Do Rational Demand Estimates Differ From Irrational Ones? Evidence from an Induced Budget Experiment

Samiran Banerjee and James H. Murphy*

August 31, 2007

Abstract

Both early and recent work have highlighted certain similarities between rational and irrational demand. We re-examine these findings using experimental choice data. After separating our subjects' choices into rational and irrational subsets based on consistency with the axioms of revealed preference, we estimate and compare demand coefficients from the resulting subsamples, finding significant differences between the two. We also predict consistency based on socio-demographics and cognitive ability, then split the sample using predicted consistency and again estimate and compare the resulting subsamples' demand coefficients. These comparisons indicate differences between rational and irrational demand and are largely consistent with successful prediction.

JEL Categories: D01, D11, D12, C93

Keywords: Rationality, Demand, Experiments, Revealed Preference, WARP

*Banerjee (corresponding author): Department of Economics, Emory University, Atlanta, GA 30322. Phone: 404-712-8168. Fax: 404-727-4639. Email: sbaner3@emory.edu. Murphy: Department of Economics, Richards College of Business, University of West Georgia, Carrollton, GA 30118. Phone: 678-839-4774. Fax: 678-839-5041. E-mail: jmurphy@westga.edu. We are grateful to Deborah J. Webb and participants at the 2006 North American ESA Meeting for their comments.

1 Introduction

Many forms of behavior, such as the random choice of Becker (1962) or the anchored locally coherent responses of Ariely et al. (2003), yield the expected inverse relationship between price and planned consumption, but may not be considered "rational" in the neoclassical sense. Predictive power rather than the full force of Hicksian neoclassical rationality may suffice for some applications (e.g., marketing researchers gauging consumer demand for a new product). Under such circumstances, applied practitioners may well question the necessity of the added efforts advocated by behavioralists (e.g., Ho et al., 2006), especially when the added variables required to construct behavioral models introduce new complexities and possibly reduce their tractability (Camerer and Loewenstein, 2004, 4).

In this paper we examine this "observational equivalence" using non-parametric rationality (i.e., revealed preference) tests to distinguish between preference-consistent and preference-inconsistent demand. As non-trivial rationality tests require choice from budget sets in which both price and endowment vary (see Bronars, 1987), our experimental subjects' responses consist of choices from experimenter-induced endowments and prices of a private good. We term such experiments as *induced budget* experiments.[1] The resulting data is directly testable for Hicksian preference consistency based on whether a subject's responses satisfy an appropriate axiom of revealed preference.

Our statistical analysis takes several tacks. After splitting the sample on the basis of preference consistency, we estimate demand for both subsamples and compare regression coefficients. We find that both subsamples show consistency with the "Law of Demand", but the data-generating process is different across the subsamples. Next, we construct and estimate a rudimentary, binary model of preference consistency based on socio-demographic data and cognitive ability and find that this model is as effective at predicting preference-consistency as running the actual revealed preference test. Finally, we parse the sample based on these predictions and compare the resulting demand estimates. Our results show that the data-generating process

[1] To contrast them from *induced value* experiments, i.e., those in which demand and supply are determined by the experimenter and the object of interest is the performance of an allocation mechanism. See Sippel (1997) and Andreoni and Miller (2002) for examples of other induced budget experiments.

between the preference-consistent versus predicted preference-consistent subsamples (and, likewise, between the preference-*in*consistent versus predicted preference-*in*consistent subsamples) are similar, but significantly different between the predicted preference-consistent and preference-inconsistent subsamples.

Section 2 below outlines revealed preference theory, formalizing the notion of rationality we consider. The experiment description is in Section 3 with a summary of the data collected in Section 4. The statistical analysis is in Section 5, followed by a discussion of our findings in Section 6.

2 Revealed Preference Theory

For the purposes of empirical testing, we seek to determine whether responses are consistent with the maximization of an underlying preference relation (or possibly, a utility function) on budgets. Revealed preference theory provides such tests. One is the Strong Axiom of Revealed Preference (SARP) originally due to Houthakker (1950), which tests for the existence of a binary relation (see Richter, 1966)—or a strictly concave utility function on finite data sets (see Matzkin and Richter, 1991)—which when maximized subject to a set of binding budgets yields *exactly* the set of observed choices. We then say the binary relation or utility function *rationalizes* the observed choices.[2]

The exposition below follows Matzkin and Richter (1991). Let the commodity space, X, be a convex subset of $\Re^k_+$, where k indexes the number of

[2]A second revealed preference test, the Generalized Axiom of Revealed Preference (GARP)—see Afriat (1967), Diewert (1973) and Varian (1982)—tests for the existence of a utility function which when maximized subject to a set of binding budgets yields a set of consumption bundles *of which the observed choices are a subset*. Under these conditions, we say the underlying utility function *subsemirationalizes* the observed choices. See Matzkin and Richter (1991, 299) for a comparison of the two axioms.

Unlike the SARP, which allows demand functions, the GARP admits demand correspondences. The GARP is weaker than the SARP: data consistent with the SARP will be GARP-consistent but not necessarily vice versa (see Richter, 1987, 169). Interpolating a continuous demand *function* using observed demand points requires consistency with the SARP (Matzkin and Richter, 1991, 301). When choice data satsifies the GARP but not the SARP, interpolating a demand function constitutes a specification error since the theoretical construct (a function) and the data used to estimate that construct (a correspondence) do not share the same underlying structure. Consequently, we focus on the SARP for the remainder of the paper.

commodities in X, and B represents a consumer's budget in $\Re_+^k$. A budget is the set of all possible consumption bundles an individual can afford given prevailing circumstances. Let $\mathcal{C}$ represent the set of all possible budgets. Furthermore, let $\mathcal{B}$ represent a finite set of budgets (i.e., $B \in \mathcal{B} \subset \mathcal{C}$). If the budget is determined by parametric prices and income, the budget is said to be *competitive* and denoted as

$$B(p, m) = \{x \in X : \; p \cdot x \leq m\}$$

where $p \in \Re_+^k$ and $m \in \Re_+$. Let $h(B)$ be the observed choice behavior of a consumer subject to budget B. Matzkin and Richter (1991) term choice behavior which completely exhausts the budget on each observed choice occasion as *exhaustive*, i.e., $p \cdot x = m$ for all $x \in h(B)$ and all $B(p, m) \in \mathcal{B}$.

If $h(B)$ maps a given budget to a single consumption bundle, then h is *single-valued* (a choice function); otherwise, h is *multi-valued* (a choice correspondence). The following definitions restate more formally the relationships between observed choices and the existence of either binary relations or utility functions rationalizing those choices.

