hat{\beta}_i$ based on (4) is given as

$$Var(\hat{e}) = \overline{p}^2\sigma^2_{\hat{\beta}_2} + \sigma^2_{\hat{\beta}_3} - 2\overline{p}\sigma_{\hat{\beta}_2\hat{\beta}_3}. \tag{10}$$

Note that $\overline{p}$ in (10) is a fixed value. Actual variance estimation is carried out by replacing the variance and covariance terms in (8), (9), and (10) with the estimated ones. In the next subsection, we investigate the performance of these variance estimates in inference.

Model (7) can be fitted using commonly available statistical packages such as SAS or R. For nonlinear mixed effects model fitting, starting values need to be provided. Starting values can be obtained from a simpler model such as nonlinear models without random coefficients or parameter estimations based on the over-parameterized model. The R codes for Model (7) and the variance estimations of the derived values are available from the authors.

3.2 Simulations

We investigate the performance of inferences based on the variance estimations in (8), (9), and (10) through an extensive Monte Carlo study (1000 simulations per scenario). For the simulation, l_i in (7) has a normally distributed random effect while a_i and b_i have fixed effects only, and the variance of ε_{ij} is constant. The fixed effect parameters for simulations are chosen to achieve the semblance of the marijuana purchasing data ($\beta_1 = 10.981, \beta_2 = -0.026, \beta_3 = 0.102$). In addition, the same 16 price points from the marijuana purchasing data are used. The true $P_{\max}$, $O_{\max}$ and e for simulations are 9.5, 37.3 and -2.4, respectively. Through the Monte Carlo study, a large sample property is investigated by constructing confidence intervals based on the normal approximation, where ε_{ij} have various underlying distributions (normal, centralized lognormal and centralized χ^2). The coverage rate of the confidence interval and magnitude of the standard errors for the nonlinear mixed effects model and over-parameterized model are compared in Table 3. For the nonlinear mixed effects model, even with a moderate sample size such as $n = 30$, a viable performance of the confidence intervals based on the variances (8), (9), and (10) is shown regardless of the different underlying distributions. On the other hand, the over-parameterized model shows unsatisfactory coverage rates throughout the various underlying distributions and sample sizes indicating that the over-parameterized model does not provide suitable inferences for the derived values. It is noteworthy that the magnitudes of the standard errors of the nonlinear mixed effects model are much smaller than those for the over-parameterized model, ensuring that the nonlinear mixed effect model is a much more efficient way to perform parameter estimation.

4. Data application

In this section, we demonstrate the actual data analysis using the marijuana purchasing data. For model fitting, we choose a small added value, 1×10^{-5} based on Strategy 3 (parallel shift). The nonlinear mixed effects model proposed in (7) considers many random coefficients and correlations between them, where not all random effects are necessary to describe the observations. In actual model fitting, for the first look at the data, one can investigate the variability of the estimated coefficients based on the over-parameterized model. The over-parameterized model can suggest the necessity of the random coefficient in the non-linear model by observing the ranges of the individually obtained parameter estimates [15]. Plots of the individual estimated values of parameters and standard errors based on the over-parameterized model are shown in Figure 4. The estimates of l_i show a large variability as opposed to fairly constant estimates of b_i and a_i, suggesting that the random effect b_{1i} in Model (7) is reasonable while the random effects b_{2i} and b_{3i} may not be necessary. The standard deviations of consumption at each price point gradually decrease from 18.01 (at $0) to 2.19 (at $160), suggesting the within-group standard errors are a decreasing function of prices. For the data analysis, the residual standard error is defined as a power function of price (i.e., $\sigma = cp_j^{\,\nu}$, where c is a constant and ν is a real number). In statistical software R, this can be easily handled using a variance function such as `varPower` in nonlinear mixed effects model fitting. The distributions of consumption at each price are not symmetric but generally skewed to the right; however, we demonstrated workable inferences based on the non-normal underlying distributions even with moderate sample sizes (e.g., $n = 30$) in the Monte Carlo study (Section 3). The fitted values of the fixed effects and standard errors (inside parentheses) based on Model (7) are $\hat{\beta}_1 = 10.163(1.369)$, $\hat{\beta}_1 = -0.037(0.006)$ and $\hat{\beta}_1 = 0.073(0.004)$. For the power function of the residual standard error, ν is estimated as -0.083.

