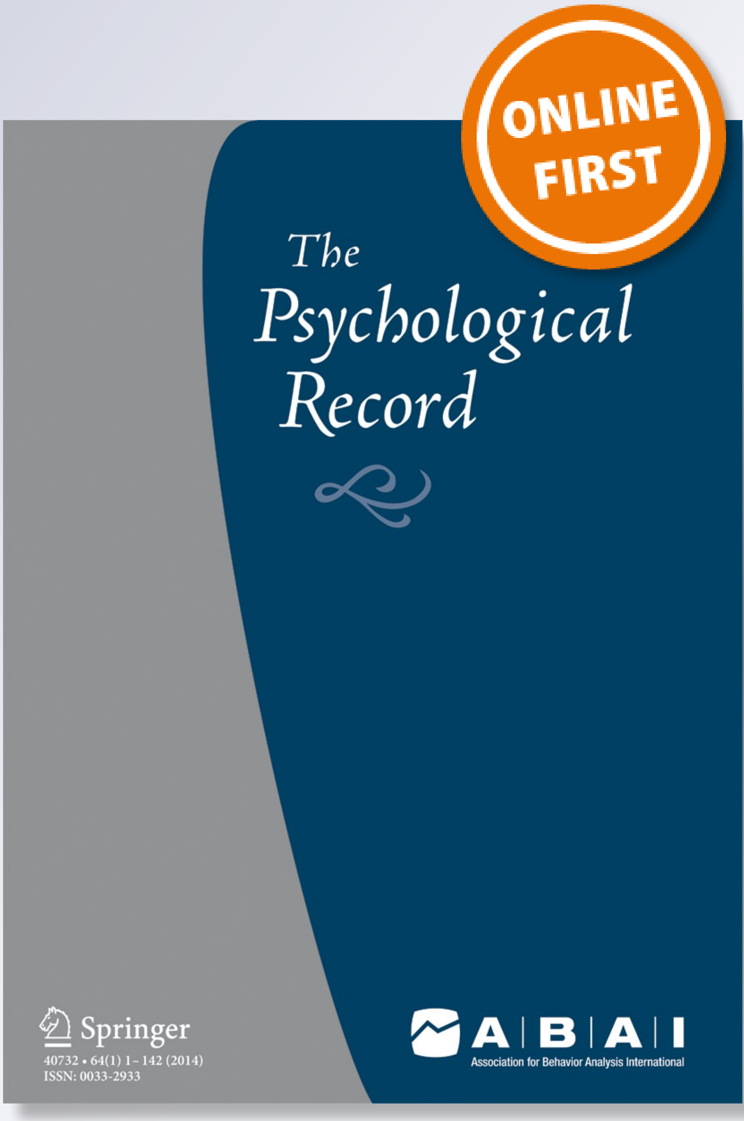
Inter-Method Reliability of Progression Sizes in a Hypothetical Purchase Task: Implications for Empirical Public Policy

Derek D. Reed, Brent A. Kaplan, Peter G. Roma & Steven R. Hursh

The Psychological Record

ISSN 0033-2933

Psychol Rec
DOI 10.1007/s40732-014-0076-1



Springer

Psychol Rec
DOI 10.1007/s40732-014-0076-1

ORIGINAL ARTICLE

Inter-Method Reliability of Progression Sizes in a Hypothetical Purchase Task: Implications for Empirical Public Policy

Derek D. Reed · Brent A. Kaplan · Peter G. Roma · Steven R. Hursh

Abstract Advances in behavioral economic research methodology have generated novel insights into public policy considerations regarding substance use and addiction. A major advancement toward this end was the development of hypothetical purchase tasks (HPTs) that permit modeling of consumption and demand without the need for actual purchases and lengthy human operant studies comprised of real reward deliveries. Psychometric research in the domain of behavioral pharmacology suggests that the HPT has adequate test–retest reliability, as well as construct and convergent/divergent validity. Research to date has exclusively examined pharmacological commodities, however, necessitating the need for technology transfer to other pressing societal issues. In the present study, we examined demand for recreational driving across three groups, each consisting of a different fuel price progression size in the HPT. Findings suggest that behavioral economic models of demand do not differ as a function of progression size, providing support to the inter-method reliability of this procedure.

Keywords Behavioral economics · Demand curve · Driving · Hypothetical purchase task · Reliability

D. D. Reed (✉) · B. A. Kaplan
Department of Applied Behavioral Science, University of Kansas, 4048 Dole Human Development Center, 1000 Sunnyside Avenue, Lawrence, KS 66045-7555, USA
e-mail: dreed@ku.edu

P. G. Roma · S. R. Hursh
Institutes for Behavior Resources, Baltimore, MD, USA

P. G. Roma · S. R. Hursh
Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Published online: 31 May 2014

Evaluation of Progression Sizes in a Fuel Hypothetical Purchase Task: Inter-Method Reliability of Empirical Public Policy

Behavioral economic considerations of real-world issues have become increasingly popular in the social sciences and public policy discussions, no doubt due to the success of popular trade books such as Thaler and Sunstein's *Nudge: Improving decisions about health, wealth and happiness* (2008) and Shafir's *The behavioral foundations of public policy* (2013). The behavioral sciences, however, have utilized behavioral economic concepts for decades due to the importance of analyzing consumption of reinforcers in the context of various environmental constraints—notably, costs associated with each unit of reinforcement (e.g., Green and Freed 1993; Hursh et al. 1989; Hursh 1980, 1984; Johnson and Bickel 2006; MacKillop et al. 2012b; Madden and Hartman 2006; Rachlin et al. 1976; Reed et al. 2009; Reed et al. 2013a, b; Roane et al. 2001). Formal demand curve analyses are used to quantify the relation between price and consumption (Hursh et al. 1988; Hursh et al. 2012; Hursh and Silberberg 2008). Such analyses plot units of reinforcement consumed as a function of the unit price (i.e., the number of responses, the amount of money, the degree of effort, etc., necessary to access the reinforcer). The primary aim of such analysis is to evaluate the degree to which consumption is defended against price increases. Consumption is considered *inelastic* when it remains relatively undisturbed across increasingly high prices. However, consumption is considered *elastic* when it demonstrates price sensitivity (i.e., consumption decreases substantially as prices increase).