Definition 1. A choice function $h(B)$ is *rational* if there is a binary relation R defined on X such that for all $B(p, m) \in \mathcal{B}$, it is true that

$$h(B) = \{x \in B(p, m) : \; xRy \text{ for all } y \in B(p, m)\}.$$

Observed choices are *regular rational* if they are rationalizable by a reflexive, total and transitive binary relation;[3] *utility rationality* refers to rationalization by a real-valued function $u : X \to \Re$ such that

$$h(B) = \{x \in B(p, m) : \; u(x) \geq u(y) \text{ for all } y \in B(p, m)\}$$

for all $B \in \mathcal{B}$. Matzkin and Richter (1991) define *special rationality* as rationalizability by a continuous, strictly concave and strictly monotonic utility function. We now introduce the axioms of revealed preference.

Definition 2. Let a binary relation S on X be defined for all x and y in X as:

$$xSy \iff \exists B_{B \in \mathcal{B}} \; x \in h(B), x \neq y, x, y \in B.$$

[3] A binary relation, R, is reflexive if for every $x \in X$, xRx; total if for any $x, y \in X$ with $x \neq y$, either xRy, or yRx, or both; and transitive if for any $x, y, z \in X$, xRy and yRz imply xRz.

Then a choice function h satisfies the *Weak Axiom of Revealed Preference* (WARP) if S is asymmetric, i.e., $xSy \Rightarrow \neg(ySx)$.

Thus xSy denotes that x was (directly) chosen over another bundle y when both x and y were available in some budget B. The WARP states that if x was chosen when y was available, then it is never the case that y will be chosen when x is available.

Definition 3. Let the binary relation H be the transitive closure of S, i.e., xHy iff there is a sequence $z_1, \ldots, z_n$ in X such that $xSz_1Sz_2S...Sz_nSy$. Then a choice function h satisfies the *Strong Axiom of Revealed Preference* (SARP) if H is asymmetric, i.e., $xHy \Rightarrow \neg(yHx)$.

The SARP embodies the notion that if bundle x is chosen over bundle y, either directly (xSy) or indirectly (xHy), then y will never be directly or indirectly chosen over x.

We can now state the results established by Matzkin and Richter (1991, 290-291):

Theorem 1: *Let h be an exhaustive choice function defined over a finite subset $\mathcal{B} = \{B(p_1, m_1), \ldots, B(p_n, m_n)\}$ of $\mathcal{C}$. Then the following statements are equivalent:*

(a) h satisfies the SARP;

(b) h is special rational, i.e., there exists a continuous, strictly concave, and strictly monotone function u rationalizing h on $(X, \mathcal{B})$, and furthermore, u can be chosen to be defined on all $\Re^k_+$ and be generically C^∞;

(c) h is regular rational.

Thus, if a consumer exhausts her available income in all observed choice situations over finitely many competitive budgets and her choices satisfy the SARP, then this consumer's choices are rationalizable by a reflexive, total and transitive preference relation. Moreover, this consumer's preference relation can be represented by a continuous, strictly concave and strictly monotonic utility function.

Rose (1958) showed that when choices are made over two goods, testing observed choices for consistency with the SARP requires only testing for WARP consistency, i.e., the WARP and the SARP are equivalent in the two-good case. Therefore, when the choice set consists of two goods, Rose's Theorem and Theorem 1 imply that the WARP serves as an operational test for the regular rationality of single-valued choice.[4]

3 Experiment

Subjects were recruited from undergraduate economic principles and graduate MBA courses at Georgia State University. The experiment was conducted manually in classrooms. Prior to participating, all subjects read, signed and dated a consent form, returned this form to the experimenter, and received a $2.00 participation fee.[5]

A total of 194 subjects (127 undergraduate and 67 graduate) were recruited in seven separate sessions. A random draw separated subjects into two groups, and the data we analyze in this paper resulted from the choices of one of these two groups which, after excluding computational errors, contained 69 subjects.[6] All subjects in this group received a set of written instructions, which were also read aloud to them. Subjects, facing the 10 budget scenarios listed in Table I below, indicated how many Hershey's Kisses candies they would purchase from each Table I budget.

[Insert Table I about here.]

Subjects received a written description of this good prior to starting the choice task and were instructed that they could buy the candy in whole units

[4]Exhaustive choice behavior over two-commodity budgets which does not satisfy the WARP cannot be rationalized by a reflexive, total and transitive binary relation. But this does not preclude rationalizability by binary relations in general. Kim and Richter (1986) allow us to strengthen Theorem 1 above by providing conditions that single-valued observed choice must meet in order to be rationalized by *any* binary relation in the two-good case. If a choice function is continuous and satisfies the Desirability Hypothesis—roughly, the requirement that the demand bundle is unbounded when one of the prices goes to zero—then applying their Proposition 1 and Theorem 7 (see Kim and Richter, 1986, 332 and 336) leads to the conclusion that observed choices that do not satisfy the WARP cannot be rationalized by *any* binary relation. Thus we have a stark contrast: demand that satisfies the WARP can be rationalized by a nice utility function while demand that violates the WARP cannot be rationalized by any preferences whatsoever.

[5]All relevant materials used in the experiment are available upon request from the authors.

[6]The two groups were separated in different classrooms for the remainder of the experiment to reduce the potential for contamination across treatments.

only.[7] We selected this good because of its properties as a *private* good—it is both rival and excludable—and, consequently, any observed deviations from rationality are unlikely to invoke the critique of survey methods of Diamond et al. (1993) as applied to *public* goods. In addition, we sought a good that was (1) familiar to subjects to increase the likelihood of well-formed preferences, (2) of small value to reduce potential income effects, and (3) whose per-unit price would not be one readily observed outside the experimental environment.[8]

Any portion of the endowment not used for purchases was converted to cash at a fixed rate (10 cents per token). After all subjects completed all ten choices, a subject volunteer drew a number from one to ten, and subjects received the actual consumption bundle they chose in the correspondingly numbered scenario, e.g., if the number six were drawn, all subjects in that session received the bundle they chose from the sixth budget they faced.[9] Thus, the linkage between subject response and experimental payoff is both explicit and tangible, increasing the likelihood of salient responses and experimental control (Smith, 1982). All subjects also recorded the time they began and ended their choices using a large clock placed at the front of the room, then completed a cognitive ability test[10] followed by a socio-demographic

[7]The description read:

> "The Good you may purchase, *here and now*, is a Hershey's Kisses candy. Each piece is individually wrapped in foil and contains approximately 5 grams of milk chocolate. Milk chocolate contains sugar, milk, cocoa butter, chocolate, soya lecithin, an emulsifier, and vanillin.
>
> Since each piece is individually wrapped, you may only buy Hershey's Kisses in whole units, for example one piece, two pieces, six pieces or eighteen pieces, not one and one third pieces or two-thirds of a piece."

The technical information concerning the item's contents paraphrases that listed on the manufacturer's packaging.

[8]This last point is driven by a desire to avoid loss of experimental control through *field price censoring* (see Harrison et al., 2004), which occurs when a subject's responses in the experimental environment are altered by knowledge of the price of the same good in the field. At the time of the experiment, this good was sold predominantly in half or one pound bags; consequently, neither a per-piece price nor a secondary market for individual candies would have been readily available to our subjects.