Figure 5 shows the fitted curves based on Models (5), (6), and (7). Model (7) follows the consumption trend more closely than other models by improving curve fitting toward the end of the price range. The residual standard error of Model (7) is much smaller (3.996) than that of Model (5) (8.594), indicating a better curve fit than the nonlinear model without random effects. The better fit of Model (7) toward the higher price gives rise to the least elastic curve among the fitted curves. Estimated values (standard errors in parentheses) of $P_{\max}$, $O_{\max}$ and e are 13.214 (0.692), 46.635 (6.474) and -1.747 (0.092), respectively. Note that, for the over-parameterized model, the estimated values (standard errors in parentheses) of $P_{\max}$, $O_{\max}$ and e are 17.422 (5.276), 64.200 (22.029) and -4.018 (0.755), respectively, where the standard errors are calculated based on 57 participants due to out-of-boundary estimations as explained in Section 2.

We note that, even if the nonlinear mixed effects model improves curve fitting compared to the other models, there is still a lack of fitting at the tail ends of the curves (Figure 5). This problem is a consequence of using the conventionally accepted theoretical model based on the concept of the first order linear relationship between elasticity and price. We note that Bickel et al. [16] discuss some limitations of the behavioral economic demand curves. If desired, other models can be used or one may try a segmented regression [17-18]. However, we argue that the conventional behavioral demand curve may be sufficient to investigate $P_{\max}$, $O_{\max}$ and elasticity, which are mostly obtained within the low-price range, where fitting is satisfactory. More specifically, $P_{\max}$ and $O_{\max}$ are the price point to have the elasticity -1 and the corresponding expenditure, respectively. These values are mostly obtained within the low-price range (e.g., $\hat{P}_{\max} = 13.214$ with the marijuana purchasing data). Also, $\hat{\text{e}}$ is obtained at the mean price, i.e., 23.46, where the curve is well fitted to the data.

5. Discussion

A common practice in the analysis of behavioral economic demand curves is to replace zero values in consumption and/or price with arbitrary small values and then to fit a linear model based on the log-transformed data. As shown earlier, there are many different ways that zero values are replaced by a small value. These different strategies and a choice of arbitrary small values can lead to hugely different model fitting results and derived values. We have observed that the nondiscretionary log transformation of the data points leads to poor model fitting in general. Although one may argue that the log transformation can stabilize the variance, we found that the magnitude of the variances can change dramatically by adding different values to zeros.

It is also a common practice to fit the model for each individual, resulting in an over-parameterized model. Because the over-parameterized model individually fits the data from each participant, it may not properly evaluate the variability from the population [15]. It is also noted that individual fitting often fails to provide realistic derived economic values such as $P_{\max}$ and $O_{\max}$, inflating the corresponding standard errors.

We proposed to use nonlinear mixed effects models that provide robust and better coefficient estimations of the demand curve than the conventional approach. We also provided convenient analytical formulas of variance estimates for the derived economic values once a model is fitted, which makes the nonlinear mixed effects model more appealing to practitioners.

We conclude this article by suggesting that the framework presented here can be extended to other demand curve models that use log transformation. The effects that these models can have with log transformation in terms of the robustness of parameter estimates and overall fitting to the data should be examined. If the log transformation may compromise the robustness of parameter estimates, we recommend considering an approach without log transformation as suggested in this article. For example, adapting the demand curve proposed by Allen [19], Hursh and Silberberg [20] discuss the following demand curve for behavioral-economic data analysis:

$$\log C = \log C_0 + k(e^{-\alpha p} - 1), \tag{11}$$

where the log of consumption (C) exponentially decreases at the rate of α as price (p) increases. Hursh and Silberberg [20] show that Model (11) can explain demand curves for both necessary commodities (e.g., food) and unnecessary commodities. A major advantage of Model

(11) is noted that the value of α works as a single 'essential value' indicating the rate of change in elasticity of demand. When fitting Model (11), the log transformation is recommended [20]; however, it may cause unstable fitting similarly to Model (1). Instead, taking exponential for the both sides of (11), we can fit

$$C = C_0^* e^{k(p_*^{-\alpha} - 1)},$$

where $p_* = e^p$ and C_0^* is the parameter related to C_0 in (11), and the random effect can be incorporated similarly to Model (7).

The approach presented in this article will have implications for models in other research areas as well, which use log-transformation for their fitting.

Acknowledgements

This research is supported by the National Institute on Drug Abuse at National Institute of Health (grant number R01-DA027606, Principal Investigator: Dr. R. Lorraine Collins).