The contemporary quantitative model of demand (Hursh and Silberberg 2008; Hursh et al. 2012) assumes an exponential function and quantifies the (a) overall level of elasticity across the entire range of prices (termed α; lower α values are indicative of relative inelasticity) and (b) level of consumption at zero cost (termed Q_0; considered a measure of maximum or baseline demand intensity). The essential value of the reinforcer is quantified as the inverse of α, such that higher values represent commodities of relatively greater demand. Another quantitatively derived metric of demand is P_{max}, which quantifies the price at which consumption becomes elastic (this metric is inversely related to α). A major advantage of the exponential model is that the equation controls for reinforcer magnitude as well as qualitative differences in reinforcement within the Q_0 parameter. Thus, the model allows differing intensities of demand to feature equal levels of elasticity, thereby permitting credible quantitative comparisons between qualitatively different reinforcers or identical reinforcers of differing sizes. The exponential demand equation is presented in greater detail below.

Demand curve analyses originated in highly controlled operant laboratories using nonhuman subjects working for primary reinforcers in closed economies (Hursh 1980, 1984). This procedure became increasingly fruitful for examining demand for qualitatively and quantitatively different drugs using self-administration (e.g., Hursh and Winger 1995; Ko et al. 2002; Spiga et al. 2005). Given public interests in human drug research, experiments began translating the procedures for human operant studies using plunger pulls as operants to access drug reinforcers (e.g., Johnson and Bickel 2006; Shahan et al. 1999). These human operant approaches to the behavioral economic study of drug addiction revolutionized human behavioral pharmacology (e.g., Bickel et al. 1993; Bickel et al. 2012). Such procedures, however, carry inherent medical risk to participants, require large amounts of time for participants to demonstrate stable responding at each drug price, and are highly expensive (payment for participants, purchasing drugs, equipment costs). Human behavioral pharmacology studies are thereby typically limited to grant funded projects and employ specialized populations of participants with a history of substance abuse.

In an effort to circumvent some of the safety and cost-prohibitive features of human operant demand curve studies, Jacobs and Bickel (1999) created a hypothetical purchase task (HPT) for use with opioid-dependent participants. In their HPT, Jacobs and Bickel asked participants to imagine they were purchasing cigarettes and heroin during their time of heaviest abuse prior to entering treatment, and that they were not receiving buprenorphine, naltrexone, or antabuse, and could not access the commodities in any other capacity. The HPT consisted of an ascending series of monetary prices to access one unit of reinforcement (one cigarette or one bag of heroin). Prices ranged from \$0.01 to \$1,120.00 per unit. Participants completed three HPTs: one for cigarettes alone, one for heroin alone, and one in which cigarettes and heroin were available concurrently. The obtained results appeared visually similar to demand curves derived from human operant studies and were well accounted for by quantitative models of demand (Hursh et al. 1988). As expected given the sample of opioid-dependent participants, demand for heroin was greater than demand for cigarettes, providing some degree of construct validity. Demand measured via individual HPTs did not differ from that of the concurrent HPT, suggesting some degree of criterion validity.

Murphy and MacKillop (2006) adapted Jacobs and Bickel's (1999) HPT to measure demand for alcohol in a standard college-aged sample of participants. In the Murphy and MacKillop HPT, the vignette was specific to alcohol and provided a price range of \$0 (free) to \$9.00 per drink. The researchers also asked participants to complete the Daily Drinking Questionnaire (DDQ; Collins et al. 1985) to determine the levels of drinking emitted by the participants each week, as well as the Rutgers Alcohol Problem Inventory (RAPI; White and Labouvie 1989). Results were similar to those obtained by Jacobs and Bickel (1999) in that problematic drinkers featured greater intensity of demand for alcohol than less problematic drinkers. Demand curves for alcohol across all participants appeared similar to those obtained in human operant studies and were accounted for by quantitative models of demand (Hursh et al. 1988). These findings provide additional construct and concurrent validity to the HPT. Most importantly, the Murphy and MacKillop (2006) findings suggest that HPT can yield valid data from non-clinical populations, providing a low-cost alternative to operant demand analyses. In one follow-up study, MacKillop and Murphy (2007) found that this alcohol HPT was predictive of alcohol treatment success, providing predictive validity to this form of demand assessment. In another study, MacKillop et al. (2012b) demonstrated that the cigarette HPT may serve as a viable model for public policy interventions aimed at assessing tax revenue and price elasticity when factoring an excise tax into the HPT prices.

Subsequent studies have integrated the procedures from Jacobs and Bickel's (1999) and Murphy and MacKillop's (2006) HPT and demand curve analyses to examine hypothetical cigarette consumption in non-clinical samples (MacKillop et al. 2008; Murphy et al. 2011). Specifically, MacKillop et al. (2008) found that cigarette demand was higher in college-aged participants with mild to moderate nicotine dependence than those participants with minimal dependence, suggesting both convergent and divergent validity in the HPT. A 2011 replication using adolescents (Murphy et al. 2011) reproduced the convergent and divergent validity findings. An additional contribution of the 2011 study was the inverse correlation between intensity of demand and self-reported motivation to change smoking habits; this correlation provides some clinical utility to the HPT. Madden and Kalman (2010) found that demand for cigarettes in a HPT following one week of treatment was predictive of long-term cessation, also providing some degree of predictive validity.