[9]See Hey and Lee (2005) for a discussion of the incentive properties of this random lottery mechanism and Hey and Orme (1994) for an example of its application to choice experiments.

[10]We administered the Wonderlic Personnel Test (WPT), a timed, 50-item scale in

survey.

To reassure subjects the procedures were carried out as described, one subject (designated the auditor) transported subjects' sealed choices to off-site research assistants; observed the research assistants as they checked subjects' responses for compliance with the experiment's rules and provided the appropriate payment; transported the sealed envelopes containing payment back to the experiment site; and verified the procedures had been carried out as described.[11,12]

4 Data

Addressing our research questions requires information concerning subjects' demand, the WARP consistency of their choices and socio-demographics as well as a number of other controls.

4.1 Demand

Note that each budget's price and endowment were set prior to the experiment (see Table I) and were thus exogenous to the consumer. Subjects successfully following the procedures described in Section 3 completely exhausted their endowment as any portion not employed to purchase candies was redeemed for cash. Consequently, the choice behavior we analyze is both competitive and exhaustive, as Theorem 1 requires.

For a given budget and consumer, let the variable **HK**, a non-negative integer, denote the number of candies purchased; **PRICE** represent the normalized per-unit price of a candy and **ENDOW** be the available endowment. The specific combinations of **PRICE** and **ENDOW** which comprised the experiment's budgets are those listed in Table I—all subjects faced all ten

which subjects provide written answers to a series of verbal, quantitative and spatial items. Dodrill (1981) found that age-adjusted WPT scores are comparable to the Wechsler Adult Intelligence Scale, a common test of cognitive ability (IQ). See the Wonderlic Personnel Test User's Manual (1992) for details.

[11] This verification procedure is due to Andreoni and Miller (2002).

[12] After completing this series of 10 choices, subjects in this treatment also completed a second set of choices; however, in this paper, we report only the first set. Subjects were not aware a second series would be administered when completing the first series; thus, we are confident the choices we analyze here were not affected by subsequent procedures.

Table I budgets, though the order differed (a random draw determined the sequence for each subject).

4.2 WARP Consistency

We constructed an algorithm which reports a WARP violation for each instance of symmetry of the S-relation (see Definition 2) and coded the variable **WARP** as 0 if a subject's responses contained one or more WARP violations. If a subject's responses contained no WARP violations, we coded **WARP** 1. Hence, if **WARP** equals zero, the corresponding choices are WARP inconsistent whereas a value of one represents WARP-consistent choice.

4.3 Socio-demographics

The variable **AGE** represents each subject's self-reported age. To incorporate ethnicity, we coded **ETHN** 0 for subjects reporting Caucasian ethnicity and 1 otherwise. We coded **GENDER** 1 for male subjects and 0 if female. Subjects indicated their before-tax household income by checking the appropriate income range (e.g., up to \$9,999, \$10,000 to \$19,999, ..., \$100,000 or more), and we report the midpoint of this range for each subject as **HHINC**. To proxy educational attainment, we coded **MBA** 1 if a subject was an MBA student (and consequently had previously received a bachelor's degree) and 0 otherwise. For married subjects we coded **MRRD** 1 and 0 otherwise.

4.4 Control Variables

Following guidelines in the Wonderlic Personnel Test User's Manual (1992), we constructed **AACOG**, each subject's age-adjusted Wonderlic Personnel Test score as a proxy for cognitive ability. The time which elapsed while an individual completed her decisions is represented by **TIME** (in minutes to the hundredth decimal place). To control for possible interviewer effects, we coded **INT** 0 for sessions conducted by Experimenter One and 1 for sessions conducted by Experimenter Two. Finally, while we randomized the sequence of budgets for each subject, we also constructed a series of nine dummy variables, **PER1** through **PER9**, to control for any residual ordering effects. Following experimental practice, we the take the last budget encountered to be the reference point; hence, we omit **PER10**.

Table II below contains descriptive statistics for both dummy (part a) and integer-valued and continuous (part b) variables.

[Insert Table II about here.]

5 Statistical Analysis and Results

5.1 Do demand estimates from rational and irrational choices differ?

We looked for differing data-generating processes between WARP-consistent or WARP-inconsistent subsamples by estimating the demand for candies for both subsamples and testing for significant differences between the two sets of coefficients. First, we split the sample based on WARP consistency (**WARP** = 1) or inconsistency (**WARP** = 0). The dependent variable, the number of candies selected (**HK**), takes only non-negative integer values and may be overdispersed. Moreover, observed zero values could result from two data-generating processes: (1) the consumer never purchases this product regardless of economic circumstances (e.g., she may be allergic to chocolate or dislike this particular product), or (2) the budget she faced was not conducive to purchase on the observed occasion, but she might purchase a positive amount given a more favorable budget. Consequently, we estimated Negative Binomial (NB) or Zero Adjusted Negative Binomial (ZANB) regressions in which the number of candies a subject chooses from a given budget is a function of the budget's price (**PRICE**), endowment (**ENDOW**), order (**PER*i***) and the socio-demographic and control variables described above.[13]

For the NB model, let w_i represent a vector of explanatory variables for consumer i, μ_i represent the corresponding exponential conditional mean, i.e., $\mu = e^{\beta' w}$ (subscripts omitted), and α capture the variance's deviation from the conditional mean. Specifically, if the mean's variance, ν, is a quadratic function of the mean itself, we have the NB2 model of Cameron and Trivedi (1986) where $\nu_i = \mu_i + \alpha\mu_i^2$ and the probability that a consumer selects j candies is

$$\text{prob}(\mathbf{HK} = j|\mu, \alpha) = \frac{\Gamma(\mathbf{HK} + \alpha^{-1})}{\Gamma(\mathbf{HK} + 1)\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)^{\alpha^{-1}}\left(\frac{\mu}{\alpha^{-1} + \mu}\right)^{\mathbf{HK}}, \quad (1)$$

[13] We used a Dell Inspiron 7000 and LIMDEP 7.0 to perform the statistical analysis (Greene, 1998). Our LIMDEP code and raw data are available upon request.

where $\Gamma(\cdot)$ is the gamma function, and $\alpha \geq 0$ (indicating overdispersion). The ZANB specification is based on a splitting model representing two differing data-generating processes. Let φ represent the probability a consumer never purchases candies and $f(0)$ represent the probability a situational consumer selects no candies; then prob$(\mathbf{HK} = 0) = \varphi + (1 - \varphi)f(0)$. Furthermore, let $F(\cdot)$ represent the logistic cumulative density function. For the ZANB models we report in Table III, we captured "sometimes purchaser" behavior by equation (1) (and so $f(\cdot)$ can be modeled as NB2) and estimated φ as $F(\tau\mu_i)$. For a given μ, τ parameterizes the splitting process as larger values of τ imply larger estimates of φ.

[Insert Table III about here.]