References

1. Mackillop J, Murphy JG, Tidey JW, Kahler CW, Ray LA, Bickel WK. Latent structure of facets of alcohol reinforcement from a behavioral economic demand curve. *Psychopharmacology (Berl)* 2009; 203:33-40.

2. Epstein LH, Dearing KK, Roba LG. A questionnaire approach to measuring the relative reinforcing efficacy of snack foods. *Eating Behaviors* 2010; 11:67-73.

3. Hursh SR, Raslear TG, Shurtleff D, Bauman R, Simmons L. A cost-benefit analysis of demand for food. *Journal of the Experimental Analysis of Behavior* 1988; 50:419-40.

4. Murphy JG, MacKillop J. Relative reinforcing efficacy of alcohol among college student drinkers. *Experimental and Clinical Psychopharmacology* 2006; 14:219-27.

5. Spiga R, Martinetti MP, Meisch RA, Cowan K, Hursh S. Methadone and nicotine self-administration in humans: a behavioral economic analysis. *Psychopharmacology (Berl)* 2005; 178: 223-31.

6. Wade-Galuska T, Winger G, Woods JH. A behavioral economic analysis of cocaine and remifentanil self-administration in rhesus monkeys. *Psychopharmacology (Berl)* 2007; 194:563-72

7. Lindstrom ML, Bates DM. Nonlinear mixed effects models for repeated measures data. *Biometrics* 1990; 46:673-87.

8. Davidian, M. and Giltinan, D.M. *Nonlinear Models for Repeated Measurement Data*. Chapman & Hall: London, 1995

9. R. Lorraine Collins et al., A Behavioral Economic Assessment of Demand for Marijuana. submitted, 2012.

10. Keene ON. The log transformation is special. *Statistics in Medicine* 1995; 14:811-9.

11. Jacobs EA, Bickel WK. Modeling drug consumption in the clinic using simulation procedures: demand for heroin and cigarettes in opioid-dependent outpatients. *Experimental and Clinical Psychopharmacology* 1999; 7:412-26.

12. Murphy JG, MacKillop J, Tidey JW, Brazil LA, Colby SM. Validity of a demand curve measure of nicotine reinforcement with adolescent smokers. *Drug and Alcohol Dependence* 2011;113:207-14.

13. Rasmussen EB, Reilly W, Hillman C. Demand for sucrose in the genetically obese Zucker (fa/fa) rat. *Behavioural Processes* 2010; 85:191-7.

14. Madden GJ, Smethells JR, Ewan EE, Hursh SR.Tests of behavioral-economic assessments of relative reinforcer efficacy: economic substitutes. *Journal of the Experimental Analysis of Behavior* 2007;87:219-40.

15. Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. Springer: New York, 2000

16. Bickel WK, Marsch LA, Carroll ME. Deconstructing relative reinforcing efficacy and situating the measures of pharmacological reinforcement with behavioral economics: a theoretical proposal. *Psychopharmacology (Berl)* 2000; 153: 44-56.

17. McGee V.E, Carleton W.T. Piecewise regression. *Journal of American Statistical Association* 1970; 65: 1109-1124.

18. Vexler A., Gurevich G.Guaranteed local maximum likelihood detection of change point in nonparametric logistic regression. *Communications in Statistics (Theory and Methods)* 2006; 35: 711-726

19. Allen RGD. *Mathematical analyisis for economists*. Macmillan: New York, 1962.

20. Hursh SR, Silberberg A. Economic demand and essential value. *Psychological Review* 2008; 115:186-98.

Figure 1: Fitted curves of Model (1) in log-log scale by various log transformation strategies. The solid line, dotted line and dashed line correspond to Strategies 1, 2 and 3, respectively. The added δ is 0.01 for all three curves. Note that the curves for Strategies 2 and 3 are almost identical.