In an effort to further demonstrate the construct validity of the HPT, MacKillop et al. (2010, 2012a) devised a procedure to examine demand in the presence of commodity-specific and neutral cues. Previous research has suggested that cravings increase consumption, and that environmental cues can modulate such cravings (Carter and Tiffany 1999; Laibson 2001). Thus, the goal of MacKillop and colleagues' cue-eliciting procedures was to evaluate whether demand is modulated through environmental cues related to the commodity. For example, MacKillop et al. (2010) exposed alcohol drinkers to two conditions: one in which alcohol cues were present and the other with neutral cues (related to water). Cues consisted of photographs of the commodity, an audiotape describing the commodity, and the presence of the commodity itself for proprioceptive and olfactory cues. Indeed, hypothetical demand for alcohol was higher in the presence of alcohol-specific cues, suggesting that the HPT features construct validity. These results have been reproduced with cigarettes as the target commodity (the neutral cue consisted of pencils) for nicotine-dependent participants (MacKillop et al. 2012a).

In addition to validity studies on the psychometric properties of HPT, two studies highlight the temporal stability of the measures obtained from such studies of demand. Murphy et al. (2009) found excellent reliability at a 2-week test–retest window for the alcohol HPT using the exponential demand model (Hursh and Silberberg 2008). Few et al. (2012) reproduced these findings with a 1-week test–retest window for the cigarette HPT. These two studies indicate that HPT for cigarettes and alcohol are reliable, in addition to being valid. The extant literature on HPT described above collectively suggests that the HPT is a psychometrically sound measure of demand for commodities of addiction.

To date, the HPT procedure has been exclusively employed for pharmacological study. This is unfortunate given the wide degree of applicability that the HPT can offer the behavioral sciences. As proposed by Hursh and Roma (2013), the HPT may be useful in examining public policy issues, such as excise taxes, that cannot be manipulated in applied contexts and that feature too limited of a range in actuarial studies of archival market data. The extant HPT research on substance use has yielded interesting information on the interaction between price structuring and dependence, which has obvious implications for public policy consideration. Namely, a one-size-fits-all approach to excise taxes will likely affect only a subsample of the population. However, an additional benefit of the HPT is the ability to model the amount of money spent at the peak level of consumption; such information provides an estimate of the tax revenue generated through the use of the excise tax (see MacKillop et al. 2012b).

Beyond substance abuse, as indicated by Reed et al. (2013a, b), demand for oil in the North American marketplace appears incredibly inelastic using the behavioral economic exponential model of demand, but this may only be a function of a limited range of market prices that is too short to result in elastic demand. Employing HPT for commodities such as fuel may be one way to model the effects of oil prices that have not yet been experienced by North American consumers. Moreover, no studies have examined the inter-method reliability of the HPT with respect to various price progression sizes, that is, whether the "essential value" of a commodity remains constant regardless of the prices queried within a specified range. The progression sizes used in the extant HPT literature have not been systematically examined to determine whether the progression sizes matter in generating rapid HPT demand curves. Thus, the goals of the present study were two-fold. First, in an effort to expand HPT technology to issues of public importance beyond substance abuse, we sought to determine whether data obtained using a HPT for a non-pharmacological commodity—in this case, fuel—could be adequately explained by contemporary exponential models of demand (Hursh and Silberberg 2008). Second, we sought to systematically evaluate various progression sizes at the group level to determine whether markers of exponential demand (e.g., demand elasticity, intensity, P_{max}, breakpoint) significantly varied as a function of progression size.

Methods

Participants, Setting, and Equipment

Seventy-six undergraduates aged 18–32 (M=19.88 years, SD=2.30) enrolled in an introductory psychology course were recruited and received extra credit for participation. Participants were seated at individual computers in a quiet 9 m x 6 m testing lab featuring 20 desks arranged around the perimeter of the room, each equipped with a Dell Optiplex computer (running Windows 7) connected to a local area network and the internet. A Dell 43.18-cm-wide aspect flat-panel monitor was connected to each computer and placed at eye-level on each desk (approximately 0.5 m from each participant's face). Approximately ten participants participated during each session.

Experimental Preparation

Upon entering the lab, the experimenter randomly assigned the participants to one of three groups: $0.25 ($n$=26), $0.50 ($n$=26), or $1.00 ($n$=25) fuel price progression sizes. The experimental task was preloaded on the screen when participants took their seat to begin the study. All participants were presented with and answered driving-related questions pertaining to a hypothetical vignette (adapted from Murphy and MacKillop 2006) displayed on the computer screen using Qualtrics® Research Suite (see http://wwww.qualtrics.com). Each vignette was displayed on a different page and read:

> In the questions that follow, we would like you to pretend to make decisions about driving. The driving

decisions that you make will appear on the next several pages. You may decide to drive as much or as little as you'd like.

As you answer the questions, assume that you drive an average 4-door mid-sized passenger car that achieves an average fuel efficiency of 20 miles per gallon. In each question, assume that you're making a decision whether to drive to a very fun vacation with some friends. You plan to be at the vacation spot for a full 3 days, not counting travel time to and from your destination. Thus, the time you spend traveling to and from your destination does not factor into whether you'd prefer to drive or find an alternative mode of transportation.

You will be presented with different scenarios asking how far you'd be willing to drive for the vacation. The number of miles that you enter on the page should be the maximum you'd be willing to drive to go on the vacation. For example, if you're willing to drive a maximum of 500 miles, then we will assume that a trip requiring 501 or more miles of driving would result in you seeking an alternative mode of transportation or not going on the vacation at all. Finally, assume that you alone are paying for the fuel to go on the trip. You cannot ask your friends for money to pay for the fuel. Assume that you have the same income/savings that you have now and NO ACCESS to any fuel other than that offered at these prices. Please answer the questions thoughtfully and honestly.