For each model in Table III, we report the NB or ZANB regression coefficients and the overdispersion parameter α.[14] In the case of ZANB models, we also report the splitting parameter τ and Vuong's statistic, which is asymptotically standard normal (Vuong, 1989). Large positive values support the ZANB specification, large negative values favor the NB specification, and small positive or negative values represent indeterminancy. Note in Table III that Vuong's test fails to unambiguously support ZANB over NB for the WARP-inconsistent subsample. Consequently, we include both NB and ZANB regression estimates of the WARP-inconsistent data in Table III and performed our comparisons using both NB and ZANB specifications to check the robustness of our conclusions.

To assess whether the same data-generating process underlies our subsamples, we conducted a series of Wald tests comparing regression coefficients of demand for WARP-consistent and WARP-inconsistent choice. Let superscripts indicate the subsample considered; we test the null, $H_0 : \beta^{\text{WARP}} = \beta^{\text{NOTWARP}}$, versus the alternative, $H_a : \beta^{\text{WARP}} \neq \beta^{\text{NOTWARP}}$, based on various subsets of the regression results reported in Table III. First, we constructed a Wald test comparing all 22 ZANB regression coefficients from both subsamples including the regression constant and both the overdispersion parameter, α, and splitting parameter, τ (from the first and second set of estimates in Table III). The observed value, 34.25, exceeds the critical value of 33.92 of the χ^2 distribution with 22 degrees of freedom at the .05 level; con-

[14] The coefficients may be interpreted as a semi-elasticity, whereby a one unit increase of the jth regressor yields a β_j proportionate change in the conditional mean (Cameron and Trivedi, 1998, 81).

sequently, we reject the null of equality between the two sets of parameters at the .05 level. Dropping the intercept coefficient (but including α and τ) and retesting yields the same conclusion (observed value is 34.25 with 21 degrees of freedom) as does comparing NB coefficients from the WARP-inconsistent sample to the ZANB coefficients from the WARP-consistent sample.[15] We summarize the tests comparing WARP-consistent and WARP-inconsistent demand estimates below in Table IV, finding in all cases considered that the null is rejected at levels of significance of 0.10 or less.

[Insert Table IV about here]

Consequently, we conclude that the data-generating process between WARP-consistent and WARP-inconsistent choice in this data set differs.[16]

5.2 Is rationality predictable?

Let **WARP** be the dependent variable in a binomial probit model of the form

$$\text{prob}(\mathbf{WARP} = 1) = \Phi(\beta' w), \tag{2}$$

where w is a vector of explanatory variables, β represents a set of parameters measuring the effect of variations in w on the probability of consistency with the WARP and $\Phi(\cdot)$ is the standard normal cumulative distribution function. Specifying w as (1, **AACOG**, **GENDER**, **MBA**, **TIME**) yields the estimates presented in Table V.[17]

[Insert Table V about here.]

[15] Recall Vuong's statistic only weakly favored the ZANB model specification for the WARP-inconsistent subsample.

[16] As a check of whether WARP consistency affects the conditional mean, we pooled the subsamples, respecified our demand model to explicitly incorporate WARP consistency as an explanatory variable (in the guise of the variable **WARP** described in Section 4.2) and estimated a ZANB demand model using the pooled data. We test the null, $H_0 : \beta^{\text{POOLED}}_{\text{WARP}} = 0$, against the alternative, $H_a : \beta^{\text{POOLED}}_{\text{WARP}} \neq 0$, where superscripts again denote the subsample considered and subscripts indicate the specified regression coefficient. Based on the results presented in Table III's ZANB Pooled Model, we fail to reject the null at the 0.10 level; therefore, WARP consistency fails to exert a statistically significant influence on the number of candies chosen.

[17] Table V presents a parsimonious form of a more extensive model which also included **AGE**, **HHINC**, **INT** and **MRRD** as explanatory variables in addition to **AACOG**, **GENDER**, **MBA** and **TIME**. A likelihood ratio test failed to reject the hypothesis that $\beta = 0$ for **AGE**, **HHINC**, **INT** and **MRRD** (observed value = 2.904; p-value = 0.574).

In addition to the index coefficients, Table V contains the corresponding marginal effects of the dummies **GENDER** and **MBA** as well as the probability derivatives of **AACOG** and **TIME**.[18]

Table V's model is significant at the 0.10 level (the likelihood ratio test of the hypothesis that all non-constant index coefficients are jointly equal to zero has a p-value of 0.076). Letting a predicted probability greater than 50 percent denote predicted WARP consistency, the model successfully predicts WARP consistency or inconsistency for 47 of our 69 subjects (68.1%). Yet, only a subject's gender (**GENDER**) yields a statistically significant relationship with WARP consistency (p-value = 0.054 for a two-sided test), while cognitive ability (**AACOG**), educational attainment (**MBA**) and elapsed time (**TIME**) do not.[19]

To assess the relative effectiveness of the model's predictions, let the direct examination for WARP violations employed to code **WARP** constitute one method and statistical prediction constitute a second. Using Cochran's nonparametric test for related observations, we fail to reject the null that the two methods yield the same probabilities of WARP consistency (observed value = 0.182; critical value is drawn from the χ^2 distribution with one degree of freedom; $p = 0.6696$).[20] Consequently, we conclude that the statistical model is as effective in determining WARP consistency as direct inspection.[21]

[18]The formula for the reported marginal effect of **GENDER** is

$$\frac{1}{t}\sum_{i=1}^{t}\{[\Phi(\beta' w_i|\mathbf{GENDER}=1)] - [\Phi(\beta' w_i \mathbf{GENDER}=0)]\},$$

where t represents the number of observations in our sample. We used a similar formula to calculate the marginal effect of **MBA**. The marginal effect of **AACOG** was calculated as $(1/t)\beta_{\mathbf{AACOG}}\sum_{i=1}^{t}[\phi(\beta' w_i)]$, where ϕ represents the standard normal probability density function. A similar calculation was used for **TIME**.

[19]We note, however, that **TIME** is *negative* and significant at the 0.1 level for a one-sided test. This is consistent with several conjectures, e.g., a subject may have well-formed preferences and know immediately what she wants or, alternatively, she may experience a prolonged, but ultimately unsuccessful attempt to research her preferences.

[20]Versus the alternative that they differ (Conover, 1980, 199-201). More formally, let prob($WARP_i$) denote the probability that subject i's responses contain no WARP violations and pprob($WARP_i$) represent the prediction of the same event; then the null hypothesis may be stated as prob($WARP_i$) = pprob($WARP_i$) for each i, $i = 1, \ldots, t$, while the alternative hypothesis is prob($WARP_i$) $\neq$ pprob($WARP_i$) for some i.