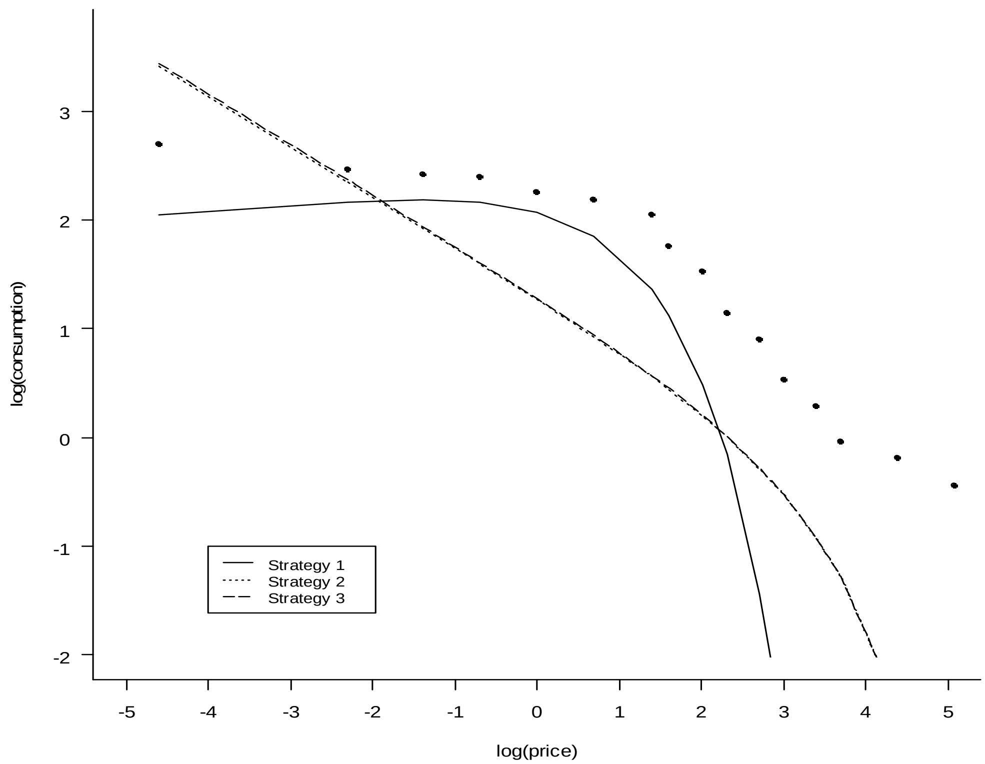

Figure 2: Fitted curves of Model (1) in log-log scale by various δ values. The left and right panels use Strategies 2 and 3 for log transformation, respectively. The bullet points are the log of the mean consumption for each price. The different curves are fitted using different δ as indicated in the figure legend.