At the bottom of each page below the vignette, a text box allowed participants to report their answers to the following question: *How many miles would you be willing to drive for the vacation if fuel was* ***<u>$X.XX per gallon?</u>*** The cost of fuel progressively increased across subsequent questions for the three groups; however, progression sizes increased differentially across the groups. Prices for the three groups progressed in \$0.25, \$0.50, or \$1.00 increments starting at \$0.00 (free) and ending at \$15.00. That is, for the \$0.25 group the values included \$0.00, \$0.25, \$0.50, \$0.75, \$1.00, \$1.25…\$15.00. For the \$.50 group the values included \$0.00, \$0.50, \$1.00, \$1.50, \$2.00, \$2.50…\$15, and for the \$1.00 group the values included \$0.00, \$1.00, \$2.00, \$3.00, \$4.00, \$5.00…\$15. Qualtrics® Research Suite validation tools required participant responses to be numerical with a minimum value of 0 for each trial. The next fuel price trial was presented when a valid response was inputted and the participant clicked a submit button.

Data Analysis

Given the purposes of the present study were to determine if the current data could adequately be explained using contemporary models of demand and if differences exist in various measures of demand when using different progression sizes, data were analyzed using Hursh and Silberberg's (2008) exponential model of demand (Eq. 1):

$$\log Q = \log Q_0 + k\left(e^{-\alpha\cdot(Q_0\cdot C)} - 1\right) \qquad (1)$$

where Q equals consumption, or how many miles participants reported they would drive, Q_0 equals baseline consumption or intensity of demand (i.e., the level of consumption when the cost of fuel is free), k is the scaling constant equal to the number of log units of consumption (k=5), and C equals cost, or the price of fuel per gallon. Finally, α is the parameter solved for and describes the change in elasticity, or how consumption changes as a function of cost across the entire range of prices (note that for the remainder of the paper, we use α and "elasticity" interchangeably). The parameter Q_0 may be solved for when consumption is not empirically observed when cost is free; however, because we directly obtained Q_0 by asking participants how many miles they would be willing to drive if fuel was free, we constrained Q_0 to the reported consumption at that price for each individual and for group medians. For graphing purposes, we replaced the lowest cost (C; free) with \$0.01. Note that this approach would affect curve fitting if Q_0 was unconstrained (i.e., a fitted parameter).

The inverse of parameter α is termed "*essential value*" because it quantifies the reinforcing efficacy of the item or product being consumed across a range of prices. Items or commodities with a lower essential value are characterized by relatively high values of α which can be seen by a relatively steep negatively accelerating function. Items or commodities with a higher essential value are characterized by relatively low values of α which can be seen by a relatively shallow negatively decelerating function. Essential value derived from the exponential demand function is especially useful as it can be used to compare reinforcing efficacy across qualitatively *different* reinforcers and different magnitudes of the *same* reinforcer.

All quantitative analyses were conducted using GraphPad Prism version 6.03 for Windows, GraphPad Software, La Jolla, California, USA (www.graphpad.com). Specifically, the exponential demand functions were obtained using The Exponential Model of Demand GraphPad Prism template available from Institutes for Behavior Resources. (www.ibrinc.org/).

Results

As an initial analysis, we sought to determine whether intensity of demand (Q_0; consumption at zero cost or free) differed as a function of HPT progression size. Any differences in intensity of demand across groups would be problematic considering that Q_0 was observed during the first trial of the HPT before participants experienced the independent variable of differing fuel progression sizes. Such differences would suggest heterogeneous groups and would confound any additional quantitative analyses. Median intensity of demand was equivalent across all three groups (1,000 miles) prohibiting the need for a formal statistical evaluation. We also compared individual participants' intensity of demand across the three groups. In this analysis, and all subsequent analyses in the study, we used the Kruskal-Wallis nonparametric analysis of variance if at least one group featured non-normally distributed data (according to the D'Agostino and Pearson omnibus normality test), unless otherwise stated (i.e., relevant data were normally distributed). There were no significant differences between groups with respect to intensity of demand (H=3.87, p=0.14). Thus, intensity of demand appears to be reliable across the different HPT progression sizes. The lack of differentiation between groups at the initial cost (free at the first trial) indicates that the three groups were relatively homogeneous before experiencing contact with the fuel progression prices.

Median consumption (in miles driven) values were extracted from each of the three groups for use in Eq. 1 model fitting. The left panel of Fig. 1 depicts consumption in miles as a function of fuel price per gallon across the three progression sizes. A general and systematic decrease in consumption as fuel prices increased was observed for all three groups. Across all three groups, elasticity (α) was comparable. For the \$0.25 group, $\alpha=5.2\cdot10^{-5}$ (R^2 for overall model fit = 0.60), for the \$0.50 group, $\alpha=6.0\cdot10^{-5}$ (R^2=0.68), and for the \$1.00, α= $4.9\cdot10^{-5}$ (R^2=0.90). We compared differences in demand elasticity across the three groups using the sum of squares F test (due to normally distributed data for all groups). No significant differences were found (F [2,49] = 0.65, p=0.52). To determine if there were any differences in demand elasticity at the individual level, we obtained individual best fit values of α and compared them across the three groups using the Kruskal-Wallis test. Results indicated there were no significant differences across groups (H=2.095, p=0.35); see the right panel of Fig. 1 for a visual representation of the differences in elasticity across groups. These results suggest that HPTs of differing progression sizes yield similar levels of elasticity when analyzing both aggregate median and individual values of consumption.