[21]Recall that Becker (1962) modeled irrational choice as the random selection of a con-

5.3 Are demand estimates from predictions similar to those derived from rational choices?

Splitting our subjects into predicted WARP-consistent (**PREDICT** = 1) or predicted WARP-inconsistent (**PREDICT** = 0) subsamples, we estimated ZANB or NB models (Table VI) and tested for statistically significant differences in regression coefficients between: (1) predicted WARP-consistent and predicted WARP-inconsistent subsamples (Table VII), (2) WARP-consistent and predicted WARP-consistent subsamples (Tables VIII), (3) WARP-consistent and predicted WARP-inconsistent subsamples (Table IX), (4) predicted WARP-consistent and WARP-inconsistent subsamples (Table X) and (5) WARP-inconsistent and predicted WARP-inconsistent subsamples (Table XI).[22]

[Insert Tables VI through XI about here.]

These comparisons are largely consistent with successful prediction of WARP consistency or inconsistency, supporting the conclusions drawn from Cochran's test in subsection 5.2. The Wald tests fail to find differences in the data-generating processes between predicted and actual WARP-consistent subsamples (Table VIII) as well as between predicted WARP-inconsistent and WARP-inconsistent subsamples (Table XI). Moreover, the Wald tests support significant differences between the predicted WARP-consistent and predicted WARP-inconsistent subsamples (Table VII) and WARP-consistent and predicted WARP-inconsistent subsamples (Table IX). Only in the case of predicted WARP-consistent versus WARP-inconsistent subsamples (Ta-

sumption bundle from a budget. To calculate a benchmark of the probability that direct inspection distinguishes rational from irrational choice, we followed Bronars (1987) and simulated 100,000 such consumers. Each simulated consumer drew a set of ten bundles, one from each of the Table I budgets. For each budget, the simulated consumer chose how many candies to purchase (integer-valued, ranging from 0 to the maximum affordable amount for each budget in Table I) from the uniform distribution; any remaining endowment was converted to cash. The choices of 89,276 (89.28 %) of these simulated consumers violated the WARP (i.e., contained at least one WARP violation); 10,724 reported WARP-consistent choices. Thus, the WARP correctly identifies nearly 9 out of 10 instances of random choice from the budgets listed in Table I.

[22]We also pooled our data and included predicted WARP consistency as an explanatory variable. Table VI's last column reports ZANB results for the pooled sample including the indicator of predicted WARP consistency, **PREDICT**, as an explanatory variable. We fail to find a statistically significant effect of predicted WARP consistency on the number of candies chosen.

ble X) are the Wald tests ambiguous in supporting our expectations: the ZANB/ZANB tests fail to detect expected differences between coefficient subsets while the ZANB/NB tests support expected differences.

6 Discussion and Conclusion

This study yields a mixed picture regarding rational versus irrational demand. Consistent with the findings of Becker (1962) and observational equivalence, all demand models estimated yield negative price coefficients, whether based on WARP-consistent, WARP-inconsistent or pooled subsamples. Additionally, because our WARP-based rationality test distinguishes choice rationalizable by a regular preference relation from that which is not,[23] finding negative and significant price coefficients for both rational and irrational choice shows that consistency with the Law of Demand does not depend on the specific form of irrational behavior assumed (e.g., Becker's impulsive or inert consumers). Furthermore, we find that both rational and irrational choice yield a positive and significant endowment coefficient (as would be expected in the case of normal goods). Thus, the study supports and generalizes observational equivalence with regards to the price coefficient and extends it to endowment effects.

In the case of applications requiring preference consistency, e.g., calculating preference-based welfare measures (e.g., Just et al., 1982), or where behavioral modeling is unwieldy, the ability to predict rationality may prove a useful alternative. While at the early, exploratory stages, Cochran's test suggests the rudimentary model of equation (2) successfully differentiated rational from irrational response. This interpretation is further supported by Wald tests which largely conform to expectations: we failed to find significant differences between the WARP-consistent and predicted WARP-consistent subsamples (Table VIII) as well as between the WARP-inconsistent and predicted WARP-inconsistent subsamples (Table XI). Where we would expect differences, those expectations were also confirmed, as was the case for the predicted WARP-consistent and the predicted WARP-inconsistent subsamples (Table VII) and between the WARP-consistent and predicted WARP-inconsistent subsamples (Table IX). Only in the comparison between pre-

[23] Alternatively, if we invoke demand continuity and the Desirability Hypothesis—see footnote 4—our WARP test distinguishes between choice that is regular rational versus choice that cannot be rationalized by *any* binary relation.

dicted WARP-consistent versus WARP-inconsistent subsamples (Table X) did we find only partial support for expected differences.

Several caveats are in order in considering our findings. First, our subjects faced only budgets consisting of small-valued endowments and chocolate candies. Whether the results from this study can be generalized to other goods remains to be explored. Second, when a subject's choices satisfy the WARP, the results of Matzkin and Richter (1991) allow us in theory to extend the rationalizing utility function over the entire commodity space. Empirically, however, the effects of using differing sets of budgets on the resulting utility parameters, and hence the robustness of these estimates, remains an open question. Finally, while exhibiting considerable variability in socio-demographics, some variables did not include the range of values one might wish for a large-scale valuation study; thus, our conclusions may be subject to differing sampling frames.

Subject to these caveats, we have shown that the theoretical distinction between rational and irrational demand has empirical significance, which lends support to the extension of behavioral methods into applied disciplines. Further, the study demonstrates the vetting of experimental choice data with nonparametric rationality tests, a general approach which offers a high degree of construct validity vis-à-vis the preference maximization model of choice. Clearly much work remains to forge a tool for applied economic analysis; yet, further research may eventually yield a preference-based valuation method that is similar in spirit to statistical analysis based on moment generating functions.

References

Afriat, S. (1967). The Construction of Utility Functions from Expediture Data, *International Economic Review* **8**, 67–77.

Andreoni, J. and J. Miller (2002). Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism, *Econometrica* **70**, 737–753.

Ariely, R., G. Loewenstein and D. Prelec (2003). "Coherent Arbitariness": Stable Demand Curves Without Stable Preferences, *Quarterly Journal of Economics* **118**, 73–105.

Becker, G. (1962). Irrational Behavior and Economic Theory, *Journal of Political Economy* **70**, 1–13.

Bronars, S. (1987). The Power of Nonparametric Tests of Preference Maximization, *Econometrica* **55**, 693–698.

Camerer, C. and G. Loewenstein (2004). Behavioral Economics: Past, Present and Future, in Advances in Behavioral Economics, eds. C. Camerer, G. Loewenstein and M. Rabin, Princeton University Press, New York.

Cameron, A. and P. Trivedi (1986). Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators, *Journal of Applied Econometrics* **1**, 29–53.

Conover, W. (1980). Practical Nonparametric Statistics, 2nd edition, J. Wiley and Sons, New York.

Diamond, P., J. Hausman, D. Leonard and M. Denning (1993). Does the Contingent Valuation Measure Preferences? Experimental Evidence, in Contingent Valuation: A Critical Assessment, ed. J. Hausman, Elsevier North Holland, Amsterdam.

Diewert, E. (1973). Afriat and Revealed Preference Theory, *Review of Economic Studies* **40**, 419–425.