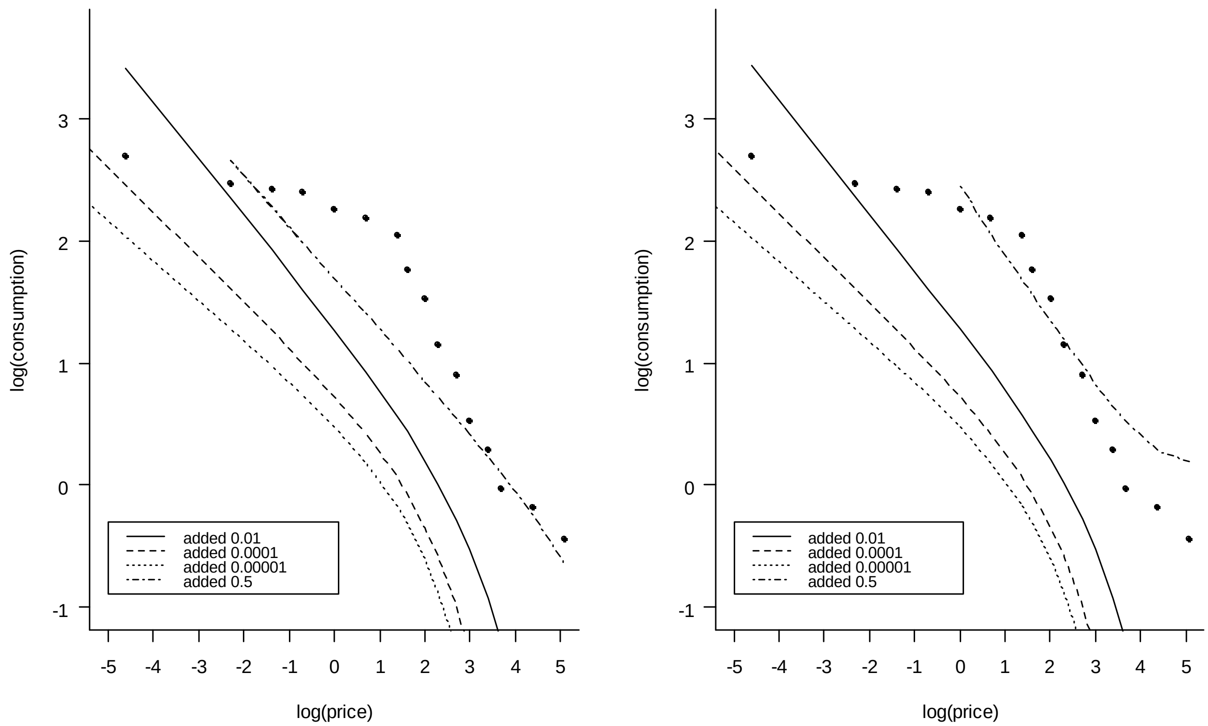

Figure 3: Fitted curves of Model (5) in log-log scale by various δ values. The left and right panels use Strategies 2 and 3 for log transformation, respectively. The bullet points are the log of the mean consumption for each price. The different lines are fitted using different δ as indicated in the figure legend.

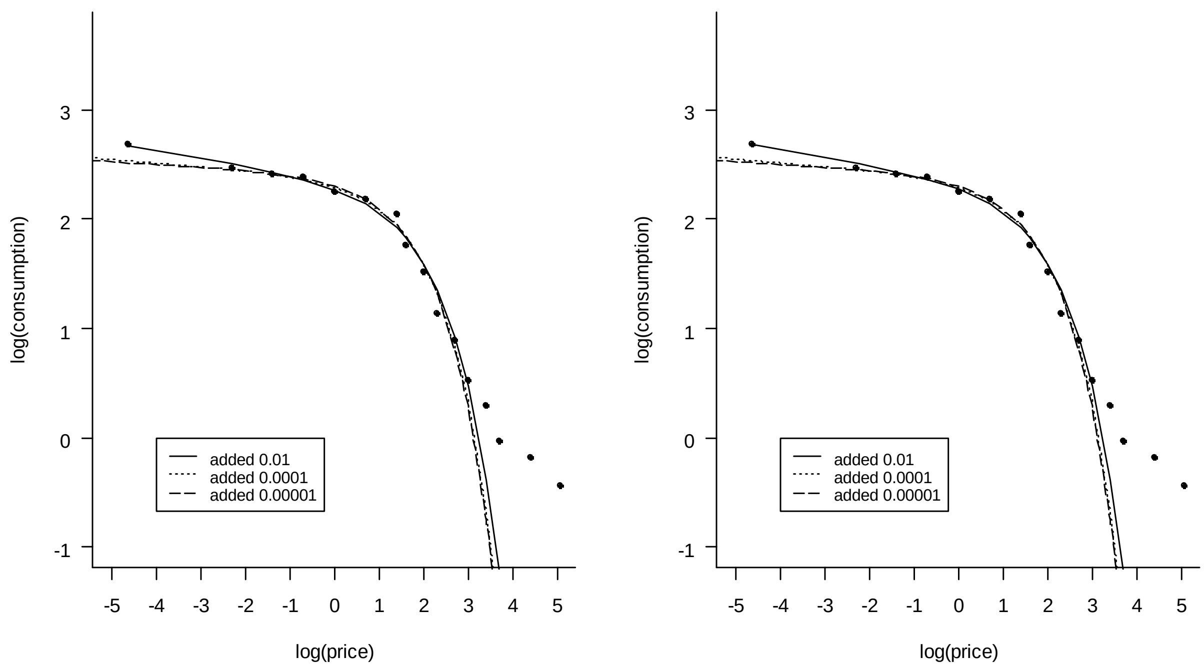

Figure 4: Parameter estimates for each individual based on the over-parameterized model. The numbers in x-axis correspond to participants' identification numbers (a total of 59 participants). The dot and the half of the bar indicate the estimated value and the magnitude of its standard error.