Given that all participants were exposed to whole dollar prices of fuel (\$1.00, \$2.00, \$3.00, etc.), we analyzed participants' responses within each progression size at each \$1.00 increment (beginning at \$0.01 [free]) by extracting the median consumption for use in Eq. 1 model fitting. Figure 2 depicts consumption (in miles driven) as a function of price at \$1.00 increments. Across all three groups, essential value was more similar than in the previous analysis and fits were better. For the \$0.25 group, $\alpha=5.1\cdot10^{-5}$ (R^2=0.76) and for the \$0.50 group, $\alpha=5.4\cdot10^{-5}$ (R^2=0.78). Results from the \$1.00 progression group remained at $\alpha=4.9\cdot10^{-5}$ (R^2=0.90). We compared these differences in demand elasticity across the three groups and no significant differences were found (F [2,22] = 0.15, p= 0.86). Additionally, we compared participants' individual demand elasticity values across the three groups when \$1.00 increments were extracted. Results from the Kruskal-Wallis test yielded no significant differences (H=2.09, p=0.35); see the right panel of Fig. 2 for a visual representation of the differences in elasticity across groups when using only the \$1.00 increment values. This analysis suggests that the \$1.00 progression size is an efficient and valid means of assessing demand across a range of prices and yields estimates of demand that do not differ from model fits using costs between and including whole dollar amounts.

In addition to comparing how demand elasticity differed as a function of progression size, we also examined to what extent normalized P_{max} values differed. To obtain P_{max} values,

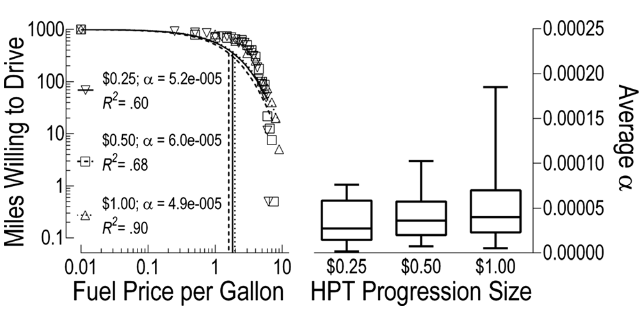


Fig. 1 *Left panel* depicts exponential demand for recreational driving mileage for each of the three progression size groups. *Vertical lines* correspond to normalized P_{max} (see text for details). *Right panel* depicts box and whisker plots (5–95 percentile) of elasticity values (α) obtained from individual participants' from all three progression size groups

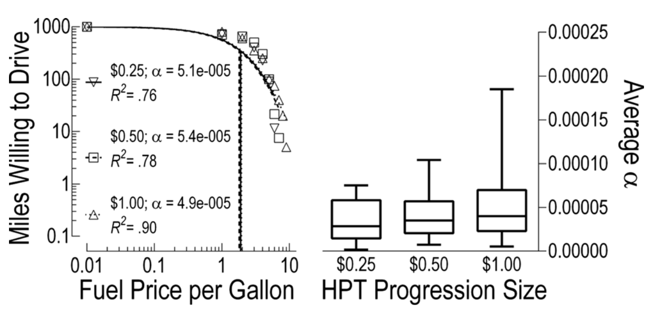


Fig. 2 Exponential demand data from all three progression size groups using only \$1.00 price increments. *Left panel* depicts exponential demand for recreational driving mileage for each of the three groups. *Vertical lines* correspond to normalized P_{max} (see text for details). *Right panel* depicts box and whisker plots (5–95 percentile) of elasticity values (α) obtained from individual participants' from all three progression size groups. Note that the data comprising the elasticity values for the \$1.00 progression size groups are identical to those depicted in Fig. 1. Despite the similarities between the box and whisker plots in Figs. 1 and 2, these data are indeed slightly different for the \$0.25 and \$0.50 progression size groups

we used the following normalized P_{max} equation (Eq. 2; Hursh in press):

$$P_{max} = m \Big/ \left(Q_0 \cdot \alpha \cdot k^{1.5}\right) \quad (2)$$

where m=0.083k+0.65. The remaining parameters are identical to those used in Eq. 1. Normalized P_{max} is derived from Hursh and Silberberg's (2008) exponential model of demand and provides a quantitative description of the point at which consumption changes from inelastic to elastic.

At the group level, using median values, normalized P_{max} values were 1.83, 1.59, and 1.94 for the \$0.25, \$0.50, and \$1.00 progression size groups, respectively. When \$1.00 progression size increments were isolated for analysis, median group normalized P_{max} values were 1.87 and 1.76 for the \$0.25 and \$0.50 groups, respectively (normalized P_{max} values for the \$1.00 group remained at 1.94). No significant differences in normalized P_{max} values between the three progressions size groups were detected when analyzing across participants' P_{max} values (H=3.83, p>0.14). When \$1.00 progression size increments were isolated and normalized P_{max} values were obtained at the individual level, we again found no significant differences between the groups (H=4.50, p=0.10). The left and right panels of Fig. 3 depict differences in P_{max} when using all fuel prices and those obtained using analyses isolated to \$1.00 increments, respectively. Collectively, these data suggest that demand assessed via each HPT was unaffected by progression size amounts.

As a final measure of demand, we evaluated the breakpoint—the smallest price value that yields zero consumption (i.e., miles willing to drive)—for each participant in each of three progression size groups. Note that this breakpoint analysis is independent of any theoretical model of demand functions (i.e., there is no use of nonlinear regression and curve-fitting to derive the values). Breakpoints were derived first using all experienced fuel prices and then evaluated using only the \$1.00 increments for each group. In each analysis, a number of participants displayed some willingness to drive at the highest fuel price (\$15.00 per gallon), thereby preventing analysis of a breakpoint (five participants from the

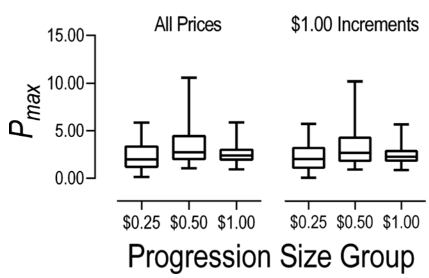


Fig. 3 Box and whisker plot (5–95 percentile) of normalized P_{max} values for each progression size using all prices (*left panel*) and prices isolated at \$1.00 increments (*right panel*). Note that the data for the \$1.00 progression size group are identical in both panels and presented concurrently to permit comparison to the other groups in both price conditions