Dodrill, C. (1981). An Economical Method for the Evaluation of General Intelligence in Adults, *Journal of Consulting and Clinical Psychology* **49**, 668–673.

Greene, W. (1998). LIMDEP Version 7.0 User's Manual, Econometric Software, Inc., Plainview, NY.

Harrison, G., R. Harstad and E. Rutström (2004). Experimental Methods and Elicitation of Values, *Experimental Economics* **7**, 123–140.

Hey, J. and J. Lee (2005). Do Subjects Remember the Past?, *Applied Economics* **37**, 9–18.

Hey, J. and C. Orme (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data, *Econometrica* **62**, 1291–1326.

Ho, T., N. Lim and C. Camerer (2006). Modeling the Psychology of Consumer and Firm Behavior with Behavioral Economics, *Journal of Marketing Research* **43**, 307–331.

Houthakker, H. (1950). Revealed Preference and the Utility Function, *Econometrica* **17**, 159–174.

Just, R., D. Heth and A. Schmitz (1982). Applied Welfare Economics and Public Policy, Prentice Hall, Englewood Cliffs, NJ.

Kim, T. and M. Richter (1986). Nontransitive-Nontotal Consumer Theory, *Journal of Economic Theory* **38**, 324–363.

Matzkin, R. and M. Richter (1991). Testing Strictly Concave Rationality, *Journal of Economic Theory* **53**, 287–303.

Richter, M. (1966). Revealed Preference Theory, *Econometrica* **34**, 635–645.

Richter, M. (1987). Revealed Preference Theory, in The New Palgrave: A Dictionary of Economics, eds. J. Eatwell, M. Milgate and P. Newman, MacMillan Press, London.

Rose, H. (1958). Consistency of Preference: The Two-Commodity Case, *Review of Economic Studies* **25**, 124–125.

Sippel, R. (1997). An Experiment on the Pure Theory of Consumer's Behavior, *Economic Journal* **107**, 1431–1444.

Smith, V. (1982). Microeconomic Systems as an Experimental Science, *American Economic Review* **72**, 923–954.

Varian, H. (1982). The Nonparametric Approach to Demand Analysis, *Econometrica* **50**, 945–973.

Vuong, Q. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, *Econometrica* **57**, 307–334.

Wonderlic Personnel Test User's Manual (1992). Wonderlic Personnel Test Inc., Libertyville, IL.

TABLE I. Budget Scenarios

Budget Scenario Number[a]	Endowment (in tokens)	Dollar Equivalent of Endowment[b]	Price Per Unit (in tokens)	Maximum Affordable Units
1	3.2	0.32	0.2	16
2	3.2	0.32	0.4	8
3	4	0.40	0.2	20
4	4.8	0.48	0.24	20
5	4.8	0.48	0.4	12
6	4.8	0.48	0.6	8
7	5.6	0.56	0.6	9.33
8	6.4	0.64	0.8	8
9	8	0.80	0.4	20
10	8	0.80	1	8

[a]For reference purposes only. Does not refer to the sequencing of scenarios. A random draw determined the sequence of budgets each subject faced.

[b]At 10 tokens per dollar.

TABLE II. Descriptive Statistics

a. Dummy Variable Frequencies[a]

Variable	Number of Observations Coded "0"	Percentage of Observations Coded "0"	Number of Observations Coded "1"	Percentage of Observations Coded "1"
ETHN	30	43.5	39	56.5
GENDER	34	49.3	35	50.7
INT	55	79.7	14	20.3
MBA	36	52.2	33	47.8
MRRD	48	69.6	21	30.4
WARP	32	46.4	37	53.6

b. Summary Statistics for Integer-Valued and Continuous Variables[a]

Variable	Mean	Standard Deviation	Minimum Observed Value	Maximum Observed Value
AACOG	23.84	7.11	8	43
AGE	27.07	5.49	19	45
HHINC	47,101.45	31,076.35	5,000.00	100,000.00
HK	5.26	5.04	0	20
TIME	8.54	3.18	1.83	15.00

[a]All statistics reported are based on the sample of 69 observations except HK, where the reported statistics summarize the 690 choices of our 69 subjects.

Table III. MLE Negative Binomial and Zero Adjusted Negative Binomial Regression Estimates. Dependent Variable: HK.

Variable	WARP Consistent ZANB	WARP Inconsistent		Pooled ZANB
		ZANB	NB	
Constant	1.411+++ (0.311)	1.987+++ (0.338)	1.993+++ (0.387)	1.717+++ (0.218)
PRICE	-1.992*** (0.213)	-1.722*** (0.209)	-1.859*** (0.240)	-1.868*** (0.146)
ENDOW	0.099*** (0.032)	0.166*** (0.032)	0.178*** (0.037)	0.129*** (0.022)
AACOG	-0.001 (0.010)	-0.006 (0.011)	-0.007 (0.013)	-0.005 (0.007)
AGE	0.030+++ (0.011)	0.004 (0.012)	0.006 (0.014)	0.014++ (0.007)
ETHN	-0.142 (0.093)	-0.739+++ (0.146)	-0.849+++ (0.165)	-0.359+++ (0.071)
MBA	-0.143 (0.132)	-0.119 (0.224)	-0.154 (0.254)	-0.021 (0.100)
HHINC	-0.000++ (0.000)	-0.000+++ (0.000)	-0.000+++ (0.000)	-0.000+++ (0.000)
GENDER	0.104 (0.101)	0.141 (0.097)	0.145 (0.112)	0.136++ (0.064)
MRRD	0.099 (0.133)	0.168 (0.143)	0.186 (0.165)	0.062 (0.086)
INT	0.425+++ (0.134)	0.497+++ (0.131)	0.536+++ (0.150)	0.484+++ (0.092)
PER1	0.107 (0.167)	0.266 (0.196)	0.320 (0.220)	0.193 (0.126)
PER2	0.153 (0.161)	0.446++ (0.187)	0.503++ (0.212)	0.275++ (0.122)
PER3	0.045 (0.167)	0.186 (0.185)	0.227 (0.210)	0.166 (0.123)
PER4	0.042 (0.171)	0.230 (0.183)	0.282 (0.205)	0.144 (0.117)
PER5	0.246 (0.153)	0.302 (0.195)	0.349 (0.220)	0.302++ (0.122)
PER6	0.094 (0.171)	0.209 (0.184)	0.246 (0.207)	0.156 (0.124)
PER7	-0.066 (0.156)	0.205 (0.179)	0.230 (0.201)	0.092 (0.118)
PER8	-0.007 (0.171)	0.193 (0.177)	0.227 (0.199)	0.121 (0.120)
PER9	-0.006 (0.168)	0.360++ (0.174)	0.410++ (0.197)	0.187 (0.120)
WARP	-----	-----	-----	-0.060 (0.063)
α	0.249+++ (0.043)	0.262+++ (0.056)	0.345+++ (0.054)	0.282+++ (0.036)
τ	-1.001+++ (0.105)	-1.955+++ (0.353)	-----	-1.284+++ (0.102)
Vuong statistic	4.102***	1.562*	-----	4.007***
LnL	-934.894	-811.170	-814.829	-1769.812

*One-sided test significant at the 0.10 level; **One-sided test significant at the 0.05 level; ***One-sided test significant at the 0.01 level.
+Two-sided test significant at the 0.10 level; ++Two-sided test significant at the 0.05 level; +++Two-sided test significant at the 0.01 level.