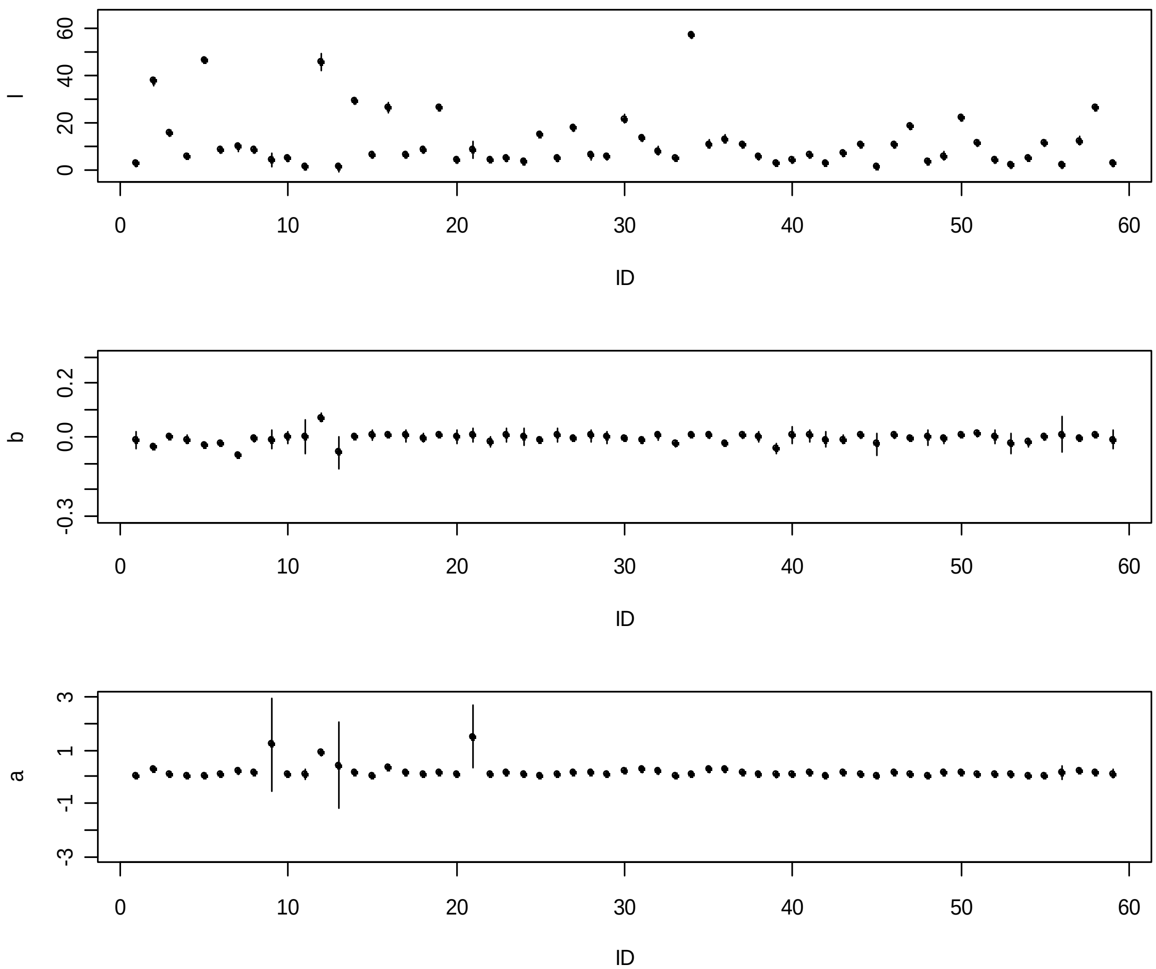

Figure 5: Comparisons of fitted curves based on Model (5) (dashed line), Model (6) (dotted line) and Model (7) (solid line) in the actual scale (left plot) and log-scale (plot). The bullet points indicate the mean consumption for each price. The grey colored points are the actual observations. When the data points are overlapped, slight noises are added at plotting. In the right figure, the points at the bottom of the plot are log-transformed zero consumption values ($\delta = 0.00001$).