Fig. 4 Box and whisker plot (5–95 percentile) of breakpoints for each progression size using all prices (*left panel*) and prices isolated at \$1.00 increments (*right panel*). Note that the data for the \$1.00 progression size group are identical in both panels and presented concurrently to permit comparison to the other groups in both price conditions

\$0.25 group, and seven from the \$0.50 and \$1.00 groups). Participants with no demonstrated breakpoint were excluded from subsequent breakpoint analyses. Using all experienced prices, median breakpoints were \$6.00, \$7.00, and \$9.00, for the \$0.25, \$0.50, and \$1.00 progression size groups, respectively; however, group differences were not significant (H=3.38, p=0.19; see left panel of Fig. 4). When \$1.00 progression size increments were isolated for analysis, median breakpoints were \$6.00, \$6.00, and \$9.00, for the \$0.25, \$0.50, and \$1.00 progression size groups, respectively, again with no significant differences between groups (H=2.77, p=0.25; see right panel of Fig. 4). These data suggest that behavioral economic measures of demand independent from the exponential model yield reliable findings.

Discussion

Demand for recreational driving mileage was assessed in three groups of undergraduate students across a range of fuel prices. Fuel price progressions differed across the three groups to determine the inter-method reliability of the HPT. The inter-method reliability of the HPT was adequate across each level of analysis for each demand parameter, extending the literature on the reliability of the HPT beyond temporal stability (Few et al. 2012; Murphy et al 2009). Specifically, elasticity of demand (α), P_{max}, and intensity of demand (Q_0) were reliable across all three progression sizes using both median and individual participants' best-fit parameters. However, model fits were generally better using the larger progression size (median R^2 values for the \$0.25, \$0.50, and \$1.00 were 0.87 [IQR=0.05], 0.82 [IQR=0.07], and 0.93 [IQR=0.02], respectively), but variability in elasticity was greater (as evidenced in Fig. 1).

Beyond adding to the growing literature on the psychometric properties of behavioral economic HPTs for demand analyses, the present study suggests that such HPTs are amenable for products outside the context of substance use. At both the median and individual levels, the exponential model of demand adequately fit the HPT data and indicated some degree of elasticity. This study indicates that whole dollar prices yield demand curves equivalent to those with a larger range of costs including partial dollar amounts. The high degree of reliability across these analyses indicates that the whole dollar progression is an efficient solution to quickly assess demand for fuel in the context of recreational driving. These findings address Hursh and Roma's (2013) call for novel demand curve research of public policy concerns and open the door for future investigations on ways to modulate demand using behavioral economic concepts. Such behavioral economic solutions could inform eventual public policy decisions affecting the entire population of drivers. The ability to quickly obtain reliable behavioral economic metrics of demand results may also reduce participant fatigue—associated with a potentially long and redundant task (e.g., answering 60 near identical questions in the $0.25 progression size)—which may, in turn, result in more accurate and valid responses. While this theory was not formally investigated in this study, our finding that larger progression sizes yielded better model fits may be an indication to support this hypothesis; future research should explicitly explore this notion. Finally, the present study suggests that the HPT may be a viable solution to real-world problems for which actual price ranges are insufficient in detecting demand elasticity. In the present case, literature on the demand for fuel appears strongly inelastic using governmental data on actual prices in both economic (Havranek et al. 2012) and behavioral economic (Reed et al. 2013a, b) analyses. As suggested by Hursh and Roma (2013), the HPT may be a way to investigate higher prices than those possibly or historically experienced in the real market.

This study represents an initial assessment of the psychometric properties of a HPT for a commodity unrelated to substances of abuse. While the inter-method reliability assessment offers some further confidence in the psychometric properties of HPTs, restricting our analyses to a novel commodity limits our confidence in the generalizability of our findings. Indeed, providing inter-method reliability of progression sizes in a cigarette or alcohol HPT (e.g., Jacobs and Bickel 1999; Murphy and MacKillop 2006) would make a much more meaningful contribution to the behavioral economic literature. We are optimistic, however, that this study provides a general demonstration of the utility of HPTs for public policy issues outside the domain of addiction or pharmacology.

Future research on the psychometric properties and external validity of the recreational driving HPT should be informed by previous work on cigarette and alcohol HPTs. For example, demonstrating differences in demand for fuel between college-aged recreational drivers and professional drivers/commuters could provide evidence of divergent validity (cf. MacKillop et al. 2008). Such findings would inform policy makers of potential differences in behavioral strategies targeting the reduction of carbon emissions associated with driving (e.g., Pritchard 2010). Furthermore, assessment of the temporal stability (i.e., test–retest reliability) of non-pharmacological HPTs would offer some degree of generalizability between HPTs of differing commodities (e.g., Few et al. 2012; Murphy et al. 2009).

In sum, the present study suggests that HPTs can be adequately translated from the behavioral pharmacology literature to the applied behavioral economics of public policy issues (see Hursh and Roma 2013). These findings open the door for many interesting simulations of behavioral economic approaches to public policy interventions that are difficult to test using real-world/experiential procedures. For example, the medical community has voiced overwhelming support for an excise tax on sugar-based beverages to drive down demand without affecting supply (cf. New York City's large soda ban; Brownell and Frieden 2009). Existing reviews of actual excise tax rate influences on soda consumption have featured too narrow a range of tax rates to yield evidence of demand elasticity (e.g., Powell et al. 2009; Sturm et al. 2010), necessitating the use of HPTs to identify therapeutic tax levels. Likewise, the passage of the Patient Protection and Affordable Care Act of 2010 mandates the inclusion of a 10 % excise tax on indoor tanning. While this approach to curbing indoor tanning in an attempt to prevent skin cancer was championed by many economists and preventative health professionals (Jain et al. 2012; Williams 2012), the tax rate was not based on empirical studies and the legislation does not stipulate whether the consumer or the salon pays the tax. A simulation of such factors using HPTs could answer these questions and serve to inform or improve public policy initiatives aimed at reducing cancer risk conferred by indoor tanning. Despite an increasingly dense literature on behavioral economic considerations surrounding substance use, the field of behavioral economics has not yet begun to meet its potential in contributing to other pressing societal issues. The refinement of HPT analogs to reliably simulate policy-level interventions is one small, yet necessary, step toward this goal.