Table IV. WARP-Consistent/WARP-Inconsistent Subsample Wald Tests

WARP-Consistent Model Specification	WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (*p*-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	34.25 (22)	Reject null at .05 level (0.0463)
ZANB	ZANB	Yes	Yes	No	30.19 (21)	Reject null at 0.10 level (0.0882)
ZANB	ZANB	Yes	No	No	30.19 (20)	Reject null at 0.10 level (0.0669)
ZANB	ZANB	No	Yes	Yes	34.25 (21)	Reject null at 0.05 level (0.0341)
ZANB	ZANB	No	Yes	No	29.85 (20)	Reject null at 0.10 level (0.0723)
ZANB	ZANB	No	No	No	29.85 (19)	Reject null at 0.10 level (0.0537)
ZANB	NB	Yes	Yes	N/A	36.76 (21)	Reject null at 0.05 level (0.0179)
ZANB	NB	Yes	No	N/A	34.00 (20)	Reject null at 0.05 level (0.0261)
ZANB	NB	No	Yes	N/A	33.29 (20)	Reject null at 0.05 level (0.0313)
ZANB	NB	No	No	N/A	31.06 (19)	Reject null at 0.05 level (0.0398)

TABLE V. Binomial Probit Maximum Likelihood Estimates.[a, b]
Dependent Variable: WARP

Independent Variable	Index Function			Variable Mean	Marginal Effect[c]
	Estimated Coefficient (Standard Error)	β/s.e.	P[|Z|>z]		
Constant	0.144 (0.919)	0.156	0.876		
AACOG	0.027 (0.031)	0.877	0.380	23.841	0.010
GENDER	0.648 (0.336)	1.931+	0.054	0.507	0.241
MBA	-0.484 (0.436)	-1.111	0.266	0.478	-0.165
TIME	-0.092 (0.058)	-1.593*	0.111	8.539	-0.033
Log-likelihood[d]	-43.407* (0.076)				
Pseudo-R^{2e}	0.089				

[a]Standard errors of estimated coefficients shown in parentheses. Number of observations is 69.

[b]*One-sided test significant at the 0.10 level; +two-sided test significant at the 0.10 level.

[c]We compute the marginal effect of a dummy by finding the predicted probability for each observation when the dummy is satisfied (coded 1) and when it is not (coded 0); then, we calculate the difference in predicted probabilities for each observation and report the sample average. For continuous or integer-level variables, we evaluate the probability derivative using each observation and report the sample average.

[d]Number in parenthesis in this section is the p-value of the test of the null hypothesis that all estimated coefficients except the intercept's are jointly equal to zero.

[e]1-(LnL/LnL_0) where LnL is the natural log of the unrestricted likelihood function, and LnL_0 is the natural log of the likelihood function when all coefficients except the intercept's are restricted to zero.

Table VI. MLE Negative Binomial and Zero Adjusted Negative Binomial Regression Estimates. Dependent Variable: HK.

Variable	Predicted WARP Consistent ZANB	Predicted WARP Inconsistent		Pooled ZANB
		ZANB	NB	
Constant	1.833+++ (0.308)	3.555+++ (0.520)	3.683+++ (0.593)	1.723+++ (0.222)
PRICE	-1.919*** (0.185)	-1.776*** (0.266)	-1.922*** (0.283)	-1.863*** (0.146)
ENDOW	0.120*** (0.027)	0.147*** (0.042)	0.152*** (0.045)	0.130*** (0.022)
AACOG	-0.008 (0.008)	-0.026 (0.017)	-0.033+ (0.018)	-0.007 (0.007)
AGE	0.018++ (0.008)	-0.058++ (0.024)	-0.055++ (0.027)	0.014++ (0.007)
ETHN	-0.185++ (0.076)	-0.761+++ (0.175)	-0.861+++ (0.190)	-0.357+++ (0.071)
MBA	-0.010 (0.110)	0.342 (0.280)	0.398 (0.313)	-0.005 (0.107)
HHINC	-0.000+++ (0.000)	-0.000+ (0.000)	-0.000++ (0.000)	-0.000+++ (0.000)
GENDER	-0.010 (0.135)	-0.191 (0.252)	-0.276 (0.270)	0.104++ (0.108)
MRRD	0.079 (0.092)	0.343+ (0.197)	0.380+ (0.223)	0.075 (0.085)
INT	0.257+ (0.134)	1.109+++ (0.215)	1.169+++ (0.232)	0.487+++ (0.095)
PER1	0.122 (0.155)	0.388+ (0.221)	0.427+ (0.245)	0.194 (0.127)
PER2	0.074 (0.156)	0.576+++ (0.223)	0.639++ (0.250)	0.278++ (0.123)
PER3	0.113 (0.151)	0.233 (0.222)	0.277 (0.252)	0.166 (0.123)
PER4	0.087 (0.158)	0.293 (0.213)	0.332 (0.233)	0.145 (0.117)
PER5	0.267+ (0.149)	0.322 (0.219)	0.392 (0.247)	0.301++ (0.123)
PER6	0.131 (0.161)	0.221 (0.210)	0.260 (0.234)	0.156 (0.125)
PER7	0.047 (0.138)	0.149 (0.215)	0.192 (0.240)	0.094 (0.117)
PER8	0.111 (0.150)	0.151 (0.201)	0.203 (0.223)	0.122 (0.120)
PER9	0.093 (0.145)	0.297 (0.225)	0.344 (0.252)	0.189 (0.119)
PREDICT	-----	-----	-----	0.026 (0.115)
α	0.204+++ (0.033)	0.392+++ (0.087)	0.425+++ (0.072)	0.279+++ (0.036)
τ	-1.048+++ (0.099)	-3.144++ (1.511)	-----	-1.270+++ (0.100)
Vuong statistic	4.432***	0.6113	-----	4.036***
LnL	-1007.602	-732.369	-733.423	-1770.321

*One-sided test significant at the 0.10 level; **One-sided test significant at the 0.05 level; ***One-sided test significant at the 0.01 level.
+Two-sided test significant at the 0.10 level; ++Two-sided test significant at the 0.05 level; +++Two-sided test significant at the 0.01 level.