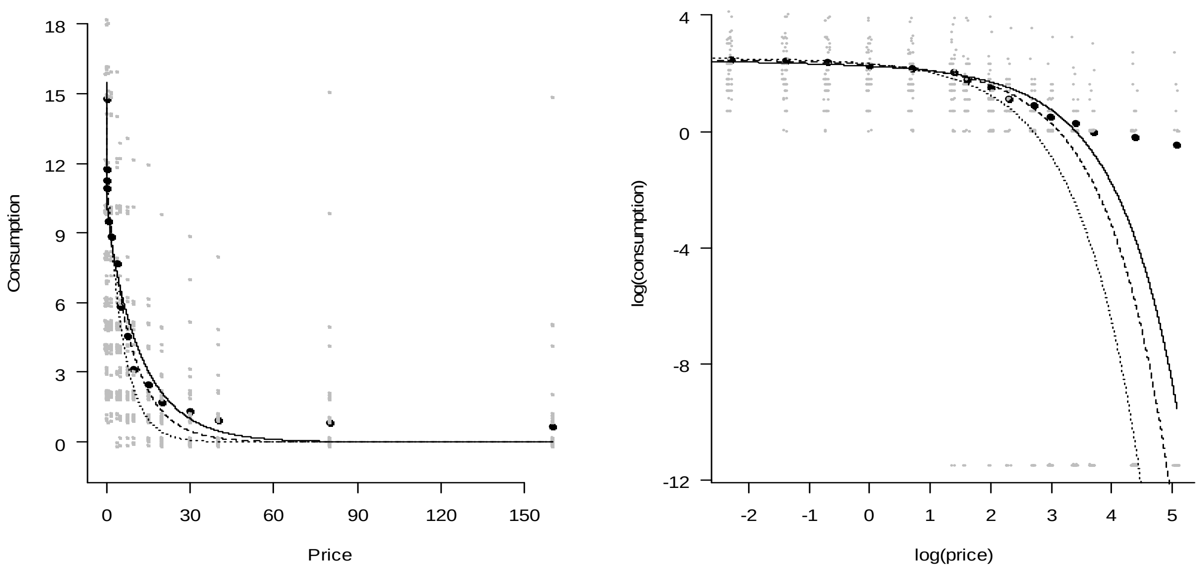

Table 1: The comparisons of estimated values of $P_{\max}$, $O_{\max}$, and e obtained using Model (1) with various δ and log transformation strategies.

Strategy	Derived value	$\delta = 0.01$	$\delta = 0.0001$	$\delta = 0.00001$
Strategy 2	$P_{\max}$	24.795	12.248	10.107
	$O_{\max}$	11.931	5.631	4.169
	e	-0.971	-1.582	-1.893
Strategy 3	$P_{\max}$	24.706	12.248	10.107
	$O_{\max}$	11.982	5.631	4.169
	e	-0.973	-1.582	-1.893

Table 2: The comparisons of estimated $P_{\max}$, $O_{\max}$, and e obtained using Model (5) with various δ and log transformation strategies.

Strategy	Derived value	$\delta = 0.01$	$\delta = 0.0001$	$\delta = 0.00001$
Strategy 2	$P_{\max}$	11.155	9.800	9.552
	$O_{\max}$	39.114	37.789	37.523
	e	-2.025	-2.348	-2.419
Strategy 3	$P_{\max}$	11.131	9.800	9.552
	$O_{\max}$	39.204	37.790	37.523
	e	-2.029	-2.348	-2.419

Table 3. Coverage rates of the 95% confidence intervals and the magnitude of the standard errors for the derived values in the Monte Carlo study. The residuals are generated based on the centralized underlying distribution indicated in the distribution column.

Distribution	n	Proposed model			Over-parameterized model		
		$P_{\max}$	$O_{\max}$	e	$P_{\max}$	$O_{\max}$	e
Normal (0,1)	30	0.951 (0.233)	0.935 (3.121)	0.950 (0.058)	0.835 (2.110)	0.757 (3.165)	0.953 (0.3243)
	60	0.942 (0.164)	0.947 (2.237)	0.946 (0.0412)	0.716 (1.841)	0.781 (2.328)	0.945 (0.318)
	100	0.946 (0.127)	0.943 (1.736)	0.946 (0.032)	0.623 (1.261)	0.743 (1.851)	0.918 (0.265)
Lognormal($0,0.75^2$)	30	0.950 (0.267)	0.934 (3.142)	0.943 (0.067)	0.770 (1.051)	0.834 (3.150)	0.941 (0.448)
	60	0.949 (0.189)	0.945 (2.248)	0.937 (0.047)	0.697 (5.231)	0.792 (3.847)	0.919 (0.392)
	100	0.942 (0.146)	0.949 (1.753)	0.935 (0.037)	0.638 (1.535)	0.697 (1.934)	0.858 (0.403)
$\chi^2_{df=1}$	30	0.947 (0.330)	0.932 (3.208)	0.939 (0.082)	0.759 (4.278)	0.805 (3.535)	0.936 (0.510)
	60	0.944 (0.232)	0.943 (2.303)	0.941 (0.058)	0.711 (2.471)	0.734 (2.689)	0.899 (0.476)
	100	0.946 (0.146)	0.946 (1.757)	0.940 (0.037)	0.637 (1.737)	0.655 (2.057)	0.847 (0.483)