References

Bickel, W. K., DeGrandpre, R. J., & Higgins, S. T. (1993). Behavioral economics: a novel experimental approach to the study of drug dependence. *Drug and Alcohol Dependence, 33*, 173–192. doi:10.1016/0376-8716(93)90059-Y.

Bickel, W. K., Jarmolowicz, D. P., MacKillop, J., Epstein, L. H., Carr, K., Mueller, E. T., & Waltz, T. J. (2012). The behavioral economics of reinforcement pathologies: Novel approaches to addictive disorders. In H. J. Shaffer (Ed.), *APA addiction syndrome handbook*. Washington: APA Books.

Brownell, K. D., & Frieden, T. R. (2009). Ounces of prevention—the public policy case for taxes on sugared beverages. *New England Journal of Medicine, 360*, 1805–1808. doi:10.1056/NEJMp0902392.

Carter, B. L., & Tiffany, S. T. (1999). Meta-analysis of cue-reactivity in addiction research. *Addiction, 94*, 327–340. doi:10.1046/j.1360-0443.1999.9433273.x.

Collins, R. L., Parks, G. A., & Marlatt, G. A. (1985). Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. *Journal of Consulting and Clinical Psychology, 53*, 189–200. doi:10.1037/0022-006X.53.2.189.

Few, L. R., Acker, J., Murphy, C., & MacKillop, J. (2012). Temporal stability of a cigarette purchase task. *Nicotine & Tobacco Research, 14*, 761–765. doi:10.1093/Ntr/Ntr222.

Green, L., & Freed, D. E. (1993). The substitutability of reinforcers. *Journal of the Experimental Analysis of Behavior, 60*, 141–158. doi:10.1901/jeab.1993.60-141.

Havranek, T., Irsova, Z., & Janda, K. (2012). Demand for gasoline is more price-inelastic than commonly thought. *Energy Economics, 34*, 201–207. doi:10.1016/j.eneco.2011.09.003.

Hursh, S. R. (1980). Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior, 34*, 219–238. doi:10.1901/jeab.1980.34-219.

Hursh, S. R. (1984). Behavioral economics. *Journal of the Experimental Analysis of Behavior, 42*, 435–452. doi:10.1901/jeab.1984.42-435.

Hursh, S. R. (in press). Behavioral economics and the analysis of consumption and choice. In F. K. McSweeney& E. S. Murphy (Eds.), *Wiley-Blackwell handbook of operant and classical conditioning.* Oxford: Wiley-Blackwell.

Hursh, S. R., & Roma, P. G. (2013). Behavioral economics and empirical public policy. *Journal of the Experimental Analysis of Behavior, 99*, 98–124. doi:10.1007/s00213-008-1120-0.

Hursh, S. R., & Silberberg, A. (2008). Economic demand and essential value. *Psychological Review, 115*, 186–198. doi:10.1037/0033-295x115.1.186.

Hursh, S. R., & Winger, G. (1995). Normalized demand for drugs and other reinforcers. *Journal of the Experimental Analysis of Behavior, 64*, 373–384. doi:10.1901/jeab.1995.64-373.

Hursh, S. R., Raslear, T. G., Shurtleff, D., Bauman, R., & Simmons, L. (1988). A cost-benefit-analysis of demand for food. *Journal of the Experimental Analysis of Behavior, 50*, 419–440. doi:10.1901/jeab.1988.50-419.

Hursh, S. R., Raslear, T. G., Bauman, R., & Black, H. (1989). The quantitative analysis of economic behavior with laboratory animals. In K. G. Grunert & F. Olander (Eds.), *Understanding economic behavior* (pp. 117–165). New York: Kluwer Academic Publishers, Theory and Decision Library.

Hursh, S. R., Madden, G. J., Spiga, R., DeLeon, I. G., & Francisco, M. T. (2012). *The translational utility of behavioral economics: The experimental analysis of consumption and choice*. Washington: American Psychological Association.

Jacobs, E. A., & Bickel, W. K. (1999). Modeling drug consumption in the clinic using simulation procedures: demand for heroin and cigarettes in opioid-dependent outpatients. *Experimental and Clinical Psychopharmacology, 7*, 412–426. doi:10.1037/1064-1297.7.4.412.

Jain, N., Rademaker, A., & Robinson, J. K. (2012). Implementation of the federal excise tax on indoor tanning services in Illinois. *Archives of Dermatology, 148*, 122–124. doi:10.1001/archderm.148.1.122.

Johnson, M. W., & Bickel, W. K. (2006). Replacing relative reinforcing efficacy with behavioral economic demand curves. *Journal of the Experimental Analysis of Behavior, 85*, 73–93. doi:10.1901/jeab.2006.102-04.

Ko, M. C., Terner, J., Hursh, S., Woods, J. H., & Winger, G. (2002). Relative reinforcing effects of three opioids with different durations of action. *Journal of Pharmacology and Experimental Therapeutics, 301*, 698–704. doi:10.1124/jpet.301.2.698.

Laibson, D. (2001). A cue-theory of consumption. *Quarterly Journal of Economics, 116*, 81–119. doi:10.1162/003355301556356.