Table VII. Predicted WARP-Consistent/Predicted WARP-Inconsistent Subsample Wald Tests

Predicted WARP-Consistent Model Specification	Predicted WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (*p*-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	49.73 (22)	Reject null at 0.01 level (0.0006)
ZANB	ZANB	Yes	Yes	No	45.64 (21)	Reject null at 0.01 level (0.0014)
ZANB	ZANB	Yes	No	No	45.12 (20)	Reject null at 0.01 level (0.0011)
ZANB	ZANB	No	Yes	Yes	43.78 (21)	Reject null at 0.01 level (0.0025)
ZANB	ZANB	No	Yes	No	41.57 (20)	Reject null at 0.01 level (0.0031)
ZANB	ZANB	No	No	No	41.24 (19)	Reject null at 0.01 level (0.0022)
ZANB	NB	Yes	Yes	N/A	60.36 (21)	Reject null at 0.01 level (0.0000)
ZANB	NB	Yes	No	N/A	52.40 (20)	Reject null at 0.01 level (0.0001)
ZANB	NB	No	Yes	N/A	51.44 (20)	Reject null at 0.01 level (0.0001)
ZANB	NB	No	No	N/A	45.69 (19)	Reject null at 0.01 level (0.0005)

Table VIII. WARP-Consistent/Predicted WARP-Consistent Subsample Wald Tests

WARP-Consistent Model Specification	Predicted WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (p-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	7.10 (22)	Fail to Reject Null at 0.10 Level (0.9989)
ZANB	ZANB	Yes	Yes	No	6.81 (21)	Fail to Reject Null at 0.10 Level (0.9985)
ZANB	ZANB	Yes	No	No	6.32 (20)	Fail to Reject Null at 0.10 Level (0.9984)
ZANB	ZANB	No	Yes	Yes	5.85 (21)	Fail to Reject Null at 0.10 Level (0.9995)
ZANB	ZANB	No	Yes	No	5.81 (20)	Fail to Reject Null at 0.10 Level (0.9991)
ZANB	ZANB	No	No	No	5.26 (19)	Fail to Reject Null at 0.10 Level (0.9992)

Table IX. WARP-Consistent/Predicted WARP-Inconsistent Subsample Wald Tests

WARP-Consistent Model Specification	Predicted WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (p-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	45.35 (22)	Reject null at 0.01 level (0.0024)
ZANB	ZANB	Yes	Yes	No	43.52 (21)	Reject null at 0.01 level (0.0027)
ZANB	ZANB	Yes	No	No	43.32 (20)	Reject null at 0.01 level (0.0019)
ZANB	ZANB	No	Yes	Yes	42.66 (21)	Reject null at 0.01 level (0.0035)
ZANB	ZANB	No	Yes	No	41.92 (20)	Reject null at 0.01 level (0.0028)
ZANB	ZANB	No	No	No	41.72 (19)	Reject null at 0.01 level (0.0019)
ZANB	NB	Yes	Yes	N/A	53.74 (21)	Reject null at 0.01 level (0.0001)
ZANB	NB	Yes	No	N/A	48.68 (20)	Reject null at 0.01 level (0.0003)
ZANB	NB	No	Yes	N/A	47.81 (20)	Reject null at 0.01 level (0.0005)
ZANB	NB	No	No	N/A	44.15 (19)	Reject null at 0.01 level (0.0009)

Table X. Predicted WARP-Consistent/WARP-Inconsistent Subsample Wald Tests

Predicted WARP-Consistent Model Specification	WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (p-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	29.42 (22)	Fail to reject null at 0.10 (0.1333)
ZANB	ZANB	Yes	Yes	No	27.83 (21)	Fail to reject null at 0.10 (0.1451)
ZANB	ZANB	Yes	No	No	27.54 (20)	Fail to reject null at 0.10 (0.1207)
ZANB	ZANB	No	Yes	Yes	28.81 (21)	Fail to reject null at 0.10 (0.1186)
ZANB	ZANB	No	Yes	No	26.56 (20)	Fail to reject null at 0.10 (0.1482)
ZANB	ZANB	No	No	No	26.30 (19)	Fail to reject null at 0.10 (0.1221)
ZANB	NB	Yes	Yes	N/A	40.38 (21)	Reject null at 0.01 level (0.0067)
ZANB	NB	Yes	No	N/A	33.59 (20)	Reject null at 0.05 level (0.0290)
ZANB	NB	No	Yes	N/A	34.70 (20)	Reject null at 0.05 level (0.0218)
ZANB	NB	No	No	N/A	29.14 (19)	Reject null at 0.10 level (0.0638)

Table XI. WARP-Inconsistent/Predicted WARP-Inconsistent Subsample Wald Tests

WARP-Inconsistent Model Specification	Predicted WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (p-value)
		Intercept	α	τ		
ZANB	ZANB	Yes	Yes	Yes	14.13 (22)	Fail to reject null at 0.10 level (0.8969)
ZANB	ZANB	Yes	Yes	No	13.17 (21)	Fail to eject null at 0.10 level (0.9026)
ZANB	ZANB	Yes	No	No	12.37 (20)	Fail to reject null at 0.10 level (0.9029)
ZANB	ZANB	No	Yes	Yes	13.13 (21)	Fail to reject null at 0.10 level (0.9038)
ZANB	ZANB	No	Yes	No	12.72 (20)	Fail to reject null at 0.10 level (0.8892)
ZANB	ZANB	No	No	No	11.85 (19)	Fail to reject null at 0.10 level (0.8918)
ZANB	NB	Yes	Yes	N/A	18.87 (21)	Fail to reject null at 0.10 level (0.5936)
ZANB	NB	Yes	No	N/A	15.51 (20)	Fail to reject null at 0.10 level (0.7465)
ZANB	NB	No	Yes	N/A	16.59 (20)	Fail to reject null at 0.10 level (0.6796)
ZANB	NB	No	No	N/A	13.68 (19)	Fail to reject null at 0.10 level (0.8023)

Table XI, continued. WARP-Inconsistent/Predicted WARP-Inconsistent Subsample Wald Tests

WARP-Inconsistent Model Specification	Predicted WARP-Inconsistent Model Specification	Ancillary Variables Included			Wald Statistic (d.f.)	Conclusion (p-value)
		Intercept	α	τ		
NB	ZANB	Yes	Yes	N/A	10.44 (21)	Fail to reject null at 0.10 level (0.9726)
NB	ZANB	Yes	No	N/A	10.29 (20)	Fail to reject null at 0.10 level (0.9626)
NB	ZANB	No	Yes	N/A	10.44 (20)	Fail to reject null at 0.10 level (0.9595)
NB	ZANB	No	No	N/A	10.29 (19)	Fail to reject null at 0.10 level (0.9455)
NB	NB	Yes	Yes	N/A	12.24 (21)	Fail to reject null at 0.10 level (0.9331)
NB	NB	Yes	No	N/A	10.92 (20)	Fail to reject null at 0.10 level (0.9482)
NB	NB	No	Yes	N/A	11.59 (20)	Fail to reject null at 0.10 level (0.9295)
NB	NB	No	No	N/A	10.56 (19)	Fail to reject null at 0.10 level (0.9377)