MacKillop, J., & Murphy, J. G. (2007). A behavioral economic measure of demand for alcohol predicts brief intervention outcomes. *Drug and Alcohol Dependence, 89*, 227–233. doi:10.1016/j.drugalcdep.2007.01.002.

MacKillop, J., Murphy, J. G., Ray, L. A., Eisenberg, D. T. A., Lisman, S. A., Lum, J. K., & Wilson, D. S. (2008). Further validation of a cigarette purchase task for assessing the relative reinforcing efficacy of nicotine in college smokers. *Experimental and Clinical Psychopharmacology, 16*, 57–65. doi:10.1037/1064-1297.16.1.57.

MacKillop, J., O'Hagen, S., Lisman, S. A., Murphy, J. G., Ray, L. A., Tidey, J. W., & Monti, P. M. (2010). Behavioral economic analysis of cue-elicited craving for alcohol. *Addiction, 105*, 1599–1607. doi:10.1111/j.1360-0443.2010.03004.x.

MacKillop, J., Brown, C. L., Stojek, M. K., Murphy, C. M., Sweet, L., & Niaura, R. S. (2012a). Behavioral economic analysis of withdrawal- and cue-elicited craving for tobacco: an initial investigation. *Nicotine & Tobacco Research, 14*, 1426–1434. doi:10.1093/Ntr/Nts006.

MacKillop, J., Few, L. R., Murphy, J. G., Wier, L. M., Acker, J., Murphy, C., & Chaloupka, F. (2012b). High-resolution behavioral economic analysis of cigarette demand to inform tax policy. *Addiction, 107*, 2191–2200. doi:10.1111/j.1360-0443.2012.03991.x.

Madden, G. J., & Hartman, E. C. (2006). A steady-state test of the demand curve analysis of relative reinforcer efficacy. *Experimental and Clinical Psychopharmacology, 14*, 79–86. doi:10.1037/1064-1297.14.1.79.

Madden, G. J., & Kalman, D. (2010). Effects of bupropion on simulated demand for cigarettes and the subjective effects of smoking. *Nicotine & Tobacco Research, 12*, 416–422. doi:10.1093/Ntr/Ntq018.

Murphy, J. G., & MacKillop, J. (2006). Relative reinforcing efficacy of alcohol among college student drinkers. *Experimental and Clinical Psychopharmacology, 14*, 219–227. doi:10.1037/1064-1397.14.2.219.

Murphy, J. G., MacKillop, J., Skidmore, J. R., & Pederson, A. A. (2009). Reliability and validity of a demand curve measure of alcohol reinforcement. *Experimental and Clinical Psychopharmacology, 17*, 396–404. doi:10.1037/a0017684.

Murphy, J. G., MacKillop, J., Tidey, J. W., Brazil, L. A., & Colby, S. M. (2011). Validity of a demand curve measure of nicotine reinforcement with adolescent smokers. *Drug and Alcohol Dependence, 113*, 207–214. doi:10.1016/j.drugalcdep.2010.08.004.

Patient Protection and Affordable Care Act, Pub. L. No. 111-148, §2702, 124 Stat. 119, 318–319 (2010).

Powell, L. M., Chriqui, J., & Chaloupka, F. J. (2009). Associations between state-level soda taxes and adolescent body mass index. *Journal of Adolescent Health, 45*, S57–S63. doi:10.1016/j.jadohealth.2009.03.003.

Pritchard, J. (2010). Virtual rewards for driving green. *Behavior Analyst, 33*, 185–187.

Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. (1976). Economic demand theory and psychological studies of choice. In G. Bower (Ed.), *The psychology of learning and motivation* (vol. 10, pp. 129–154). New York, NY: Academic Press.

Reed, D. D., Luiselli, J. K., Magnuson, J. D., Fillers, S., Vieira, S., & Rue, H. C. (2009). A comparison between traditional economical and demand curve analyses of relative reinforcer efficacy in the validation of preference assessment predictions. *Developmental Neurorehabilitation, 12*, 164–169. doi:10.1080/17518420902858983.

Reed, D. D., Niileksela, C., & Kaplan, B. A. (2013a). Behavioral economics: a tutorial for behavior analysts in practice. *Behavior Analysis in Practice, 6*, 34–54.

Reed, D. D., Partington, S. W., Kaplan, B. A., Roma, P., & Hursh, S. (2013b). Behavioral economic analyses of demand for fuel in North

American transportation. *Journal of Applied Behavior Analysis, 46*, 651–655. doi:10.1002/jaba.64.

Roane, H. S., Lerman, D. C., & Vorndran, C. M. (2001). Assessing reinforcers under progressive schedule requirements. *Journal of Applied Behavior Analysis, 34*, 145–166. doi:10.1901/jaba.2001.34-145.

Shafir, E. (2013). *The behavioral foundations of public policy*. Princeton: Princeton University Press.

Shahan, T. A., Bickel, W. K., Madden, G. J., & Badger, G. J. (1999). Comparing the reinforcing efficacy of nicotine containing and denicotinized cigarettes: a behavioral economic analysis. *Psychopharmacology, 147*, 210–216. doi:10.1007/s002130051162.

Spiga, R., Martinetti, M. P., Meisch, R. A., Cowan, K., & Hursh, S. (2005). Methadone and nicotine self-administration in humans: a behavioral economic analysis. *Psychopharmacology, 178*, 223–231. doi:10.1007/s00213-004-2020-6.

Sturm, R., Powell, L. M., Chriqui, J. F., & Chaloupka, F. J. (2010). Soda taxes, soft drink consumption, and children's body mass index. *Health Affairs, 29*, 1052–1058. doi:10.1377/hlthaff.2009.0061.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol, 50*, 30–37.

Williams, S. N. (2012). A tax on indoor tanning would reduce demand in europe. *British Medical Journal, 345*, e6550. doi:10.1136/bmj.e6550.